

Introduction	15
CHAPITRE 1	
LA RECONNAISSANCE DES OBJETS SUR LA BASE D'INDICES BIDIMENSIONNELS OU TRIDIMENSIONNELS	19
A. APPROCHES THÉORIQUES SUR LA RECONNAISSANCE VISUELLE DES OBJETS	20
 1. Perception de la forme et des objets	20
1.1. La distinction figure-fond.....	20
1.2. L'organisation perceptive.....	21
1.3. Modèles théoriques sur la reconnaissance des objets.....	22
 2. Rôle des informations de surface sur la reconnaissance des objets.....	26
2.1. Place fonctionnelle des informations de surface dans les modèles théoriques	26
2.2. Rôle fonctionnel de la couleur pour une reconnaissance d'objet.....	27
2.3. La texture.....	28
 3. Conclusion.....	38
B. LES INDICES NECESSAIRES À LA PERCEPTION VISUELLE EN 3D	39
 1. Les indices binoculaires et monoculaires dynamiques.....	40
1.1. La disparité	40
1.2. Les indices oculomoteurs	42
1.3. Les indices de mouvement	43
 2. Les indices picturaux	44
2.1. La perspective	46
2.2. L'interposition.....	51
2.3. Les ombres et ombrages	53
 3. La combinaison d'indices	63
C. PRÉSENTATION DU CORTEX VISUEL DU PRIMATE ET DE L'HOMME.....	64
 1. Introduction	64
 2. De l'œil au cerveau	65
2.1. La rétine et le corps genouillé latéral	65
2.2. Le cortex visuel primaire (V1)	66
2.3. L'aire V2	68
 3. Les deux voies visuelles	68
 4. L'aire V4	70
4.2. Rétinotopie	73
4.3. Les connexions corticales.....	75
 5. Le cortex inférotemporal	80
5.1. Position anatomique	80
5.2. Connexions corticales	80
5.3. Connexions sous-corticales.....	81
5.4. Organisation rétinotopique	81

1.1.	Les variables de 1 ^{er} ordre	122
1.2.	Les variables de 2 nd ordre	127
1.3.	Calcul des paramètres selon le type de stimuli	129
2.	Contenu fréquentiel des images	128
2.1.	Méthode de Van der Schaaf et van Hateren	128
2.2.	Les filtres linéaires de type GABOR.....	129
E.	ANALYSE DES DONNÉES	131
1.	Décharges neuronales	131
1.1.	Cellules individuelles	131
1.2.	Les analyses de population.....	131
2.	Mesures de latences.....	132
2.1.	Méthode améliorée de Maunsell et Gibson (1992)	132
2.2.	Calcul des dérivés secondes	134
2.3.	Spike Train Analysis (STA)	135

CHAPITRE 3

BASES NEURONALES DU CODAGE DES TEXTURES NATURELLES DANS L'AIRE V4 DU SINGE ÉVEILLÉ

A.	INTRODUCTION ET PRESENTATION DE L'ETUDE.....	138
B.	RÉTINOTOPIE DES AIRES V4 ET POPULATIONS NEURONALES	140
1.	Trajectoires des électrodes	140
2.	Champs récepteurs et rétinotopie.....	141
3.	Activité moyenne	143
4.	Populations d'étude.....	144
C.	SÉLECTIVITÉ AUX TEXTURES NATURELLES - SET A : 12 TEXTURES, 2 PHASES ET 3 ANGLES D'INCIDENCE	146
1.	Base de données	146
2.	Mise en évidence d'une sélectivité aux textures naturelles.....	146
2.1.	Exemples de cellules et critère de sélectivité	146
2.2.	La sélectivité est préservée si on déplace le stimulus dans le CR du neurone..	150
2.3.	Synthèse sur la sélectivité des neurones TS aux textures naturelles	151
3.	Influence des paramètres statistiques des images	152
3.1.	Paramètres de 1 ^{er} ordre : moyenne, variance, skewness et kurtosis de la luminance, RMS, et SNR	152
3.2.	Paramètres de 2 nd ordre : Energie, entropie, inertie et homogénéité.....	154
4.	Analyses multidimensionnelles sur les réponses de la population TS	155
4.1.	Analyse hiérarchisée par clusters	155

4.2. Analyse multidimensionnelle de type MDS.....	158
5. Influence de la phase de l'image	162
5.1. Préférence des cellules TS pour les stimuli originaux	162
5.2. Histogramme de population et analyse par rang	166
5.3. Synthèse	168
6. Influence de l'angle d'incidence de la source d'illumination	169
6.1. Sélectivité à l'angle d'incidence	169
6.2. Index de sélectivité à l'angle d'incidence (DMI).....	170
6.3. Synthèse sur l'influence de la direction d'illumination	175
7. Approche computationnelle pour une classification de textures à l'aide d'une pyramide de filtres GABOR.....	172
7.1. Energie de convolution des filtres	172
7.2. Analyses multidimensionnelles de l'énergie de la pyramide de filtres	172
7.3. Comparaison entre les réponses de la population TS et l'énergie de la pyramide de filtres GABOR.....	177
7.4. Synthèse	178
8. Conclusion générale	179
D. ETUDE SYSTÉMATIQUE DE L'INFLUENCE DU SNR ET DE LA LUMINANCE MOYENNE - SET B : 65 TEXTURES NATURELLES.....	180
1. Population d'étude	180
2. Cellules TS	180
3. Particularités de ce nouveau set de stimuli	181
4. Analyses de population	182
4.1. Analyse hiérarchisée par cluster.....	182
4.2. Analyse multidimensionnelle (MDS).....	186
5. Comparaisons entre les réponses de la population TS et les énergies de convolution des filtres GABOR.....	188
5.1. Test de Procrustes.....	188
5.2. Factorisation non négative (NNF).....	189
6. Synthèse des analyses sur le set B	189
E. ETUDE SYSTÉMATIQUE DE LA SELECTIVITÉ AUX TEXTURES NATURELLES : SET C : 41 TEXTURES NATURELLES, 2 ANGLES D'INCIDENCE, 2 STIMULI CONTRÔLES.....	190
1. Population d'étude	190
2. Analyses de population	191
2.1. Analyse hiérarchisée par clusters	191
2.2. Analyse multidimensionnelle de type MDS.....	193
3. Comparaison entre les réponses de la population TS et les sorties des filtres GABOR	196

4. Influence de la phase	197
4.1. Résultats du test ANOVA	197
4.2. Analyse par rang.....	198
5. Un nouveau contrôle : les réseaux orientés.....	199
5.1. Population d'étude.....	199
5.2. Analyse hiérarchisée par cluster.....	200
6. Influence de l'angle d'incidence.....	203
F. DISCUSSION	206
1. Synthèse des principaux résultats	206
2. Localisation des sites d'enregistrements	207
3. Sélectivité aux textures naturelles.....	208
3.1. Sélectivité générale	208
3.2. Paramètres des textures	209
3.3. Dimensions de la classification des textures	210
3.4. Comparaison avec les études computationnelles	211
4. Invariance à la direction d'illumination.....	212
5. Conclusion.....	213

CHAPITRE 4

CODAGE DES OMBRAGES DANS L'AIRE V4 DU SINGE VIGILE**215**

A. MISE EN ÉVIDENCE D'UN CODAGE DE LA PROFONDEUR SUR LA BASE D'OMBRA	GES AU SEIN DE L'AIRE V4	216
1. Base de données	216	
2. Etude de la sélectivité 3D	216	
2.1. Résultats du test ANOVA	216	
2.2. Index de dissociation.....	219	
3. Réponses des cellules ICS aux stimuli random.....	221	
3.1. Analyse par rang.....	221	
3.2. Analyse hiérarchisée par clusters	223	
3.3. MDS	224	
3.4. Réponses individuelles des cellules 3D aux stimuli random.....	227	
4. Réponses des cellules ICS aux stimuli postérés.....	228	
4.1. Population de base.....	228	
4.2. Analyses par rang.....	228	
4.3. Analyse hiérarchisée par cluster.....	229	
4.4. MDS	232	
4.5. Réponses individuelles des cellules 3D aux stimuli postérés.....	234	

5. Synthèse des principaux résultats.....	236
B. SÉLECTIVITÉ AUX FORMES TRIDIMENSIONNELLES	237
1. Résultats du test ANOVA	237
2. Index de sélectivité (SMI)	239
3. Synthèse.....	239
C. SÉLECTIVITÉ À LA DIRECTION D'ILLUMINATION.....	241
1. Population d'étude	241
2. Résultats ANOVA	241
3. Analyse de population.....	244
D. DISCUSSION.....	246
1. Synthèse des principaux résultats.....	246
2. Mise en évidence d'un codage différentiel pour les stimuli 3D	247
2.1. Sélectivité générale	247
2.2. Codage des ombrages comme indice tridimensionnel	250
3. Sélectivité à la forme 3D présentée	252
4. Reconnaissance des objets et direction d'illumination.	252
CHAPITRE 5	
SHORT VISUAL LATENCIES IN AREA V4 OF THE MACAQUE MONKEY	255
A. ABSTRACT.....	256
B. INTRODUCTION.....	257
C. MATERIAL AND METHODS	259
1. Animals.....	259
2. Task and stimuli	259
3. Recording	260
4. Measurements of response latencies.....	260
D. RESULTS	262
E. DISCUSSION.....	264
1. Location of recording sites	264
2. Latency measurement	265
3. Discrepancies in procedures (animals, stimuli, anesthesia).....	265
F. CONCLUSION.....	267
G. ACKNOWLEDGMENTS	268

H. REFERENCES CITED	269
FIGURES	271
TABLE 5-1	272
ANALYSES COMPLEMENTAIRES SUR LES LATENCES DE NEURONES	273
1. Une autre méthode : Spike Train Analysis (STA).....	273
2. Correction de l'erreur aléatoire.....	275
CHAPITRE 6	
SYNTHÈSE DES PRINCIPAUX RÉSULTATS ET DISCUSSION GÉNÉRALE	277
A. PERTINENCE DE L'ETUDE	278
B. CODAGE D'INDICES TRIDIMENSIONNELS AU SEIN DE V4	279
Synthèse générale	279
De nouvelles propriétés fonctionnelles de V4	279
Codage réel de la 3D.....	281
Rôle de la texture pour une reconnaissance d'objet	283
Rapidité des informations au sein de V4.....	287
V4 centre d'intégration ?	285
C. PERSPECTIVES	287
BIBLIOGRAPHIE.....	289
GLOSSAIRE	307
ANNEXES	309

Introduction

Nous vivons dans un monde en trois dimensions dans lequel notre capacité à percevoir la profondeur est fondamentale. Bien que, dans une scène visuelle, tous les objets soient définis en trois dimensions, il est possible de les identifier sur la base d'images en deux dimensions ; les images projetées à nos rétines sont bidimensionnelles. Une des principales interrogations sur notre vision consiste à comprendre comment à partir d'images en deux dimensions, nous pouvons extraire des informations sur la profondeur des objets, et ainsi appréhender cette troisième dimension. Cette interrogation est résumée par Gibson en 1950 dans *Perception of the visual world*, où il écrit que «l'environnement physique est en 3 dimensions ; il est projeté sur la surface rétinienne en deux dimensions ; mais il est quand même perçu en 3 dimensions. Comment la perte de cette troisième dimension est-elle restaurée dans la perception ?».

La présence d'indices binoculaires et monoculaires nous renseigne sur la profondeur présente dans une image réelle ou une œuvre picturale. La contribution de ces indices nous permet de déterminer la structure tridimensionnelle de l'objet afin de le reconnaître, voire de nous aider à sa préhension. Ces indices contiennent de l'information qui, s'ajoutant à l'image 2D projetée sur la rétine, vont nous permettre de percevoir les objets en 3D. Sur une photographie ou un dessin, images 2D, les objets présents nous paraissent définis en trois dimensions. Cette perception tridimensionnelle repose sur la présence d'indices monoculaires statiques dit « picturaux » au sein-même de l'image tels que la perspective, la disposition des objets les uns par rapport aux autres ou encore les conditions d'illuminations.

Au sein de notre cortex visuel et de celui du macaque, il est établi que les aires visuelles situées le long d'un axe occipito-temporal traitent de la reconnaissance des objets. L'aire visuelle V4 fait partie de cette voie ventrale impliquée dans la reconnaissance des objets sur la base d'indices 2D et 3D. Ces attributs intrinsèques sont codés le long de cette voie visuelle et V4 participe activement aux traitements de ces attributs, notamment au niveau de la couleur de l'objet et de sa forme 2D externe (Pasupathy et Connor, 1999; Kusunoki *et al.*, 2006). Plusieurs études ont également démontré que l'aire V4 traite des indices tridimensionnels tels que le gradient de texture ou la disparité (Hanazawa et Komatsu, 2001; Hinkle et Connor, 2001). De toute évidence, cette aire visuelle participe aux traitements de plusieurs attributs de l'objet.

Le travail que j'ai accompli durant ma thèse consiste en l'étude électrophysiologique de la perception d'indices monoculaires statiques. Il est bien établi en psychophysique humaine que des indices tels que la texture et les ombrages permettent d'appréhender la structure tridimensionnelle d'une image en vision monoculaire. Le premier indice auquel je me suis intéressé, la texture d'un objet, peut se définir comme un pattern de micro-ombrages permettant une perception locale de la profondeur cantonnée à la surface de l'objet. Outre la couleur, la texture est également une propriété matérielle essentielle à sa reconnaissance (Price et Humphreys, 1989; Rossion et Pourtois, 2004). Comme le premier indice, le second indice est également fonction des conditions d'illumination. En effet, en condition naturelle, une source lumineuse crée sur toute la surface de l'objet un gradient d'illumination qui définit des macro-ombrages. Ce gradient d'illumination nous permet une perception tridimensionnelle de la forme globale de l'objet grâce aux macro-ombrages. Les bases neuronales de ces indices monoculaires sont peu étudiées par rapport à celles impliquées dans le traitement des indices binoculaires de profondeur comme la stéréopsie. De plus, le processus de reconnaissance d'objets sur la base des indices monoculaires statiques est très mal connu que ce soit au sein de notre système visuel ou celui des primates. Durant cette thèse, j'ai entrepris d'étudier les bases neuronales de deux indices monoculaires statiques ; la texture naturelle et les ombrages. Les enregistrements électrophysiologiques ont été effectués au sein de l'aire visuelle V4 du singe macaque vigile dont le système visuel se rapproche le plus de celui de l'Homme.

La première étude que nous avons réalisée a pour objectif de déterminer les bases neuronales de l'encodage des textures naturelles au sein de l'aire V4. Cette étude démontrera que les neurones de V4 sont influencés par les textures naturelles et déterminera quels paramètres des textures naturelles sont importants pour les neurones de V4 afin d'opérer à une classification des textures indépendamment des conditions d'illumination. Des études computationnelles ont déjà démontré qu'une classification de texture invariante à la direction d'illumination était possible en utilisant comme outil mathématique un ensemble de filtres GABOR.

La seconde étude concerne le traitement des ombrages au sein de l'aire V4. Des études électrophysiologiques, de neuro-anatomie et de neuroimagerie ont démontré que chez le singe et l'Homme, l'aire V4 participe au traitement de la profondeur en encodant la disparité, l'interposition ou le gradient de texture. Nous avons voulu tester si un indice monoculaire de la profondeur tel que les ombrages est également traité dans cette aire V4.

Introduction

Enfin, la troisième étude porte sur les latences de réponses de l'aire V4. Peu d'études se sont intéressées précisément aux latences de réponses des neurones de cette aire chez des animaux vigiles. L'étude qui sert de référence en la matière est celle de l'équipe de Schmolesky (1998) qui a enregistré des latences moyennes chez le singe anesthésié de 70 ms. Au cours de nos enregistrements cellulaires pour les deux précédentes études, nous avons remarqué que les latences de réponses étaient relativement courtes.

Chapitre 1

La reconnaissance des objets sur la base d'indices bidimensionnels ou tridimensionnels

Ce chapitre reprend les données bibliographiques sur la reconnaissance des objets. Il est composé de 4 sous-parties. La première partie décrit les modèles théoriques de la reconnaissance des objets, les indices 2D permettant une identification des objets ainsi que le rôle des informations de surface pour une reconnaissance des objets. La seconde partie présente les principaux indices visuels nécessaires à la perception de la troisième dimension. La troisième partie décrit le cortex visuel du macaque, modèle animal utilisé lors des expérimentations, mais aussi le cortex visuel de l'Homme en s'attachant à définir les homologies entre les structures corticales visuelles. Enfin la quatrième partie détaille le codage neuronal des indices 2D et 3D nécessaires à la reconnaissance de l'objet. Un intérêt particulier sera porté sur la contribution de l'aire V4 dans le codage de ces indices.

A. APPROCHES THEORIQUES SUR LA RECONNAISSANCE VISUELLE DES OBJETS

Les objets se différencient les uns des autres grâce à plusieurs caractéristiques qui leur sont propres comme leur couleur, la composition de leur surface ou encore leur forme. De nombreux travaux ont porté sur la contribution de ces attributs sur la reconnaissance d'un objet dans notre environnement visuel. Dans ce chapitre, je détaillerai les principales théories actuelles sur la reconnaissance des objets. Elles se sont principalement intéressées à la reconnaissance de l'objet à partir de sa forme externe. Le rôle fonctionnel des informations de surface dans la reconnaissance des objets a été peu étudié et s'est principalement focalisé sur l'influence de couleur. Un intérêt particulier dans cette partie sera porté sur une propriété intrinsèque de l'objet autre que la couleur ; la texture dont nous avons entrepris d'étudier les bases neuronales.

1. Perception de la forme et des objets

Le terme « objet » définit une entité dans l'espace tridimensionnel environnant le sujet. L'objet a généralement une fonction précise et peut être reconnu grâce à une étiquette verbale. Quant au terme « forme », il renvoie à des informations structurales perceptibles grâce à un traitement sensoriel visuel. Dans cette première partie, nous traiterons de trois aspects relatifs à la forme et aux objets : la distinction figure-fond, l'organisation perceptive et la reconnaissance d'objets, en présentant les principaux modèles théoriques sur le sujet.

1.1. *La distinction figure-fond*

Pour extraire un objet de la scène visuelle, les informations visuelles doivent se structurer pour nous permettre de discriminer une forme des éléments du fond. Notre perception organise alors le pattern visuel selon une différentiation figure-fond particulière. Ce processus mental est décrit par Edgar Rubin en 1921. Il est fondamental pour toute perception. Les termes « figure » et « fond » sont apparus dans le lexique commun, mais ils sont souvent simplement compris comme différenciant l'objet – la figure – qui se détache des objets qui se situent en arrière plan – le fond –. Pour permettre cette extraction, il nous faut définir les contours de cet objet. Ces contours peuvent être perçus grâce à une différence de couleur, de luminosité, de texture entre l'objet et le fond visuel. Le contour présente également d'autres fonctions que délimiter la figure du fond ; il permet, par exemple, de souligner les détails de la figure voire ses reliefs.

La perception du contour nous permet de différencier une figure du fond visuel mais il nous reste à déterminer à quel élément appartient ce contour. Autrement dit, de quel côté du contour se situe la figure. Plusieurs théories psychophysiques ont été proposées pour expliquer comment s'organise notre perception visuelle.

1.2. L'organisation perceptive

La première étape du système visuel est de détecter la structure générale de la scène en regroupant des indices communs afin de segmenter la scène en plusieurs sous-régions. La première approche théorique pour expliquer ce procédé est une approche globaliste énoncée par les psychologues allemands de la Gestalt¹.

a) Les lois de la Gestalt

Les expériences psychophysiques de ces scientifiques allemands basées sur des configurations de points ont permis de dégager plusieurs principes sur l'organisation perceptive. L'ensemble de ces principes est connu sous le nom de « *loi de prégnance* ».

- Le premier principe est la loi de **proximité** ; les éléments d'une configuration qui sont les plus rapprochés les uns des autres, ont tendance à être perçus comme formant une unité.
- Le principe de **similitude** s'applique aux éléments présentant une même forme, une même couleur ou une même texture (éléments avec une même orientation).
- Le principe de **fermeture** désigne le fait qu'un contour interrompu ait tendance à être perçu comme une figure fermée.
- Pour les éléments en mouvement, le principe du **destin commun** indique que si les éléments définissent un mouvement dans une même direction et à la même vitesse, les éléments sont globalement perçus comme une entité structurale.

D'après Wertheimer (1958), un des psychologues allemands, ces principes élémentaires de la perception ne nécessitent aucun apprentissage. Ces principes nous permettent ainsi de percevoir une forme privilégiée la meilleure possible, à savoir une forme simple, régulière et symétrique.

¹ *Forme* en allemand. La *Gestalttheorie* est une théorie de la perception fondée par trois psychologues allemands : Max Wertheimer, Wolfgang Köhler et Kurt Koffka.

b) L'approche écologique de Gibson (1979)

La perception visuelle d'après Gibson repose sur l'ensemble des informations disponibles de l'espace ambiant qu'il définit comme le réseau optique. D'après cette théorie, les informations sensorielles visuelles sont la structure même du réseau optique tout entier. L'organisation perceptive existe déjà dans l'environnement naturel. Pour lui, la distinction figure-fond se limite uniquement à des images bidimensionnelles (un dessin, par exemple) mais non à des situations réelles dans l'environnement. En effet, l'environnement autour de l'objet définit grâce à sa texture un contexte particulier permettant d'en extraire la taille apparente et la distance de l'objet. Une notion intéressante dans son approche écologique est le rôle primordial du mouvement dans la structuration de la scène visuelle. L'association entre les mouvements de l'espace ambiant et les déplacements du sujet favorise la détection des invariants et permet l'extraction de la structure (Gibson, 1979).

1.3. Modèles théoriques sur la reconnaissance des objets

Une principale interrogation concernant notre capacité à reconnaître les objets a amené les psychologues, les neurophysiologistes et les mathématiciens à proposer plusieurs modèles théoriques sur la reconnaissance des objets. Ces scientifiques ont cherché à déterminer les critères importants pour la reconnaissance des objets. L'un des principaux outils utilisés pour élucider ce problème a été l'intelligence artificielle avec la mise en place d'algorithmes mathématiques permettant de simuler la perception visuelle (Marr, 1982; Hummel et Biederman, 1992). Un modèle parfait mimerait notre système visuel en reconnaissant l'objet malgré ses variations de taille, d'orientation, d'éclairage ou de position auxquelles il peut être confronté dans notre environnement visuel. Dans cette description des modèles théoriques, je ne m'intéresserai qu'à deux principaux types de modèles de la reconnaissance des objets : les modèles structuraux et le modèle d'alignement.

a) Les modèles structuraux

Ce type de modèle est fondé sur deux postulats :

- La forme physique de l'objet est segmentée lors du traitement perceptif sous la forme de primitives volumétriques simples. Ces composantes sont communes à tous les objets et sont conservées en mémoire pour le codage des objets.
- En raison du caractère volumétrique des primitives, les représentations sont invariantes aux modifications paramétriques de la forme comme son orientation, sa taille ou sa localisation spatiale.

Marr et Nishihara (1978) furent les premiers à proposer un modèle de ce type. Ils définissent notre système visuel comme une organisation hiérarchique basée sur 3 niveaux d'intégration (Marr et Nishihara, 1978) :

- Un premier niveau bidimensionnel, l'ébauche primitive. Le système visuel construit un premier niveau de description de l'image en ne tenant compte que des orientations des éléments et des différences de luminance.
- Le second niveau, niveau $2D \frac{1}{2}$, est centré sur l'observateur. C'est un niveau de description de la surface visible des objets quand on regarde l'objet dans une direction donnée. En conséquence, ce niveau varie avec la position de l'observateur par rapport à l'objet et la forme de l'objet est modifiée lorsque l'observateur change de position.
- La représentation 3D centrée sur l'objet. L'objet est défini dans son propre système de coordonnées où sont répertoriées son orientation, la longueur et l'arrangement spatial des composantes de l'objet et la nature géométrique de ces composantes (cylindres, cônes...). Ces composantes sont invariantes à l'orientation.

L'argument principal de ce paradigme est la structure hiérarchique de traitement et de représentation. Marr a posé un cadre de traitement rigoureux qui a inspiré un nombre important de travaux, autant sur la vision bidimensionnelle que tridimensionnelle. Biederman (1987) corrobore cette théorie en y rajoutant des données expérimentales de tests psychophysiologiques chez l'Homme absentes dans le modèle de Marr (Biederman, 1987). Son modèle est basé sur une description structurelle de l'objet. Il représente les objets sous la forme de primitives ou « géons » ; c'est le modèle de « reconnaissance par les composantes » (Recognition By Components, RBC). Biederman décrit 24 types de géons différents. La forme complexe d'un objet peut ainsi se résumer en l'assemblage de 3 au 4 géons.

Sur la base de cette théorie, Hummel et Biederman (1992) proposent un modèle qui constitue une référence en psychologie de la vision. Ce modèle permet une reconnaissance invariante à la position, à la taille ou à la rotation des objets. Un même objet pourra être reconnu selon divers points de vues tant que l'extraction de la même description structurelle de plusieurs vues différentes de l'objet reste possible (Hummel et Biederman, 1992).

L'interrogation selon laquelle notre système visuel utilise une représentation centrée sur l'objet ou dépendante du point de vue a été le sujet de nombreuses études soulevant une controverse (pour revues, voir (Biederman et Gerhardstein, 1993; Tarr et Bulthoff, 1995). Une première théorie considère que la représentation d'un objet est conçue comme un ensemble de

caractéristiques - les invariants - de l'objet indépendant des vues de ce même objet (Biederman, 1987; Wallis et Rolls, 1997). La seconde théorie prône une reconnaissance des objets dépendante du point de vue en se basant sur des résultats de tests psychologiques ou des enregistrements neurophysiologiques (Logothetis et Pauls, 1995; Tarr *et al.*, 1998b). Tarr et son équipe (1998) ont proposé à des sujets humains trois tâches différentes de reconnaissance d'objets (*dénomination*, ‘*match to sample*’, ‘*sequential matching*’) avec des stimuli constitués de géons similaires à ceux utilisés par Biederman. Ils démontrent que la reconnaissance d'objet dépend du point de vue et considèrent donc que la représentation d'un objet est liée à des vues spécifiques à l'objet. La reconnaissance est exprimée comme une fonction des vues spécifiques des images qui ont été déjà présentées (Tarr *et al.*, 1998b; Edelman *et al.*, 1999). Les résultats physiologiques de l'équipe de Logothetis confirment cette hypothèse. Logothetis et ses collaborateurs font apprendre aux singes une série d'images d'objets divers (trombones, objets sphériques) et vont leur présenter ces mêmes objets mais avec une orientation différente. Les singes reconnaissent davantage les objets si le point de vue est identique à celui présenté au moment de l'apprentissage. Cette même équipe enregistrera par la suite des neurones au sein du cortex inférot temporal sensibles au point de vue des objets (Logothetis *et al.*, 1994; Logothetis et Pauls, 1995).

Pour conclure, Foster et Gibson (2002) proposent un modèle structural alternatif en se basant sur les résultats d'expériences de psychophysique chez l'Homme. Ce modèle repose sur 2 principaux arguments : la structure globale de l'objet est invariante au point de vue tandis que la reconnaissance de certaines parties de l'objet est dépendante du point de vue (Foster et Gilson, 2002).

En résumé, les modèles structuraux présentent de nombreux avantages mais aussi quelques imperfections.

Points positifs :

- Ce système rend possible une reconnaissance de l'objet invariante en fonction de la distance, de la position ou de l'orientation de celui-ci. Toutefois, ce point porte à controverse (Tarr et Bulthoff, 1995; Tarr *et al.*, 1998b).
- Un autre point intéressant est le principe même de déstructuration d'un objet en composantes. Ces primitives sont communes à tous les objets et nécessiteraient une mise en mémoire minime.
- La description des métriques des formes et principalement des composantes.

Points négatifs :

- La non prise en compte des informations liées à la surface (texture, couleur ou luminance) qui sont considérées comme secondaires dans la reconnaissance des objets (Biederman et Ju, 1988).
- La décomposition en géons d'une forme complexe tridimensionnelle est beaucoup plus difficile s'il y a présence d'ombres que si l'objet est représenté simplement par des traits noirs.
- L'origine des primitives n'est pas expliquée.
- Un mouvement qui déformerait certaines parties de l'objet par occlusion est ignoré par ce modèle.

b) Le modèle de l'alignement

Les modèles basés sur la décomposition de l'objet ne prennent pas en compte d'importantes informations comme la localisation dans l'image des attributs des objets. Ainsi, des méthodes géométriques comme la théorie de l'alignement proposé par Ullman en 1989 utilisent ce type d'informations pour identifier les objets. Le principe de base de ces modèles repose sur le fait que nous avons en mémoire des représentations de l'objet. La reconnaissance d'un objet se résumerait en un appariement entre l'entrée visuelle de l'objet et une représentation mémorisée (Ullman, 1989).

Le modèle d'Ullman comprend deux composantes essentielles :

- La première concerne la comparaison de l'objet à identifier avec une image mémorisée. Il s'agit de l'**alignement**.
- La seconde implique l'existence d'un traitement bidirectionnel. Selon Ullman, les images d'objets, qu'elles soient situées au niveau de notre entrée visuelle ou qu'elles soient mémorisées peuvent subir des transformations.

Ces deux types de traitement seraient effectués par les voies ascendantes et descendantes du cortex. Une même zone corticale serait le siège de ces traitements lorsqu'il y a appariement entre l'entrée visuelle et l'image mémorisée. Les attributs de l'objet (la forme, la couleur ou la texture) seraient traités par les voies ascendantes et les différentes représentations mémorisées d'un objet par les voies descendantes.

Points positifs et négatifs :

Un élément important dans ce type de modèle est la présence d'indices pertinents pour une reconnaissance d'objets. En effet, le principe d'alignement se résume en la mise en correspondance entre l'entrée visuelle et la représentation en mémoire. Cet appariement peut se réaliser grâce à la présence d'indices pertinents. Par exemple, Ullman (1989) démontre qu'un ajustement est possible si on se base uniquement sur trois points spatialement distincts dans l'image. Les propriétés matérielles de l'objet (couleur ou texture) pourraient constituer un élément déterminant pour une reconnaissance d'objet.

La résolution temporelle reste le principal inconvénient de ce modèle car il ne s'accorde pas avec les résultats démontrant la rapidité d'une catégorisation ou d'une identification d'objets chez l'Homme ou chez l'animal au vu de la vitesse de conduction ou des latences de réponses des neurones des aires visuelles (Fabre-Thorpe *et al.*, 1998a; Fabre-Thorpe *et al.*, 1998b; Fabre-Thorpe *et al.*, 1998c).

2. Rôle des informations de surface sur la reconnaissance des objets

D'après Regan, notre système visuel reconnaît les objets en se basant sur différents attributs comme la forme, la texture, la couleur ou un mouvement particulier. Il utilise tous ces attributs et les combine afin d'obtenir une représentation de l'objet (Regan, 2000). Cependant, les modèles théoriques se sont principalement concentrés à proposer une reconnaissance de l'objet basée sur sa forme externe et ces théories ne s'accordent pas sur la place fonctionnelle qu'elles attribuent aux propriétés matérielles.

2.1. Place fonctionnelle des informations de surface dans les modèles théoriques

Les modèles structuraux considèrent que la représentation de l'objet est uniquement basée sur une description de sa forme et d'autres soulignent que les informations liées à la surface de l'objet participent également à cette représentation (Tarr *et al.*, 1998b; Tanaka *et al.*, 2001b). Selon le modèle de Biederman, les informations de surface ne seraient que secondaires. La théorie de Marr propose que les propriétés de surface d'un objet comme ses variations de luminosité ou sa texture soient importantes pour la mise en place de l'ébauche primitive et la construction de la représentation 2D^{1/2} (Marr et Nishihara, 1978; Marr, 1982; Biederman, 1987). Toutefois, la représentation 3D se base uniquement sur les éléments de la forme de l'objet. Les informations de surface seraient alors *écartées* de la représentation de l'objet. Les modèles structuraux suggèrent donc que les informations de surface telles que la

couleur et la texture ne facilitent pas la reconnaissance d'un objet puisqu'elle est basée uniquement sur la définition des contours de l'objet.

Les modèles d'alignement proposent que l'encodage des objets soit traité selon le point de vue de l'observateur. D'après ce type de modèle, les représentations de l'objet ne contiennent pas uniquement comme information la forme mais aussi d'autres indices comme les propriétés de surface. Toutefois, aucun de ces modèles n'explique le rôle de ce type d'information dans le traitement des objets. Enfin, le débat autour du rôle fonctionnel des informations de surface dans la reconnaissance des objets s'est principalement focalisé sur l'influence de la couleur.

2.2. *Rôle fonctionnel de la couleur pour une reconnaissance d'objet*

Le système visuel du macaque rhésus est très proche de celui de l'Homme pour le traitement de la couleur. Tout comme l'Homme, sa rétine comporte 3 types de cônes et si l'on excepte des différences mises récemment en évidence dans la sensibilité spectrale (Dobkins *et al.*, 2000), les résultats des expériences chez l'humain ou chez le singe peuvent être regroupés.

Une controverse existe et soulève plusieurs interrogations sur le rôle fonctionnel de la couleur: la couleur est-elle intégrée dans la représentation perceptuelle de l'objet ou est-elle traitée séparément ou seulement lorsque la reconnaissance de l'objet est ambiguë ? Par exemple, elle pourrait être traitée comme une information sémantique complémentaire ; un fruit rouge en forme de triangle représente une fraise parce que la fraise est rouge. A l'aide de tests psychophysiques, Price et Humphreys (1989) ont demandé à des sujets de nommer des objets qui leur étaient présentés. Ils ont remarqué que l'ajout d'informations de surfaces congruentes permet une meilleure reconnaissance que si les objets sont seulement définis par leur contour. Cette reconnaissance est plus rapide avec l'ajout de la couleur (Rossion et Pourtois, 2004). Ce type d'observation amène à penser que la couleur ou d'autres informations de surface joue un rôle dans la reconnaissance de l'objet seulement si sa couleur est hautement diagnostique. La couleur jouerait également un rôle primordial dans la reconnaissance de l'objet lorsque sa forme est moins diagnostique (c'est à dire si les objets présentent une forme similaire). C'est le cas par exemple pour les fruits. L'absence de couleur entraîne une augmentation des temps de réaction dans une tâche de catégorisation d'aliments mais aucune modification dans une tâche de détection d'animaux chez le singe et chez l'Homme (Delorme *et al.*, 2000). L'utilisation de la couleur dépend donc de la catégorie de l'objet. Certains objets ont une couleur diagnostique mais ont aussi une forme diagnostique,

dans ce cas, la couleur jouerait un rôle minimal dans la reconnaissance de l'objet (Biederman et Ju, 1988).

Chez l'Homme, des expériences ont démontré l'importance de la couleur dans des tâches de détection et de reconnaissance d'objets, notamment lorsque le sujet doit nommer l'objet présenté (Ostergaard et Davidoff, 1985; Wurm *et al.*, 1993; Boucart et Humphreys, 1997). Toutefois, l'importance de la couleur comme indice pour la reconnaissance des objets porte à controverse. En effet, Biederman et Ju (1988) ont démontré que l'absence de couleurs ne réduisait pas les performances dans une tâche de catégorisation de cibles (Biederman et Ju, 1988).

D'après la revue de Tanaka *et al.* (2001), l'information de couleur intervient dans la reconnaissance d'un objet en trois étapes (Tanaka *et al.*, 2001b). Premièrement, la couleur permet une segmentation de la scène visuelle afin d'en extraire les objets 3D (Regan 2000). Puis la couleur peut aider à la reconnaissance d'un objet en tant qu'élément constitutif de la représentation de l'objet, du moins pour les objets qui ont une couleur diagnostique. Ce point est largement débattu entre les théories structurales et les théories d'alignement. Enfin, la couleur aiderait également à un niveau sémantique hiérarchiquement élevé pour le traitement des objets.

2.3. La texture

a) Généralité

La texture est un attribut des objets très présent dans la scène visuelle, facile à reconnaître mais difficile à définir. En effet, le terme texture est utilisé pour définir plusieurs types de surfaces d'un seul objet mais sert également à définir une surface composée de plusieurs objets. En fait, tout dépend de l'échelle où on se situe. Par exemple, un épi de maïs pris indépendamment dans un champ, est perçu comme un objet entier avec sa propre texture (grains de maïs) alors que l'ensemble des épis de maïs crée une texture qui définit le champ.

La texture est donc partout dans notre environnement, au niveau du sol (une plage ou un pré), si nous levons les yeux vers le ciel (feuillage d'une forêt ou toit d'une maison...) mais également au niveau des objets (un toit, un mur) ou des êtres vivants (des pelages d'animaux comme les zèbres ou tigres).



Figure 1-1 : Photographie en noir et blanc de 3 guépards dans la steppe africaine. La texture permet de percevoir plusieurs strates de la steppe et de distinguer les guépards.

Par exemple, sur cette photographie de guépards (Figure 1-1), la différence de luminance entre la steppe en haut de l'image et le sol où se trouvent les guépards entraîne une séparation au niveau de la perception de ces deux patterns texturés. La perception des guépards par rapport au fond visuel dépend également de la texture du pelage tacheté de ces bêtes. Les contours des animaux se définissent aussi grâce à la texture de leur pelage. Le sol sableux au premier plan présente également une texture différente de la steppe où se trouvent les animaux et est perçu différemment.

b) Définition d'une texture

Nous reconnaissons une texture quand nous la percevons en tant que propriété matérielle de l'objet et dans notre vie quotidienne, la notion de texture est utilisée pour traduire un aspect homogène de la surface d'un objet. Nous caractérisons qualitativement la texture à l'aide de nombreux adjectifs comme granuleuse, fine, grossière, rugueuse, lisse, tachetée, marbrée, régulière ou irrégulière (Bhushan *et al.*, 1997) mais il nous est difficile de la paramétriser (comme par exemple la couleur d'un objet avec sa longueur d'onde). La difficulté pour paramétriser une texture est telle qu'il existe un nombre important de définitions différentes (voir revue de (Coggins, 1982)).

Cette multitude de définitions montre que le terme « texture » n'a pas de définition unique mais que celle-ci dépend de son contexte d'utilisation. Une définition assez générale serait de caractériser une texture comme un ensemble de primitives disposées de façon régulière dans une surface. Le terme de « primitive » renvoie à un pattern plus ou moins important de pixels de niveaux de gris similaires dans l'image. Il s'agit en fait du motif de base de la texture, défini par le terme «**texton**» par Bela Julesz. Il fut le premier à utiliser ce

terme au cours d'une étude sur les paramètres statistiques d'une texture (Julesz, 1981) (voir paragraphe A.2.3.d).

Haralick propose une définition élargie en décrivant une texture comme un phénomène à deux dimensions: la première dimension concerne les **primitives** à partir desquelles est formée la texture et la seconde est fonction de l'**organisation spatiale** des primitives (Haralick, 1979). Toutefois, une autre approche permettrait de définir la texture à partir de deux types d'informations essentielles contenues dans l'image: les **contours** qui marquent les frontières entre les régions homogènes et l'**aspect de surface**. Mais, la description d'une texture peut changer avec l'échelle d'observation, c'est à dire en changeant la résolution et/ou la distance d'observation.

Toutefois, ces définitions s'accordent sur le fait qu'on distingue deux grands groupes de textures:

- Les **macrotextures** ou textures périodiques qui présentent un aspect régulier, sous formes de motifs répétitifs spatialement disposés de manière régulière et ordonnée (ex: peau de lézard, mur de brique) ;
- Les **microtextures** présentant des primitives "microscopiques" distribuées de manière aléatoire (ex: sable, laine tissée, herbe).

Plusieurs bases de données répertorient des textures naturelles : Brodatz a mis en place un catalogue de textures largement utilisées pour valider les différentes méthodes d'analyse de texture (Brodatz, 1966). Un ensemble de textures naturelles (microtextures) a été diffusé plus récemment sur Internet que nous avons utilisé au cours de nos expériences (CUREt database, <http://www1.cs.columbia.edu/CAVE/////////software/curet/index.php> (Dana *et al.*, 1999).

Ces bases de données ont permis à plusieurs équipes de recherche s'intéressant à la vision artificielle de tester leurs outils mathématiques sur des analyses d'images.

c) L'analyse computationnelle des textures

(1) Les 4 principaux axes de recherche

La vision artificielle d'images de textures se découpe en quatre principaux axes de recherche que sont la classification, la segmentation, la synthèse de texture et l'extraction de formes à partir des textures.

Le premier axe concerne le fait de segmenter les différentes régions de l'image sur la base de la texture. Notre système visuel a la faculté d'extraire généralement sans difficulté les bordures produites par les textures. Par exemple sur la Figure 1-2A, il est possible d'identifier 4 régions comportant des textures différentes le béton, le plâtre, le papier froissé et le papier de verre. La texture est ici l'indice visuel le plus important pour séparer les régions différentes, donc les segmenter. Ce processus d'identification est appelé **segmentation de textures**. La finalité de ce processus est de produire une carte où chaque texture différente est répertoriée (Figure 1-2B). Une part importante des études computationnelles vise à trouver des algorithmes de **segmentation de texture** visant à identifier les différentes parties de l'image où la texture reste constante (Figure 1-2B).

La Figure 1-3 illustre le principe d'une **classification de textures**. Dans cette figure, il est possible de regrouper les 13 photographies en deux classes de textures naturelles : les galets (2, 4, 8, 10 et 11) et l'écorce des arbres (1, 3, 5, 6, 7, 9, 12 et 13).

Du point de vue computationnel, une classification de textures comprend deux phases importantes: une phase d'apprentissage et une phase de reconnaissance. Pendant la *phase d'apprentissage*, une série d'images texturées est présentée à un modèle d'analyse mathématique. Ce modèle va caractériser le contenu de chaque texture par le calcul d'une gamme de paramètres spécifiques comme la structure spatiale, le contraste, la rugosité ou l'orientation. Dans la *phase de reconnaissance*, une nouvelle texture sera présentée au même modèle qui en extraira une série de paramètres propre à cette texture. Ces paramètres sont ensuite comparés à ceux des images qui ont servi pendant la phase d'apprentissage pour établir des classes de textures.

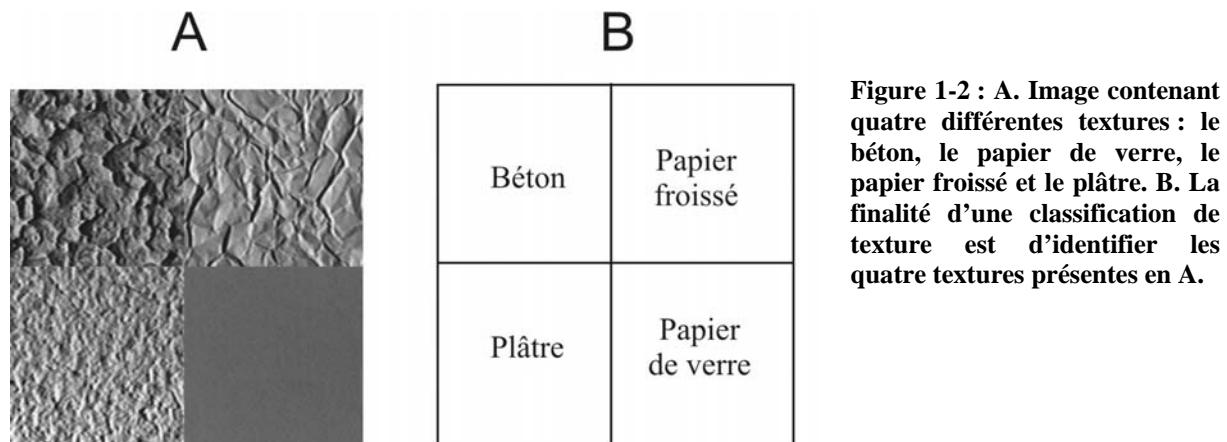


Figure 1-2 : A. Image contenant quatre différentes textures : le béton, le papier de verre, le papier froissé et le plâtre. B. La finalité d'une classification de texture est d'identifier les quatre textures présentes en A.

Le troisième axe de recherche concerne la **synthèse de texture**. Il consiste à créer un pattern texturé proche d'un modèle en utilisant divers algorithmes et outils mathématiques.

Enfin, le dernier problème concerne l'**extraction d'une forme** à partir d'une texture. En effet, la texture permet de définir l'orientation ou la forme d'une surface. Ce postulat implique que la texture serait la même en chaque point de la surface et que les variations intrinsèques de la texture sont un indice pour percevoir la forme de l'objet. Ce dernier problème a fait l'objet de nombreuses recherches, notamment par Gibson qui décrit la texture comme un indice déterminant pour extraire la forme tridimensionnelle d'un objet, au même titre que les ombres et la stéréoscopie (Gibson, 1950a).



Figure 1-3 : 13 photographies de différentes textures naturelles. Les photographies 2, 4, 8, 10 et 11 représentent une texture naturelle : « Galets ». Les photographies 1, 3, 5, 6, 7, 9, 12 et 13 représentent une texture naturelle : « Écorce ».

(2) Les outils mathématiques

Des 4 axes de recherche portant sur la texture, les trois premiers ont favorisé le développement de plusieurs méthodes d'analyses. Ces méthodes d'analyse de texture utilisent principalement trois outils mathématiques différents : les banques de filtres linéaires, les ondelettes et le modèle de Markov.

(a) Les filtres

Le but des méthodes d'analyses d'images consiste à extraire d'une image présentant une unique texture (cette unicité doit absolument être confirmée par un observateur), un ensemble de paramètres censés être représentatifs de la texture qu'on regroupe sous la forme d'un vecteur, nommé « vecteur de caractéristiques ». Ce vecteur se doit d'être peu sensible à des transformations d'images qui laissent la texture inchangée. Plusieurs méthodes de calcul sont exploitées mais elles sont plus ou moins dérivées d'un même principe : le vecteur de caractéristiques représente la sortie d'une banque de filtres. L'outil mathématique le plus communément utilisé est un ensemble – banque ou pyramide – de filtres linéaires. Cette approche computationnelle a été utilisée avec succès pour des segmentations de textures (Fogel et Sagi, 1989; Malik et Perona, 1990), pour des reconnaissances de textures (Rubner et Tomasi, 1998), pour la synthèse de textures (Zhu *et al.*, 1998) mais aussi pour une classification de texture.

Un problème supplémentaire concerne la classification de texture lorsque la direction d'illumination ou le point de vue varie. Ainsi, Leung et Malik (2001) travaillent sur la reconnaissance des surfaces de plusieurs matériaux (textures naturelles comme le béton et le plâtre issues de la CURet) sur la base de leur apparence « *texturale* ». Dans cette étude, les auteurs proposent un vocabulaire basé sur les propriétés photométriques et géométriques des textures. Ils reprennent le terme de « textons » émis originellement par Julesz et définissent les matériaux 3D à l'aide d'histogrammes de textons 2D. Ils parlent alors de « **textons 3D** » (Leung et Malik, 2001).

En suivant un même procédé, Varma et Zisserman présentent un algorithme (VZ algorithme) composé d'une banque de filtres qui réalise une classification des textures invariante à la rotation (Varma et Zisserman, 2002). De même, Hayman et ses collaborateurs (2004) s'intéressent à la reconnaissance de textures naturelles lorsque plusieurs facteurs varient tels que la direction d'illumination ou la distance (Hayman *et al.*, 2004). Dans un

premier temps, ils testent les performances d'un modèle (SVM²) dans une tâche de classification de texture et les comparent avec celles obtenues avec le VZ algorithme. Avec l'outil mathématique proposé par Varma et Zisserman, le taux de réussite s'élève à 97.66% alors qu'avec le SVM, on obtient un taux de réussite de 98.36%. Toutefois, dans une tâche de reconnaissance de textures selon différentes distances, les performances de deux modèles chutent (moins de 50%) lorsque les modèles n'ont jamais été entraînés avec ces images. Les résultats de ces études démontrent que l'utilisation d'une banque de filtres linéaires permet une classification invariante à la direction d'illumination et à la rotation des textures naturelles.

Un autre type d'outil utilise des éléments constitutifs de filtres appelés “ondelettes discrètes”. Une transformation en ondelettes discrètes permet de choisir un ensemble complet de filtres de décomposition en sous-bandes. Les ondelettes présentées par Mallat (1989) utilisent une décomposition en octaves (Mallat, 1989). Toutefois, des travaux suggèrent qu'une décomposition plus serrée que l'octave soit plus efficace pour des analyses de textures. Les ondelettes proposées par Daubechies demeurent une des gammes d'ondelettes les plus utilisées pour les analyses de textures grâce à leurs performances pour séparer les fréquences et leur efficacité de calcul (Daubechies, 1992). Plus récemment, Portilla et Simoncelli (2000) ont proposé un nouveau modèle très performant pour une synthèse de textures reposant sur la transformation en ondelettes (Portilla et Simoncelli, 2000). Ce modèle est basé sur des calculs statistiques relatant les corrélations entre des paires de filtres selon une fréquence spatiale ou une orientation donnée. Les textures synthétisées qu'elles soient naturelles ou artificielles ressemblent fortement aux textures modèles.

(b) Les champs aléatoires de Markov

Li (1994) propose une méthode alternative pour analyser des textures en développant un modèle basé sur un autre type d'outil : les champs aléatoires de Markov (Markov random field ; MRF) (Li, 1994). Ce modèle est basé sur des interactions statistiques entre des pixels proches. Les MRF sont des outils mathématiques intéressants pour l'analyse de texture car ils comportent naturellement des descriptions de dépendance spatiale entre pixels. Ils sont particulièrement adaptés pour la synthèse et la modélisation de textures. Il existe également une abondante littérature sur l'emploi de MRF pour segmenter des textures (Won et Derin, 1992; Andrey et Tarroux, 1998).

² SVM : Support Vector Machine

De ces trois types d'outils mathématiques, il semble que la pyramide de filtres GABOR soit la plus performante dans une tâche de classification de textures et que les deux autres outils soient davantage utilisés pour une synthèse ou une segmentation de textures.

d) Données psychophysiques

La plupart des expériences de psychophysique sur les textures se sont intéressées à la segmentation de textures et ont cherché à caractériser les attributs de la texture nécessaires à une extraction d'un objet du fond visuel. Toutefois, d'autres études ont eu pour objectifs la reconnaissance d'objets sur la base des textures ou la reconnaissance tactile des textures. Les études concernant la texture comme indice tridimensionnel pour coder l'orientation d'une surface seront traitées dans le paragraphe B.2.1.b)(2)

(1) Segmentation de textures

Beck (1972) émet l'hypothèse que la segmentation est basée sur des attributs simples des éléments de la texture comme la luminance, la taille, la couleur ou l'orientation (Beck, 1972). Marr (1976) y ajoute également la forme de l'élément de texture comme indice important (Marr, 1976). Le principe de **similitude** des lois de la Gestalt s'applique alors pour différencier deux textures entre elles. Par exemple, un groupement par similarité d'orientation permet la segmentation spontanée d'une région texturée d'une autre région texturée (Figure 1-4A). Les travaux les plus souvent cités dans la littérature sont ceux de Bela Julesz (Julesz, 1975; Julesz, 1981; Julesz, 1986; Julesz et Kroese, 1988). Il s'est particulièrement intéressé à déterminer comment s'effectue une discrimination entre deux textures identiques en luminosité, contraste ou couleur. Pour expliquer ce processus, il a entrepris d'analyser les textures à l'aide de méthodes statistiques³.

Julesz définit alors les **textons** comme des éléments de base d'une texture regroupant des paramètres similaires comme la taille, l'orientation, les contours. Les textures de la Figure 1-4B ne peuvent pas être segmentées car leurs textons sont trop similaires, notamment au niveau de leur forme carrée. A l'inverse, les textures de la Figure 1-4C sont facilement segmentées grâce au nombre de coins différents entre les textons.

L'interprétation traditionnelle de la segmentation a été remise en question par Northdurft. Il indique que la segmentation n'est pas due uniquement à un groupement par similarité mais grâce aux contrastes d'orientation des éléments situés au niveau des frontières.

³ Je décrirai ces méthodes d'analyse statistique dans le chapitre 2

D'après Northdurft (1985), la séparation entre deux textures peut se faire par un gradient de structure (Nothdurft et Li, 1985).

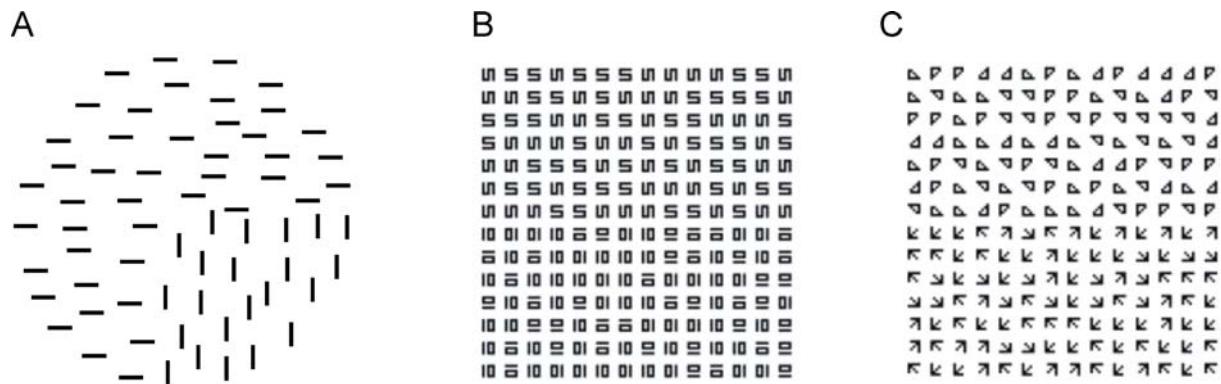


Figure 1-4 : A. Textures orientées. Le quadrant inférieur droit du cercle se discrimine des autres quadrants du fait de l'orientation verticale des éléments. B et C. Deux exemples de patterns de textures utilisés par Bela Julesz dans des tests de discrimination. Aucune segmentation n'est possible dans le pattern présenté en B (les textons carrés sont trop similaires) alors que les sujets peuvent facilement discriminer deux textures dans le pattern texturé C ; les textons diffèrent par leur nombre de coins ; la partie supérieure contient des textons définis par 3 coins, celle inférieur par 4 coins.

Enfin, les résultats des études psychophysiques sur la segmentation s'accordent sur le fait que la texture est un indice permettant la perception de contours à l'intérieur d'une image. Toutefois, d'autres indices comme la luminance, la couleur ou la vitesse peuvent également nous aider dans la définition des contours. Rivest et Cavanagh (1996) ont démontré que la détermination d'un contour dans une image découle d'un compromis entre la texture et d'autres indices comme la luminance et la couleur. La contribution de plusieurs indices permet une meilleure définition des contours que si un seul indice était présent dans l'image (Rivest et Cavanagh, 1996).

(2) Reconnaissance d'objets sur la base de texture

Rossion et Pourtois (2004) ont testé l'influence des propriétés matérielles (couleur et texture) dans une tâche de reconnaissance d'objets (Rossion et Pourtois, 2004). Les sujets doivent nommer le plus rapidement possible l'objet qui leur est présenté. Les stimuli regroupent 3 catégories d'objets (des animaux, des fruits ou légumes et des objets manufacturés) définis selon 3 conditions : (1) uniquement par leurs bordures externes et internes, (2) bordures + texture (3) bordures + texture + couleur. Les performances des sujets (taux de réussite et temps de réaction moyens) sont comparées entre les différentes conditions. Je ne traiterai ici que des comparaisons entre les conditions 1 et 2. Les résultats de cette étude démontrent que l'apport de la texture permet une meilleure reconnaissance des objets ; le taux

de réussite des sujets augmente significativement. Toutefois, leur temps de réaction moyen est identique à celui de la condition (1) même si les sujets reconnaissent plus rapidement les animaux *texturés* que les animaux sans texture. En fait, lorsque les objets sont reconnus rapidement sans texture, l'ajout de la texture ralentit leur reconnaissance. L'ajout de la texture permet une meilleure dénomination notamment pour des items qui définis sans texture étaient plus longs à reconnaître par les sujets. Un possible rôle fonctionnel de la texture dans une reconnaissance d'objets avait été démontré préalablement par Price et Humphreys (1989) dans une tâche de classification superordonnée où les sujets devaient nommer les objets qui leur étaient présentés (Price et Humphreys, 1989). Les performances des sujets sont améliorées si une texture est présente à la surface de l'objet.

L'étude de Rossion et Pourtois (2004) démontre donc que la texture permet une meilleure reconnaissance des objets et présente deux effets opposés sur les temps de réaction. Elle ralentit la reconnaissance d'objets rapidement identifiés sans textures et accélère la représentation d'un objet qui n'est pas facilement reconnaissable uniquement ses contours externes. Or, dans notre environnement visuel, les objets n'apparaissent jamais totalement définis par leur contour externe comme c'est le cas avec les images sans texture. Leur texture permettrait donc une meilleure reconnaissance bas-niveau de la forme de l'objet (Regan, 2000).

(3) Reconnaissance tactile des textures

Autre qu'une information visuelle, la texture d'un objet est également une information d'ordre tactile. De nombreuses expériences de psychophysique se sont intéressées à la perception des textures sur la base d'informations tactiles. Yoshida (1968) fut le premier à étudier la perception tactile de surfaces d'objets (bois, verre, papier, ivoire, coton, soie ou laine) à l'aide d'une analyse multidimensionnelle de type MDS. Yoshida met alors en évidence 4 dimensions pour décrire les textures : chacune est caractérisée par une paire d'adjectifs qualificatifs antinomiques que les sujets vont habituellement utiliser pour décrire la texture qu'ils touchent. Les sujets utilisent les adjectifs « dur/doux », « épais/léger », « froid/chaud » et « rugueux/lisse » (Yoshida, 1968). Hollins et al (1993) ont également utilisé une analyse multidimensionnelle afin de paramétriser des textures naturelles sur la base du toucher. Ils ont identifié 3 dimensions dont deux principales : la rugosité (« rugueux/lisse ») et la dureté (« doux/dur »). La nature de la troisième n'est pas clairement élucidée ; elle serait dépendante de l'appréciation des sujets mais pourrait correspondre à l'élasticité de la surface (Hollins *et al.*, 1993). Plus récemment, une étude sur la perception

tactile des sièges automobiles signale qu'une description de texture n'excède pas 4 dimensions (Picard *et al.*, 2003). La première dimension concerne la douceur de la surface (les sujets utilisent par exemple les adjectifs « râche/doux »), la seconde la finesse du matériau (« fin/épais »), la troisième fait état du relief présent dans la texture (« lisse/bosselé ») et enfin la quatrième dimension est relative à la souplesse du matériau (« souple/dur ou rigide »).

Peu d'études se sont intéressées à la perception à la fois tactile et visuelle des textures naturelles. Jones et O'neil (1985) ont montré que la précision du jugement de la rugosité d'un matériau est équivalente entre le toucher et la vision même si les appréciations visuelles sont plus rapides que les appréciations tactiles (Jones et O'Neil, 1985). De plus, Heller (1989) démontre également une similarité entre les précisions de jugement de surfaces abrasives notamment au niveau du caractère lisse ou pas de surfaces relativement rugueuses. Toutefois, les sujets jugent mieux des matières plutôt fines (comme la soie) en la touchant qu'en la voyant (Heller, 1989).

3. Conclusion

Au cours de ce chapitre, nous avons vu qu'un objet se discrimine du fond visuel grâce à la définition de sa forme 2D, la catégorisation de sa couleur et la segmentation de sa texture. D'un point de vue théorique, les modèles se basent sur l'architecture du cortex visuel pour expliquer le traitement de la forme de l'objet mais accordent peu d'importance aux propriétés de surface pour une reconnaissance de l'objet alors que la texture ou la couleur sont des informations pertinentes pour extraire (ou segmenter) l'objet du fond visuel. Les indices décrits dans ce chapitre (forme, couleur et texture) vont nous permettre de définir l'objet uniquement en deux dimensions alors que dans notre entourage visuel, les objets sont définis en trois dimensions. C'est l'apport d'indices visuels supplémentaires qui va nous permettre d'extraire cette troisième dimension et de percevoir de la profondeur dans la forme de l'objet. Le paragraphe suivant présente les indices visuels nécessaires à la perception de la troisième dimension.

B. LES INDICES NECESSAIRES A LA PERCEPTION VISUELLE EN 3D

La perception tridimensionnelle d'un objet repose sur la présence d'indices binoculaires et monoculaires nous renseignant sur sa profondeur (Tableau 1-1). Ces indices contiennent de l'information qui, s'ajoutant à l'image 2D projetée sur la rétine, vont nous permettre de percevoir les objets en 3D. La disparité est l'indice de perception de la profondeur jugé le plus important dans notre système visuel. Cet indice est basé sur la parallaxe binoculaire définie comme la différence entre les images d'un même objet projetées sur chaque rétine. Le degré de disparité dépend également de l'angle de vergence, angle formé par l'axe d'optique de chaque œil convergeant vers un objet donné. Cet angle est également fonction de la distance entre l'objet et les yeux. Si la distance augmente, l'angle de vergence diminue et la perception de la profondeur devient alors de plus en plus difficile. L'angle de vergence constitue le second indice binoculaire.

Les autres indices importants sont perceptibles avec un seul œil. Ces indices monoculaires de la profondeur peuvent être séparés en deux catégories en fonction de la présence ou pas d'un mouvement au niveau de l'image. Si l'image est en mouvement, le flux optique, la parallaxe du mouvement et les phénomènes d'accrétion et de déletion⁴ nous renseignent sur la profondeur de l'image. Tous ces indices monoculaires sont dits *dynamiques*. Parmi les indices monoculaires *statiques*, on distingue l'accommodation (ou netteté), la perspective, l'interposition et l'éclairage. Au niveau de la terminologie, les trois derniers indices sont définis comme des indices picturaux car ils sont présents au sein d'images bidimensionnelles comme des peintures, des dessins ou des photographies. Ces indices statiques (ou picturaux) seront décrits dans la seconde partie de ce chapitre. La première partie décrira les indices binoculaires, les indices oculomoteurs et les indices monoculaires dynamiques de la perception tridimensionnelle.

⁴ les termes d'accrétion et de déletion sont utilisés pour définir un mouvement lorsqu'un objet se rapproche (accrétion) ou s'éloigne (déletion) de l'observateur.

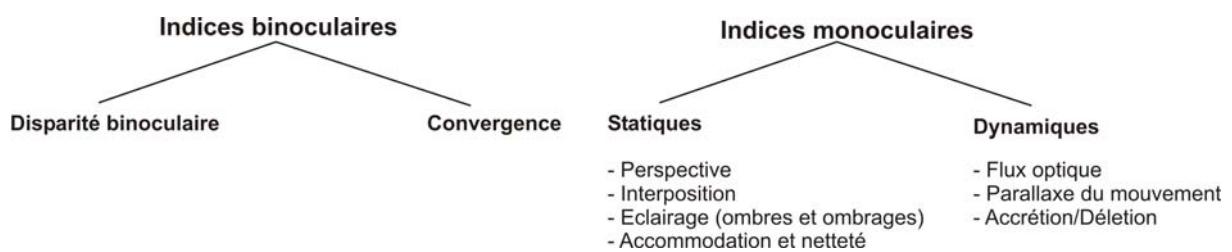


Tableau 1-1 : les indices binoculaires et monoculaires nécessaires pour une perception tridimensionnelle

1. Les indices binoculaires et monoculaires dynamiques

1.1. La disparité

a) Définition

Lorsque nous explorons une scène visuelle, les images projetées sur chacune de nos rétines sont légèrement différentes ou disparates, du fait de l'écartement entre les deux yeux (6 cm en moyenne). Cette différence entre les images rétinianes de l'œil gauche et de l'œil droit est appelée **disparité binoculaire** et fournit des informations qui, une fois exploitées, nous permettent de reconstruire la troisième dimension, autrement dit, la profondeur. En effet, l'information transmise par cette disparité nous renseigne sur la différence de distance entre deux points : quand la distance des objets en profondeur par rapport à nous diffère alors leurs images rétinianes diffèrent également. Un objet plat placé dans un même plan, à équidistance de nos yeux ne génère pas de disparité rétinienne. Cette disparité binoculaire permet une perception de la profondeur via la vision stéréoscopique.

b) Aspects psychophysiques

Afin de démontrer que la disparité rétinienne est bien un indice de profondeur, il est nécessaire de prouver que la différence présente entre les images rétinianes de chaque œil peut à elle seule produire une impression de profondeur. Charles Wheatstone, en 1838, fut le premier dans une publication parue dans le *philosophical transactions of the Royal Society* à s'intéresser à la vision stéréoscopique (Wheatstone, 1838). Il créa un appareil, le stéréoscope, permettant de présenter à chaque œil des images légèrement différentes et dont la fusion produite par le système visuel induit une impression de profondeur. Le dispositif présenté par Wheatstone utilisait deux miroirs permettant de renvoyer une image à chaque œil. La paire d'images constitue un stéréogramme. Avec ce dispositif, Wheatstone a ainsi projeté des images bidimensionnelles sur chaque œil. Grâce à la fusion des images, les sujets percevaient de la profondeur. Il en a conclu que la perception de la profondeur était la conséquence de la disparité entre les deux vues. Il a également démontré que la disparité rétinienne est un indice

plus performant pour la perception de la profondeur que les indices picturaux de l'image. Plus récemment, David Brewster (1781-1868) proposa un autre système pour induire une vision stéréoscopique en utilisant des lentilles réfractantes. Ces lentilles en face de chaque œil vont rendre les rayons lumineux parallèles, comme s'ils provenaient d'une grande distance. Les directions de regard de chaque œil restent parallèles comme s'ils fixaient un objet très éloigné, de telle manière que chaque œil voit une figure différente.

La Figure 1-5 présente deux exemples de stéréogrammes classiques qui, projetés dans un stéréoscope, produisent une impression de profondeur. Dans la Figure 1-5A, l'image de gauche est projetée sur l'œil gauche et celle de droite sur l'œil droit. Le sujet perçoit alors le petit carré en avant du grand carré. La perception de profondeur peut être simplement renversée en inversant les stéréogrammes droit et gauche. En 1959, Bela Julesz crée des stéréogrammes définis par des patterns de points aléatoires ou RDS⁵ (Figure 1-5B). Ces RDS, projetés sur chaque œil, créent des figures (ici un carré) en avant ou en arrière du plan. Le carré central ne peut pas être perçu avec un seul œil. C'est la fusion binoculaire que va opérer notre système visuel qui engendre la perception du carré. La découverte des RDS est une avancée importante pour la compréhension des mécanismes nécessaires à l'extraction de la profondeur car la disparité binoculaire y est le seul indice présent.

Ce procédé de vision stéréoscopique est naturellement appliqué lorsque nous regardons une scène visuelle du fait de la présentation d'images disparates à nos deux yeux ; l'écartement horizontal entre nos deux yeux va générer ces disparités positionnelles⁶.

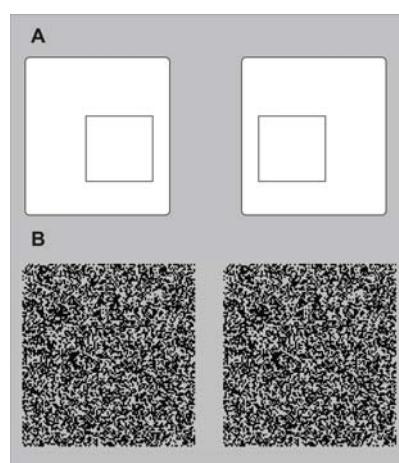


Figure 1-5 : Exemples de stéréogrammes. A. Stéréogrammes représentant deux carrés. L'image de gauche est projetée à l'œil gauche et celle de droite à l'œil droit ; le plus petit carré (à l'intérieur sur l'image 2D) apparaît en avant du grand carré. B. Stéréogrammes (RDS) de B. Julesz.

⁵ Random Dot Stereograms (RDS)

⁶ La différence entre les images rétinianes droite et gauche génère une disparité binoculaire qui englobe des différences entre les positions, les orientations, les tailles et les perspectives des éléments. Je ne décrirai que le phénomène de disparités de position ou *disparités positionnelles*. Ce sont des disparités élémentaires à partir desquelles peuvent être retrouvées les autres disparités comme les disparités de taille, de perspective ou de déformation. Les disparités positionnelles se définissent comme la différence de localisation sur les deux rétines des éléments à décrire. On distingue la **disparité positionnelle horizontale** et la **disparité positionnelle verticale**.

Deux principaux modèles théoriques s'attachent à expliquer la perception de la profondeur à partir des disparités positionnelles. Le premier type de modèle considère la disparité horizontale et verticale comme deux paramètres parfaitement indépendants ; la disparité horizontale serait à la base du percept stéréoscopique alors que le rôle de la disparité verticale serait plus global et renseignerait sur la direction du regard et la distance de fixation (Mayhew, 1982; Mayhew et Longuet-Higgins, 1982). Pour l'autre type de modèle, les composantes horizontale et verticale de la disparité ne sont pas dissociées et contribuent à la mise en place d'une représentation régionale des disparités. Cette configuration régionale conduit à l'extraction du gradient de disparité permettant la perception du relief ou l'orientation d'une surface (Koenderink et van Doorn, 1976; Liu *et al.*, 1994).

1.2. *Les indices oculomoteurs*

La vergence et l'accommodation sont considérées comme des indices oculomoteurs car ils sont fonction des mouvements des yeux. Par ailleurs, la vergence est un indice binoculaire alors que l'accommodation est un indice monoculaire. Je ne ferais que décrire succinctement ces deux phénomènes.

a) La vergence

Si le mouvement d'un œil était totalement indépendant de l'autre œil, l'image d'un objet serait projetée dans des régions complètement distinctes des deux rétines, ce qui aboutirait à une vision double ou diplopie. Lorsque notre regard se fixe sur un point de l'espace, la direction des deux yeux tend normalement à converger vers ce même point. La vergence est définie comme l'angle formé par les deux yeux fixant un point donné dans l'espace. Si l'objet est très éloigné, les directions de chaque œil tendent à être parallèles et l'angle de vergence sera nul. L'angle est relativement petit pour des objets lointains et augmente au fur et à mesure que l'objet se rapproche. Ainsi, la vergence des deux yeux peut constituer un indice de profondeur à condition que l'information concernant le degré de convergence soit transmise à notre système visuel. Ce principe de vergence a été initialement décrit par Descartes (1673) dans la *Dioptrique* où il fait l'hypothèse d'une possible participation de la vergence comme indice de profondeur, notamment pour évaluer la distance relative d'un objet.

b) L'accommodation

Lorsque nous explorons notre espace visuel, les cristallins de nos yeux modifient de manière réflexe leur épaisseur pour assurer une bonne mise au point des objets positionnés à

différentes distances. Ainsi, cette épaisseur du cristallin ou état d'**accommodation** pourrait être utilisé par notre système visuel comme un indice de profondeur. Un autre critère lié à l'accommodation peut nous renseigner également sur la distance des objets : les objets lointains du plan de fixation produisent des images floues. La netteté d'une région de l'image nous renseigne sur sa distance. Les artistes peintres ou les photographes utilisent la méthode de flou pour donner une impression de profondeur à leur œuvre (voir Figure 1-6). Les photographes fixent le focus de l'appareil sur un objet afin qu'il apparaisse le plus net possible, par exemple ici la forêt au premier plan. En augmentant l'ouverture du diaphragme, les alentours de l'objet sont floutés et donnent une impression de profondeur ; la montagne apparaît plus lointaine. Le flou n'est pas ici l'unique indice de perception de profondeur puisque l'interposition et la perspective aérienne interviennent également pour apporter une impression de profondeur. Ces deux indices seront décrits dans la partie sur les indices picturaux (voir paragraphe B.2). Toutefois, les images d'objets flous peuvent être un indice de profondeur ambigu. En effet, un objet en avant de notre plan de fixation peut apparaître flou comme s'il était éloigné. Des images floues servent, en fait, d'indices de profondeur seulement si les contours des objets sont parfaitement définis. Par exemple, Marshall et al (1996) démontrent que le flou d'une texture combinée avec le flou des bordures d'une région texturée induit une perception de profondeur non ambiguë (Marshall *et al.*, 1996).

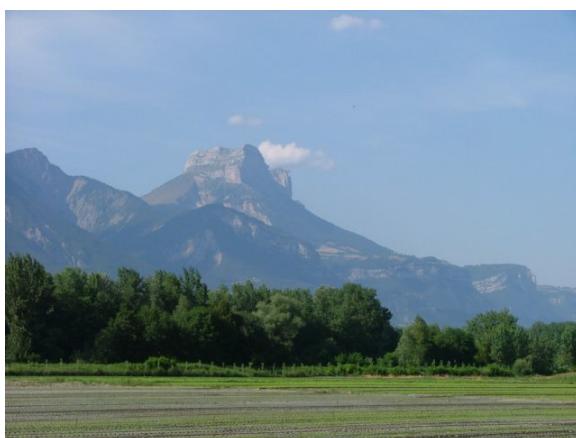


Figure 1-6: Photographie de la Chamechaude prise à Grenoble (juin 2005). Le flou en arrière de la clairière entraîne une impression de profondeur. Ce n'est pas l'unique indice de profondeur puisque l'interposition (les arbres devant les montagnes) nous renseignent également sur la profondeur.

1.3. Les indices de mouvement

Comme nous pouvons percevoir avec une précision correcte la profondeur avec un seul œil, il existe forcément d'autres indices nous renseignant sur la distance présente dans notre espace visuel, notamment via des informations dérivées du mouvement. Le principe de la **parallaxe du mouvement**, l'indice de profondeur apporté par le mouvement, a été initialement décrit par Helmholtz (Helmholtz, 1910). Dans une scène visuelle, chaque objet est défini par une direction particulière du regard dépendante de notre positionnement par rapport

à la scène. Si nous changeons de position et que nous fixons toujours un même objet, la direction du regard vers cet objet change également en fonction de la distance de l'objet. La direction du regard d'un observateur vers un objet proche change énormément si l'observateur est en mouvement alors que la direction du regard vers un objet éloigné change peu. L'importance du changement de direction pourrait ainsi nous indiquer la distance de l'objet. Gibson (1950) a décrit la parallaxe du mouvement comme un important indice de profondeur (Gibson, 1950a). On distingue habituellement cinq différents types de parallaxe du mouvement :

- La parallaxe absolue : l'image d'un point dans l'espace se déplace sur la rétine si l'objet bouge et que l'œil reste en position stationnaire ou si l'œil bouge et l'objet reste immobile. L'importance du changement de direction visuelle est inversement proportionnelle à la distance de l'objet.
- Le fait que les images d'un ou plusieurs objets augmentent en taille quand ils approchent de l'observateur (phénomène de *looming*).
- Parallaxe linéaire : lorsqu'un objet ou un ensemble d'objets se déplace latéralement ou est vu par une personne se déplaçant latéralement, les images des parties les plus distantes se déplacent moins vite que les images des parties les plus proches. Cet effet est aussi appelé perspective de mouvement.
- L'effet cinétique de profondeur. Un objet paraît être en profondeur uniquement s'il est en mouvement.
- L'effet stéréocinétique. Un mouvement de rotation de plusieurs objets définis dans un même plan induit une impression de profondeur.

Les trois derniers effets seront décrits en annexe.

2. Les indices picturaux

Les indices décrits précédemment sont soit des indices tridimensionnels binoculaires (vergence et disparité) ou soit des indices monoculaires dynamiques (la parallaxe du mouvement). Ces paramètres sont, dans notre vie quotidienne, de forts indices de la profondeur cependant il nous est également possible de percevoir de la profondeur en regardant simplement une image bidimensionnelle comme une photographie ou un tableau d'un artiste peintre. La présence d'indices particuliers sur ces représentations 2D nous renseigne sur la profondeur. Depuis la Renaissance, les artistes peintres ont été les premiers à

découvrir et à utiliser ces indices dans leurs œuvres en trompe-l’œil. Ils ont parfaitement réussi à reproduire des scènes visuelles et nous apporter une impression de profondeur en utilisant d’autres types d’informations que les indices binoculaires ou monoculaires dynamiques. Prenons comme exemple, l’œuvre de Gustave Caillebotte, *rue de Paris, temps de pluie* (voir Figure 1-7). Des indices tels que la perspective (perspective linéaire : immeuble du fond, gradient de texture : pavés au sol), le flou (effet de brume au loin) ou l’interposition (l’homme moustachu devant le lampadaire) nous font percevoir de la profondeur dans ce tableau. Ces informations sont dites **picturales** car elles sont présentes dans des images comme des peintures, des dessins ou des photographies. Ces informations sont plus fonction de la représentation des objets dans une scène et des surfaces sur lesquelles ils reposent que de la position de l’observateur et de son regard. Elles peuvent être perçues par un observateur statique, utilisant un seul œil. Elles sont donc définies comme des **indices monoculaires statiques**.



Figure 1-7 : Rue de Paris, temps de pluie. Œuvre de Gustave Caillebotte (1848-1894). (Image tirée de www.ibiblio.org/.../auth/caillebotte/rainy.jpg). Des indices picturaux comme la perspective (linéaire : immeuble du fond, gradient de texture : pavé au sol), flou ou interposition nous font percevoir de la profondeur dans cette image 2D.

Les indices picturaux ne servent pas à percevoir de la profondeur uniquement dans des peintures ou des photographies mais ont également une fonction dans la perception de la profondeur réelle. Il semble que ces indices picturaux soient nécessaires pour définir une distance relative voire absolue entre des objets, ce qui nous permettrait de mieux percevoir la profondeur dans le monde réel. Ces indices picturaux nous renseignent également sur la

surface des objets afin de mieux les reconnaître. Quels sont ces indices picturaux que les peintres ont utilisés si efficacement dans leurs œuvres et comment contribuent-ils à une perception tridimensionnelle dans notre vie quotidienne ? Nous distinguerons et décrirons parmi ces indices picturaux : la perspective, l'éclairage (ombres et ombrages) et le recouvrement partiel ou interposition d'un objet. Dans cette partie, nous nous intéresserons particulièrement aux informations de surface nous renseignant sur la profondeur que sont les gradients de texture et les gradients d'illumination générant des ombres et ombrages.

2.1. La perspective

a) Généralités

Parmi tous les indices picturaux, la perspective semble être l'indice le plus important pour percevoir de la profondeur dans une image 2D. La perspective se définit comme l'art de représenter les objets en trois dimensions sur une surface plane en tenant compte de l'éloignement et de leur position dans l'espace par rapport à l'observateur (Hachette 2004). Plusieurs lois sont nécessaires pour représenter un objet en 3D et vont s'appliquer à différents types de perspectives que ce soit chez les artistes peintres ou chez les scientifiques.

- Du côté des artistes*

Les principes de la perspective linéaire ont été étudiés pendant la Renaissance italienne, à l'orée du Quattrocento, par un architecte florentin, Filippo Brunelleschi (1377-1446). Il en fait la démonstration par une petite expérience réalisée sur la place San Giovanni à Florence en 1415 à l'aide d'un miroir et d'un dessin du baptistère de Florence réalisé selon des principes stricts de perspective : lignes convergentes, point central, etc.... Son dessin est disposé sur un support dans lequel est percé un trou pour apercevoir l'image du baptistère réfléchie par le miroir. Le dessin se superpose parfaitement à l'édifice réel et crée une illusion parfaite de profondeur. Toutefois, la véritable explication scientifique de la perspective linéaire fut énoncée par Léon Battista Alberti (1406-1472) dans son traité *Della pictura* (1435) où il décrit que l'œil humain se placerait au sommet d'une pyramide visuelle et la surface plane du tableau où se forme l'image en constituerait la base. Les œuvres picturales usant de perspective selon Alberti deviennent alors des fenêtres ouvertes sur le monde. Les artistes peintres de la renaissance vont utiliser ces principes pour donner à une surface plane l'illusion de la profondeur et du relief. La fresque de la sainte Trinité du peintre florentin Masaccio (1401-1428) constitue le premier exemple d'œuvre picturale réalisée en respectant strictement les principes de la perspective géométrique découverts par Brunelleschi. Auparavant, plusieurs œuvres avaient été réalisées selon quelques principes de perspective

mais ne respectaient pas toutes les règles de perspective. Par exemple, dans le tableau de La dernière cène de Duccio Di Buoninsegna (1255-1319), des premiers essais de perspective ont été réalisés mais sans aucune théorie : les images perpendiculaires au tableau, telles les poutres du plafond ou les lignes de la nappe, ne convergent pas vers un même point de fuite⁷ (Figure 1-8).



Figure 1-8 : La dernière cène de Duccio Di Buoninsegna. Les lignes orange et rouge sont rajoutées au tableau afin de souligner les erreurs de perspective linéaire. Les droites convergent vers deux points de fuite différents. Les flèches noires indiquent d'autres erreurs de perspective comme les rayures sur la nappe.

- *Du côté des scientifiques*

Gibson (1950), par plusieurs expériences psychophysiques, a entrepris de déterminer quels étaient les indices monoculars importants pour une perception tridimensionnelle (Gibson, 1950a). La Figure 1-9 présente un des indices importants décrits par Gibson ; le gradient de perspective dans lequel l'écartement entre les lignes parallèles change graduellement dans la direction verticale. La perspective visuelle, dans toutes ces formes, survient en fait lorsqu'un objet concret s'éloigne en distance et que de ce fait son image rétinienne diminue en taille. On distingue différents types de perspective : la perspective linéaire, la perspective de taille, la hauteur dans le champ et le gradient de texture (Figure 1-9).

La perspective linéaire fait simplement référence au fait que des droites parallèles qui s'éloignent dans l'espace tridimensionnel tendent à converger vers un même point (le point de fuite) et se rejoignent à l'infini.

⁷ Le point de fuite est le point de concours des droites qui représentent les perpendiculaires au tableau.

La perspective de taille est un autre élément de la perspective ; un objet voit sa taille diminuer plus il est éloigné dans la scène visuelle, par exemple dans la Figure 1-9, les colonnes diminuent en taille le long du couloir.

Gibson (1950) décrit un aspect différent de la perspective de taille ; le gradient de densité de texture. Un gradient de texture est présent dans cette figure au niveau du sol, la densité des carreaux au sol augmente le long du couloir. Un autre aspect de la perspective est la compression. Elle concerne la représentation de distances égales dans la troisième dimension ; par exemple ici, l'écart entre deux lignes horizontales du sol diminue si la perspective s'éloigne.



Figure 1-9 : Exemple de dessin où toutes les formes de perspectives sont représentées : perspective linéaire, perspective de taille, le gradient de texture. (Tiré de www.janushead.org)

Les peintres de la renaissance ont également utilisé deux autres aspects de la perspective ; la perspective de détail et la perspective aérienne. La perspective de détail concerne le fait que des détails des objets ne soient plus perceptibles avec l'éloignement à cause des limitations de l'acuité visuelle. Cette perspective du détail est liée à un autre indice de profondeur qu'est le flou. Quant à la perspective aérienne, elle souligne la tendance des objets éloignés à devenir bleutés à cause des impuretés de l'atmosphère (par exemple la montagne Chamechaude dans la Figure 1-6).

Dans la partie suivante, je ne décrirai que les expériences de psychophysiques effectuées sur les gradients de texture. Les travaux sur les autres principes de perspectives (linéaire, taille, hauteur dans le champ) font l'objet d'une excellente revue dans le livre de Ian P. Howard (Howard, 2002).

b) La perspective de texture : le gradient de texture

(1) Généralités

La Figure 1-10 montre un gradient de texture dans lequel à la fois la taille et le nombre d'éléments renseignent sur la perception de profondeur ; la taille des éléments est plus importante et la densité est plus faible dans le bas de l'image, donnant l'impression d'une bordure plus proche de nous. Les **gradients de textures** ont été décrits par Gibson (1950) comme un des plus importants indices picturaux de profondeur, avec le gradient de perspective. Dans notre environnement visuel, de nombreux objets de taille similaire définissent une texture ; par exemple, une plage de sable ou un champ d'herbe auront l'apparence d'une texture plutôt qu'un assemblage d'objets. Les gradients de textures (variations en taille et densité) s'installent quand un pattern qui varie avec la distance apporte des indices importants pour la perception de la profondeur et de la forme tridimensionnelle. On distingue plusieurs types de gradients selon la texture présente.

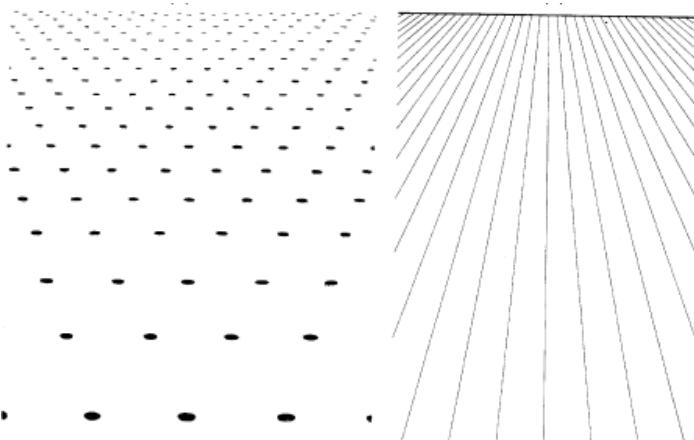


Figure 1-10 : Gradients de texture définis, à gauche par des points, à droite par des lignes convergentes (Gibson, 1950).

Dans un premier temps, si nous considérons une surface présentant une texture **homogène**, les textons sont alors identiques en forme, taille et densité quelle que soit leur position dans la surface⁸. Le gradient de texture présenté dans la Figure 1-10A est défini par une texture isotrope et 3 indices présents dans la texture interviennent pour l'impression de perspective. Ces 3 indices ont été définis par Gibson (Gibson, 1950a) et peuvent être séparés expérimentalement.

- La **taille** des éléments de texture : la taille apparente d'un élément de texture décroît selon la distance le séparant de l'œil. L'échelle de la taille des éléments de texture est

⁸ Une texture est également dite **isotrope** [0] si les éléments de texture sont identiques en forme, taille et densité pour toutes les orientations.

plus fiable comme indice de profondeur quand les éléments de texture dans une surface ont la même taille réelle.

- **La densité** des éléments de textures : la densité des éléments de textures augmente avec la distance. Le gradient de densité de texture indique alors l'orientation ou la courbure de la surface.
- **La compression ou aspect-ratio** : le rapport entre la largeur et la hauteur des éléments d'une surface permet de se renseigner sur son inclinaison.

(2) Données psychophysiques

La contribution du gradient de texture dans la perception de la profondeur a été étudiée dans des expériences psychophysiques sur le jugement de l'orientation tridimensionnelle d'une surface plane mais aussi sur l'appréciation de l'état de courbure d'un objet 3D.

(a) Juger l'inclinaison d'une surface

Grâce à de nombreux tests, Gibson a démontré que le gradient de texture est un indice important pour définir l'orientation tridimensionnelle d'une surface (Gibson, 1950a). Il fait varier pour cela le gradient de texture dans une image et remarque qu'une variation du gradient de texture entraîne une impression de profondeur. Dans une autre expérience, Gibson (1950b) présente à des sujets des surfaces texturées pouvant être inclinées par rapport à un axe horizontal. Les sujets voient cette surface à travers une ouverture circulaire dans un plan frontal et indiquent l'inclinaison de la surface qu'ils perçoivent grâce à un clavier tactile. L'inclinaison perçue augmente lorsque que le gradient de texture augmente (Gibson, 1950b). Cette expérience a été réalisée avec des textures artificielles mais Gruber et Clark observent le même type de résultat avec des textures naturelles inclinées vues à travers une ouverture carrée (Gruber et Clark, 1956). Ces deux études indiquent également un autre résultat intéressant ; le plan de l'inclinaison est sous estimé par les sujets. Les auteurs expliquent ce résultat par le fait qu'une surface isolée inclinée tend à se normaliser vers un plan frontal ; les sujets perçoivent l'ouverture sur la surface frontale comme les bordures de la surface texturée. Quant à la précision du jugement de l'inclinaison perçue d'une surface texturée, Zimmerman et ses collaborateurs rapportent qu'elle est proche de 2° (Zimmerman *et al.*, 1995).

Dans un autre type d'expérience, Tozawa et ses collaborateurs (2001) ont étudié les effets combinés des gradients de disparité et de texture sur la perception de l'inclinaison d'une surface chez l'Homme. Lorsque les deux indices sont présents au sein des stimuli

binoculaires, l'impression d'inclinaison de la surface est plus importante que si seul l'indice de gradient de texture est présent (voir revue de (Sakata *et al.*, 2005)).

(b) Juger l'état de courbure d'un objet 3D

Autre que l'inclinaison, le gradient de texture peut également nous renseigner sur la courbure d'une surface. Todd et ses collaborateurs (2001) demandent à des sujets d'évaluer la courbure d'une forme tridimensionnelle définie par un gradient de texture. Les résultats de cette étude montrent une parfaite corrélation entre les jugements des sujets et l'état réel de courbure de l'objet ; les estimations des concavités et des convexités sont très précises mais la magnitude du relief est souvent sous estimée (entre 38 et 75%). Le gradient de texture permet donc à lui seul à définir l'état de courbure d'un objet (Todd *et al.*, 2001).

Alors que ces études psychophysiques suggèrent que les humains perçoivent la profondeur grâce à un gradient de texture, il existe peu de données qui prouvent que les singes ont la même capacité. Tsutsui et al (2002), ont, dans un premier temps, entraîné les singes à une tâche de comparaison d'orientations (« delayed match-to-sample ») puis à une tâche de généralisation au cours de laquelle les singes doivent reconnaître l'orientation d'une surface définie par différents gradients de texture ou de disparité. Les résultats de cette étude montrent que les singes utilisent ces deux indices (texture et disparité) de manière équivalente pour percevoir l'orientation tridimensionnelle d'une surface (Tsutsui *et al.*, 2002).

2.2. *L'interposition*

a) *Généralités*

Lorsque plusieurs objets sont présents au sein d'une même scène visuelle, un autre indice pictural peut nous renseigner sur la profondeur ; l'interposition ou disposition des objets les uns par rapport aux autres. L'interposition fournit des informations sur l'ordre de profondeur mais pas sur l'importance de l'intervalle de profondeur. Deux cas distincts d'interposition peuvent être considérés :

- Occlusion d'une partie d'un objet par d'autres parties du même objet.
- Occlusion d'un objet par un autre objet.

Le second cas semble le plus prompt à nous renseigner sur la profondeur. Le triangle de Kanizsa (1976) schématisé les deux phénomènes perceptuels présents dans l'interposition d'un objet par d'autres objets (Figure 1-11) (Kanizsa, 1976). Le triangle blanc paraît complet alors que certaines parties de ses cotés ne sont pas visibles du fait d'une même luminance entre le triangle et le fond de l'image ; on définit ce phénomène comme la **compléition**

modale. D'autre part, chaque disque noir paraît être complet et placé en arrière du triangle blanc ; puisque la partie du disque masquée est inaccessible à une modalité sensorielle, on parlera dans ce cas de **compléTION amodale** (décrit initialement par (Guzman, 1968). Ces deux exemples de compléTION sont heuristiques puisque nous percevons une bordure qui n'est physiquement pas présente.

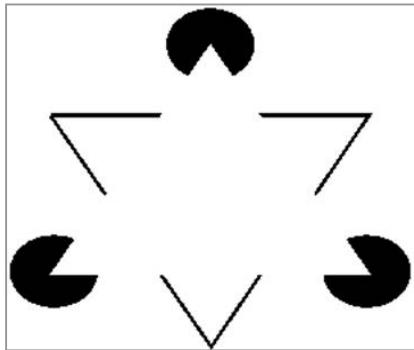


Figure 1-11 : Triangle de Kanizsa (1976). Le triangle blanc semble être complet alors que ces bords ne sont pas définis physiquement. C'est un exemple de compléTION modale. Les disques noirs en arrière du triangle paraissent complets, c'est un exemple de compléTION amodale.

b) Données psychophysiques

L'interposition d'un objet par un autre objet semble être un indice puissant de profondeur. Des expériences ont été entreprises pour tester son influence sur notre impression de profondeur par rapport aux autres indices binoculaires comme la disparité. Les conclusions de ces expériences sont unanimes ; l'interposition domine habituellement la stéréoscopie. Pour le tester, on présente à des sujets humains des stéréogrammes constitués de deux barres, une horizontale et une verticale se superposant (Figure 1-12A). La disparité rétinienne conduit à la perception d'une barre verticale devant la barre horizontale, comme le fait le pattern d'interposition. Mais si ces deux indices sont en conflit (par exemple, la disparité indique que la barre verticale est devant alors que l'interposition indique que la barre horizontale est devant), les sujets perçoivent davantage la barre horizontale en avant de la barre verticale alors que la disparité indique le contraire. L'interposition, indice monoculaire pictural, domine alors la stéréoscopie (Schriever, 1924).

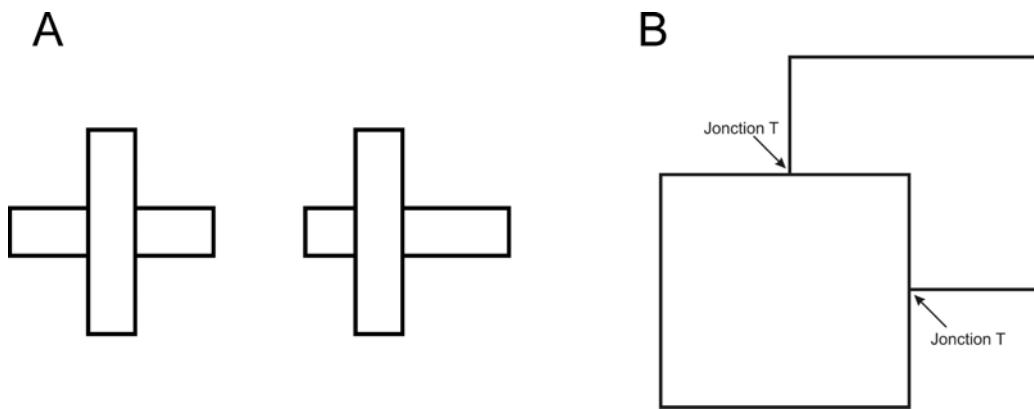


Figure 1-12 : A. Un exemple de stéréogrammes montrant une interposition. L’interposition indique que la barre verticale est en avant de la barre horizontale et cela quels que soient les indices de disparité. B. Deux carrés en complétion modale. Le carré de gauche nous paraît être en avant du carré de droite du fait de la présence des deux jonctions T.

Une interrogation se pose alors sur comment notre système visuel perçoit la profondeur grâce à l’interposition. Un élément important est la bordure de l’objet masqué par l’objet devant lui. Dans la Figure 1-12B par exemple, deux bordures du carré de droite sont cachées par le carré de gauche. Pour chaque objet, les arrangements des bords, coins et surface obéissent à certaines lois regroupées sous le terme “*loi de syntaxe pour les formes*”. Huffman (1971) a élaboré une liste exhaustive des bords et angles qui interviennent dans des phénomènes d’interposition ; on y retrouve les définitions des jonctions T, Y ou L (Huffman, 1971). Dans la Figure 1-12B, les jonctions T sont les indices importants pour nous faire percevoir la forme de gauche en avant de celle de droite.

2.3. *Les ombres et ombrages*

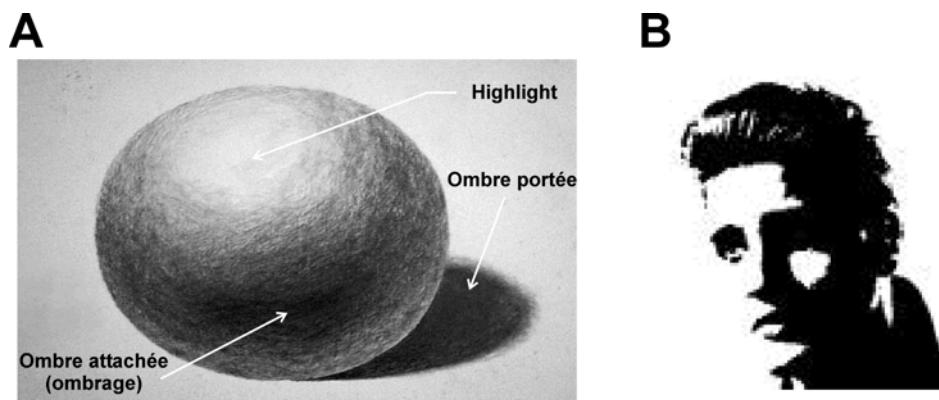


Figure 1-13 : A. Schéma définissant ombres et ombrages. Une sphère est éclairée par au-dessus, par une source diffuse d’illumination créant des ombrages et des ombres portées B. Mooney face représentant le visage d’Elvis Presley (tiré de www.graphics.cornell.edu/.../elvis_thresh.gif). Seuls les ombrages nous font percevoir le visage d’Elvis.

L'indice pictural basé sur l'ombrage et les ombres est un facteur critique de la perception de la profondeur à la fois dans les images et dans la vie quotidienne. Par exemple, sur la Figure 1-13B, le visage d'Elvis dessiné uniquement en noir et blanc nous apparaît tridimensionnel, et cela en l'absence d'autres indices. La disposition des ombrages et des zones éclairées crée cette perception tridimensionnelle. En 1957, un psychologue, Craig Mooney, utilise ce type de stimuli afin d'étudier la perception de visages chez des enfants. Ce type de stimuli, plus connus sous le nom de *Mooney faces* sont souvent utilisés dans des expériences psychophysiques de perception de visages.

a) *Quelques définitions*

(1) Ombres et ombrages

Lorsqu'une scène visuelle est éclairée, la lumière en chaque point de l'image est définie par la nature de la source lumineuse (direction d'illumination et intensité) et la disposition des objets. Cette source lumineuse va générer l'apparition d'ombres portées sur la surface qui l'entoure (plus communément appelée ombres) et d'ombrages sur l'objet lui-même. La Figure 1-13A résume les différents ombres et ombrages ainsi que les termes utilisés pour les définir. Le point de la sphère où la lumière est la plus réfléchie est défini comme le point de réflexion (*highlight*). Autour de ce point se crée un gradient d'illumination, fonction de la position de la source lumineuse et de la surface de l'objet : ce sont les *ombrages* ou *ombres attachées* ou *chiaroscuro*. L'objet éclairé va projeter sur les surfaces avoisinantes des *ombres portées* ou *ombres projetées*. Toutes ces variations de luminosité apportent une perception tridimensionnelle de l'objet. Surface, ombre portée ou ombrage sont autant de facteurs contribuant à cette perception, chacun a été testé séparément dans différentes expériences de psychophysique.

(2) Luminosité, albédo, surface lambertienne ou spéculaire

La luminosité définit en chaque point d'une image un gradient de luminosité dépendant de la nature et de la disposition des objets dans la scène visuelle et des sources lumineuses elles-mêmes. La surface des objets est un élément important pour la qualité du gradient de luminosité de l'objet. Premièrement, une surface peut être définie par sa réflectivité ou sa capacité à réfléchir la lumière incidente. De cette réflectivité découle un paramètre important : l'**albédo** ou “blancheur” de la surface. L'albédo est une mesure de la réflectivité de la surface, il est défini par le rapport entre la quantité de lumière réfléchie et la quantité de lumière incidente. Ainsi, une surface noire吸 une maximum de lumière et

aura un albédo de 0. Une surface nous paraîtra blanche avec un albédo de 0.8, une surface ayant un albédo de 1 étant un miroir parfait.

La réflectivité permet de définir plusieurs types de surface. Le type de réflectivité le plus simple concerne les surfaces lambertiennes. Une surface est dite **lambertienne** si elle réfléchit la lumière de façon égale dans toutes les directions, comme par exemple une feuille blanche de papier. Il n'y a donc pas de point de réflexion (highlight) sur une surface lambertienne contrairement à une surface spéculaire. Une surface est dite **spéculaire** si elle réfléchit la lumière préférentiellement dans une direction par rapport à d'autres directions. La direction préférée est celle pour laquelle l'angle de réflexion est égal à l'angle d'incidence. Ce type de surface crée des points de réflexion (highlight) où se réfléchit préférentiellement la lumière. Un miroir peut être considéré comme un exemple extrême d'une surface spéculaire. Plusieurs études ont démontré que les surfaces spéculaires sont un indice important pour estimer la forme d'un objet (Todd et Mingolla, 1983; Norman *et al.*, 2004). Toutefois, la majorité des expériences sur les ombres et ombrages utilisent plutôt des surfaces lambertiennes afin de limiter la perception tridimensionnelle de la forme de l'objet uniquement au jeu de lumière et d'ombres.

b) Données psychophysiques

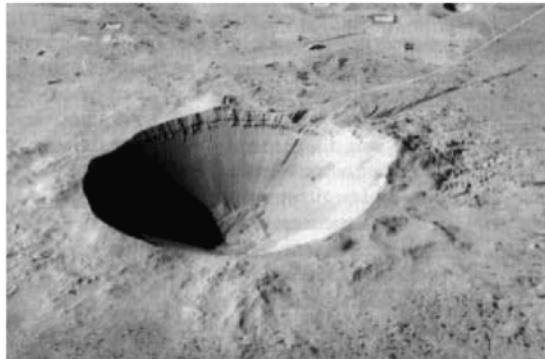
(1) Généralités

L'interprétation du gradient d'ombrage comme une courbure de la surface dépend de plusieurs hypothèses à propos de l'uniformité de l'illumination (source unique) et de la réflectivité de la surface. L'ombrage, seul, est un indice ambigu. Sans une information nous renseignant sur la direction de l'illumination, les ombres ne permettent pas de distinguer des formes concaves de formes convexes. Un exemple de cette ambiguïté est illustré sur la Figure 1-14.

Sur cette photographie de la surface de la Lune, la forme représentée nous paraît concave et nous percevons un cratère. Cette perception tridimensionnelle est due à la position supposée de la source lumineuse. Sur l'illustration A, nous présumons que la lumière provient d'une source lumineuse située en haut à gauche de la photographie ; le versant gauche du cratère apparaît plus sombre que son versant droit. Toutefois, la perception d'un cratère est directement influencée par la direction supposée de l'illumination car ce pattern d'ombres peut être également expliqué en positionnant une source lumineuse sur la droite de la photographie. Le versant droit plus lumineux serait alors un versant d'une forme convexe (une bosse). Un autre argument concernant la position supposée de l'illumination est le fait

que dans l'illustration B, vous percevez une bosse plutôt qu'un creux alors qu'il s'agit de la même illustration mais retournée de 180°. Pourquoi supposons-nous que la source lumineuse provient d'en haut à gauche sur cette photographie, ce qui influence considérablement notre perception tridimensionnelle?

A



B



Figure 1-14 : A. Photographie d'un cratère de la Lune. Sur cette image, nous percevons un cratère, forme concave, car nous supposons que la source lumineuse est située en haut à gauche de la photographie. Ce qui expliquerait que le versant gauche du cratère apparaît plus sombre que son versant droit. B. L'impression de profondeur est totalement inversée, alors qu'il s'agit de la même image mais retournée de 180°.

(2) Direction supposée de l'illumination

Le fait que les humains supposent qu'une direction d'illumination provient d'au-dessus dans l'interprétation d'une image, a été, dans un premier temps, rapporté par Philip Gmelin of Wurtenburgh (1744) dans la Royal society of London et a été repris plus tard par un américain, David Rittenhouse en 1786. Le principe se résume comme suivant : dans une surface plane vue selon un plan vertical frontal, une forme dont la partie supérieure est plus lumineuse que la partie inférieure apparaît convexe alors que l'inverse entraîne la perception d'une forme concave (voir Figure 1-15). Brewster (1826) conclut que les sujets résolvent l'ambiguïté due au pattern d'ombrages en supposant que la source d'illumination vient d'au-dessus (« light from above assumption »). Cette direction supposée de l'illumination a été démontrée chez l'Homme grâce de nombreuses études, notamment les travaux de Ramachandran (Ramachandran, 1988; Kleffner et Ramachandran, 1992).

Peu d'études sur le sujet ont été entreprises chez l'animal, j'en détaillerai une sur le poulet. Pour son étude, W. Hershberger entraîna deux groupes de poussins : le premier à picorer des monticules plutôt que des creux et le second à faire l'inverse, en les renforçant avec de la nourriture quand ils faisaient le bon choix (Hershberger, 1970). Les monticules et les creux sont exempts de toutes ombres ; si bien que la perception de la profondeur ne

dépend que des indices binoculaires ou de la parallaxe. Après cet apprentissage, on présentait ensuite des photographies de monticules et des creux uniquement définis par des ombrages. Si le poussin avait appris à picorer des monticules, il picorait alors les photographies sensées représenter des monticules et se dirigeait vers des photographies où les ombres sont situées en bas. Ceux qui avaient été entraînés à picorer dans des creux, picoraient la photo avec l'ombre en haut. Un résultat surprenant dans cette étude est le fait que la moitié des poussins ait été élevée depuis leur naissance dans des cages où la lumière venait d'en dessous. Si un quelconque apprentissage de la signification des ombres était survenu, il aurait du être en sens opposé de ce qu'il serait en condition naturelle. Les résultats de cette étude sont les mêmes quelle que soit la situation d'élevage des poussins : les poussins qui ont appris se nourrir dans des formes convexes picorent les disques dont la partie supérieure est plus lumineuse. Selon Hershberger, le fait que la lumière vienne d'au-dessus serait un indice inné, génétiquement programmé.

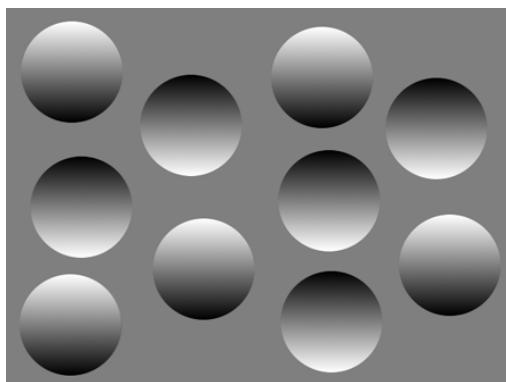


Figure 1-15 : Stimuli utilisés dans les études de Ramachandran (1988). Le pattern d'ombrages induit une impression de profondeur

Chez l'Homme, Yonas et al (1979) ont cherché à déterminer si la position de la source de lumière et /ou la gravité influencent la perception de patterns ombragés chez des enfants de 4 à 8 ans (Yonas *et al.*, 1979). Leurs résultats soulignent que les enfants de 4 ans utilisent comme référentiel la position de leur tête plutôt que le système gravitationnel pour interpréter le relief d'une surface alors que les enfants de 7 ans utilisent à la fois comme référentiel le système gravitationnel et la position de la tête pour interpréter ces patterns ombragés. Enfin, chez les adultes, la perception des ombres et l'hypothèse de la direction de l'illumination sont basées principalement sur la position de la tête, donc grâce aux coordonnées rétiniennes (Howard *et al.*, 1990; Kleffner et Ramachandran, 1992).

Ramachandran (1988) a étudié notre capacité visuelle à voir des formes tridimensionnelles à partir d'ombrages (Shape From Shading ou SFS) (Ramachandran, 1988). D'après lui, les ombres révèlent la nature tridimensionnelle des formes. Pour le tester, il crée

des figures simples, disposées de telle façon que seules de légères nuances dans la répartition des ombres donnent une impression de relief. La Figure 1-15 présente les disques partiellement ombrés utilisés. Ce sont des stimuli ambigus car ils peuvent être perçus comme des formes protubérantes -convexes- ou comme des formes creuses -concaves-. Cette ambiguïté provient du fait que notre cerveau, ignorant la position de la source lumineuse, interprète ces ombres de différentes façons. Les sujets peuvent mentalement modifier le pattern perçu. Une observation importante est que si nous inversons mentalement le relief d'un des disques, le relief de tous les autres s'inverse également. Ramachandran démontre également l'unicité d'une source lumineuse en utilisant un assemblage de deux rangées de disques deux à deux symétriques. Si les disques d'une des rangées apparaissent convexes alors ceux de l'autre rangée semblent automatiquement concaves, cela signifie que notre système visuel considère que la scène est éclairée par une source lumineuse unique.

Kleffner et Ramachandran (1992) décrivent également plusieurs expériences de psychophysique en utilisant ces mêmes stimuli et rapportent d'autres caractéristiques de l'extraction de forme 3D à partir d'ombrages. Ils confirment que l'extraction d'une forme à partir d'ombrage nécessite au moins deux postulats: une seule source lumineuse illumine la scène visuelle et cette source lumineuse éclaire d'en haut (coordonnées rétinianes). Notre système visuel paraît donc plus sensible à un gradient d'illumination vertical (du haut vers le bas). Cette sensibilité se retrouve également lorsque des ombrages sont impliqués dans une perception de mouvement. Les sujets déterminent plus précisément la direction d'un mouvement quand la direction des ombres est verticale plutôt qu'horizontale. Ces expériences (Ramachandran, 1988; Kleffner et Ramachandran, 1992) ont été effectuées chez des humains en position normale et ce postulat pourrait être fonction du référentiel gravitationnel ou de la position de la tête. Qu'en est-il si les sujets ont à juger une forme présentant des ombrages la tête penchée sur le côté ou la tête en bas ?

Howard et son équipe (1990) testent des sujets avec les stimuli de Ramachandran selon trois conditions différentes : (1) la tête bien droite cohérente avec l'axe de gravité, (2) la tête en bas (incohérente avec l'axe de gravité) et (3) la tête penchée de 90° (les axes de la tête et de la gravité sont orthogonaux). Dans cette troisième condition, les sujets ne se basent pas sur l'axe gravitationnel comme référentiel si l'axe défini par les patterns d'ombrage est aligné avec celui de la tête, et inversement. L'axe de la tête comme référence pour déterminer la direction de l'illumination est dominant lorsque les axes de la tête et de la gravité sont opposés. Toutefois, si l'axe de la tête est incohérent avec celui de la gravité, les sujets

montrent un biais sur la direction d'illumination car ils supposent que l'illumination est au-dessus avec l'axe de gravité comme référence. Les adultes utilisent donc l'axe de la tête comme référence et plus précisément l'axe rétinien (Howard *et al.*, 1990).

La perception d'une forme définie par des ombrages dépend donc de la direction de l'illumination supposée dans un référentiel basé sur la position de la tête. Toutefois, Jenkin et al (2004) ont testé si d'autres facteurs peuvent influencer la perception d'une forme à partir d'ombrages lorsque les référentiels définis par la vision, la gravité ou le corps sont modifiés. Ils placent des sujets humains dans un dispositif permettant d'altérer tous ces référentiels (the York Tilted room ; il s'agit d'une pièce construite à 90° par rapport à l'orientation normale où tous les objets de la pièce sont orientés de façon à coïncider avec l'orientation de la pièce). Les sujets se disposent dans la pièce selon différentes orientations afin de démontrer si l'orientation de la pièce (donc du repère visuel), du corps (principalement de la tête) ou de la gravité influe sur la perception des ombrages. Les résultats de cette expérience montrent que la direction d'illumination dépend à la fois des indices basés sur la vision, sur la position de la tête et de la gravité et aucun de ces indices ne domine les autres (Jenkin *et al.*, 2004).

Pour résumer, en l'absence d'une source précise d'illumination, la lumière est supposée venir d'en haut (Ramachandran, 1988; Kleffner et Ramachandran, 1992) et va engendrer, sur la surface des objets, des gradients d'illuminations. Ces gradients d'illuminations créent des ombrages et permettent la perception tridimensionnelle d'un objet sur la base de ces ombrages. Si ce gradient d'illumination a pour direction « au-dessus vers le dessous » alors il contribue à une meilleure perception de la forme tridimensionnelle de l'objet mais si ce gradient se situe dans un sens autre que celui supposé, il n'est pas interprété comme générant des ombrages et ne contribue pas à la perception de la surface de l'objet. La perception d'une forme 3D à partir d'ombrages dépend donc non seulement du gradient de luminance mais aussi de la direction supposée de l'illumination. Toutefois, une étude a démontré l'apparition d'un biais sur la position supposée de la source lumineuse ; elle serait positionnée au-dessus vers la gauche plutôt qu'au-dessus (Sun et Perona, 1998). Les raisons de ce biais vers la gauche restent inconnues mais ce biais peut s'expliquer par le fait que la position de la source lumineuse en haut à gauche est la meilleure pour écrire ou dessiner pour les droitiers. Si la source de lumière se situe juste au-dessus, cela risquerait de générer des ombres gênantes de la main ou de la tête de celui qui a réalisé l'œuvre.

(3) Influence de la direction d'illumination sur la reconnaissance de visages ou d'objets

Deux modèles importants ont été proposés pour expliquer le rôle de la direction de l'illumination sur la reconnaissance des objets ou de visages. Selon un premier modèle, le système visuel pourrait extraire des contours indépendants de l'illumination comme par exemple les traits du visage : la reconnaissance ne serait donc pas affectée par les variations de direction de l'illumination. Cette première hypothèse reprend les théories de Biederman (1987) et de Marr (1978) sur la reconnaissance d'objets (Marr et Nishihara, 1978; Biederman, 1987) (cf paragraphe A.1.3.a). D'autres études proposent un modèle alternatif où la direction de l'illumination serait codée dans la représentation du visage ou de l'objet car il est difficile de ne pas en tenir compte. Ce serait donc une source d'information importante pour reconnaître la forme tridimensionnelle d'un objet (Poggio et Edelman, 1990; Gauthier et Tarr, 1997).

Les conclusions de plusieurs travaux corroborent la seconde hypothèse. En effet, plusieurs études ont démontré des déficits dans la reconnaissance d'objets ou de visages lorsque l'illumination change de direction (Hill et Bruce, 1996; Braje *et al.*, 1998; Tarr *et al.*, 1998a; Troje et Bulthoff, 1998). Toutes ces études ont utilisé un protocole de *match to sample* dans lequel deux stimuli sont présentés de façon séquentielle et les sujets doivent déterminer s'il s'agit du même objet ou pas. Les sujets mettent davantage de temps à reconnaître s'il s'agit d'un même visage ou d'un même objet si la direction de l'illumination varie. Cette augmentation de temps de réaction pourrait refléter le temps nécessaire pour coder une autre direction d'illumination ou pour analyser un pattern différent d'ombrages. Toutefois, plusieurs études signalent une possible implication du changement de présentation du stimulus (*display change*) pour expliquer ces résultats. Un sujet aurait tendance à reconnaître plus rapidement deux objets identiques illuminés par une même direction car leurs représentations ne changent pas alors que les représentations d'objets identiques illuminés selon différents angles varient. Cette différence de représentation pourrait expliquer l'augmentation du temps de réaction lorsque deux objets ou visages sont illuminés différemment. Pour exclure ce possible biais, il faut modifier la position de présentation des stimuli dans un test *match to sample*. Biederman et Bar (1999) ont alors testé ce *display change* dans un test de reconnaissance d'objets selon différents points de vue. Ils ne trouvent pas de moins bonne reconnaissance d'objets si le point de vue change (Biederman et Bar, 1999). Plus récemment, Nederhouser *et al* (2001) ont également testé le rôle de la direction d'illumination sur la reconnaissance d'objets. Ils ont utilisé le même protocole que Tarr et al

(1998) mais en incorporant un changement de position de présentation du second stimulus. Leurs résultats montrent que lorsque la position entre les deux objets diffère, le changement de direction de l'illumination n'altère pas la reconnaissance des objets (Tarr *et al.*, 1998a; Nederhouser *et al.*, 2001).

Ces deux études (Bierderman et Nederhouser) démontrent que la mauvaise reconnaissance des objets ou visages vus ou illuminés différemment décrite dans les études de Braje *et al.* (1998) ou Tarr *et al.* (1998) pourrait être expliquée par ce *display change*. Toutefois, d'autres études, en changeant systématiquement la position de présentation du stimulus, ont démontré des effets de l'illumination ou de la position de l'observateur sur la reconnaissance des objets. Par exemple, Hayward et Williams (2000) ont fait varier le point de vue et la position des deux objets dans leur test et démontrent une influence du point de vue sur la reconnaissance des objets (Hayward et Williams, 2000). De la même manière, Braje et al (1998) dans une tâche de reconnaissance de visages montrent des déficits lorsque la direction d'illumination varie (Braje *et al.*, 1998). Cependant, il est à noter que les variations de position dans ces deux études ne sont pas aussi importantes que celles présentes dans les études de Bierderman (1998, 1999) ou de Nederhouser (2001).

Braje et al (2003) ont conduit une série d'expériences sur la reconnaissance de visages illuminés différemment (Braje, 2003). Dans une première expérience, ils démontrent que les sujets n'éprouvent aucune difficulté à reconnaître un même visage s'il est illuminé différemment lorsque la position de présentation des stimuli varie. Toutefois, ils notent que si les sujets ont à fixer des visages différemment illuminés mais présentés à la même position, la reconnaissance des visages dépend de la direction d'illumination. Ces résultats démontrent l'effet du *display change*. Dans une seconde expérience, ils vont présenter ces mêmes stimuli mais en position parafovéale et ne trouvent pas de déficits dans la reconnaissance des visages. Son étude conclut donc que si les visages sont présentés en périphérie, les sujets n'auront aucun mal à reconnaître s'il s'agit de la même personne même si la direction d'illumination diffère. Notre vision périphérique peut ainsi expliquer le fait que l'effet de l'illumination soit éliminé si la position du stimulus est modifiée. Par contre, un effet de la direction de l'illumination est présent dans notre vision fovéale.

Dans ses séries d'expériences, Braje a également testé l'influence des ombres portées sur la reconnaissance des visages et démontre que la présence d'ombres portées n'affecte pas celle-ci. Il retrouve, ainsi le même type de résultat qu'il avait démontré dans une tâche de reconnaissance d'objets (Braje *et al.*, 2000).

Pour résumer, la direction d'illumination influe la reconnaissance d'un même objet ou visage dans une tâche de dénomination d'objets (Braje *et al.*, 1998; Tarr *et al.*, 1998b). Cette altération peut être expliquée par un *display change*. La position des objets dans notre champ visuel est un facteur important puisque la direction d'illumination n'a pas d'effet sur notre vision périphérique mais a un effet sur notre vision centrale (Braje, 2003)

(4) Ombrages et perception de la profondeur

Les artistes peintres sont devenus des experts dans l'utilisation d'ombres et d'ombrages pour construire des percepts tridimensionnels à partir d'images bidimensionnelles. Ils ont compris bien avant les scientifiques, l'importance des ombres et ombrages dans la perception tridimensionnelle d'un objet. Ils utilisent pour cela la technique de *chiaroscuro*. Les œuvres de Rembrandt (1606-1669) ou de Georges Delatour (1593-1652) exploitent ce jeu entre les ombres et lumière pour donner une impression de profondeur. Grâce à des expériences de psychophysique, des équipes ont cherché à caractériser la précision de jugement de profondeur basé sur les ombrages et ont développé pour cela plusieurs outils. J. Todd décrit ces différentes méthodes dans une revue (Todd, 2004).

Dans une expérience sur l'estimation de la courbure d'une surface, Phillips et Todd (1996) présentent à des sujets des formes 3D simples comme des cylindres et des ellipses. Les sujets doivent choisir parmi un ensemble de courbes 2D celle qui décrit le mieux la courbure de l'objet 3D. Les performances des sujets sont assez faibles avec ce type de méthode (Phillips et Todd, 1996). Dans une précédente expérience, Todd et Mingolla (1983) ont présenté à des sujets des cylindres non texturés illuminés par une source lumineuse positionnée en haut qui va générer plusieurs degrés de courbure (Todd et Mingolla, 1983). Pour chaque cylindre, les sujets jugent le niveau de convexité de la surface à l'aide de courbes 2D et déterminent également la direction de la source lumineuse. Si les ombrages sont sur une surface lambertienne, les sujets sous-estiment la courbure du cylindre mais estiment parfaitement la direction de l'illumination.

Une autre méthode consiste à placer deux points assez proches sur une surface. Les sujets estiment l'état de courbure d'une surface en indiquant lequel des 2 points paraît être le plus avancé. Ainsi, Langer et Bülthoff (2000) demandent aux sujets de juger la profondeur de deux points sur une surface éclairée par une lumière diffuse. En général, si un des deux points se situe dans une partie plus sombre de la surface par rapport à l'autre point, il paraît être plus profond (Langer et Bulthoff, 2000). Toutefois cette stratégie n'est plus appliquée lorsque les

points se situent proche des versants d'une « vallée » car le fond d'une « vallée » est plus lumineux que ses versants.

Jan J. Koenderink et Andrea von Doorn (1992) ont développé une autre méthode d'analyse de la forme 3D qui permet aux sujets de décrire l'orientation des surfaces. Cette méthode est dite « méthode de jauge » (Tissot gauge). Elle consiste à ajuster une ellipse tangentielle à la surface à décrire et à positionner une droite orthogonale à cette ellipse au centre de celle-ci. Koenderink et ses collaborateurs valident cette méthode avec des descriptions de formes tridimensionnelles du même type que celles que nous utiliserons dans le protocole sur les ombres et ombrages (voir Chapitre 2). La description de l'état de courbure de la surface est plus précise avec cette méthode qu'avec les autres méthodes d'estimation de la profondeur (Koenderink *et al.*, 1996a; Koenderink *et al.*, 1996b).

Ces méthodes de jugement permettent de paramétriser la perception de la profondeur sur la base des ombrages et démontrent que le gradient d'illumination sur un objet est un indice important pour la description de sa forme tridimensionnelle. Je ne décrirai pas le rôle des ombres détachées dans la perception de la profondeur mais la revue de P. Mamassian traite de ce sujet (Mamassian *et al.*, 1998).

3. La combinaison d'indices

Une fois la description de chaque indice tridimensionnel faite, nous pouvons nous demander comment vont se combiner tous ces indices pour nous faire percevoir de la profondeur dans la scène visuelle. Il est logique de penser que la combinaison de plusieurs indices permet une meilleure perception tridimensionnelle (Gibson, 1950a). Ces indices semblent également véhiculer des informations dépendantes de la distance de l'objet. Ainsi, pour une distance relativement faible de l'objet, les indices oculomoteurs et la vision stéréoscopique nous renseignent davantage sur la profondeur qu'un indice pictural pris isolément (pattern d'ombrage ou perspective par exemple). Toutefois, la présence de tous les indices picturaux dans une scène visuelle ou dans une image améliore notre perception de la profondeur. Par exemple, si des indices picturaux sont rajoutés à des stéréogrammes, alors la perception de la profondeur sera plus efficace.

C. PRÉSENTATION DU CORTEX VISUEL DU PRIMATE ET DE L'HOMME

Dans les parties précédentes, j'ai décrit les indices 2D et 3D nécessaires à la reconnaissance d'un objet dans notre champ visuel. Tous ces indices perceptifs sont traités par notre cerveau afin de nous permettre de reconnaître l'objet et d'estimer sa forme tridimensionnelle. Cette thèse a pour objectif d'étudier les bases neuronales de deux de ces indices : la texture et les ombrages. Au cours de ce chapitre, je vais donc détailler le fonctionnement du système visuel du macaque, primate non-humain le plus utilisé lors d'expérimentation animale et sujet de mes expériences, car son cortex visuel est relativement proche de celui de l'Homme. Toutefois, certaines données bibliographiques concernent le système visuel humain, notamment pour situer l'aire homologue de V4 au sein du cortex visuel humain. Je détaillerai dans le chapitre suivant le codage neural de la reconnaissance des objets ainsi que le traitement de la profondeur.

1. Introduction

Les premières cartes décrivant l'organisation du cortex visuel chez le singe ont été éditées suite aux travaux de Von Bonin et Bailey (Von Bonin, 1947). Le cortex visuel est constitué de plusieurs sous-régions - aires - comprenant une seule représentation plus ou moins complète du champ visuel. L'organisation visuelle corticale chez le singe macaque, primate non-humain le plus utilisé lors d'expérimentation animale, a pu être explorée grâce à de nombreuses approches. Les cartes se basent sur la connectivité, la cytoarchitecture, la myéloarchitecture, les propriétés fonctionnelles et les organisations rétinotopiques des différentes aires corticales visuelles. Elles sont devenues plus complexes et plus précises au fur et à mesure des recherches scientifiques sur le cerveau simien. Ainsi, de nombreuses revues plus ou moins récentes traitent de la répartition des aires visuelles corticales chez le macaque (Desimone et Gross, 1979; Maunsell et Newsome, 1987; Felleman et Van Essen, 1991). Plus récemment, la technique d'imagerie fonctionnelle (IRMf) a permis de mettre en place une cartographie plus précise des aires visuelles chez le macaque (Brewer *et al.*, 2002; Orban *et al.*, 2004; Van Essen, 2004). Je ne détaillerai pas chacune de ces cartographies mais elles me permettront de mieux définir les aires corticales de la voie visuelle ventrale.

Dans le cerveau du singe macaque, les aires visuelles correspondent aux aires 17, 18 et 19 de Brodmann, l'aire 17 étant l'aire visuelle primaire V1 (aussi appelé cortex strié) et les aires 18 et 19, le cortex visuel extrastrié. Tandis que les limites de V1 sont bien définies (Van

Essen *et al.*, 1984), les limites des aires se situant dans le cortex extrastrié paraissent plus difficiles à définir (Van Essen, 2004). Malgré cette difficulté, trois études proposent une architecture détaillée du cortex visuel (Ungerleider et Desimone, 1986; Felleman et Van Essen, 1991; Lewis et Van Essen, 2000).

Dans leurs travaux, Felleman et Van Essen (1991) ont dénombré près de 30 aires visuelles différentes. Ils se basent sur des données anatomiques et physiologiques pour délimiter les différentes aires tout comme l'avaient fait Ungerleider et Desimone (1986). Le schéma édité par Lewis et Van Essen (2000) diffère puisqu'il se base sur l'analyse architectonique de plusieurs hémisphères, avec une cartographie générée par enregistrement des cartes individuelles. Plusieurs similitudes sont à noter au niveau des cartographies même si les techniques utilisées font apparaître quelques différences. Premièrement, au niveau de la terminologie, des dénominations différentes peuvent décrire une même aire corticale. Par exemple l'aire 17 de Brodmann est équivalente à l'aire V1. Plusieurs régions corticales sont également considérées, par certains auteurs, comme une seule et même aire visuelle alors que d'autres auteurs préfèrent « morceler » cette région en plusieurs subdivisions, en définissant principalement une région ventrale et dorsale (ex : V3v/V3d, V4d/V4v ou LIPd/LIPv).

Durant cette présentation du système visuel du primate, je décrirai succinctement les premiers relais visuels (répine, corps genouillé latéral) et plus en détail les aires visuelles corticales intervenant dans la voie de traitement occipito-temporale : les aires V1, V2, V4 et IT.

2. De l'œil au cerveau

2.1. La rétine et le corps genouillé latéral

L'organisation laminaire de la rétine des primates est semblable à celle des autres mammifères. La couche interne de la rétine est constituée des deux principales catégories de photorécepteurs : les cônes et les bâtonnets. Comme chez l'humain, les cônes sont de trois types différents selon leur propriété d'absorption spectrale. On distingue les cônes avec une absorption spectrale maximale respective pour les longueurs d'ondes correspondantes au rouge, au vert et au bleu (Abramov et Gordon, 1994). Au niveau des couches les plus externes de la rétine se trouvent les cellules ganglionnaires. Ces cellules ganglionnaires se différencient selon leur morphologie et vont projeter vers des sites distincts. Les principales cibles des projections rétiniennes sont le thalamus (le corps genouillé latéral - CGL - et le pulvinar) et le colliculus supérieur. La majorité des projections rétiniennes se font vers le

CGL (90%), le reste vers le colliculus supérieur et le pulvinar, soit de manière directe ou soit de manière indirecte via le colliculus supérieur (Perry et Cowey, 1984; Perry *et al.*, 1984).

Les projections rétiniennes vers le CGL se répartissent dans les différentes couches du CGL décrites par Kaas et ses collaborateurs chez le primate de l'ancien monde (Kaas *et al.*, 1978):

- 2 couches magnocellulaires M dans lesquelles les cellules ont un corps cellulaire de grande taille,
- 4 couches dorsales parvocellulaires P caractérisées par de neurones de taille réduite.

D'autres cellules se répartissent entre les différentes couches magnocellulaires et parvocellulaires et sont regroupées sous le terme « koniocellulaire ». Les efférences du CGL convergent principalement vers l'aire V1 même si d'autres aires corticales (V2, V4) reçoivent également des projections directes du CGL (Wilson et Cragg, 1967). Leur proportion par rapport aux projections du CGL vers V1 est relativement faible (Bullier et Kennedy, 1983).

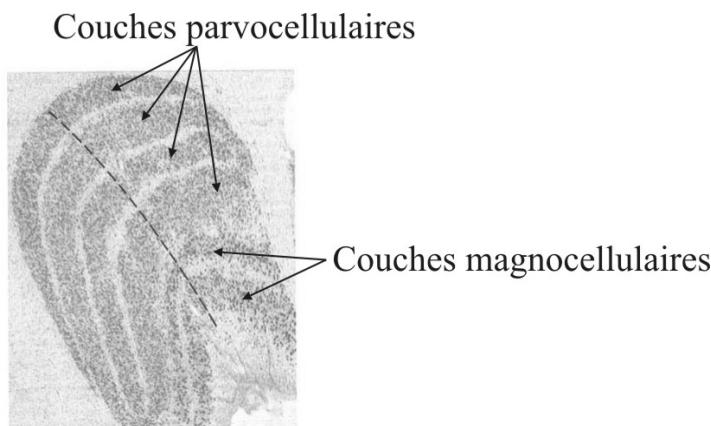


Figure 1-16 : Couches parvocellulaires et magnocellulaires du CGL. (Illustration de Hubel 1995).

2.2. Le cortex visuel primaire (V1)

L'aire V1 est constituée de différentes couches cytologiques parallèles à la surface du cortex ; les couches 1, 2/3, 4B, 4C α , 4C β , 5 et 6. La couche 4 est appelée couche granulaire, les couches 2/3 sont dites supragranulaires et les couches 5 et 6 infragranulaires. La couche 4C reçoit presque la totalité des axones provenant du CGL. La partie ventrale 4C \forall reçoit des projections des couches magnocellulaires alors que les axones issus des couches parvocellulaires projettent vers la partie inférieure de la couche 4C (4C β et la couche 4a). Les neurones de la couche 4C se projettent ensuite sur les couches 2, 3 et 4B d'une part, et sur les couches 5 et 6 d'autre part (Hubel et Wiesel, 1972) (Figure 1-17). Les neurones de la couche

4C sont organisés en colonnes selon leur dominance oculaire⁹. Les couches 2/3 sont caractérisées par la présence de *blobs*, petites structures cylindriques à forte concentration en cytochrome oxydase (Horton et Hubel, 1981). Entre ces blobs, les régions sont définies comme les *interblobs*. Les blobs se situent au centre des colonnes de dominance oculaire. Les couches granulaires et supragranulaires de l'aire V1 présentent donc 3 subdivisions anatomiques importantes: la couche 4B, les blobs et les interblobs.

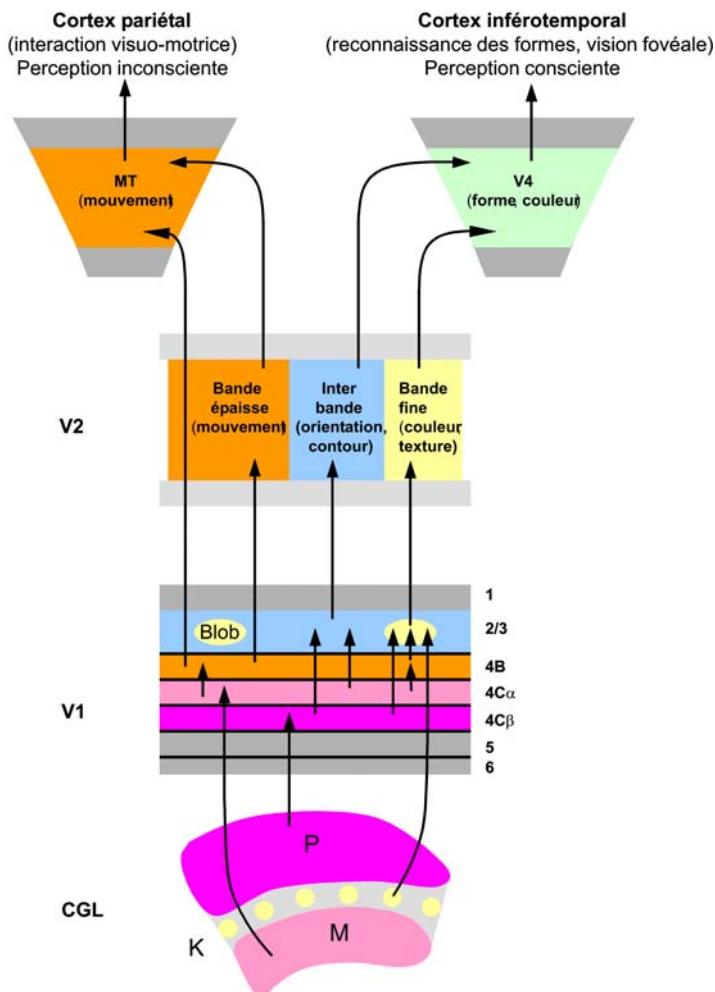


Figure 1-17: Schéma récapitulatif des diverses projections entre le corps genouillé latéral (CGL), les différentes couches de V1, V2 et les aires corticales V4 et MT. (Tirée de (Bullier et Barone, 1997)

Les afférences provenant du CGL sont également organisées de façon rétinotopique. Les parties supérieures du champ visuel correspondent à des cellules projetant ventralement par rapport à la fissure de calcarine alors que les parties basses du champ visuel correspondent à des cellules projetant dorsalement. La surface de la rétine se projette ainsi sur toute la surface de V1. Il existe un facteur important d'amplification corticale, qui fait que la surface

⁹ Dominance oculaire : la couche 4C présente une alternance de bandes de neurones successivement connectées à l'œil droit puis à l'œil gauche mise en évidence par Hubel et Wiesel (1974).

corticale correspondant à la représentation de la fovéa est beaucoup plus importante que celle dévolue à la périphérie.

2.3. L'aire V2

Comme nous avons pu le voir, les connexions semblent être compartimentées au niveau de l'aire V1. L'utilisation d'une technique de coloration (Cytochrome Oxydase, CO) a également montré la présence de subdivisions anatomiques au sein de l'aire V2, sous la forme de bandes fines ou épaisses exprimant fortement la CO. Ces bandes réactives à la CO sont séparées par des interbandes (Tootell *et al.*, 1983). Les connexions entre les aires V1 et V2 se font selon une organisation particulière : la couche 4B de V1 se projette sur les bandes épaisses, les interblobs sur les interbandes et les blobs sur les bandes fines (Livingstone et Hubel, 1984b; Livingstone et Hubel, 1987). Toutefois, cette ségrégation des connexions ne semble pas aussi exclusive puisqu'une étude récente indique que les interblobs se connectent à la fois aux interbandes et aux bandes épaisses (Sincich et Horton, 2005a; Sincich et Horton, 2005b).

Des études en électrophysiologie ont cherché à vérifier si ces subdivisions anatomiques correspondent également à des subdivisions fonctionnelles. Ainsi, au niveau de V1, les neurones des blobs semblent peu ou pas sélectifs à l'orientation mais présentent une sélectivité à la couleur. A l'inverse, les neurones des interblobs sont plus sélectifs à l'orientation mais peu ou pas sélectifs à la couleur (Livingstone et Hubel, 1984b; Ts'o et Gilbert, 1988). Ces spécificités fonctionnelles se retrouvent également au niveau de V2. Les bandes fines regroupent des neurones plus sélectifs à la couleur qu'à l'orientation. Les interbandes présentent des neurones sélectifs à l'orientation (DeYoe et Van Essen, 1985). Les bandes épaisses contiennent des cellules sensibles à la direction d'un stimulus en mouvement et à la stéréoscopie (Shipp et Zeki, 1985; Hubel et Livingstone, 1987). Toutefois, les études plus récentes remettent en cause les spécificités fonctionnelles des subdivisions anatomiques (Kiper *et al.*, 1997; Sincich et Horton, 2005a).

L'existence de ces subdivisions anatomiques et fonctionnelles ainsi que l'étude de leurs connexions ont conduit plusieurs équipes à caractériser différentes voies de traitement des informations visuelles.

3. Les deux voies visuelles

Dans leur travail de synthèse des études portant sur les conséquences de lésions situées soit dans le cortex inférotemporal, soit dans le cortex pariétal, Miskin et al (1983) et

Goodale et Milner (1992) suggèrent l'existence de deux voies de traitement de l'information visuelle appelées voie dorsale - occipito-pariétale - et voie ventrale - occipito-temporale - (Mishkin *et al.*, 1983; Goodale et Milner, 1992). Les lésions du cortex pariétal entraînent des déficits au niveau de la localisation spatiale des objets, alors que les lésions du cortex inférot temporal sont la cause de mauvaises identifications et reconnaissances des objets. Le système visuel se partagerait donc en deux systèmes : une voie ventrale traitant les informations relatives aux propriétés intrinsèques des objets et une voie dorsale pour le traitement des informations spatiales relatives aux objets.

De nombreux travaux vont confirmer l'existence de ces deux voies grâce à l'étude des connexions cortico-corticales chez le primate (Morel et Bullier, 1990; Baizer *et al.*, 1991). La voie dorsale regroupe les aires V1, V2, MT, V3A, PO puis les aires du cortex pariétal postérieur (MST, LIP, VIP) et de la partie rostrale du sillon temporal supérieur (STS). La voie ventrale débute également par les aires V1 et V2 mais se prolonge avec les aires V3 et V4 puis TEO et les différentes aires du cortex inférot temporal. Ces deux voies ne sont pas complètement indépendantes puisqu'il existe des connexions entre les différentes aires de ces deux voies : les aires MT et V4 sont, par exemple, reliées entre elles (Maunsell et van Essen, 1983). Cependant, il semble que la proportion de ces connexions inter-voies soit moindre par rapport aux connexions intra-voies (Morel et Bullier, 1990).

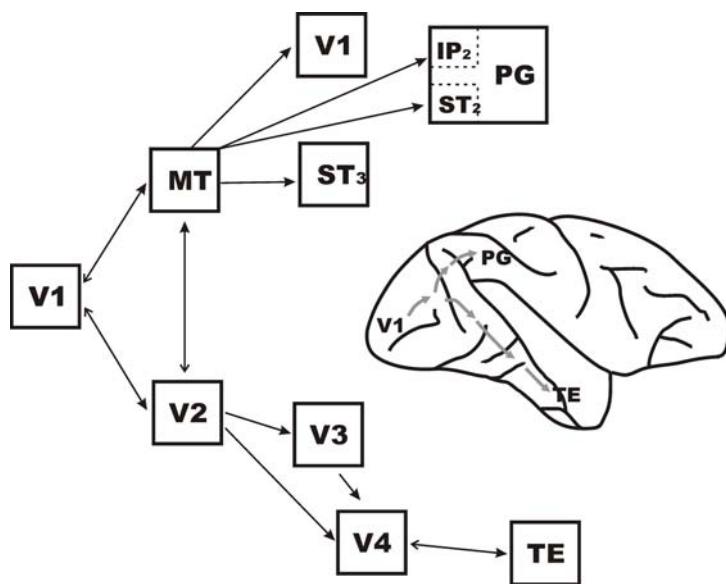


Figure 1-18: Les deux voies visuelles d'après Ungerleider (1985).

4. L'aire V4¹⁰

L'aire V4 occupe une position intermédiaire dans la voie visuelle ventrale entre le cortex inférotemporal et l'aire V2. La voie ventrale a pour fonction principale la reconnaissance des objets. Tous les attributs nécessaires à l'identification de l'objet (couleur, forme...) sont donc traités le long de cette voie. De nombreux travaux ont donc été effectués sur les bases neuronales de ces attributs. Ils feront l'objet d'une revue bibliographique dans la partie suivante (voir paragraphe D) mais auparavant intéressons-nous à la position anatomique de l'aire V4 au sein du cortex visuel simien et humain, ses caractéristiques rétinotopiques ainsi que sa connectivité avec les autres aires visuelles.

a) Chez le singe

L'aire V4 a été initialement décrite chez le singe de l'ancien monde dans les travaux de Cowey (Cowey, 1964), Zeki (Zeki, 1969; Zeki, 1973b; Zeki, 1974b) et Almann et Kaas (Allman et Kaas, 1971; Allman et Kaas, 1974). Toutes ces études s'accordent sur le fait que l'aire V4 se situe entre les sillons luné, inféro-occipital (IOS) et temporal supérieur (STS) (Figure 1-19). Chaque hémisphère cérébral contient une aire V4.

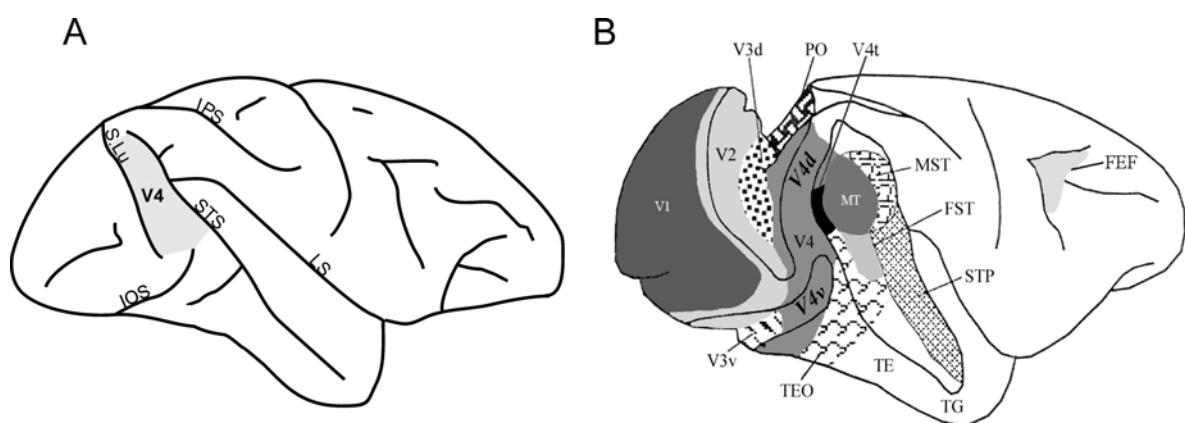


Figure 1-19: A. Schéma d'une vue latérale d'un cerveau de Macaque. IPS : Sillon inféro-pariétal, STS : Sillon temporal supérieur, IOS : Sillon inféro-occipital, LS : Sillon latéral, S.Lu : Sillon luné. La limite dorsale de V4 n'est pas clairement définie. B. Aires visuelles chez le Macaque, vue latérale, schéma adapté de (Bullier *et al.*, 1994). Les deux sous régions V4v et V4d y sont représentées.

En se basant sur les connexions anatomiques et sur les propriétés des champs récepteurs des cellules, V4 se dissocie en deux subdivisions contiguës (V4 dorsale : V4d et V4 ventrale : V4v). V4d est bordée postérieurement par V2, V3d et V3A, et au niveau

¹⁰ Durant mon doctorat, j'ai travaillé sur l'aire visuelle V4. Cette partie sera beaucoup plus détaillée que les descriptions des autres aires visuelles car V4 est le site exclusif de mes enregistrements cellulaires.

antérieur par V4t. V4v est bordée au niveau postérieur par V2 et V3v et au niveau antérieur par TEO (Boussaoud *et al.*, 1990; Felleman et Van Essen, 1991). Zeki a, au cours de ses multiples travaux, isolé d'autres sous-structures de V4 comme – V4 complexe – et -V4A- (Zeki, 1977; Shipp et Zeki, 1995). Ces termes sont rarement utilisés dans les expériences faites sur V4 et les auteurs parlent généralement de V4 en distinguant seulement les deux sous parties ; V4v et V4d. Quant à l'aire V4t (t pour transition entre V4 et MT), elle a été décrite par Desimone et Ungerleider comme une partie de V4 latérale, postérieure à MT dans le sillon temporal supérieur (Desimone et Ungerleider, 1986). Elle reçoit des connexions directes de MT et a une myélinisation relativement claire et distincte de V4. L'aire V4t semble donc être une aire distincte de V4 et correspondrait à une partie de l'aire sensible à la couleur décrite par Zeki en 1977 d'après Gattass (Zeki, 1977; Gattass *et al.*, 1988).

b) Chez l'Homme

(1) V4v

Bien qu'il soit difficile de faire des homologies entre la neuroanatomie humaine et celle du singe, les progrès en neuroimagerie fonctionnelle ont permis de mieux caractériser le cortex visuel humain. Plusieurs aires paraissent homologues à celles identifiées chez le macaque. Cette homologie se réfère à divers critères : (i) les propriétés fonctionnelles, (ii) la rétinotopie (iii) les connexions intracorticales (iv) les données histologiques et biochimiques (v) la topographie. En se basant, sur au moins trois de ces critères, plusieurs aires visuelles corticales ont pu être déclarées homologues chez l'Homme et le singe : V1, V2, V3 (V3d), Vp (V3v) et V3A (Sereno *et al.*, 1995; DeYoe *et al.*, 1996; Hadjikhani *et al.*, 1998; Wandell *et al.*, 1999). Les deux dernières études proposent également que V4v soit présente au sein du cortex visuel chez l'Homme, relativement proche de celle décrite chez le singe. Cette hypothèse semble se confirmer puisque la rétinotopie de V4v est équivalente chez les deux espèces (Fize *et al.*, 2003). D'autres travaux se sont intéressés aux propriétés fonctionnelles de V4 observées chez le singe afin de définir son homologue chez l'Homme. Zeki avait défini l'aire V4 du singe comme un centre de la couleur. La localisation d'une aire homologue chez l'Homme a donc été étudiée chez des patients achromatopsiques ou, en imagerie cérébrale, dans des tâches de perception de patterns colorés. Trois études en tomographie par émission de positons (TEP) montrent une activation de l'aire occipitale ventromédiane dans des tâches de perception de couleur chez les humains (Corbetta *et al.*, 1991; Zeki *et al.*, 1991; Gulyas et Roland, 1994). Cette structure a été déclarée homologue à l'aire V4 (Zeki *et al.*, 1991). Les lésions corticales causant une achromatopsie correspondent à la localisation de zones activées

pendant une stimulation colorée (Damasio *et al.*, 1980). La structure corticale lésée se trouve entre le gyrus fusiforme et le gyrus lingual et correspond à la région activée dans les études de neuroimagerie.

(2) V4d

Concernant V4d, des interrogations subsistent sur sa localisation chez l'Homme et plusieurs terminologies ont été utilisées pour définir une aire homologue à V4d chez l'humain. En IRMf, Malach *et al* (1995) ont isolé une aire LO - Lateral Occipital - (Malach *et al.*, 1995). Comme V4d, cette région est située postérieurement à l'aire MT et s'active lorsque des images d'objets sont présentées aux sujets. Des résultats similaires ont été observés dans des études en TEP (Kanwisher *et al.*, 1996). Dans des études plus récentes, l'aire LO est considérée comme une subdivision dorso-caudale d'une aire plus complexe, l'aire LOC - Lateral Occipital Complex - (Grill-Spector *et al.*, 1999). L'autre subdivision ventro-latérale de LOC située dans le gyrus fusiforme appelée pFs/Loa recouvrirait l'aire FFA - Fusiform Face Area - décrite par Kanwisher et son équipe comme une aire sensible aux visages (Kanwisher *et al.*, 1997). L'activation de LOC semble invariante à différents points de vue des objets et aux variations de position ou de taille des objets (Grill-Spector *et al.*, 1999; Kourtzi et Kanwisher, 2000). Cette aire serait donc une aire de haut niveau spécialisée dans les représentations d'objets (Grill-Spector *et al.*, 2001). Ces caractéristiques fonctionnelles, décrites chez l'Homme, suggèrent qu'elle serait plutôt l'homologue de l'aire IT du macaque (Malach *et al.*, 1995; Tootell *et al.*, 1996).

Plus récemment, l'étude entreprise par Tootell et Hadjikhani (2001) porte plus particulièrement sur la localisation de l'homologue de V4d chez l'Homme en IRMf (Tootell et Hadjikhani, 2001). Ces auteurs définissent la rétinotopie, la topographie et la fonction d'une aire homologue chez l'Homme qu'ils nomment - V4d topologue - en se basant sur la topologie faite chez le macaque rhésus. Selon eux, l'aire V4d topologue regrouperait les aires LO centrale – LOc - et LO périphérique – LOp -. LOc aurait une représentation centrale du champ visuel et LOp, une représentation périphérique.

D'autres études définissent une autre région homologue à l'aire V4d, l'aire LOS (Orban *et al.*, 1999; Denys *et al.*, 2004). Cette région serait une sous partie de l'aire LO décrite initialement dans l'étude de Malach (1995), postérieure au LOC. Plusieurs études suggèrent que l'aire LOS présente plusieurs spécificités fonctionnelles différentes. Elle serait sensible au mouvement et à la forme de l'objet (Murray *et al.*, 2003; Orban *et al.*, 2003). Les sous-régions LOS sensibles à ces deux propriétés ne sont pas coextensives (Denys *et al.*,

2004). Une autre sous-région KO - Kinetic Occipital - particulièrement sensible à la définition de bordures en mouvement fait également partie de la région LOS et constituerait une subdivision fonctionnelle.

En résumé, il semble que la subdivision ventrale de V4 se retrouve chez les deux espèces car elles ont une même organisation rétinotopique et des propriétés fonctionnelles similaires (Zeki *et al.*, 1991; Fize *et al.*, 2003). La partie dorsale de l'aire V4 aurait évolué différemment entre les singes et les humains ; l'organisation rétinotopique diffère, par exemple, entre les deux genres. Plusieurs études ont donc cherché à définir une aire V4d homologue chez l'Homme sans s'accorder sur la terminologie. Pour Tootell et Hadjikhani, l'aire V4d humaine se rapprocherait de l'aire V4d topologue qu'ils décrivent dans leur étude alors que pour d'autres auteurs, l'aire LOS (subdivision de l'aire LO initialement décrite par Malach) serait une aire potentiellement homologue (Orban *et al.*, 2003; Denys *et al.*, 2004; Orban *et al.*, 2004).

4.2. Rétinotopie

a) Rétinotopie chez le singe

L'aire V4 est subdivisée en régions dorsale et ventrale dans lesquelles nous retrouvons une représentation rétinotopique des champs visuels inférieur et supérieur, respectivement. L'organisation rétinotopique de l'aire dorsale de V4 paraît être plus complexe et variable suivant les études (Maguire et Baizer, 1984; Gattass *et al.*, 1988; Boussaoud *et al.*, 1990; Van Essen *et al.*, 1990).

(1) V4v

L'aire V4v contient une représentation de l'hémichamp visuel supérieur controlatéral. L'organisation rétinotopique est relativement simple puisque les représentations des méridiens verticaux et horizontaux coïncident parfaitement avec les deux bordures antérieure et postérieure de V4v. La représentation du méridien vertical délimite la bordure postérieure, entre V3 et V4v ; la représentation du méridien horizontal délimite la bordure antérieure de l'aire V4v et sépare l'aire V4v de l'aire TEO (Boussaoud *et al.*, 1991).

(2) V4d

La rétinotopie de V4d a été décrite dans peu d'études et les différents auteurs montrent certains points de désaccord sur celle-ci surtout concernant la représentation des méridiens verticaux et horizontaux. L'étude électrophysiologique ayant traité de ce sujet la plus souvent citée est sans aucun doute celle réalisée par Gattass et ses collaborateurs en 1988 (Gattass *et*

al., 1988). La Figure 1-20 montre une cartographie de V4 provenant de cette étude, en vue aplatie, les sillons étant ouverts. Les auteurs décrivent une représentation fovéale latéralement, au niveau de la bordure entre V4d et V2. La représentation du méridien vertical suit le fond du sillon luné et délimite la séparation entre V3A et V4d. La représentation du méridien horizontal sépare l'aire V4d de V4T.

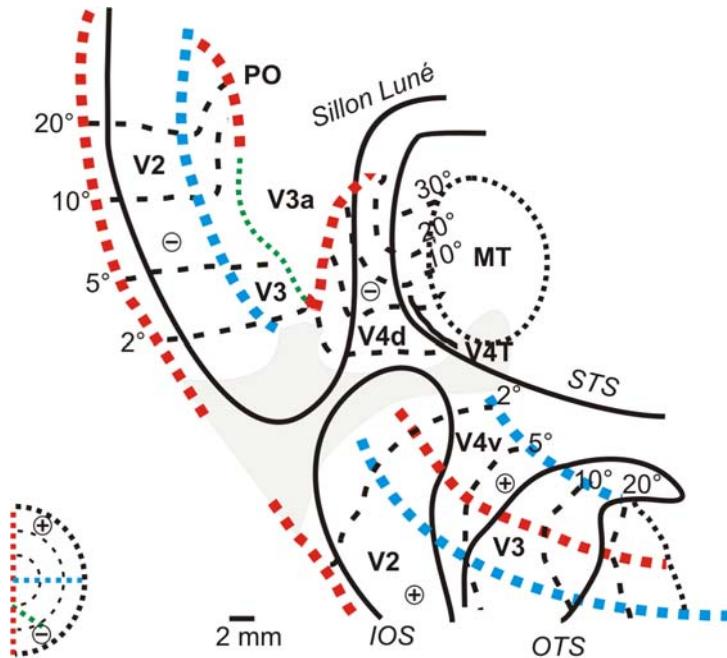


Figure 1-20: Carte rétinotopique des aires visuelles V2, V3, V3A et des parties ventrale et dorsale de V4, décrivant les représentations des champs visuels supérieur (+) ou inférieur (-), le méridien horizontal (rouge), le méridien vertical (bleu). Le bord antérieur de la partie dorsale de V3 est représenté en vert. Schéma édité à partir des cartes de Gattass et al (1988) et de Fize et al (2003)

Fize et son équipe se sont également intéressés à la rétinotopie de l'aire V4 chez le macaque vigile mais en utilisant l'IRMf (Fize *et al.*, 2003). Ils se sont attachés à mieux définir le bord antérieur de V4d car c'est sur cette limite que la controverse existe. Ils observent une transition entre les méridiens horizontaux et verticaux correspondants aux résultats décrits par Gattass mais pas à ceux publiés par Brewer chez le singe anesthésié en IRMf (Brewer *et al.*, 2002). En effet, ils décrivent la présence de la représentation du méridien horizontal au niveau de la limite rostrale de V4d, résultat en accord avec ceux publiés dans deux précédentes études (Van Essen et Zeki, 1978; Maguire et Baizer, 1984).

L'organisation rétinotopique des deux sous régions ventrale et dorsale de V4 diffère donc chez le macaque (Fize *et al.*, 2003) alors que chez le singe du nouveau monde, les deux subdivisions présentent une même organisation rétinotopique (Pinon *et al.*, 1998). De toute évidence, les parties dorsale et ventrale de V4 ont donc évolué différemment pendant l'évolution des espèces. Tootell et Hadjikhani, de par les asymétries présentes entre V4v et V4d, suggèrent que ces subdivisions puissent correspondre à deux aires visuelles distinctes (Tootell et Hadjikhani, 2001).

b) Rétinotopie chez l'Homme

L'aire V4v chez l'Homme semble avoir une même organisation rétinotopique et topographique que celle observée chez le macaque (Sereno *et al.*, 1995; DeYoe *et al.*, 1996; Tootell *et al.*, 1997; Tootell et Hadjikhani, 2001). Mais aucune étude ne montre une homologie rétinotopique entre l'aire V4d du macaque et d'autres aires chez l'Homme. Par exemple, L'aire LOS humaine et V4d simienne diffèrent dans leur organisation rétinotopique (Orban *et al.*, 1999; Denys *et al.*, 2004). La région LO semble avoir une organisation rétinotopique bien définie ; elle coderait pour le champ visuel inférieur (Grill-Spector *et al.*, 2000). Une région de LO s'active pour des stimuli centraux (LOc) et une autre région pour des stimuli périphériques (LOp) (Mendola et Corkin, 1999; Tootell et Hadjikhani, 2001). Quant à l'aire LOC, elle ne présente pas d'organisation rétinotopique particulière. Elle est activée à la fois par des stimuli dans le champ controlatéral ou dans le champ ipsilatéral (Grill-Spector *et al.*, 1998). Seule l'aire « V4 topologue » mentionnée dans l'étude de Tootell et Hadjikhani est décrite selon des homologies observées au niveau de l'organisation rétinotopique chez le singe (Tootell et Hadjikhani, 2001).

4.3. Les connexions corticales

L'étude des connexions cortico-corticales a également permis de considérer l'aire V4 comme une aire à part entière du cortex extra-strié du macaque.

a) Les différents types de connexions

Les projections entre les différentes aires corticales sont classées en plusieurs groupes :

- Connexions en avant (feedforward, FF): elles acheminent l'information d'une aire corticale hiérarchiquement moins élevée vers une aire de plus haut niveau (ex : V1 vers V4). Elles sont responsables du transfert de l'information provenant de la rétine sur l'ensemble des aires corticales visuelles. L'inactivation d'une région de V1 entraîne, par exemple, un arrêt ou une forte baisse d'activité des neurones d'autres aires corticales supérieures dont les CR sont situés dans la zone représentée par la région inactivée (Girard et Morrone, 1995).
- Connexions en retour (feedback, FB) : les informations sont véhiculées d'une aire de haut niveau vers une aire bas niveau et auraient une fonctionnalité différente que celles véhiculées en FF. Selon Hupé et al (1998), elles pourraient notamment être le support de la segmentation des objets par rapport au fond (Hupe *et al.*, 1998).

Ces deux types de connexions (FF et FB) diffèrent également au niveau de leur organisation laminaire. Les neurones des couches 1/2/3 (couches supragranulaires) sont à l'origine des connexions FF projetant dans la couche 4 des aires de plus haut niveau. Les neurones des connexions FB sont issus des couches infragranulaires (5 et 6) et projettent en dehors de la couche 4.

- Connexions latérales : elles s'effectuent entre des aires se situant à un même niveau hiérarchique, des couches 2/3 et 5/6 vers toutes les autres couches.

b) Hiérarchie des aires visuelles

La description de ces projections a permis de proposer plusieurs modèles d'organisation hiérarchique des aires visuelles. Une organisation hiérarchique postule qu'il existe des niveaux successifs de traitement de l'information visuelle. La distinction entre les niveaux se fait d'après le type de connexions (FB ou FF) échangées entre les aires. Une aire envoie des projections de type FF vers des aires d'un niveau supérieur et en reçoit de type FB. Au même niveau, les aires corticales échangent des connexions latérales. Ce concept d'une organisation hiérarchique des aires corticales visuelles a été proposé, dans un premier temps, par Maunsell et Van Essen. Leur système hiérarchique comprend 13 aires corticales organisées selon 6 niveaux hiérarchiques. Au premier niveau se situe V1, puis au second V2 puis V3, VP. Les aires V4 et MT occupent le quatrième niveau ; les aires 7a, MST, VIP, Poa et IT le cinquième, et enfin l'aire FEF le dernier (sixième) niveau. L'augmentation de la taille et la complexité des champs récepteurs le long de cette organisation confirmeraient cette hiérarchie (Maunsell et van Essen, 1983).

Lorsque de nouvelles aires corticales et leurs connexions anatomiques ont été identifiées, ce schéma d'organisation hiérarchique a été complété. Ainsi, Felleman et Van Essen répertorient une trentaine d'aires corticales chez le primate et proposent une organisation hiérarchique plus élaborée (Figure 1-21) (Felleman et Van Essen, 1991). Ce modèle situe les aires corticales à d'autres étages hiérarchiques que le système précédent de Maunsell et van Essen. Par exemple, les aires V4 et MT y occupent le cinquième rang.

Cette répartition arbitraire a conduit le groupe de M. Young à proposer, par une approche computationnelle, une hiérarchie plausible parmi les 150 000 combinaisons possibles issues des données de Felleman et van Essen (Hilgetag *et al.*, 1996). L'aire V4 y occuperait le 5^{ème} rang. Par la suite, plusieurs études apportent des modifications dans les niveaux hiérarchiques. Par exemple, Barone et ses collaborateurs placent les aires V4 et MT à des niveaux différents (respectivement 5^{ème} et 6^{ème} rang) (Barone *et al.*, 2000).

La Figure 1-21 reprend l'organisation des aires corticales et décrit les différentes connexions. Ce modèle hiérarchique présente toutefois certaines imperfections, provenant du fait qu'il ne tient pas compte des aspects fonctionnels ou de l'aspect temporel du traitement de l'information visuelle. En effet, l'étude des latences des réponses des neurones visuels des différentes aires remet en cause cette hiérarchie (Bullier et Nowak, 1995; Nowak *et al.*, 1997). Ce point sera discuté dans le chapitre 5. Dans les paragraphes suivants, nous nous intéresserons aux connexions de V4 avec les autres aires visuelles.

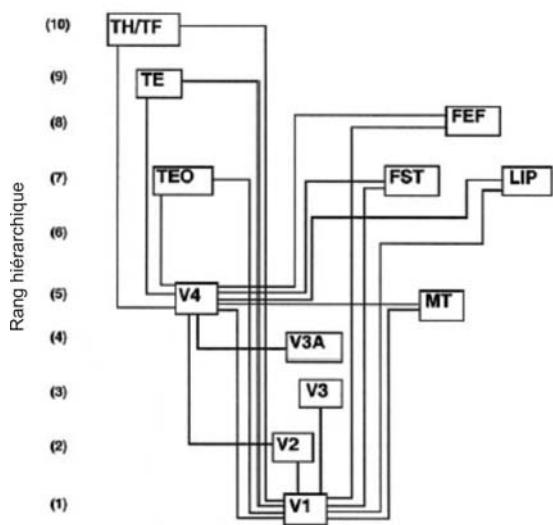


Figure 1-21: Hiérarchie des aires visuelles chez le primate. Schéma édité par Felleman et Van Essen (1991). V4 occupe le cinquième rang de cette hiérarchie.

c) Projections afférentes à l'aire V4 chez le macaque

Dans cette description des projections afférentes à V4, je ne décrirai que les projections cortico-corticales principales au sein de la voie visuelle ventrale mais il est à noter que V4 reçoit également des projections directes de structures sous-corticales telles que le pulvinar ou le CGL (Fries, 1981; Shipp, 2001).

(1) Connexions FF

Les projections FF afférentes à V4 sont issues par définition des aires visuelles de plus bas niveau : V1, V2, V3A. L'aire V1 ne projette vers V4 qu'au niveau de la représentation de la fovéa et en très faible proportion (Zeki, 1978a; Yukie et Iwai, 1985). Les projections de V1 vers V4 ont également été étudiées par Nakamura et ses collaborateurs avec l'injection de deux différents traceurs rétrograde dans V4 et dans TEO (Nakamura *et al.*, 1993). Les neurones projetant vers V4 se retrouvent uniquement dans la représentation de la fovéa de V1, à la fois dans les blobs et les interblobs.

Les principales projections FF sur V4 proviennent de l'aire V2, issues à la fois des bandes fines et des interbandes (DeYoe et Van Essen, 1985; Shipp et Zeki, 1985; Shipp et

Zeki, 1989). Les subdivisions de V2 (bandes fines et interbandes) projettent massivement vers des compartiments distincts de V4 même s'il est possible que des projections issues des interbandes de V2 projettent vers un compartiment majoritairement destiné aux bandes fines; l'inverse a également été démontré (Xiao *et al.*, 1999). De même, Felleman et son équipe remarquent une répartition ségrégée des cellules de V2 projetant vers V4. Ils rapportent la présence de mini clusters de neurones marqués à l'intérieur même d'une seule bande fine (Felleman *et al.*, 1997b). Des subdivisions de V4 reçoivent donc des projections de sous compartiments spécifiques des bandes fines de V2.

Quant aux projections issues de V3A, les couches supragranulaires semblent massivement projeter vers V4 alors qu'il semble que l'aire V3 projette peu ou pas du tout vers V4 (Van Essen *et al.*, 1986; Felleman *et al.*, 1997a).

L'aire FEF projette également vers V4 mais une controverse existe sur le caractère FF ou FB de ces projections. D'après Felleman et Van Essen, les projections issues de FEF vers V4 ont été classées comme des connexions FB mais cette classification s'est faite en l'absence de données quantitatives. L'étude de Barone et al (2000) dénombre une quantité importante de neurones rétromarqués dans les couches supragranulaires du FEF suite à des injections dans V4 suggérant davantage la présence de connexions FF vers V4 (Felleman et Van Essen, 1991; Barone *et al.*, 2000).

V4 reçoit à la fois des projections parvocellulaires et magnocellulaires (Ferrera *et al.*, 1992). Dans une étude ultérieure, Ferrera et al., (1994) entreprennent de déterminer la part respective des entrées en provenance de ces deux principales voies (Ferrera *et al.*, 1994). Ils ont bloqué sélectivement projections parvocellulaires ou magnocellulaires vers V4 à l'aide d'injections de GABA. Ils observent une diminution de 40% de la réponse moyenne des neurones de V4 à un stimulus visuel dans les deux cas. Un blocage de chacune des couches M ou P du CGL peut diminuer les réponses d'une même cellule de V4, suggérant que les projections P ou M dans V4 convergent.

(2) Connexions FB

Il est admis que l'aire V4 reçoit des projections FB dont la proportion est plus importante que les entrées FF de l'aire V1 (Morel et Bullier, 1990; DeYoe *et al.*, 1994; Barone *et al.*, 2000). Ces projections FB sont principalement issues du cortex inférot temporal (aire TE et TEO), de l'aire MT, de l'aire FST, de l'aire FEF, des aires TH/TF et du cortex pariétal (FST et LIP) (Felleman et Van Essen, 1991; Rockland *et al.*, 1994). Les plus grandes proportions des connexions feedback sont issues du cortex inférot temporal ; l'aire TEO, 18%

des projections afférentes de V4 et l'aire TE, 11%. A une proportion moindre (10%), l'aire V4 reçoit des projections de MT. Les projections des autres aires n'excèdent pas les 5% ; 4% pour les aires TH/TF, 1-2% pour l'aire FST, 1-2% pour l'aire LIP et -1% pour le FEF (Kennedy et Barone, 1999).

d) Projections efférentes de V4

(1) Connexions FF

L'aire V4 projette principalement vers des aires visuelles de haut niveau situées dans le cortex inférotemporal, l'aire TEO et l'aire TE (Desimone *et al.*, 1980; Boussaoud *et al.*, 1991). L'aire TEO vers laquelle V4 projette massivement se trouve au niveau de la partie temporo-occipitale du cerveau du singe (Boussaoud *et al.*, 1991). Elle est aussi appelée PIT par Felleman et Van Essen mais il n'est pas certain que ces deux aires coïncident exactement (Felleman et Van Essen, 1991). L'aire V4 projette également vers l'aire FST (Boussaoud *et al.*, 1990), l'aire LIP et les aires TH/TF (Felleman et Van Essen, 1991).

(2) Connexions FB

Des connexions FB de V4 ont été étudiées dans l'étude de Rockland et de ses collaborateurs grâce à des injections de traceurs rétrogrades dans V4 (Rockland *et al.*, 1994). Ces injections mettent en évidence plusieurs zones de marquages à travers le cortex, situées dans les aires V1, V2 et V3. La majorité des connexions FB projettent vers V2, il y a en moins vers V3 et V1. Certains neurones issus de V4 projettent à la fois vers V2 et V1. Les projections de V4 vers V2 ne sont pas restreintes à une sous-unité fonctionnelle particulière de V2 mais sont réparties à travers les bandes fines et épaisses de V2 (Zeki et Shipp, 1989). Quant aux projections vers V1, elles représentent approximativement 10% des neurones des couches supragranulaires (Barone *et al.*, 2000).

L'étude des connexions de V4 indique que cette aire occupe une position intermédiaire au sein de la voie visuelle ventrale ; les principales projections de V4 proviennent de l'aire V2 et l'aire V4 projette massivement vers le cortex inférotemporal (aire TEO). Toutefois, V4 n'est pas uniquement connectée aux aires visuelles ventrales et présente également des projections FF, FB ou latérales avec des autres aires de la voie dorsale notamment les aires MT, FST et LIP. Toutefois, leurs proportions sont minimes en comparaison avec les projections des aires ventrales (Kennedy et Barone, 1999).

5. Le cortex inférotemporal

Nous savons que le cortex inférotemporal constitue l'étape finale de la voie visuelle ventrale et que l'aire V4 projette massivement vers l'aire TEO. Pour ces 2 principales raisons, il m'est apparu nécessaire de décrire la position anatomique du cortex inférotemporal ainsi que ces principales connexions.

5.1. Position anatomique

Le cortex inférotemporal (IT) s'étend du sillon inféro-occipital (IOS) jusqu'à l'extrémité du STS. Le cortex inférotemporal regrouperait les aires 20 et 21 décrites par Brodmann et correspondrait à l'aire TE décrite initialement par Von Bonin (Von Bonin, 1947). Cette aire TE sera divisée par la suite en 2 sous-régions : une région antérieure, l'aire TE et une région plus postérieure l'aire TEO. L'aire TEO a été abondamment décrite par l'équipe de Boussaoud (Boussaoud *et al.*, 1991; Distler *et al.*, 1993).

En se basant sur la topographie et l'organisation laminaire des projections, Felleman et Van Essen (1991) divisent IT en 3 aires : PIT, CIT et AIT (Felleman et Van Essen, 1991). Chacune de ses sous-parties présente une région ventrale et une région dorsale. L'aire PIT correspondrait à l'aire TEO et les aires CIT et AIT à l'aire TE.

5.2. Connexions corticales

Les études du groupe de Distler et de Boussaoud décrivent les connexions de l'aire TEO (Boussaoud *et al.*, 1991; Distler *et al.*, 1993). L'aire TEO reçoit des projections FF principalement des aires V2, V3 et V4. Ces projections sont organisées topographiquement; les représentations du champ visuel supérieur, inférieur, central ou périphérique sont connectées avec les représentations correspondantes des aires inférieures. Les projections FF issues des aires V3A, V4t et MT sont plus rares. Chacune de ces aires reçoit également des projections FB de TEO (Distler *et al.*, 1993; Rockland, 1994). L'aire TEO projette également vers les aires TEm, TEa et IPa, lesquelles projettent en FB vers TEO. Les aires TEm, TEa et IPa et leurs connexions ont été décrites dans les études de Seltzer et Pandya (Seltzer et Pandya, 1978). Elles se situent le long du sillon temporal supérieur (STS). Ces sous-régions du STS présentent des caractéristiques physiologiques qui leur sont propres (Baylis *et al.*, 1987). L'aire TEO reçoit également des connexions FB de l'aire parahippocampique (TH) et des aires TG et 36 (Distler *et al.*, 1993). L'aire TEO projette également en connexion FB vers V4 (18% des projections afférentes de V4) (Kennedy et Barone, 1999).

L'aire TE projette vers les aires TH/TF, STP, FEF¹¹, et l'aire 46 (Barbas et Mesulam, 1981; Shiwa, 1987; Barbas, 1988). L'aire TE est également connectée directement ou indirectement avec des structures limbiques ; l'amygdale (connexions directes, Turner, 1980) et à l'hippocampe soit directement (Yukie et Iwai, 1988) ou soit indirectement via le gyrus parahippocampique (voir revue de Van hoesen, 1982). L'aire TE ne projette pas directement vers le cortex enthorhinal mais via le cortex périrhinal et parahippocampique. V4 reçoit également des connexions FB de TE (11% des projections afférentes).

5.3. Connexions sous-corticales

Les aires TE et TEO sont également connectées à des structures sous-corticales telles que des noyaux du thalamus, des noyaux de l'hypothalamus, le locus coeruleus, la formation réticulaire, le noyau basal de Meynert et le noyau du raphé (partie ventrale et dorsale). Ces 2 aires sont également inter-connectées avec des structures sous corticales plus antérieures telles que le pulvinar et la région ventrale du claustrum. Elles projettent également de manière non réciproque vers le striatum. L'aire TEO projette seule vers le colliculus supérieur et l'aire TE vers le noyau médial dorsal du thalamus (Webster *et al.*, 1993).

5.4. Organisation rétinotopique

L'aire TEO a une organisation visuotopique assez grossière. Elle cartographie la totalité du champ visuel controlatéral. Les cellules ont des CR plus grands que ceux de V4 (Boussaoud *et al.*, 1991). L'aire TE ne présente pas d'organisation visuotopique particulière. Les cellules ont de larges CR, ipsi, contro ou bilatéral. La plupart incorporent la fovea (Gross *et al.*, 1972). La taille des CR augmente le long d'un axe postero-antérieur traversant IT en fonction de l'excentricité des champs. Dans TEO, les CR ont un diamètre de 1.5° à 55° selon l'excentricité¹² et dans TE entre 30° à 50° (Boussaoud *et al.*, 1991).

6. Résumé

Cette description anatomique du cortex visuel simien nous permet de positionner l'aire V4 au sein de la voie visuelle ventrale. Elle se situe entre l'aire V2 dont elle reçoit ses principales connexions et le cortex inférot temporal, issue finale de la voie ventrale vers laquelle elle projette massivement. La voie ventrale a pour fonction principale la reconnaissance des objets. Tous les attributs nécessaires à l'identification de l'objet (couleur, forme, texture...) sont donc traités le long de cette voie. Plusieurs études

¹¹ TH/TF= aires para-hippocampique, aire STP= Superior Temporal Polysensory et FEF= Frontal Eye Field

¹² Taille_{CR}= 0.76X Excentricité_{CR} + 4.8°

neurophysiologiques (lésions, neuroimagerie ou enregistrements cellulaires) ont cherché à localiser les bases neurobiologiques impliquées dans le traitement des propriétés de l'objet. Dans le paragraphe suivant, nous nous intéresserons à la contribution de l'aire V4 dans le codage des attributs 2D intrinsèques de l'objet (couleur, forme et texture) tout en résitant ce codage au sein de la voie visuelle ventrale mais nous nous intéresserons également à l'implication de l'aire V4 dans l'encodage d'indices tridimensionnels.

D. BASES NEUROPHYSIOLOGIQUES DE LA RECONNAISSANCE DES OBJETS

Les objets sont partout dans notre entourage visuel et se différencient les uns des autres grâce à plusieurs caractéristiques qui leur sont propres comme leur couleur, la composition de leur surface ou encore leur forme. Deux types d'informations 2D nous permettent donc de reconnaître un objet : sa forme 2D externe définie par des contours séparant la figure -l'objet- du fond et son contenu interne -sa couleur ou sa texture-. Des études anatomo-fonctionnelles, de neuroimagerie ou de neurophysiologie, ont permis de déterminer les substrats et les corrélats neuronaux de tous ces attributs notamment le long de la voie visuelle ventrale. L'encodage de ces attributs sera traité dans la première partie de ce chapitre et un intérêt particulier sera porté sur la contribution de l'aire V4 dans le codage de ces paramètres.

Dans la seconde partie de ce chapitre, nous nous intéresserons au traitement de la profondeur et plus précisément aux bases neuronales des indices tridimensionnels binoculaires et monoculaires (perspective de gradient, interposition, ombres et ombrages).

1. Reconnaissance des objets sur la base d'indices bidimensionnels

1.1. Encodage de la forme 2D de l'objet

a) Traitement de la forme de l'objet dans V4

(1) Lésions corticales chez le singe.

De nombreuses études ont porté sur les effets des lésions de la voie temporale sur la reconnaissance des objets. La plupart se sont concentrées sur le traitement de la forme chez des animaux lésés au niveau du cortex inférot temporal, issue finale de la voie ventrale (Dean, 1976). Dans ce paragraphe, je traiterai des effets des lésions de l'aire V4 sur le traitement perceptif de la forme 2D.

Les études qui traitent de ce sujet diffèrent selon les tâches exécutées par des animaux lésés ; allant de tâches de discrimination simple de barres orientées à des tâches de discrimination de formes 3D. De Weerd et ses collaborateurs mettent en évidence une baisse des performances des animaux lésés dans une tâche de discrimination lorsque les stimuli, barres ou réseaux orientés, diffèrent par leur orientation (De Weerd *et al.*, 1996; De Weerd *et al.*, 1999). D'autres études ont porté sur les performances des animaux lésés dans des tâches

de discrimination de formes 2D plus complexes. Les animaux lésés montrent des déficits importants dans une tâche de discrimination (pop-out) selon la forme 2D (triangle, carré ou cercle) du stimulus (Schiller, 1993). Les animaux devaient effectuer une saccade oculaire vers la forme 2D différente des distracteurs. Les pourcentages de réussite des animaux lésés chutent et leur latence de saccade augmente significativement montrant ainsi que les animaux lésés ne reconnaissent pas la forme 2D à discriminer. Dans un autre type de protocole, Walsh et son équipe traitent des performances d'animaux lésés bilatéralement au niveau de V4 dans une tâche de discrimination entre deux formes 2D simples, apprise avant l'opération. Les performances des animaux lésés diminuent pendant le réapprentissage de la tâche de discrimination par rapport à celles observées chez les animaux sains. De plus, la discrimination entre deux formes identiques mais qui diffèrent par leur orientation est également altérée chez ces animaux lésés (Walsh *et al.*, 1992). Schiller démontre également que des transformations de la forme à discriminer (en taille, occlusion ou en faisant disparaître certains de ces contours) affectent fortement les performances des animaux lésés dans une tâche de discrimination (Schiller, 1995). Toutefois, une plasticité importante et relativement rapide peut se produire à la suite d'une lésion. Une inactivation réversible (par refroidissement ou *cooling*) a l'avantage d'éviter ce phénomène et établit qu'on observe un déficit d'identification de formes simples (Girard *et al.*, 2002).

L'implication de l'aire V4 dans le traitement de la forme a également été étudiée dans des tâches où les animaux sont entraînés à discriminer des objets 3D selon plusieurs angles de vision (ou « points de vue »). Merigan et Pham ont entraîné des animaux à reconnaître un objet 3D indépendamment de son orientation (Merigan et Pham, 1998). Les animaux doivent indiquer s'il s'agit ou pas du même objet lorsqu'on leur présente deux images d'un objet 3D. Une fois la tâche parfaitement apprise, ils ont lésé unilatéralement l'aire V4 de ces animaux (avec de l'acide iboténique). Puis, ils ont retesté ces animaux lésés dans la même tâche de discrimination et remarquent que les animaux discriminent mal les formes 3D entre elles si elles sont placées dans leur quadrant visuel lésé. A l'inverse, ils réussissent parfaitement à les discriminer si elles sont placées dans leur quadrant visuel non lésé.

De toute évidence, ces expériences de lésions de V4 chez le singe ont démontré qu'une reconnaissance d'objet basée sur sa forme 2D ou 3D est altérée. Les enregistrements électrophysiologiques vont confirmer le rôle fonctionnel de l'aire V4 dans le traitement de la forme de l'objet.

(2) Électrophysiologie unitaire chez le singe

La présence de cellules sélectives à l'orientation dans l'aire V4 (Desimone *et al.*, 1985; Desimone et Schein, 1987; Schein et Desimone, 1990), et le fait que cette aire reçoive massivement des projections des interbandes de l'aire V2 (Shipp et Zeki, 1985; Zeki et Shipp, 1989; DeYoe *et al.*, 1994) suggèrent, qu'outre la couleur, des informations concernant le traitement des formes transitent vers V4.

Comme pour les expériences de lésions anatomiques chez l'animal, les premières études électrophysiologiques qui traitent de la sélectivité des neurones de V4 ont utilisé comme stimuli des formes relativement simples : des barres et des réseaux orientés. L'implication de l'aire V4 dans le traitement des formes a été supposée grâce à l'enregistrement de cellules sélectives à l'orientation de ces stimuli (Desimone *et al.*, 1985; Desimone et Schein, 1987; Schein et Desimone, 1990). Les auteurs remarquent que peu de cellules ont un pattern très sélectif à une orientation donnée, ce qui tend à montrer qu'elles répondent plus à la forme du stimulus qu'à son orientation (Desimone et Schein, 1987). Par la suite, d'autres études ont montré que les neurones de V4 semblent mieux répondre à des formes complexes qu'à des formes simples (Kobatake et Tanaka, 1994; Pasupathy et Connor, 1999; Pasupathy et Connor, 2001). Pasupathy et al. (1999) ont testé cette sélectivité en disposant d'un set très large de formes complexes définies par des contours particuliers, convexes ou concaves, situés à différents endroits dans la forme globale. Ils démontrent que les neurones de V4 répondent à un ensemble de formes complexes et non à une seule forme (Pasupathy et Connor, 1999). Ceci provient du fait que les neurones sont avant tout sensibles à la présence d'un contour spécifique dans la forme complexe. Le neurone de la Figure 1-22 répond, par exemple, à un angle convexe orienté à 90° avec une forme concave adjacente dans le sens horaire. Grâce à leurs études, ils mettent en évidence que les dimensions importantes pour le codage de ces formes complexes sont l'état de courbure de cet élément (convexe ou concave) et sa position angulaire (par exemple si celui-ci est situé en bas de la forme). Dans des travaux ultérieurs, Pasupathy et ses collaborateurs ont démontré que la sélectivité d'un neurone de V4 pour un élément spécifique dépend également des formes adjacentes à cet élément (Pasupathy et Connor, 2001). A l'inverse, des facteurs comme l'orientation des bordures, la direction du contraste et la fréquence spatiale semblent moins importants pour le codage d'une forme complexe dans V4 (Pasupathy et Connor, 1999).

Les neurones de V4 seraient donc sélectifs à une combinaison d'éléments. Ces résultats s'accordent avec les théories de Bierderman (1987) et Riensuhuber et Poggio (1999) selon lesquelles «une forme complexe est représentée par une combinaison d'éléments

simples : les primitives » (Biederman, 1987; Riesenhuber et Poggio, 1999). Ainsi, la population de neurones de V4 coderait tous les éléments d'un alphabet de primitives permettant l'analyse de n'importe quelle forme complexe.

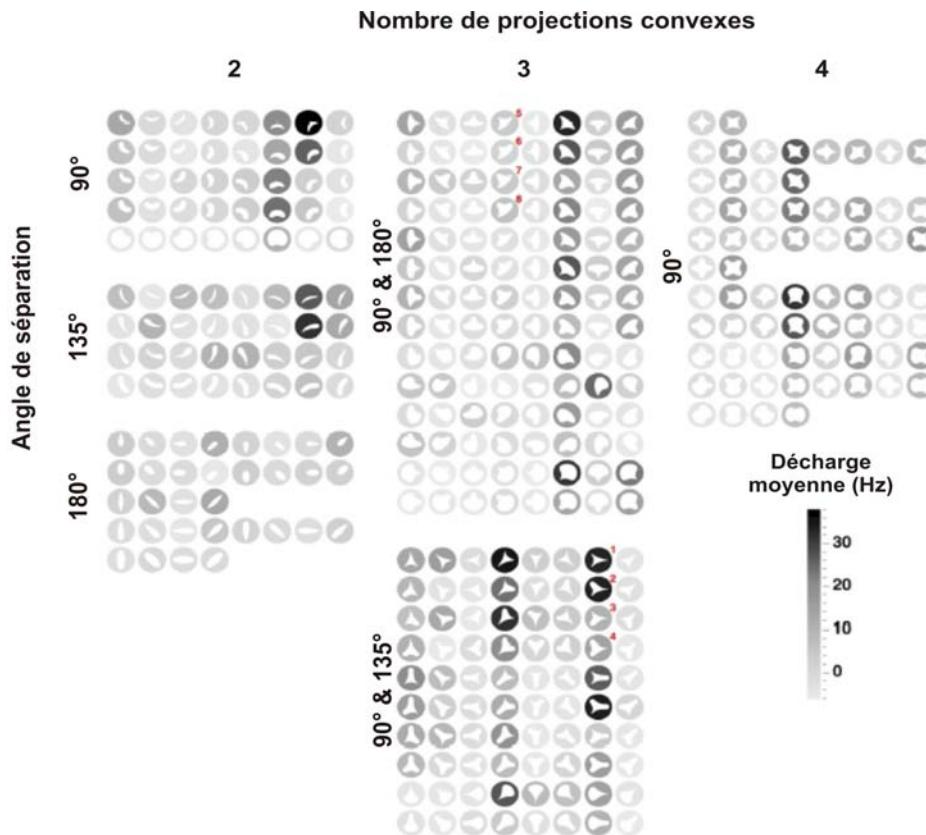


Figure 1-22 : Exemple de neurone de V4 enregistré avec le set de stimuli utilisé par Pasupathy et ses collaborateurs (Pasupathy et Connor, 2001). Chaque rond correspond à un stimulus particulier, le niveau de gris indiquant le taux de réponse du neurone pour ce stimulus. Le stimulus est dessiné au centre du rond. Les stimuli sont rassemblés sur la figure en fonction du nombre de projections convexes qu'ils contiennent et de l'angle de séparation de ces projections. Dans les différents ensembles, chaque colonne définit une orientation donnée du stimulus, et chaque ligne l'angle de séparation entre les projections. Les neurones de V4 ont, en général, une sélectivité assez large pour un ensemble de formes complexes comprenant une combinaison spéciale d'éléments constitutifs (D'après Pasupathy et Connor, 2001).

b) Traitement de la forme dans la voie visuelle ventrale

De par l'organisation hiérarchisée des aires de la voie visuelle ventrale, il a été proposé que le traitement de la forme de l'objet, essentiel pour sa reconnaissance, s'effectue également selon une certaine hiérarchie; d'un niveau simple à un niveau plus complexe (Tanaka, 2002).

Le premier niveau de codage consiste à définir les orientations des lignes, bordures et d'autres éléments de l'objet qui constituent l'objet. La sélectivité à l'orientation a été démontrée au niveau des aires V1 et V2 (Hubel et Wiesel, 1968; Baizer *et al.*, 1977; Hubel et Livingstone, 1987). Cette sélectivité à l'orientation de formes simples n'est pas uniquement

réservée aux aires V1 et V2 puisque des neurones sélectifs ont été enregistrés dans les aires de plus haut niveau hiérarchique telles que V4 (Desimone et Schein, 1987) ou PIT (Tanaka *et al.*, 1991). Cette sélectivité à l'orientation a souvent été interprétée comme une première étape de traitement pour le codage d'une forme si la reconnaissance d'un objet se fait grâce à des primitives comme le propose Biederman dans sa théorie.

Le niveau secondaire d'intégration de la forme de l'objet serait le codage de la position angulaire et de l'état de courbure des segments. Les résultats des travaux de Pasupathy et Connor démontrent que V4 serait le siège de ce niveau secondaire. De plus, les neurones de V4 sont aussi sensibles à des stimuli définis par différents patterns de courbure, et ces résultats suggèrent que l'information de courbure puisse être sommée globalement (Gallant *et al.*, 1993; Gallant *et al.*, 1996; Pasupathy et Connor, 1999; Pasupathy et Connor, 2001). Toutefois, une récente étude démontre des résultats similaires dans V2 (Hegde et Van Essen, 2006). Au niveau de IT, plusieurs études démontrent que les neurones sont plus sensibles à une structure grossièrement définie de l'objet plutôt qu'à sa forme réelle. Ainsi, Tanaka et al (1991) enregistrent les réponses des neurones de IT à des objets complexes tout en simplifiant, petit à petit, la forme de l'objet. Pour la majorité des cellules, les décharges neuronales augmentent au fur à mesure que la complexité de l'image diminue jusqu'à obtenir une forme optimale pour le neurone (c'est à dire celle qui déclenche une réponse maximale du neurone). Ainsi, par exemple, l'image d'une main peut être simplifiée jusqu'à obtenir un ensemble de 3 barres orientées (Tanaka *et al.*, 1991).

1.2. Encodage de la couleur de l'objet

Outre la forme, dont je viens de parler, les bases neuronales de 2 attributs intrinsèques de l'objet que sont la couleur et la texture font l'objet de plus amples descriptions dans la partie suivante ; la couleur car V4 a été souvent considéré dans la littérature comme le centre de la couleur et la texture puisque cette thèse propose d'en étudier l'encodage au sein de l'aire V4.

a) V4 et la couleur

(1) Lésions corticales chez le singe

Plusieurs expériences de lésion de V4 ont eu pour but d'étudier les performances des singes dans des tâches de discrimination selon les teintes¹³ des stimuli, leurs longueurs d'ondes ou encore leur catégorie de couleur. Des larges lésions de V4 provoquent une diminution des

performances dans un test de discrimination de teintes en augmentant les seuils de détection mais n’entraînent pas de déficits dans une tâche de discrimination de luminosité (c’est à dire, entre différents niveaux de gris) (Dean, 1979; Heywood et Cowey, 1987). Cependant, les stimuli de couleur utilisés dans cette expérience sont perçus uniquement grâce à la réflexion de la lumière et, lorsque l’animal bouge la tête, la composition de la lumière réfléchie est modifiée. Par conséquent, il est possible que le déficit de performances des animaux lésés soit dû à une mauvaise discrimination de la constance de la couleur et non de la longueur d’onde. Ainsi, dans une seconde étude portant sur les mêmes animaux, mais en utilisant des teintes lumineuses sur fond noir (donc avec une lumière réfléchie négligeable), les animaux lésés réussissent parfaitement la tâche de discrimination de teintes qui se rapproche, en fait, d’une tâche de discrimination de longueur d’ondes (Heywood *et al.*, 1992). Walsh et son équipe (1993) ont testé leurs performances d’animaux lésés à la fois dans une tâche de discrimination des teintes et sur la constance de la couleur. Ces deux tâches étaient parfaitement effectuées par les animaux avant l’opération. Les auteurs remarquent que les animaux lésés peuvent récupérer leurs performances initiales d’avant l’opération pour la tâche de discrimination de teintes mais dans la seconde tâche sur la constance des couleurs, les animaux montrent d’importants déficits (Walsh *et al.*, 1993). Merigan (1996) a mis en évidence chez des animaux lésés des déficits dans une tâche d’appariement de teintes (Merigan, 1996). L’équipe de V. Walsh a également testé les performances de singes lésés au niveau de V4 dans une tâche de catégorisation de 4 couleurs (rouge, vert, bleu et jaune). Les animaux lésés faisaient davantage d’erreurs que des animaux témoins non lésés et mettaient plus de temps à réussir cette tâche mais ils réussissent quand même à déterminer la catégorie de la couleur à discriminer (Walsh *et al.*, 1992).

Ces études de lésions de V4 chez le singe montrent que les singes lésés n’ont pas de déficits dans une tâche de discrimination de teintes ou de longueurs d’ondes. Au contraire, ils montrent d’importants troubles dans des tâches sur la constance de la couleur. Les résultats d’études électrophysiologiques vont confirmer le rôle fondamental de V4 dans la constance de couleur¹⁴.

¹³ La teinte d’un objet renvoie à sa couleur (verte, rouge...). Plusieurs longueurs d’ondes peuvent définir une même teinte.

¹⁴ On nomme la constance de couleur le fait que nous puissions percevoir une couleur comme identique quel que soit le type d’éclairage. Comme la lumière qui éclaire un objet particulier dans une scène visuelle éclaire également le reste de la scène, l’utilisation du contexte nous aide à extraire une information globale quant à la nature de la lumière ambiante, au sein du spectre de la lumière réfléchie par l’ensemble des objets. Autrement dit, l’ensemble des longueurs d’ondes présentes dans l’image visuelle nous permet d’identifier la nature de la lumière ambiante et de maintenir constante la couleur perçue d’un objet en particulier. Ce principe a été énoncé

(2) Électrophysiologie unitaire chez le singe

La dénomination (V4 = centre de la couleur) résulte des multiples travaux entrepris par S. Zeki. Dans ses premières études électrophysiologiques sur le singe anesthésié, Zeki a trouvé un nombre important de cellules sélectives à la longueur d'onde qui combiné avec l'absence de cellules sélectives au mouvement d'un stimulus¹⁵ l'a conduit à considérer l'aire V4 comme l'aire traitant la couleur. Cependant, très tôt, Zeki signale que cette aire ne code pas uniquement la couleur de l'objet car les neurones de V4 sont aussi sélectifs à des formes ou des orientations particulières (Zeki, 1973a). Les conclusions d'études ultérieures ne sont pas aussi unanimes sur la spécificité de V4 dans le codage de la couleur. En effet, le nombre de cellules sélectives à la couleur n'a cessé de diminuer tout au long des études électrophysiologiques entreprises sur V4. Dans sa première étude, Zeki (1973) rapportait que la totalité des neurones étaient sélectifs à la couleur soit 100%, puis 87% en 1977 (Zeki, 1977) et 68% en 1978 (Zeki, 1978b). Les études du groupe de Schein (Schein *et al.*, 1982), et celui de Tanaka (Tanaka *et al.*, 1986) rapportent des pourcentages de cellules sélectives à la couleur encore plus faibles (environ 20%).

Une explication possible à cette baisse du nombre de neurones de V4 sélectifs à la couleur concerne les critères de sélectivité utilisés par Zeki pour déterminer si une cellule est influencée ou pas par la couleur du stimulus. La première étude de Zeki était uniquement qualitative ; or la quasi-impossibilité de déterminer une sélectivité neuronale fine de façon qualitative est désormais établie. En dépit de leur sensibilité aux longueurs d'ondes, les cellules de V4 présentent des caractéristiques ne permettant pas de les classer comme « sélective à la couleur ». Premièrement, peu de cellules ont des réponses de type ON/OFF pour des couleurs dites opposées comme le vert/rouge ou bleu/jaune comme c'est le cas au niveau de la rétine ou du CGL. Deuxièmement, l'étude de Schein et Desimone (1990) indique qu'une quantité non négligeable de cellules de V4 (30%) présentent deux pics de réponse à une gamme de longueurs d'ondes donc n'ont pas de sélectivité unique (Schein et Desimone, 1990). Enfin, la majorité des cellules répondent à la lumière blanche et les cellules décrites comme sélectives à la couleur dans V1 ou V4 ne devraient pas répondre à la lumière blanche (Zeki, 1973a; Zeki, 1983b; Livingstone et Hubel, 1984a).

par Land en 1977. Il s'est interrogé sur l'importance du contexte visuel dans la perception des couleurs. Dans son expérience, il présente des stimuli composés de divers patterns de couleur de type *Mondrian*, avec des illuminations variables. Il démontre que la couleur perçue d'un patch particulier reste constante à condition que l'illumination reste homogène. En revanche, si l'illumination du patch est différente de celle du contexte, la constance de couleur disparaît.

¹⁵ La présence de cellules sensibles au mouvement lui avait permis de définir l'aire V5 Zeki, S. (1974a).

Pour résumer, les cellules de V4 semblent donc se comporter comme des filtres spectraux à bande large qui laissent passer certaines longueurs d'onde plus efficacement que d'autres mais également les informations issues de la lumière blanche. Une autre caractéristique des cellules de V4 par rapport aux cellules des aires inférieures est qu'elles ont des CR avec des contours suppressifs plus importants (Zeki, 1983b; Desimone *et al.*, 1985). L'influence du contour sur la réponse des cellules de V4 conforte une possible participation de cette aire dans le traitement de la couleur, car la couleur d'un stimulus est déterminée en comparaison avec les couleurs des régions avoisinantes, et non pas seulement par sa longueur d'onde. Zeki a enregistré chez le singe les réponses des cellules de V1 et V4, sensibles à la longueur d'ondes, en leur présentant des stimuli visuels de type «Mondrian» en faisant varier l'illumination ambiante (Zeki, 1983a). Les résultats de cette étude indiquent que les cellules de V1 sont sensibles à la longueur d'ondes des patchs qui se trouvent dans leur CR et leurs réponses varient selon la lumière ambiante. A l'inverse, les cellules de V4 ont leur sélectivité à la longueur d'onde qui n'est pas affectée par les variations de la lumière ambiante. Les réponses des neurones restent constantes malgré les changements de longueur d'ondes du patch coloré présent dans leur CR, en adéquation avec la couleur perçue. L'aire V4 jouerait donc un rôle fondamental dans la perception de la **constance** de la couleur.

L'implication de V4 dans la constance de la couleur a été récemment confirmée grâce à l'étude de Kusunoki *et al.* (2006) (Kusunoki *et al.*, 2006). Contrairement à l'étude initiale de Zeki (1983), ce travail a été effectué sur des animaux anesthésiés et vigiles alors que l'étude de Zeki n'avait été effectuée que sur des animaux anesthésiés. Ces travaux sur des animaux vigiles ont permis de corrélérer le percept de l'animal (à l'aide d'un test *match to sample*) avec les réponses des cellules. Pour la majorité des cellules de V4 (71%), leur taux de réponse reste inchangé si la perception de la couleur reste constante. Les réponses des cellules de V4 sont donc en accord avec les percepts colorés qu'a le singe et non avec les longueurs d'ondes des patchs colorés présents dans leur CR.

b) Conclusion

Une conclusion possible sur le traitement de la couleur est que les mécanismes neuronaux en amont de V4 (V1, V2 et V3) sont suffisants pour coder les 4 couleurs primaires (rouge, vert, bleu et jaune). En effet, les singes lésés au niveau de V4 se basent sur ces 4 couleurs primaires pour effectuer une tâche de catégorisation de couleurs avec des performances similaires à celles d'animaux non lésés (Walsh *et al.*, 1992). Ceci rejoint également le fait que les potentiels évoqués au niveau occipital visuel (V1, V2) soient

identiques chez des singes lésés au niveau de V4 et des animaux sains si on leur présente divers stimuli chromatiques. Le traitement de la couleur dans les aires V1 et V2 serait donc normal malgré la lésion de V4 (Kulikowski *et al.*, 1994). La sensibilité à la longueur d'onde serait traitée avant V4 dans les aires V1 et V2.

Toutefois, il est important de souligner que V4 n'est pas LE centre de la couleur. En effet, plusieurs études ont démontré que la couleur est également traitée dans d'autres aires (Gegenfurtner *et al.*, 1996; Gegenfurtner *et al.*, 1997). De plus, les neurones de V4 présentent d'autres patterns de sélectivité, notamment à la forme de l'objet (Desimone et Schein, 1987). Toutefois, il est clairement établi que l'aire V4 jouerait un rôle primordial dans la constance de la couleur afin de déterminer la couleur perçue par l'observateur. L'aire V4 permettrait donc de maintenir constante notre perception de la couleur d'un objet indépendamment des conditions d'illumination mais permet-elle de nous indiquer réellement la couleur en tant que qualia ? Aucune des études détaillées ci-dessus ne permet de conclure sur ce sujet.

1.3. Encodage de la texture

a) Lésions corticales

Plusieurs études anatomo-fonctionnelles ont démontré une possible implication de l'aire V4 dans le codage de texture. Dans leur série d'expériences, Schiller et Lee (1991) ont testé les performances des animaux lésés au niveau de V4 dans une tâche de discrimination basée sur le pattern du stimulus (pop-out). Parmi un ensemble d'échiquiers, les singes doivent reconnaître si un échiquier diffère des autres (Schiller et Lee, 1991). Les performances des singes lésés sont mauvaises dans cette tâche. Dans une étude ultérieure, Schiller démontre que les animaux lésés peuvent détecter un pattern texturé sur la base de l'orientation de ces éléments. Toutefois, leurs latences de saccades augmentent par rapport à des animaux sains (Schiller, 1993). Merigan (1996) a testé des animaux lésés dans une tâche de discrimination de patterns orientés se segmentant du fond. Les animaux doivent indiquer l'orientation du pattern texturé à segmenter. Les performances des animaux lésés chutent significativement en comparaison avec celles enregistrées sur des animaux sains (Merigan, 1996). Dans une étude ultérieure, Merigan (2000) teste les performances de singes lésés dans des tâches similaires de discrimination de patterns orientés mais en faisant varier le pattern à discriminer. Il conclut alors que la perte de l'aire V4 altère uniquement les discriminations de textures qui requièrent

un regroupement d’éléments similaires pour effectuer une tâche de pop-out¹⁶ (Merigan, 2000).

De Weerd et son équipe (1996) ont également testé les capacités de singes lésés à discriminer l’orientation de contours lorsqu’ils sont définis par différents indices comme la luminance, la couleur, le mouvement ou la texture. De façon surprenante, les performances des animaux lésés chutent considérablement pour les contours définis par des textures différentes alors que les auteurs remarquent que les animaux présentent peu ou pas de déficits pour des contours définis par les autres indices (couleur, luminance ou mouvement). Ces résultats suggèrent que l’aire V4 joue un rôle essentiel dans la définition des contours basée sur la texture (De Weerd *et al.*, 1996).

b) Électrophysiologie unitaire

Au niveau de V1, les principales études sur ce sujet ont traité de la contribution de cette aire dans les définitions des bordures des textures dans des processus de pop-out. Dans son étude, Knierim et son équipe démontrent qu’un tiers des cellules de V1 augmentent leur taux de réponse si les éléments de bordure de la texture se trouvent dans le CR de la cellule par rapport à des éléments simples d’une texture homogène (Knierim et Van Essen, 1992). Certaines cellules de V1 présentent également une sélectivité à l’orientation des éléments mais au niveau de la population, l’orientation n’intervient pas. Des informations sur la définition des bordures des textures, essentielles pour une segmentation de textures, sont donc présentes au sein de l’aire V1.

Au niveau de V2, Hegde et van Essen (2003) ont enregistré les réponses des neurones à un ensemble complet de stimuli représentant des segments complexes (angles, courbures, arc ou cercles) et des patterns de textures identiques à ceux de Gallant (1993). La population de V2 répond en moyenne davantage aux patterns texturés qu’aux segments (Hegde et Van Essen, 2003).

Au niveau de l’issue finale de la voie visuelle ventrale, l’aire IT, les neurones sont sensibles à la texture d’un objet et codent différemment suivant le pattern du stimulus présenté (Tanaka *et al.*, 1991; Komatsu et Ideura, 1993). Cette sélectivité présente dans IT a également été démontrée dans l’aire inférieure hiérarchiquement, l’aire V4 (Gallant *et al.*, 1993; Gallant *et al.*, 1996). S’intéresser au traitement de la texture d’un objet au niveau de V4 revient en fait à savoir si les neurones de V4 sont, avant tout, intéressés par le pattern d’un

¹⁶ Pop-out : dans ce type de tâche, il s’agit de remarquer si un stimulus qui diffère par exemple en luminance, orientation ou texture est présent au sein d’un ensemble de stimuli.

objet, autrement dit par son contenu. Les neurones de V4 sont sensibles au contenu d'une image dans la mesure où quelques neurones de V4 répondent par exemple sélectivement à la couleur de l'objet (travaux de Zeki, voir paragraphe A.1.1.a)(2)). Autre que la couleur, Gallant et son équipe se sont intéressés à la sélectivité des neurones de V4 pour des stimuli visuels définis par différentes fréquences spatiales. Desimone et Schein (1987) avaient déjà remarqué une sélectivité des neurones de V4 aux fréquences spatiales lors de la présentation de réseaux orientés (Desimone et Schein, 1987). Gallant a décidé de tester cette sélectivité en présentant différents patterns cartésiens (réseaux orientés) ou non cartésiens (réseaux polaires : radiaux ou concentriques ou des réseaux hyperboliques) à des singes anesthésiés tout en enregistrant les réponses cellulaires dans V4. Il remarque un codage plus important des neurones de V4 pour des stimuli non cartésiens par rapport à de simples réseaux orientés. Par ailleurs, à l'aide d'une analyse multidimensionnelle de type MDS, il a pu caractériser trois clusters de réponses de sa population de neurones dans V4. Ces clusters sont bien séparés et correspondent à chaque type de pattern utilisé (cartésiens, polaire et hyperbolique). Selon l'auteur, les cellules de V4 constituerait un stade intermédiaire entre le processus du traitement des fréquences spatiales entre les cellules de V1 et V2 et la représentation finale de la reconnaissance de l'objet dans IT. Toutefois cette étude ne permet pas de conclure s'il s'agit de sélectivité à un contenu d'un objet ou à des éléments descriptifs de son contour.

Hanazawa et al (2001) ont enregistré la réponse de neurones de V4 à des textures artificielles (Hanazawa et Komatsu, 2001). Ces textures sont composées de plusieurs éléments circulaires où la densité et la taille varient et la quasi-totalité des neurones enregistrés est sensible aux paramètres des textures (taille et densité). Toutefois, s'agit-il réellement de réponses à la texture interne de l'objet ou à un ensemble de plusieurs éléments de plus ou moins petite taille suivant que l'on éloigne plus ou moins de l'observateur ?

1.4. Homologies homme-singe

Les études de cas cliniques chez l'Homme et de neuroimagerie ont cherché à cibler les zones corticales impliquées dans la reconnaissance d'objets et à les comparer à celles du singe. Cette comparaison pour l'aire V4 est difficile car, comme nous l'avons vu dans la description anatomique, l'aire homologue de l'aire V4d du singe chez les humains n'a pas été clairement définie et les études ne s'accordent pas sur la dénomination de cette aire homologue. Les résultats présentés ici sont donc à considérer avec prudence car il est possible que les lésions des patients ou les zones corticales activées en neuroimagerie ne correspondent pas exactement à l'aire V4 du macaque. Toutefois, toutes les études détaillées

ci-dessous décrivent l'implication de la voie visuelle occipito-temporale dans le traitement des attributs de l'objet.

Les études d'altération de la reconnaissance de l'objet ont été réalisées chez des patients porteurs de lésions cérébrales principalement localisées au niveau du cortex inférot temporal. Ces patients lésés présentent généralement d'importants troubles dans la reconnaissance des objets. L'ensemble de ces déficits est réuni sous le terme d'*agnosie visuelle*. L'agnosie visuelle se définit donc comme un trouble majeur de la perception des objets mais sans déficit manifeste des fonctions visuelles élémentaires (acuité, champ visuel...). Je ne détaillerai pas les différents types d'agnosie; elles font l'objet d'une excellente revue dans le livre de M.J. Farah (Farah, 1990).

Parmi les exemples de lésions de V4 chez l'Homme, un cas clinique est souvent cité. Il s'agit d'une patiente, D.F, souffrant d'importantes lésions bilatérales au niveau de la voie visuelle ventrale à la suite d'une hypoxie provoquée par une intoxication au monoxyde de carbone (pour l'étendue des lésions, voir Milner 1991 et al.). Cette patiente présente une agnosie visuelle profonde concernant les formes. Elle est incapable de discriminer des formes géométriques simples et ne peut indiquer ni la taille, ni la forme, ni l'orientation d'un objet aussi bien verbalement qu'en faisant un signe de la main même si elle peut saisir sans difficulté des objets (Milner *et al.*, 1991; Milner et Goodale, 1993). Elle éprouve également de fortes difficultés à dessiner des objets qui lui sont présentés (c'est à dire les recopier) mais peut les dessiner grâce à un travail de mémoire. Il semble également que contrairement à la forme d'un objet, la surface semble être correctement traitée par D.F. En effet, elle reconnaît plus facilement des objets colorés ou en niveaux de gris que des objets dessinés simplement par leurs contours. James *et al* (2003) ont enregistré en IRMf les zones cérébrales activées chez D.F, pendant une tâche de reconnaissance d'objet (vision passive) mais également pendant une tâche de maniement d'objet (James *et al.*, 2003). Dans la tâche de reconnaissance d'objet, l'expérimentateur présente des dessins d'objets ou des patterns mélangés ne définissant aucun objet (*stimuli scrambled*). Les résultats de l'IRMf ne montrent aucune différence d'activation au niveau du cortex latéral occipital (LO) entre les deux conditions alors que chez des sujets sains, cette aire répond préférentiellement lorsque les objets sont correctement dessinés. La faible activation de l'aire LO peut expliquer les piétres performances de D.F à identifier des objets lorsqu'on lui présente des dessins. Toutefois, des zones d'activations sont relevées le long de la voie visuelle ventrale lors de la présentation de dessins colorés ou en niveaux de gris, tâche où D.F montre de meilleures performances. Ces

zones semblent être plus disséminées que celles enregistrées chez des témoins sains et n'incluent pas l'aire LO lésée bilatéralement chez D.F. La reconnaissance d'objets sur la base d'indices matériels comme la couleur semble donc être traitée dans des zones cérébrales distinctes d'une reconnaissance d'objet basé sur le contour. Les auteurs concluent que l'aire LO absente chez D.F permettrait la reconnaissance d'objet basée sur la forme même si l'étude d'autres cas cliniques est nécessaire pour conclure sur les contributions de cette aire dans la reconnaissance d'objets.

Gallant et son équipe (2000) ont étudié les capacités d'un patient lésé, A.R, dans différentes tâches de discriminations (Gallant *et al.*, 2000). La lésion de ce patient se situe dans son hémisphère droit entre le gyrus fusiforme et le sillon lingual. Cette région correspondrait à une aire identifiée par des études en IRMf comme la partie ventrale de l'aire V4 chez l'Homme : V4v (Sereno *et al.*, 1995; DeYoe *et al.*, 1996). A.R réussit parfaitement à discriminer des stimuli dont la luminance, l'orientation ou la vitesse varie. Il ne présente donc pas de déficits dans des tâches de perception « simple », ce qui avait déjà été remarqué dans des tâches similaires chez des patients achromatopsiques (Merigan *et al.*, 1997). Le patient présente par contre de forts déficits de performances dans des tâches de discrimination plus élaborées faisant intervenir des processus visuels intermédiaires telles que la discrimination d'orientation de bordures illusoires, de patterns non cartésiens ou de courbure en utilisant des stimuli dont ces mêmes auteurs avaient montré leur efficacité à activer les neurones de V4 chez le singe.

En neuroimagerie, deux études d'IRMf ont comparé les zones activées chez le macaque et chez l'humain dans une tâche de reconnaissance d'objets (Tsao *et al.*, 2003; Denys *et al.*, 2004). Selon la première étude de l'équipe de Tsao et de Tootell (2003), les humains possèdent une aire temporale ventrale spécifique pour la reconnaissance des visages ; l'aire fusiforme déjà décrite par Kanwisher en 1997 (Kanwisher *et al.*, 1997). D'après les auteurs, cette aire spécifique est absente chez les macaques puisque les neurones qui répondent aux visages sont plus éparpillés dans le cortex temporal, proche du sillon STS. Pour la reconnaissance d'objets chez le macaque, les zones activées regroupent l'aire V4 et le cortex inférot temporal (TE) tandis que chez l'Homme, la présentation de ces mêmes objets entraîne une activation significative de l'aire LOC. Sur le même principe, l'équipe de Denys compare les aires activées dans une reconnaissance d'objets chez le macaque et chez l'humain (Denys *et al.*, 2004). Les stimuli pour cette étude sont similaires à ceux utilisés dans l'étude de Kourtzi et Kanwisher (2000). Il s'agit d'images d'objets définies soit par un pattern en

noir et blanc ou soit par un simple contour externe. Des stimuli contrôles (“scrambled” de pixels) sont également présentés. Chez l’humain, la présentation d’objets intacts active l’aire LOC et l’aire LOS. Chez le singe, les zones activées regroupent V3, V4 et IT.

Plusieurs études ont cherché à localiser les aires impliquées dans le traitement des textures présentes dans une scène visuelle. La première étude présentée concerne la contribution des aires visuelles primaires dans la perception de contours définis par différentes textures, soit le phénomène de segmentation. La ségrégation des différents composants d’une scène visuelle se base sur la définition des contours présents dans la scène. Ces contours sont perçus grâce à une différence de luminance, de couleur ou de vitesse entre deux régions distinctes de l’image mais peuvent également être définis par des indices plus complexes comme la texture. Les aires visuelles primaires V1 et V2 semblent participer au codage de tels indices. En effet, Lamme et son équipe (1992) ont mis en évidence en enregistrant des potentiels évoqués chez l’humain et chez le singe, une participation du cortex visuel primaire dans ce processus de ségrégation de texture (Lamme *et al.*, 1992).

L’activation du cortex strié durant une perception de textures a été confirmée en TEP et en IRMf (Beason-Held *et al.*, 1998; Gulyas *et al.*, 1998). Gulyas et ses collaborateurs ont enregistré en TEP les zones d’activation lorsque des sujets humains effectuent une tâche de segmentation de textures. Des zones d’activation significatives ont été relevées dans le gyrus fusiforme et le gyrus temporal supérieur (Gulyas *et al.*, 1998). Puce et ses collaborateurs (1996) mettent en évidence chez l’Homme en IRMf une activation significative d’une zone corticale située au niveau du sillon collatéral lors de présentations de textures aléatoires. Cette zone recouvrerait une région pouvant être homologue à l’aire V4. Une des premières études en IRMf sur la perception des textures montrent une activation significative du cortex extrastrié (aire 18/19) lorsque les stimuli sont composés d’une texture définie par des rectangles noirs et blancs disposés aléatoirement (Beason-Held *et al.*, 1998). Les auteurs remarquent qu’une aire située dans une zone antérieure au gyrus fusiforme s’active particulièrement pour ce type de pattern. Plus récemment, la définition des contours d’une scène visuelle sur la base de la texture a été étudiée en IRMf par Kastner et ses collaborateurs (Kastner *et al.*, 2000). Les stimuli représentent soit une texture uniforme ou soit plusieurs textures faisant apparaître des contours. L’activité dans V4 et l’aire TEO semble plus importante lors de la présentation de textures ségrégées que dans les aires visuelles dites inférieures (V1 et V2). Nous pouvons comparer les résultats de cette étude avec ceux d’une étude effectuée sur les contours illusoires (Mendola *et al.*, 1999) puisque, lors d’une

segmentation de texture, le contour n'est physiquement pas présent. Cette étude a également démontré l'activation significative de V4v lorsque les stimuli définissent des contours illusoires. La taille relativement restreinte des CR des neurones des aires primaires ne permet pas d'expliquer le codage pour les contours des textures. Une possibilité est que cette information soit transmise aux aires visuelles primaires par des projections FB issues d'aires hiérarchiquement supérieures (Hupe *et al.*, 1998). Ces aires supérieures intégreraient des informations de régions plus vastes du champ visuel grâce à une taille plus importante de leurs CR. Ainsi, les neurones dans V4 qui ont un CR plus grand que ceux des aires visuelles inférieures (V1, V2) pourraient participer à la définition des contours présents dans une scène visuelle sur la base de leur texture (Lamme *et al.*, 1992).

2. Reconnaissance des objets sur la base d'indices tridimensionnels

2.1. Encodage des indices binoculaires de la perception tridimensionnelle

a) Encodage de la disparité

La vision stéréoscopique nécessite une fusion binoculaire des deux images rétinien-nes. Au niveau neuronal, la sélectivité à la disparité est bien établie dans le cortex primaire (Hubel et Wiesel, 1970; Poggio et Fischer, 1977; Trotter, 1995) ainsi que dans les aires extrastriées de la voie dorsale.

Au niveau de V3 et V3A, les neurones sont organisés en colonne selon leur sélectivité à la disparité binoculaire (Adams et Zeki, 2001). La présence de neurones influencés par la disparité a également été démontrée dans l'aire MT où les neurones sélectifs à la disparité sont organisés en colonnes corticales selon leur disparité préférée (DeAngelis et Newsome, 1999). Des microstimulations appliquées dans ces colonnes perturbent les décisions comportementales d'un singe effectuant une tâche de discrimination basée sur la disparité (DeAngelis *et al.*, 1998). Par ailleurs, des projections directes entre V3, V3A et l'aire CIP ont été mises en évidence (Adams *et al.*, 1997) ; ceci permettrait la transmission des informations de disparité à l'aire CIP. En effet, le gradient de disparité binoculaire peut être utilisé pour définir l'orientation d'une surface tridimensionnelle au niveau de l'aire CIP. Pour le démontrer, Taira et ses collaborateurs (2000) ont enregistré dans l'aire CIP des neurones sélectifs à cet indice qu'ils nomment SOS (surface-orientation-selective) (Taira *et al.*, 2000). Ces neurones SOS sont sélectifs à une orientation d'une surface si l'orientation est définie par le gradient de disparité. De plus, certains de ces neurones SOS répondent également

sélectivement à des gradients de textures et à des perspectives linéaires. Leurs réponses neuronales augmentent si ces indices sont combinés (Tsutsui *et al.*, 2001).

La présence de la sélectivité à la disparité binoculaire a également été démontrée au niveau de l'aire TE, issue finale de la voie ventrale. Les neurones au niveau du cortex inférot temporal sont sensibles à la disparité horizontale (Uka *et al.*, 2000), à des gradients de disparité (Janssen *et al.*, 2000), à des formes définies par des stéréogrammes (RDS) (Tanaka *et al.*, 2001a) et à des formes 3D définies par des gradients de disparité (Janssen *et al.*, 1999; Janssen *et al.*, 2001). Ces résultats suggèrent que la disparité binoculaire soit traitée le long de la voie visuelle ventrale en parallèle à la voie visuelle dorsale. Toutefois, il est nécessaire de déterminer d'où les neurones de IT reçoivent leur sélectivité à la disparité.

Watanabe et al (2002) proposent que cette sélectivité soit présente avant l'aire IT et étudient donc le codage de la disparité binoculaire dans l'aire V4 (Watanabe *et al.*, 2002). Ils se sont basés sur des observations faites par De Yoe et Van Essen (1985) au cours d'enregistrements unitaires chez le singe anesthésié montrant une sensibilité particulière de quelques neurones de V4 à des indices de disparité (DeYoe et Van Essen, 1985). Les résultats de Watanabe et de ses collaborateurs (2002) rapportent qu'une population de neurones de V4 modulent leurs réponses en fonction des changements de disparité des stimuli (Watanabe *et al.*, 2002). Hinkle et al (2001) remarquent également que la majorité des neurones de V4 enregistrés (80%) est sensible à la disparité horizontale (Hinkle et Connor, 2001). Certains neurones présentent une sélectivité à l'orientation d'une barre 3D sur la base d'indices de disparité (Hinkle et Connor, 2002). Enfin, une étude intéressante de Hegde et Van essen (2005) démontre que, contrairement aux neurones du cortex inférot temporal (Janssen *et al.*, 2001), les neurones de V4 ne codent pas la courbure d'une forme 3D sur la base de la disparité (Hedge et Van Essen, 2005).

Ces études sur la disparité dans V4 (Watanabe *et al.*, 2002 ; Hinkle et Connor, 2001), ont utilisé uniquement comme stimuli disparates des barres appelées aussi - solide figure stéréograms ou -SFS-. Contrairement à ces études, Tanabe et al (2005) ont testé la disparité dans V4 avec des RDS et rapportent également une grande sélectivité des neurones de V4 à la disparité binoculaire (Tanabe *et al.*, 2005). Toutefois, dans cette même étude, d'intéressants résultats rapportent que la sensibilité à la disparité des neurones de V4 dépend de la nature du stimulus (SFS ou RDS). De la même manière, Hegdé et Van essen (2005) ont démontré une dépendance à la nature du stimulus sur les courbes de sélectivités à la disparité (Hegde et Van Essen, 2005). Les informations concernant la forme bidimensionnelle et la disparité

binoculaire seraient donc traitées de manière interdépendante. Les résultats des études sur le codage de la disparité dans V4 suggèrent donc que V4 soit une source potentielle de disparité pour l'aire IT.

Pour le codage de la disparité, Watanabe propose que les neurones de V4 ne détectent pas la disparité binoculaire mais reçoivent cette information de neurones sélectifs à la disparité situés dans des aires de plus bas niveau. En se basant sur les connexions anatomiques de V4, les informations relatives à la disparité pourraient être transmises à V4 via les bandes fines et les interbandes de V2, principales sources de projection vers V4 (DeYoe *et al.*, 1994; Felleman *et al.*, 1997b; Xiao *et al.*, 1999). Cependant, les neurones sélectifs à la disparité sont principalement concentrés dans les bandes épaisses de V2 qui projettent massivement vers la voie dorsale et très peu de neurones des bandes fines et interbandes sont sélectifs à la disparité (Hubel et Livingstone, 1987; Peterhans et von der Heydt, 1993; Roe et Ts'o, 1995). Watanabe *et al* (2002) émettent l'hypothèse que comme l'aire V4 reçoit des projections des aires dorsales V3, V3A et MT vers lesquelles les bandes épaisses projettent massivement, les informations relatives à la disparité seraient issues des bandes épaisses de V2 via les aires V3, V3A et MT.

b) Encodage de la perspective de gradient de texture

Peu d'études électrophysiologiques se sont intéressées aux bases neuronales traitant le gradient de texture ; une a été effectuée dans l'aire CIP et une autre au sein de l'aire TE (Tsutsui *et al.*, 2002; Liu *et al.*, 2004).

Tsutsui *et al* (2002) enregistrent l'activité des neurones dans la partie caudale du sillon intrapariétal (CIP). Les stimuli sont des plans inclinés comportant un gradient de texture tel qu'un fort percept de profondeur est présent. Ils démontrent que la majorité de neurones sont sélectifs aux gradients de texture. Ces neurones codent l'inclinaison tridimensionnelle de la surface définie par le gradient de texture. Leur réponse est, par ailleurs, invariante quel que soit le pattern de texture utilisé (points ou lignes). La plupart de ces neurones sont également sensibles au gradient de disparité. Ils pourraient donc intégrer les signaux concernant la disparité et la texture pour construire une représentation généralisée de l'orientation tridimensionnelle de la surface.

En enregistrant des cellules dans une région corticale autre que dans le sillon intrapariétal, Liu *et al* (2004) ont trouvé des neurones sensibles au gradient de texture dans la partie inférieure du sillon temporal supérieur, au niveau de l'aire TE. Ces neurones peuvent coder la profondeur en utilisant les gradients de texture comme indice et sont également

sélectifs à l'inclinaison définie par ce gradient de texture. Cette sélectivité persiste si l'inclinaison est également définie par des indices de disparité. Cette étude démontre également que le codage pour une inclinaison spécifique de la surface 3D est invariant pour plusieurs types de gradients de textures comme Tsutsui l'avait démontré dans CIP. Des neurones participant au codage de l'inclinaison d'une surface sur la base du gradient de texture se trouvent donc à la fois dans la voie dorsale et la voie ventrale.

En neuroimagerie, Shikata et al (2001), en IRMf, proposent aux sujets humains une tâche de discrimination d'orientation entre des surfaces définies par des gradients de texture (Shikata *et al.*, 2001). Cette étude démontre qu'une zone corticale s'active au niveau du sillon intrapariétal lorsque que les sujets effectuent la tâche de discrimination. Ces résultats sont en accord avec les données électrophysiologiques publiées par l'équipe de Tsutsui (2002). Plus récemment, dans notre laboratoire, une équipe a démontré, en enregistrant les potentiels évoqués de sujets humains, l'implication du cortex pariétal droit dans le traitement d'indices tridimensionnels comme la perspective linéaire ou des perspectives de gradients de textures (Severac Cauquil *et al.*, 2005).

c) Synthèse sur l'encodage des indices binoculaires

L'information tridimensionnelle concernant l'inclinaison d'une surface sur la base d'un gradient de texture ou de disparité semble donc être traitée à la fois dans la voie ventrale (au niveau de l'aire IT) mais aussi dans la voie dorsale (dans l'aire CIP). D'après Liu et al (2004), ces voies de traitement des informations concernant la profondeur des surfaces sont parallèles. Les informations au niveau de l'aire CIP permettraient de reconnaître la structure tridimensionnelle de l'objet afin que nous puissions s'en saisir. En effet, une sélectivité des neurones de l'aire AIP (qui se trouve en bordure antérieure du sillon IPS) a été démontré pour des objets 3D lorsque les animaux se saisissent des objets (Murata *et al.*, 2000). Au niveau du cortex IT, les informations 3D seraient impliqués uniquement dans la reconnaissance tridimensionnelle de l'objet où il a été clairement démontré que les neurones codent la forme 3D (sa courbure), sur la base d'indice de disparité et l'orientation de sa surface. Ces arguments sont exposés dans une récente revue (Orban et al, 2006). Quant au rôle de l'aire V4 dans le traitement de ces indices tridimensionnels, il semble que l'aire V4 code l'orientation tridimensionnelle d'une barre orientée (Hinkle et Connor, 2002) mais pas la courbure 3D sur la base du gradient de disparité. Il reste à démontrer si l'aire V4, principale source de projection vers le cortex inférot temporal est sensible au gradient de perspective de texture.

2.2. Encodage des indices monoculaires de la perception tridimensionnelle

a) Interposition

Plusieurs études électrophysiologiques ont cherché à définir les corrélats neuronaux responsables de la définition de contours illusoires. Von der Heydt et Peterhans ont étudié les bases neurophysiologiques de l'apparition de contours (von der Heydt et Peterhans, 1989) ou de surfaces illusoires (von der Heydt et Peterhans, 1989). Ils ont enregistré les réponses de neurones de V1 et V2 chez le singe macaque vigile à des contours illusoires présentés dans leur CR. Un seul neurone enregistré dans V1 répond au contour illusoire qui lui est présenté alors que près de la moitié des neurones de V2 répondent à ces contours illusoires. Les neurones de V2 répondent indifféremment à la position du contour dans leur CR et cela quelle que soit sa taille indiquant que c'est bien le contour illusoire qui est codé et non les stimuli inducteurs. De plus, ces neurones ont une même sélectivité à l'orientation que le contour soit illusoire ou réel. Heitger et al (1992) proposent un modèle où des neurones de V1 de type end-stopped¹⁷ projettent vers les neurones de V2 afin qu'ils codent la présence de contours illusoires ou réels (Heitger *et al.*, 1992).

Zhou et al (2000) testent également le codage de bordures au sein des aires V1, V2 et V4. Ils rapportent que plus de la moitié des neurones de V2 et de V4 sont influencés par la présence d'une bordure dans leur CR alors que moins de 20 % des neurones de V1 montrent une modulation de leur réponse (Zhou *et al.*, 2000). Baumann et al (1997) ont enregistré des neurones sélectifs aux contours illusoires dans des aires visuelles primaires (V1, V2, V3/V3A) du singe (Baumann *et al.*, 1997). Les neurones de V1 sont sélectifs à plusieurs paramètres des contours illusoires ; par exemple, à l'orientation du contour¹⁸ (Heider *et al.*, 2002) ou à la polarité du contraste entre la figure et le fond (Baumann *et al.*, 1997; Zhou *et al.*, 2000). Les singes lésés au niveau de l'aire V4 présentent d'importants déficits dans des tâches de discrimination d'orientation de contours illusoires. Ces observations indiquent que l'aire V4 pourrait jouer un rôle essentiel dans la perception des contours illusoires (De Weerd *et al.*, 1996).

b) Encodage des ombres et ombrages

Deux notions sont essentielles pour traiter de l'encodage des ombres et ombrages par le système visuel. La première concerne le codage des ombres et ombrages pour extraire

¹⁷ Ces neurones ne déchargent que lorsque qu'une ligne se termine dans leur CR et non pas lorsque celle-ci la dépasse. Ils codent donc la terminaison des lignes.

¹⁸ L'orientation du contour illusoire est confirmée par des humains dans un test de psychophysique.

l'information tridimensionnelle présente dans l'image et la seconde, le codage d'un objet indépendamment de la direction d'illumination.

(1) Extraction de la forme 3D de l'objet à partir d'ombrage

Dans les aires V1 et V2, Lee et al (2002) ont testé la réponse des neurones à un paradigme de pop-out basé sur les ombrages (Lee *et al.*, 2002). Le pattern de stimuli correspondait à des disques ombragés similaires à ceux présentés dans les études de Ramachandran (voir). Cette étude utilise également le même type de contrôle 2D: des disques bicolores en noir et blanc. Parmi une série de disques, les singes doivent reconnaître un disque différent en terme de pattern d'ombrage. Lee et son équipe vont alors effectuer des enregistrements cellulaires dans les aires V1 et V2 à différents moments de l'apprentissage et comparent l'activité neuronale de ces deux aires dans ces différentes situations. Une augmentation significative au niveau de la réponse de la population de V1 apparaît une fois le test appris par rapport à la situation de départ. Cette augmentation est visible uniquement pour des patterns 3D. Au niveau de l'aire V2, la population de neurones semble sensible au paradigme de *Shape from shading* avant les séances d'entraînement et cette sélectivité s'accentue avec l'entraînement.

Hanazawa et Komatsu (2001) ont utilisé un même pattern de disques ombragés que Lee et son équipe mais de plus petite taille. Ils ont enregistré la réponse neuronale de cellules dans l'aire V4. Le paradigme utilisé dans cette expérience est une tâche de fixation simple. Cette étude démontre l'existence d'un biais dans la réponse de neurones à un pattern de texture illuminé selon plusieurs angles (0 à 360° par pas de 45°). Les réponses neuronales d'une majorité de cellules de V4 sont plus importantes lorsque que le gradient d'illumination est vertical (90° et 270°) (Hanazawa et Komatsu, 2001). Selon les auteurs, ces cellules participent à la perception de textures définies en 3D sur la base d'ombrages.

Sereno et ses collaborateurs (2002), en présentant à des singes anesthésiés des formes 3D définies par des ombrages, ont relevé des foyers d'activation proches du STS, dans la bande inférieure qui pourraient correspondre à des activations dans l'aire TE (Sereno *et al.*, 2002).

Chez l'Homme, Humphrey (1997) décrit que dans des tâches où les individus sont confrontés passivement à des objets illuminés selon différentes directions, le signal IRMf n'apparaît que dans les aires visuelles précoces (V1, V2 et V3) (Humphrey *et al.*, 1997). Il utilise les mêmes stimuli que Ramachandran. Contrairement aux résultats de Hanazawa et Komatsu (2001), Humphrey et ses collaborateurs démontrent qu'un gradient horizontal au

niveau des stimuli entraîne une meilleure activation des aires visuelles précoces qu'un gradient vertical. Par ailleurs, ce résultat est en désaccord avec les conclusions de l'équipe de Ramachandran où un gradient vertical d'illumination entraîne une meilleure perception tridimensionnelle qu'un gradient horizontal (Ramachandran, 1988; Kleffner et Ramachandran, 1992). Toutefois, Humphrey émet l'hypothèse que les stimuli illuminés selon un gradient vertical paraissent plus familiers et entraîneraient donc une moindre activation.

Par la suite, Taira et al (2001) ont cherché à déterminer, en IRMf, quelles aires corticales sont activées suite à une présentation de surfaces tridimensionnelles définies uniquement grâce à des ombrages (Taira *et al.*, 2001). Ils ont isolé une aire dans le cortex intrapariétal droit qui participerait à la discrimination de formes à partir d'ombrages. Cette aire semble être l'homologue de l'aire CIP chez le singe où les cellules répondent à la fois à disparité, à la perspective linéaire et au gradient de texture (Tsutsui *et al.*, 2001; Tsutsui *et al.*, 2002).

Moore et Engel (2001) ont enregistré, en IRMf, une différence d'activation au niveau de l'aire LOC, entre des présentations d'objets 2D et des présentations de volumes où seuls les ombrages apportent une impression de profondeur. Lorsque les sujets perçoivent des objets comme définis en 3 dimensions (réponse comportementale relevée), l'aire LOC s'active alors que s'ils perçoivent l'image en 2D, l'activité au niveau de LOC décroît (Moore et Engel, 2001). Cette activation significative de l'aire LOC lorsque que les objets 3D sont définis par des ombrages avait été préalablement décrite dans l'étude de Kourtzi et Kanwisher (Kourtzi et Kanwisher, 2000). Toutefois, leurs résultats diffèrent de ceux de Moore et Engel puisqu'ils ne trouvent pas de différence significative entre les activations pour des stimuli 3D et des stimuli 2D même si les régions corticales activées se superposent. Mais il est bon de souligner que les stimuli 2D de l'étude de Kourtzi et Kanwisher comportent des contours internes qui donnent une impression de profondeur.

Dans une autre étude, Humphrey et son équipe testent la capacité de la patiente D.F. à reconnaître des formes définies grâce aux ombrages dans une tâche de discrimination entre les gradients d'illumination (avec des disques de type « Ramachandran », voir **Figure 1-15**). La patiente D.F. perçoit correctement la structure tridimensionnelle des disques (donc l'orientation des gradients d'illumination) en indiquant s'ils sont concaves ou convexes.

(2) Encodage de l'invariance à la direction d'illumination

Peu d'études physiologiques ont cherché à démontrer une invariance de la reconnaissance des objets quelle que soit la direction d'illumination. Chez l'Homme,

Warrington (1985) a décrit chez des patients lésés au niveau pariétal postérieur droit une certaine difficulté à reconnaître des mêmes objets illuminés différemment (Warrington, 1985). Chez le singe, Weiskrantz et Saunders (1984) indiquent que des animaux lésés au niveau de IT et des aires préstriées présentent aussi des difficultés à reconnaître des objets illuminés différemment (Weiskrantz et Saunders, 1984).

Vogels et Bierderman (2002) se sont intéressés au codage de la direction d'illumination dans l'aire IT de singes vigiles. Ils présentent en fovea aux animaux des formes complexes diverses (des géons) illuminées selon 5 directions (de face, 45° à droite, 45° à gauche, au-dessus et en dessous). La majorité des neurones enregistrés ne montrent pas de modulation de leurs réponses selon la direction d'illumination utilisée. Ils rapportent, par contre, une grande sélectivité pour les différentes formes. Les neurones de IT semblent donc invariants à la direction de l'illumination pour leur forme préférée (Vogels et Biederman, 2002). Ces résultats sont en accord avec ceux publiés par Hietanen et son équipe (1992) qui enregistrent au niveau du sillon temporal supérieur (STS) une petite sous population de neurones invariants (n=20) à la direction d'illumination pour une présentation de visages (Hietanen *et al.*, 1992).

2.3. Conclusion sur les indices tridimensionnels.

Les études décrites dans ce paragraphe ont mis en évidence deux traitements parallèles des informations tridimensionnelles. Les travaux du groupe de Tsutsui ont démontré qu'une sous-région du cortex pariétal, l'aire CIP chez le singe, semble être un centre nerveux hiérarchiquement élevé pour la perception des structures tridimensionnelles (Taira *et al.*, 2000; Taira *et al.*, 2001; Tsutsui *et al.*, 2001; Tsutsui *et al.*, 2002; Tsutsui *et al.*, 2005). En effet des neurones sélectifs au gradient de perspective, au gradient de texture, au gradient de disparité ou au gradient d'illumination ont été enregistrés dans cette région corticale. Les études du groupe de G. Orban ont démontré que les neurones au sein du cortex inférot temporal sont influencés par des indices tridimensionnels comme la disparité ou le gradient de texture (Janssen *et al.*, 1999; Janssen *et al.*, 2000; Liu *et al.*, 2004). A l'heure actuelle, les contributions respectives de l'aire dorsale CIP et de l'aire ventrale IT dans le codage de la forme 3D de l'objet ne sont pas clairement établies. Il est possible que l'aire CIP code la profondeur de la surface d'un objet et que cette description permette ainsi de se saisir de l'objet ; l'aire AIP étant relativement proche de CIP. Quant à l'aire IT, elle permettrait la reconnaissance tridimensionnelle de l'objet. Les informations sur les indices 3D qui transitent vers IT seraient issues d'aires visuelles ventrales plutôt que dorsales du fait des connexions

anatomiques. En effet, les études de Hinkle rapportent que les neurones de V4 sont sélectifs à l'orientation 3D d'une barre (Hinkle et Connor, 2001; Hinkle et Connor, 2002) et ces mêmes neurones de V4 sont sensibles au gradient d'illumination (Hanazawa et Komatsu, 2001).

Au cours de cette introduction, nous avons établi que l'aire V4, située au sein de la voie visuelle ventrale, participe à la reconnaissance de l'objet en traitant sa forme 2D, sa couleur ou encore sa texture. Les études sur le traitement de la couleur de V4 ont démontré que l'aire V4 joue un rôle fondamental dans la constance de la couleur, propriété de surface de l'objet. Cette aire nous permet donc de reconnaître la couleur d'un objet indépendamment des conditions d'illumination. De plus, il a été clairement démontré que l'aire V4 joue également un rôle dans le traitement d'une autre propriété matérielle de l'objet, la texture. Les études sur la texture dans V4 ont porté quasi exclusivement sur la segmentation de textures (Kastner *et al.*, 2000; Merigan, 2000). Il reste à démontrer le rôle de l'aire V4 dans la perception de la texture en tant que propriété matérielle de surface d'un objet. Ceci implique d'une part la présence d'un codage spécifique de la distribution des ombrages présents sur l'image d'une texture. D'autre part, nous avons établi dans l'introduction que la reconnaissance de ces propriétés matérielles doit être indépendante des conditions d'illumination qui varient en permanence dans un environnement naturel. Il est alors intéressant de faire le parallèle avec la constance de couleur présente dans V4: nous chercherons à démontrer la présence d'un codage de la texture constant en fonction de différents angles d'illumination. Est-il possible que les neurones de V4 reconnaissent la texture naturelle d'un objet indépendamment des conditions d'illumination et participent ainsi à une **constance de la texture** ?

Outre son rôle dans le traitement d'indice 2D, plusieurs études ont démontré que l'aire V4 traite également des indices tridimensionnels tels que la disparité et les gradients de texture (Hanazawa et Komatsu, 2001; Hinkle et Connor, 2001; Hinkle et Connor, 2002). De plus, des études de neuroimagerie chez l'Homme soulignent que la perception de la forme tridimensionnelle d'un objet sur la base d'ombrages s'effectue au sein du cortex inférot temporal, principale source de projection de V4 (Kourtzi et Kanwisher, 2000; Vogels et Biederman, 2002; Moore et Armstrong, 2003). De même, lorsqu'un objet comporte des concavités éclairées par différentes sources de lumières, il peut être reconnu quelle que soit la source d'éclairage qui créent des patterns d'ombrages différents. Au vu des propriétés 3D énoncées ci-dessus et du rôle possible de V4 dans la constance de perception, notre seconde étude s'intéresse à l'influence des ombrages sur les réponses des neurones de V4.

Chapitre 2

Matériel et méthodes

Ce chapitre présente la méthodologie générale utilisée durant les différentes expérimentations effectuées au cours de cette thèse. Ce chapitre décrit dans un premier temps (A) la préparation des animaux (les phases de conditionnement et les interventions chirurgicales) et les conditions d'enregistrements des données expérimentales (B). Les stimuli visuels utilisés lors des expériences (C), les analyses statistiques effectuées sur ces stimuli visuels (D) et les analyses des données (E) sont décrits dans les autres sous-parties de ce chapitre.

A. ANIMAUX

Pour mon étude de thèse, les singes impliqués sont deux macaques rhésus (*Macaca Mulatta*), issus d'un centre d'élevage. Un mâle (Zrem, 4 ans) de 6 kg et une femelle (Tima, 8 ans) de 4 kg ont été conditionnés pour ces expériences. Les protocoles effectués sont en accord avec la législation française et les recommandations européennes sur les soins et l'utilisation des animaux de laboratoire (EUVD 86/609/EEC). La préparation des animaux pour des enregistrements cellulaires se fait en plusieurs mois et comporte 4 grandes étapes : (1) une première phase d'habituation au poste expérimental, (2) un premier acte chirurgical pour la mise en place d'un dispositif de fixation de la tête, (3) une phase de conditionnement et enfin (4) un second acte chirurgical pour la mise en place de la chambre d'enregistrement.

1. Habituation

Cette première phase consiste à habituer l'animal à l'expérimentateur et aux protocoles de sorties de cage, de mise en place dans la chaise de contention (CRIST Instruments) et de retour dans la cage. La capture de l'animal dans sa cage individuelle se fait à l'aide de cannes rigides qui permettent de l'attraper, de le tenir à distance et de le diriger dans les différentes phases de conditionnement (sortie ou entrée de cage, mise en place dans la chaise). Afin que l'animal s'habitue à la chaise de contention, il est exclusivement, dans un premier temps dans l'animalerie puis dans le poste d'expérimentation, nourri (fruit et boisson) une fois installé dans la chaise. Cette première phase d'habituation est assez rapide (1 à 2 semaines). Dès que l'animal est habitué à ces manipulations quotidiennes et ne montre aucun signe de nervosité particulier, une première intervention chirurgicale peut alors être effectuée.

2. Dispositif de fixation

L'animal subit deux interventions chirurgicales sous anesthésie générale. La première intervention consiste à implanter une pièce de fixation de tête en inox sur le sommet du crâne de l'animal et a lieu en début de conditionnement. Ce système permet de maintenir la tête fixe lors des séances d'enregistrement. L'opération a lieu sous anesthésie générale et dure entre 2 et 3 heures. Nous utilisons des instruments stérilisés et le milieu est aseptisé. Une injection intramusculaire de kétamine (20 mg/kg) est effectuée sur l'animal au préalable afin de le tranquilliser. 0.2 ml d'atropine sont également injectés afin de prévenir tout étouffement par

l'excès de salivation dû à l'injection de kétamine. Il est ensuite maintenu sous anesthésie générale à l'aide d'une perfusion intraveineuse d'alphadolone/alphaxolone (Saffan®). La tête de l'animal est placée dans un cadre stéréotaxique. Le sommet du crâne est mis à nu et 8 vis (0.8 mm) sont introduites dans l'os sans toucher la dure-mère. La barre métallique de fixation de la tête est placée au sommet du crâne de sorte que la barre soit verticale et en position médiane. L'ensemble de vis et la base de la barre métallique sont ensuite recouverts délicatement par un ciment (Palacos ®). Puis la peau est recousue. Suite à cette première intervention, des applications d'antibiotiques locales (Auréomycine) et des injections intramusculaires de pénicilline (extencilline ®) sont effectuées afin de prévenir tout risque d'infection. Après 3 semaines de repos, l'animal peut alors avoir sa tête fixée et accomplir une tâche de fixation simple.

3. Conditionnement : la tâche de fixation simple

Au début de la période de conditionnement de l'animal, le singe est placé en privation hydrique. La totalité du besoin hydrique quotidien est assurée pendant les sessions d'entraînement ou d'enregistrements, sous formes de gouttes d'eau délivrées après chaque récompense correcte. Ainsi, nous pouvons contrôler sa consommation hydrique quotidienne (1400 à 2400 essais corrects ; ce qui équivaut à un apport hydrique quotidien entre 100 et 200 ml). A la fin de semaine, une quantité équivalente d'eau absorbée dans la semaine chaque jour est délivrée aux animaux. Le poids de l'animal est également contrôlé chaque jour afin de prévenir toute privation trop sévère. Un apport supplémentaire en cage est effectué après la séance d'entraînement si besoin est.

Dans le dispositif expérimental, le singe est assis dans une chaise de contention et est placé dans l'obscurité, face à un écran situé à 57 cm de ses yeux. Le singe est maintenu « tête fixe » grâce à la pièce de fixation. Pendant le conditionnement, l'animal apprend à fixer un point central de couleur grise avec une précision de +/- 1° pendant une période de 1s et à ignorer les stimuli visuels présentés en même temps en périphérie. La tâche de fixation simple se déroule comme suit. Un essai débute lorsque l'animal fixe le point central au moins pendant 400ms. Un stimulus visuel est ensuite présenté pendant 250 ms. L'animal doit alors conserver son regard fixé sur le point central. Si la fixation n'est pas stoppée pendant la durée de présentation du stimulus, l'essai est comptabilisé comme correct. Chaque essai correct est récompensé par des gouttes d'eau délivrées par un dispositif CRIST situé au niveau de la bouche et relié à une électrovanne.

Les séances sont pilotées par un ordinateur de type PC qui contrôle les évènements de la tâche et l'enregistrement des données comportementales et électrophysiologiques grâce au logiciel CORTEX (Computerized Real Time Experiments, NIH).

4. Seconde intervention chirurgicale: la chambre d'enregistrement

Lorsque le singe a atteint un niveau de performance souhaitable (95% de fixations correctes) avec la tête fixée, il est préparé en vue des enregistrements électrophysiologiques. Une seconde intervention chirurgicale est nécessaire afin de positionner la chambre d'enregistrement. La chambre d'enregistrement consiste en un cylindre en inox, de diamètre interne de 18 mm où sera fixé le microdescendeur lors des enregistrements électrophysiologiques. L'anesthésie générale et la préparation de l'animal sont similaires à celles utilisées lors de la première intervention. La partie occipitale du crâne est découverte. L'endroit correspondant à la position de la chambre d'enregistrements est repéré grâce à des données stéréotaxiques utilisées dans une étude faite au niveau de l'aire visuelle V4 du macaque vigile (Girard *et al.*, 2002). Une trépanation du diamètre correspondant à celui de la chambre est ensuite entreprise. Des vis sont fixées dans l'os autour de la chambre et du ciment dentaire est appliqué afin de maintenir et d'assurer la solidité de la fixation de la chambre. Un cercle de silicone (Silastic) de même diamètre que la chambre est placé à même la dure-mère afin de limiter sa prolifération. La craniotomie est exploitée tant que la dure-mère qui repousse n'empêche pas la microélectrode de pénétrer. Lorsque la pénétration de microélectrode n'est plus possible, une petite intervention chirurgicale est effectuée chez l'animal anesthésié où des couches de tissu conjonctif ayant proliférés sont ôtées. Après chaque séance d'enregistrement, l'intérieur de la chambre est nettoyé à la bétadine et un antibiotique local (Auréomycine) est appliqué afin de prévenir toute infection. Le morceau de silicone stérilisé est replacé à la surface de la dure-mère et la chambre est alors refermée par un bouchon CRIST. Tous les instruments utilisés lors des soins (pinces, silastic et couvercle) sont stérilisés quotidiennement au four à chaleur sèche à une température de 140 °C pendant 4h.

5. Confirmation histologique et radiographique de l'emplacement de la chambre d'enregistrement

Afin de confirmer l'emplacement de notre chambre d'enregistrement au sein du cortex visuel, nous avons fait passer à nos 2 animaux une radiographie à rayon X de leur hémisphère

cérébral droit. Les radiographies obtenus sur Zrem n'étant pas exploitables, je ne présenterai que celle effectuée sur Tima (Figure 2-1).

On peut distinguer sur cette radiographie la fenêtre osseuse enlevée lors du second acte chirurgical pour mettre en place la chambre d'enregistrement (l'animal n'avait plus la chambre au moment de la radiographie).

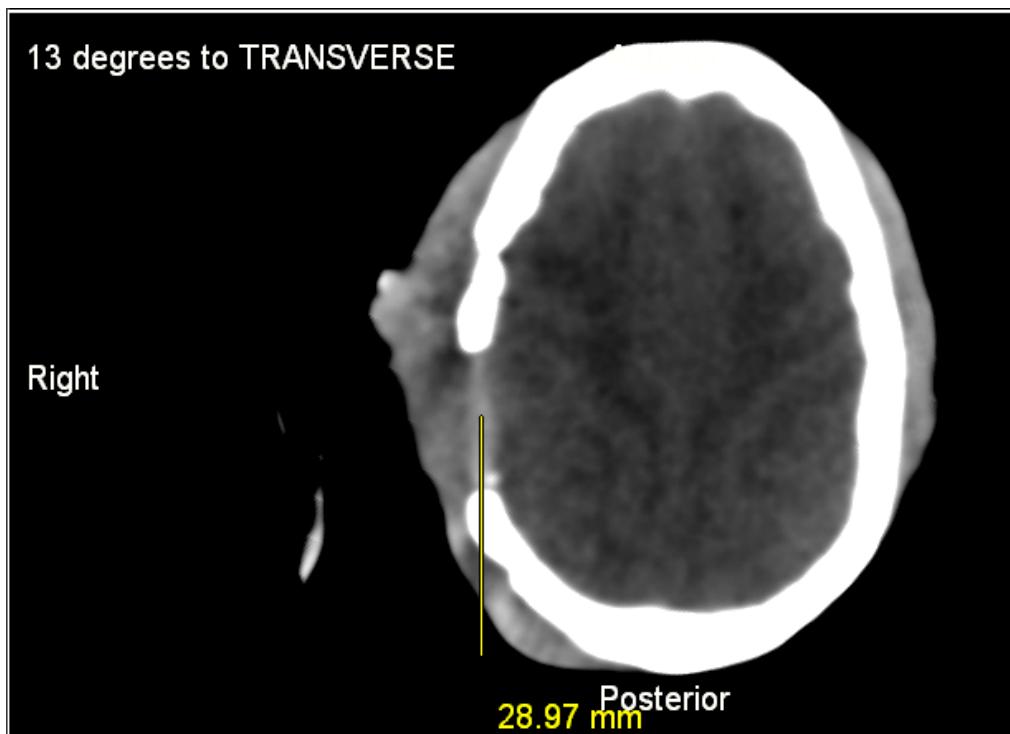


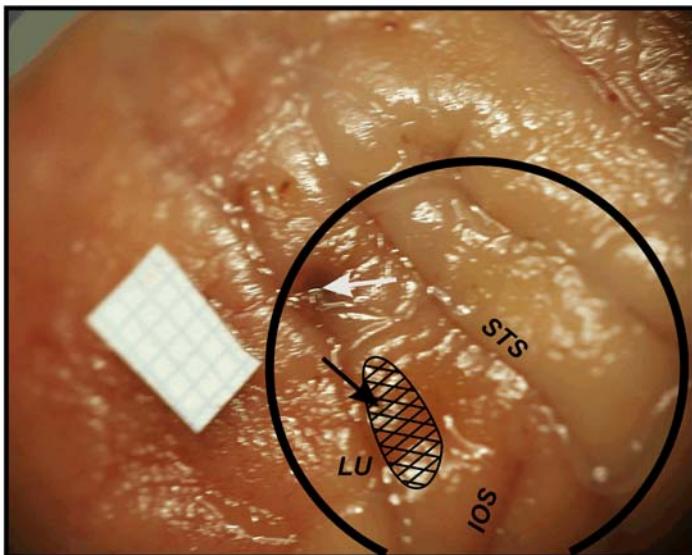
Figure 2-1 : Radiographie Scanner de Tima. Le plan de la coupe est de 13° par rapport au plan horizontal. Le périmètre crânien apparaît en blanc mettant en évidence la présence du défaut osseux où est située la chambre d'enregistrement. Le centre de ce défaut osseux est situé à 28.97 mm du bord postérieur du crâne de l'animal.

Pour le second animal, une étude histologique a été réalisée. Au niveau de la chambre d'enregistrement, nous avons injecté en plusieurs points de l'encre de tatouage. Ceci nous a permis de visualiser le positionnement de la chambre à la surface corticale (Figure 2-2A). Sur la photographie, nous pouvons apercevoir un de ces points (flèche blanche). Puis, sur le trajet des électrodes, nous avons injecté une dose de marqueur fluorescent (Dimamino Yellow, DY à 2%, flèche noire).

Pour l'euthanasie, l'animal reçoit une overdose de barbiturique (pentobarbital) et est perfusé au niveau cardiaque avec de la solution saline (0.9%) puis avec une solution de paraformaldéhyde à 4% pendant 40 minutes. Le cerveau de l'animal est ensuite prélevé et conservé dans une solution de sucre (30%). Les coupes du cerveau se font à un cryomac à –

30°C, les sections ayant une épaisseur de 40µm. Les coupes ont été réalisées frontalement. Les coupes sont ensuite colorées au Nissl et le site de marquage peut être révélé par fluorescence au microscope. Le positionnement de notre chambre d'enregistrement et la coupe histologique (Figure 2-2) confirment que nous avons bien enregistré nos cellules au sein de l'aire V4 chez cet animal.

A



B

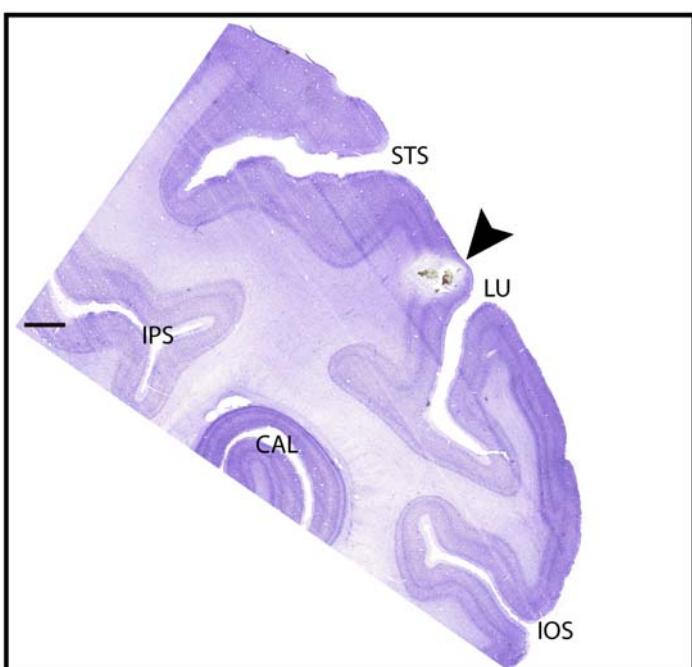


Figure 2-2 : Histologie post-mortem chez Zrem. A. Photographie de l'hémisphère droit. Le cercle noir schématise le positionnement de la chambre d'enregistrement et la zone hachurée, la région corticale de nos enregistrements. La flèche blanche indique l'emplacement d'un point d'injection d'encre de chine et la flèche noire, un point d'injection d'un marqueur fluorescent (DY, 2%). Du papier millimétré indique l'échelle. B. Coupe frontale du cerveau de Zrem. Coloration au Nissel. La flèche noire indique le site de marquage fluorescent. Les sillons sont indiqués. L'échelle est de 500µm. STS : Sillon Temporal Supérieur. LU : Sillon Luné. IOS : Sillon Inféro-Occipital. IPS : Sillon Inféro-Temporal. CALC : Fissure Calcarine.

B. ENREGISTREMENTS DES DONNEES

1. Position des yeux

Le contrôle de la position des yeux se fait grâce à une caméra infrarouge de marque ISCAN. Le calcul de la position de l'œil est basé sur un traitement numérique d'image. La caméra va capturer des images de l'œil en séquence (120 Hz). Différentes caractéristiques des images sont ensuite traitées pour en extraire la position de l'œil : les réflexions de lumière et les caractéristiques naturelles de l'œil. Les méthodes basées sur les réflexions de lumière utilisent le reflet d'une source lumineuse comme information utile. Lorsqu'un rayon lumineux est dirigé contre l'œil, il est réfléchi. La réflexion peut se faire sur 4 surfaces de l'œil : l'extérieur et l'intérieur de la cornée ainsi que l'extérieur et l'intérieur du cristallin. Ces 4 reflets sont communément appelés Images de Purkinje. Le logiciel ISCAN utilise cette image de Purkinje pour extraire la position de l'œil. La position de l'œil est ensuite transmise à l'ordinateur pilote (CORTEX) afin qu'il détermine si l'animal fixe ou pas.

2. Enregistrements électrophysiologiques

L'enregistrement extracellulaire de l'activité neuronale se fait grâce à des microélectrodes de tungstène (Thomas RERecording) de 80 μm de diamètre recouvertes d'une fine couche de verre, dénudées sur 20 à 50 μm au niveau de la pointe et d'impédance comprise entre 0,5 et 1,5 $\text{M}\Omega$. Les microélectrodes sont fixées à un microdescendeur TREC. Cet appareil permet de contrôler la descente de plusieurs microélectrodes (jusqu'à 5). Le plus souvent, seules 2 microélectrodes sont installées sur le microdescendeur au cas où une microélectrode serait défectueuse. Une canule guide stérilisée dans laquelle s'insèrent les microélectrodes permet de percer la dure-mère et de garantir la trajectoire verticale des électrodes. Les descentes sont commandées à l'aide d'un logiciel permettant une précision de 1 μm . L'ensemble est fixé à la chambre d'enregistrement sur la tête du singe. Les coordonnées rostrocaudale et médiolatérale de la pénétration sont déterminées grâce aux verniers du microdescendeur. Une fois les coordonnées choisies, l'ensemble guide-microélectrode est descendu manuellement à la surface de la dure-mère jusqu'à la pénétrer de quelques millimètres (3-4 mm) correspondant à l'épaisseur des couches méningées¹. Les microélectrodes sont préalablement rentrées dans leur canule guide afin d'éviter tout

¹ En condition normale, ces couches méningées ont une épaisseur de 1.5 à 2 mm. La craniotomie entraîne une prolifération des couches méningées. C'est pourquoi nous descendons l'électrode à une profondeur de 3-4 mm.

désagrément lors de la pénétration de la dure-mère. Une fois la dure-mère pénétrée, la microélectrode est ensuite descendue grâce au logiciel TREC jusqu'à la surface du cortex. Cette profondeur est considérée comme la référence à partir de laquelle l'électrode est descendue dans le cortex cérébral.

L'activité électrique des neurones est captée par l'électrode sous forme de signal analogique ; le signal est amplifié (X 10 000) et filtré (bande passante 0,5 à 5kHz) puis visualisé sur un oscilloscope couplé à un haut-parleur. Les potentiels d'action (PA) peuvent ainsi être reconnus à l'oreille. Lorsque le singe effectue la tâche de fixation simple, l'enregistrement de l'activité électrique corticale fait apparaître un signal électrophysiologique dense. La discrimination des neurones se fait grâce au logiciel MSD (AlphaOméga). Ce matériel permet d'isoler dans le signal électrophysiologique dense, la décharge d'un ou plusieurs neurones. Il faut pour cela, dans un premier temps, définir la forme du potentiel d'action à isoler afin de créer un modèle ou '*template*'. MSD compare continuellement l'activité enregistrée et le *template*. Il transforme ensuite les PA correspondant au *template* défini en impulsions digitales transmises à l'ordinateur pilote (CORTEX). MSD permet de contrôler en temps réel la forme du potentiel d'action discriminé. Ainsi l'unicité du neurone enregistré est facilement contrôlable.

2.1. Détermination du champ récepteur des neurones

Lorsqu'un neurone est isolé, la première étape consiste à déterminer son champ récepteur (CR). Afin de localiser le CR des neurones enregistrés, nous présentons manuellement dans le champ visuel inférieur gauche des barres orientées de différentes couleurs et en mouvement. Une fois le CR repéré, nous délimitons ses bordures pour estimer sa taille. Cette méthodologie présente comme inconvénient le fait que l'expérimentateur doive se baser sur le son du haut-parleur pour définir les limites du CR. La recherche du stimulus optimum (couleur, forme) nécessite également beaucoup de temps et d'essais pendant que le singe fixe le point central. Afin de limiter toute variabilité sur l'appréciation du CR et d'accélérer cette cartographie, nous avons utilisé un protocole inspiré de celui décrit par Jones et Palmer en 1987 basé sur le principe de la corrélation inverse et l'utilisation de M-séquence (Jones et Palmer, 1987). Ces auteurs ont créé un protocole expérimental afin de localiser rapidement les CR des neurones enregistrés dans le cortex primaire visuel du chat. Le protocole que nous avons utilisé consiste à l'affichage d'une séquence de stimuli visuels constitués de 144 carrés de même taille, noirs ou blancs. Chaque carré est défini par ses coordonnées dans le champ visuel. Chaque matrice (144 stimuli) est divisée en plusieurs

séquences d'affichage comprenant chacune 12 carrés. Une séquence correspond donc à une matrice de 12 X 12 carrés. Pour mettre en évidence de meilleures réponses des neurones, cette matrice est multipliée par deux : une avec des stimuli blancs, l'autre avec des stimuli noirs. Chaque séquence correspond donc à une matrice de 12X12X2 carrés. Nous avons pré-défini plusieurs séquences, avec des tailles de stimuli différents ($0,2^\circ$; $0,5^\circ$ et 1°). Nous avons utilisé également d'autres couleurs présentant une forte opposition chromatique (vert vs. rouge et bleu vs. jaune). Nous avons testé ces différentes séquences (taille ou couleur différentes) sur un nombre assez restreint de neurones. La majorité des séquences effectuées au début de chaque session d'enregistrement est constituée de carrés noirs ou blancs de taille 1° .

Un programme informatique créé dans CORTEX gère l'apparition des séquences. Chaque carré est flashé pendant 50 ms. Cette séquence va afficher pendant toutes les 50 ms, un carré, à une position aléatoire (dans un cadre proportionnel à la taille du stimulus) du champ visuel inférieur controlatéral, si et seulement si le singe fixe le point central. Au moins 5 essais par condition sont enregistrés pour chaque neurone. Un programme informatique créé dans Matlab permet ensuite de corrélérer, en quelques secondes, la décharge du neurone enregistré avec les localisations des carrés. Dès qu'un potentiel d'action est enregistré, nous déterminons quel est le dernier stimulus présenté lors des dernières 50 ms. La réponse du neurone est ainsi corrélée au stimulus trouvé. Nous pouvons donc mesurer les réponses des neurones par rapport à la position spatiale des stimuli. Le centre du CR est considéré comme la région du champ visuel où la réponse du neurone est maximale. L'animal fixant le point central durant toute la tâche, les CR des neurones restent constants et ne sont pas modifiés comme si l'animal bougeait ses yeux (Bremmer, 2000).

Cette cartographie en temps réel va permettre de positionner plus rapidement les CR afin de placer les stimuli en leurs centres. Ces données recueillies pour chaque neurone nous permettront également de réaliser une carte rétinotopique des aires visuelles V4 de nos deux animaux.

2.2. Séance d'enregistrements

Une fois le CR du neurone cartographié, nous plaçons les stimuli visuels au centre de celui-ci en adaptant la taille des stimuli à la taille du CR (voir paragraphe C.3). Les stimuli sont présentés de manière pseudo-aléatoire afin qu'au moins 5 répétitions d'une même image soient effectuées. Seules les cellules pour lesquelles au moins 5 répétitions ont été effectuées sont retenues pour les analyses. Au cours d'une séance quotidienne, le nombre de neurones enregistrés peut aller jusqu'à 6 neurones. En moyenne, il varie entre 1 et 2 neurones.

C. STIMULI VISUELS

Durant ma thèse, je me suis intéressé au codage d'indices picturaux au niveau de l'aire V4 du macaque. Pour cela, nous avons traité deux problématiques différentes. La première concerne les textures naturelles. Nous avons voulu tester si les neurones de V4 sont influencés par des textures naturelles aussi diverses que le béton, des brins de paille ou la peau humaine. Plusieurs sets d'images nous ont permis de mettre en évidence une sélectivité des neurones de V4 à la texture et si cette sélectivité varie selon l'angle d'incidence de l'illumination.

Dans la seconde partie, nous nous sommes intéressés à un autre indice pictural que la texture ; les ombrages. Cet indice est un puissant indice de profondeur et nous avons testé si les neurones de V4 traitent la 3D sur la base des ombrages. Pour cela, nous disposons d'images de formes tridimensionnelles (uniquement définies par les ombrages) illuminées selon différentes directions d'illumination. Les différents sets d'images utilisés seront détaillés dans le paragraphe suivant.

1. Les textures naturelles

Le set d'images pour tester l'influence des textures naturelles sur les réponses des neurones de V4 comprend plusieurs types de stimuli visuels: des stimuli originaux uniquement définis par une texture naturelle et des stimuli contrôles (stimuli *scrambled* ou réseaux orientés).

1.1. *Stimuli originaux : des photographies de textures naturelles*

Les stimuli originaux utilisés dans la première partie du projet de thèse proviennent de la base de données CUReT (Columbia Utrecht Texture Database). CUReT a été mise au point par l'équipe de Jan Koenderink (Université d'Utrecht, www1.cs.columbia.edu/CAVE/curet/). Cette base de données est une collection d'images de textures naturelles éclairées et photographiées sous différents angles incidents. Toutes les photographies utilisées pour notre projet représentent des textures naturelles photographiées de face, uniquement niveaux de gris. Au cours des enregistrements cellulaires, nous avons utilisé trois sets différents de photographies de textures naturelles.

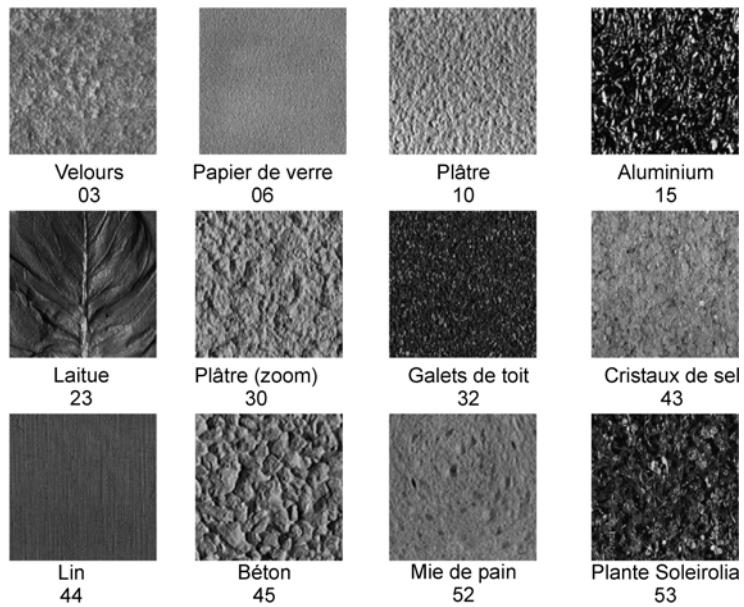
Le premier sous-ensemble de stimuli (**set A**) comprend 12 textures naturelles photographiées sous une illumination comportant 3 angles d'incidence différents (voir Figure 2-3A et B). L'angle L1 est de 22,5° par rapport à la verticale, L2 de 45° et L3 de 67,5° (Figure 2-3C). La Figure 2-3A présente les 12 textures naturelles utilisées dans le set A.

Le **set B** comprend les photographies de 65 textures naturelles illuminées par un angle L3. Ce set élargi de textures nous a permis de mieux caractériser la sélectivité aux textures naturelles des neurones de V4 (pour illustrations voir annexe).

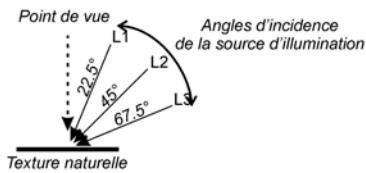
Le troisième set de stimuli (**set C**) est constitué de 41 textures naturelles dont font partie les 12 textures initialement testées dans le set A. Les photographies de ce set représentent des textures illuminées selon deux angles (L1 et L2). Nous avons limité ce dernier set de texture à 41 afin de réduire le nombre de conditions présentées à l'animal.

Les textures naturelles qui composent ces trois différents sets de stimuli sont résumées dans le Tableau 2-1 et sont illustrées en annexe.

A



B



C

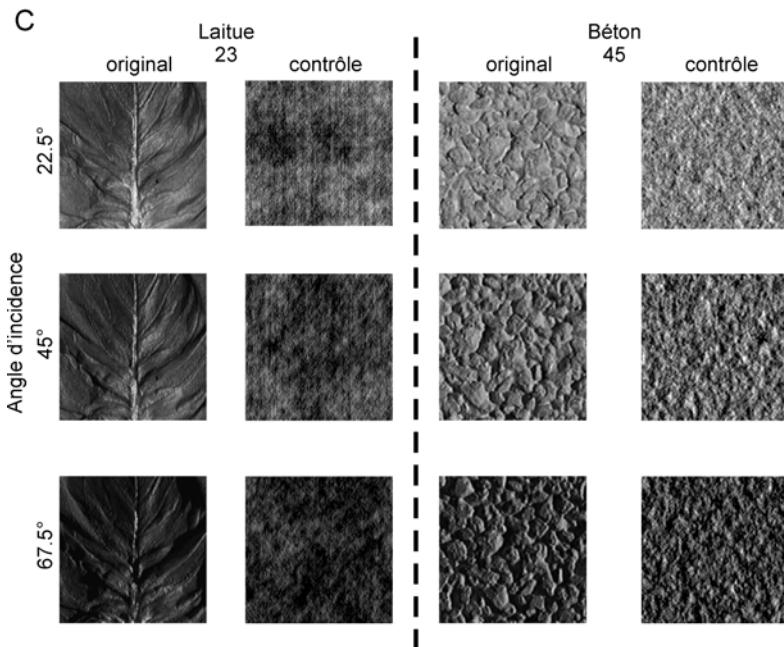


Figure 2-3 :A. Photographies des 12 textures naturelles constituant le set A. Ces photographies sont issues de la base de données CUReT. Le numéro en dessous du nom indique le numéro de la texture dans la CUReT. **B.** Schéma récapitulant les angles d'incidence d'illumination. **C.** Exemples de stimuli originaux et contrôle pour la *laitue* et le *béton*. Chaque texture naturelle est illuminée sous trois différents angles incidents ($22,5^\circ$, 45° et $67,5^\circ$). A chaque image correspond son stimulus contrôle dans lequel les phases du spectre de Fourier de l'image originale sont redistribuées aléatoirement, les amplitudes étant conservées.

Set A:

Velours (03_0)	Papier de verre (06_0)	Plâtre_a (10_0)	Aluminium (15_0)	Laitue (23_0)	Plâtre (z) (30_0)
Galets de toit (z) (32_0)	Sel (cristaux) (43_0)	Lin (V) (44_0)	Béton (a) (45_0)	Pain (mie) (52_0)	Soleirolia (53_0)

Set B :

Bitume (01_0)	Velours (03_0)	Plastique rug. (04_0)	Cuir (05_0)	Papier de verre (06_0)	Velours (07_0)
Cailloux (08_0)	Verre dépoli (09_0)	Plâtre_a (10_0)	Plâtre_b (11_0)	Papier rugueux (12_0)	Herbe artificielle (13_0)
Galets de toit (14_0)	Aluminium (15_0)	Liège (16_0)	Carreau (17_0)	Tapis_a (18_0)	Tapis_b (19_0)
Polystyrène (20_0)	Eponge (21_0)	Laine d'agneau (22_0)	Laitue (23_0)	Peau de lapin (24_0)	Tuile (25_0)
Loofa (26_0)	Fibre d'isolation (27_0)	Papier froissé (28_0)	Plâtre (z) (30_0)	Papier rug. (z) (31_0)	Galets de toit (z) (32_0)
Ardoise a (33_0)	Ardoise b (34_0)	Sphères peintes (35_0)	Roche calcaire (36_0)	Brique a (37_0)	Papier peint (V)I (38_0)
Papier peint (H) (38_1)	Peau humaine (39_0)	Paille (H) (40_0)	Paille (V) (40_1)	Brique b (41_0)	Velours b (H) (42_0)
Velours b (V) (42_1)	Sel (cristaux) (43_0)	Lin (V) (44_0)	Lin (H) (44_1)	Béton a (45_0)	Pierres (47_0)
Pain (48_0)	Béton b (49_0)	Béton c (50_0)	Maïs (V) (51_0)	Maïs (H) (51_1)	Mie de pain (52_0)
Plante Soleirolia (53_0)	Bois a (H) (54_0)	Bois a (V) (54_1)	Peau d'orange (55_0)	Bois b (V) (56_0)	Bois b (H) (56_1)
Queue de paon (57_0)	Ecorce d'arbre (58_0)	Cracker a (59_0)	Cracker b (60_0)	Mousse végétale (61_0)	

Set C :

Velours (03_0)	Plastique (04_0)	Papier de verre (06_0)	Velours (07_0)	Verre dépoli (09_0)	Plâtre_a (10_0)
Herbe artificielle (13_0)	galets de toit (14_0)	Aluminium (15_0)	Tapis_b (19_0)	Eponge (21_0)	laitue (23_0)
Papier froissé (28_0)	Plâtre (z) (30_0)	Papier rug. (z) (31_0)	Galets de toit (z) (32_0)	Ardoise b (34_0)	Sphères peintes (35_0)
Papier peint (V)I (38_0)	Papier peint (H)I (38_1)	Peau humaine (39_0)	Paille (H) (40_0)	Paille (V) (40_1)	Brique b (41_0)
Velours b (H) (42_0)	Velours b (V) (42_1)	Cristaux de sel (43_0)	Lin (V) (44_0)	Lin (H) (44_1)	Béton a (45_0)
Béton c (50_0)	Maïs (V) (51_0)	Maïs (H) (51_1)	Mie de pain (52_0)	Soleirolia (53_0)	Bois a (H) (54_0)
Bois a (V) (54_1)	Paon (queue) (57_0)	Cracker a (59_0)	Cracker b (60_0)	Mousse végétal (61_0)	

Tableau 2-1 : Liste des textures naturelles utilisées dans les 3 sets d'images A, B et C. (z) = zoom, (V) = vertical, (H) = horizontal. Plusieurs photographies peuvent représenter un même matériau (ex : béton a et béton b). Le numéro de chaque texture est indiqué en dessous du nom. Les deux premiers chiffres indiquent le numéro de la texture ; par exemple, lin : 44. Le dernier chiffre (0 ou 1) signale l'orientation de l'image. Le lin (44) est à la fois présent en position horizontale (44_1) et en position verticale (44_0). Voir annexes pour illustration.

1.2. *Stimuli contrôles : une distribution aléatoire des phases*

Pour chaque stimulus, un stimulus contrôle est créé ; il contient les mêmes amplitudes que l'image originale, seules les phases de son spectre de Fourier sont redistribuées aléatoirement (Figure 2-3B). Ces stimuli sont aussi appelés stimuli « *scrambled* ». Cette distribution aléatoire des phases a la particularité de conserver des paramètres statistiques caractéristiques de l'image, comme la luminance moyenne ou le contraste moyen mais fait disparaître toute reconnaissance possible de l'image originale. Ceci a été démontré dans le cas où l'image originale représenterait une scène naturelle (Tadmor et Tolhurst, 1993).

1.3. *Réseaux orientés (voir annexes pour illustration)*

Pour un sous-groupe de neurones ($n = 15$), nous avons remplacé les stimuli *scrambled* du set C par une série de réseaux orientés afin de caractériser les courbes de préférences à l'orientation et à la fréquence spatiale des neurones enregistrés. Les réseaux utilisés ont une fréquence spatiale s'échelonnant de 0,4 à 12,8 cycles/gré. La fréquence spatiale est augmentée d'une octave d'un stimulus à l'autre (0,42 ; 0,85 ; 1,66 ; 3,33 ; 7,15 ; 12,5 c/deg). Les différentes orientations sont au nombre de 8 s'échelonnant de 0 à 157,5° par pas de 22,5°. La taille des réseaux orientés correspond à la taille des stimuli de textures (cf.paragraphe C.3.) Ils sont présentés à la même position que les autres stimuli.

2. Formes tridimensionnelles

La seconde partie des enregistrements cellulaires concerne le traitement de la profondeur sur la base des ombrages. Pour cela, nous disposons d'images 2D représentant des formes tridimensionnelles où seul l'indice « *ombrage* » nous permet d'appréhender la 3D. Ces images sont aussi appelées « *stimuli 3D* ». Nous avons également à notre disposition plusieurs stimuli contrôles où la perception de la profondeur est absente (2D blobs, random et postérésés).

2.1. *Les stimuli 3D*

Pour le second protocole, nous disposons d'une série d'images en niveaux de gris de 8 formes différentes donnant une impression de relief grâce à la présence d'un gradient de luminance. Ces formes consistent en des sphères déformées par des convexités et des

concavités aléatoires. Elles ont été créées sur la base des travaux de Norman et Todd (Norman et Todd, 1995). Le set d'images utilisé comprend 8 formes 3D illuminées selon 4 directions : par-dessus, par la droite, par la gauche ou par en dessous, soit 32 images 2D représentant des formes 3D sur la base d'ombrages (Figure 2-4).

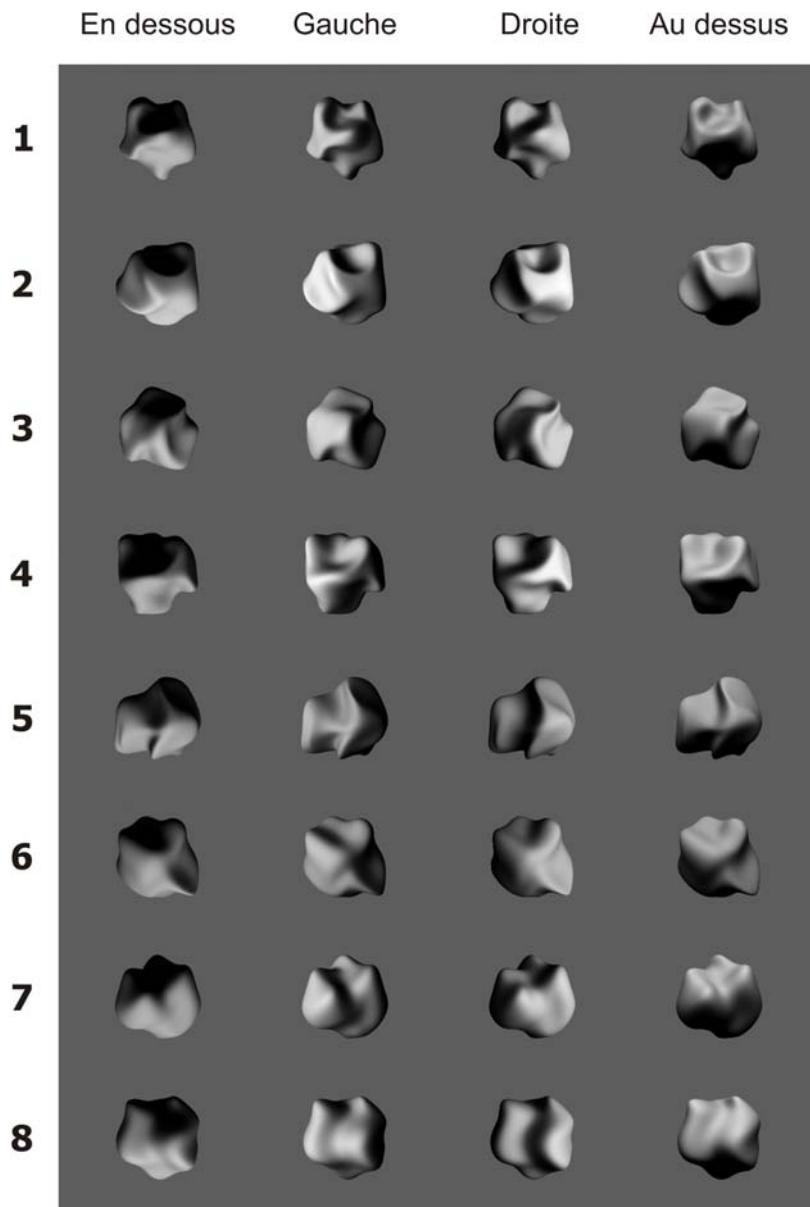


Figure 2-4 : Illustration des 32 stimuli 3D. Chaque ligne représente une forme tridimensionnelle (1 à 8) et chaque colonne indique la direction d'illumination (en dessous, à gauche, à droite et au dessus).

2.2. *Les stimuli contrôles : 2D blobs, random et postérésés*

Afin de tester l'indice « ombrage » comme indice de perception tridimensionnelle, nous avons créé plusieurs types de stimuli contrôles dont la principale caractéristique est la disparition du percept tridimensionnel. Cette perte du percept tridimensionnel pour les stimuli contrôles a été testée en psychophysique chez l'Homme par une équipe de chercheurs

travaillant sur le même projet (voir annexe). De tous les stimuli, seuls les stimuli originaux 3D apparaissent tridimensionnels pour les sujets humains (n=8). Les résultats sont présentés en annexes.

A chaque stimulus correspondent trois stimuli contrôles différents (Figure 2-5). Pour chaque type de stimuli contrôle, le contour externe propre à chacune des 8 formes tridimensionnelles est conservé. Le premier type de contrôle (Figure 2-5B) est appelé '**2D blob**'. Ils ont été créées par l'équipe de Rufin Vogels à l'université de Louvain, grâce au logiciel Studio max. Cet outil informatique a permis de créer une texture définie par une même distribution des niveaux de gris. Les stimuli 2D blob respectent donc les paramètres de 1^{er} ordre du stimulus original (luminance moyenne, contraste, etc..) ainsi que la majorité des paramètres de second ordre, hormis l'inertie (voir paragraphe D.1 pour les définitions et annexes pour illustrations). Le second type de stimulus contrôle est le type **random** présenté dans la Figure 2-5C (également créée par l'équipe de R.Vogels). Il s'agit de créer un nouveau stimulus en opérant à une redistribution aléatoire des pixels dans la forme initiale. Ce type de contrôle respecte les paramètres de premier ordre (histogramme de distribution identique) mais non de second ordre des stimuli originaux.

Enfin, le troisième type de stimuli contrôle (stimulus **postérisé**) a été créé suite à des observations concernant les stimuli contrôles 2D blobs faites durant ma thèse. En effet, alors que les stimuli 2D blobs gardent les mêmes paramètres que les images originales, la distribution des patterns noir et blanc n'est pas la même entre les stimuli originaux et les stimuli contrôles. Par exemple, le stimulus contrôle '2D blob' de la Figure 2-5B présente une tâche blanche au centre qu'on ne retrouve pas dans l'image originale. Ceci peut théoriquement poser un problème si nous enregistrons des neurones dont les champs récepteurs présentent des sous régions ON ou OFF. Afin de mieux respecter la disposition des patterns noirs et blancs présents dans l'image originale, nous avons décidé de créer un nouveau type de contrôle 2D: les stimuli **postérisés**. Ces stimuli contrôles ont été créées grâce au logiciel Adobe photoshop. Nous en avons élaboré différents types en choisissant plusieurs niveaux de gris (n = 2, 4, 6 ou 8). Nous devions choisir ensuite parmi ces exemples de stimuli, le nombre de niveaux de gris à utiliser afin d'éliminer le percept tridimensionnel, tout en conservant une même distribution de patchs noirs et blancs. Pour cela nous avons testé ces stimuli chez des sujets humains (n= 6, pour résultats voir annexes) en présentation foveale ou périphérique. Les résultats du test psychophysique indiquent que le percept tridimensionnel

des formes disparaît avec 2 niveaux de gris. En conséquence, les stimuli postérisés qui ont été utilisés présentent seulement 2 niveaux de gris.

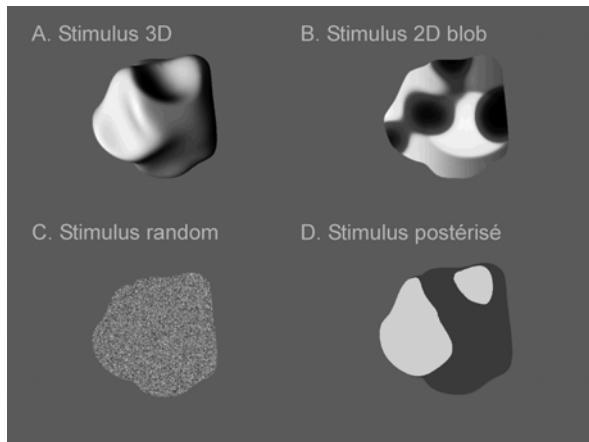


Figure 2-5: 3 types de stimuli contrôles. (A) l'objet tridimensionnel n°2 illuminé par la gauche et ses stimuli contrôles correspondant : stimulus 2D blob (B), stimulus 2D random (C) et stimulus 2D postérisé (D).

Tous les stimuli contrôles (2D blob, random et postérisés) sont illustrés en annexes.

3. Formes et tailles des stimuli

Pour le protocole de textures naturelles, les stimuli ont une forme carrée. Pour les 2 protocoles, la taille des stimuli est adaptée à la taille et à l'excentricité du champ récepteur du neurone enregistré selon la formule publiée par Gattass et col (1988) : $1^\circ + 0,625 \times \text{excentricité}$. Pour garder une certaine reproductibilité entre les enregistrements des neurones, nous avons utilisé 5 tailles différentes (2, 3, 4 et 5 DAV²). Chacun des stimuli de l'étude est présenté au moins 5 fois, dans un ordre aléatoire. D'après nos cartographies, on peut estimer qualitativement que les stimuli recouvrent en moyenne 75% du CR du neurone enregistré.

NB : Pour un nombre restreint de cellules (n=39 cellules), les stimuli du set A sont présentés en position centrale avec une forme carrée (9°X9°). Ce protocole était utilisé par l'équipe de R. Vogels lors des enregistrements dans le cortex inféro-temporal. Nous l'avons répliqué dans un premier temps, avant d'adapter la taille et la position des stimuli aux CR des cellules de V4. Les neurones enregistrés avec ces 2 types de stimuli sont regroupés ensemble et seront traités dans une même partie des résultats (Chap3, partie B).

² DAV: Degré d'angle visuel. La taille 2 correspond à un stimulus de 100 pixels, la taille 3 à 150 pixels, la taille 4 à 200 pixels.

D. ANALYSE STATISTIQUE DES IMAGES

Les stimuli décrits ci-dessus sont des textures naturelles ou des formes tridimensionnelles, chacun étant original. Afin de mieux caractériser ces stimuli visuels les uns par rapport aux autres, nous avons utilisé des méthodes d'analyses d'image (analyse de textures notamment). Les méthodes d'analyse de texture sont nombreuses et ce n'est pas l'objet d'en faire une revue exhaustive ici. Elles sont utilisées dans un grand nombre d'applications, aussi bien pour des techniques d'imagerie biomédicales que pour la segmentation d'images satellitaires. Ces méthodes d'analyse de textures ont pour but de quantifier les éléments descriptifs de la texture par des paramètres mathématiques qui nous permettent de les identifier parmi un ensemble de textures. Plusieurs méthodes, de variantes et de combinaisons de méthodes ont déjà été proposées dans la littérature et éprouvées en pratique : la méthode structurelle, la méthode basée sur l'étude des formes, les méthodes spatio-fréquentielle, la méthode fractale et les méthodes d'analyse statistique. Les méthodes d'analyse les plus communément utilisées sont les méthodes d'analyse statistique de premier ordre et de second ordre. Nous les avons utilisées pour paramétrier nos stimuli ; je ne détaillerai donc pas les autres méthodes.

Les calculs statistiques que je vais présenter concernent les images de textures naturelles mais ils ont également été effectués sur les images de formes tridimensionnelles ainsi que leurs images contrôles.

1. Méthodes statistiques d'analyse de textures

Pour analyser une texture à l'aide de méthodes statistiques, la texture est considérée comme la réalisation d'un processus stochastique stationnaire. Les méthodes statistiques étudient les relations entre un pixel et ses voisins. On distingue plusieurs ordres d'analyse statistiques (méthodes de premier, second ou troisième ordre). Plus l'ordre de la statistique est élevé, plus le nombre de pixels voisins considérés (0 à n) est important. Pour cette étude, seuls les paramètres de premier ordre et de second ordre ont été calculés.

1.1. *Les variables de 1^{er} ordre*

L'étude de l'analyse de textures s'effectue sur des images numérisées. L'analyse de texture consiste à calculer à l'aide de l'intensité des pixels de chaque image, un certain nombre de paramètres statistiques caractéristiques de la texture à étudier (Haralick, 1979).

Les paramètres statistiques sont calculés à partir de l'histogramme des intensités. Celui-ci décrit pour chaque image la fréquence d'apparition de chaque niveau de gris. La luminance moyenne, le contraste moyen (RMS), la variance, le skewness (asymétrie), le kurtosis (aplatissement) et le SNR (rapport signal bruit) sont les paramètres les plus souvent utilisés pour caractériser une texture. Les paramètres sont calculés sur des images de taille 150 X 150 pixels. Les 256 niveaux de gris varient de 0 à 1 (0 = pixel noir et 1= pixel blanc). Prenons comme exemple trois textures issues du panel de stimuli ; *le plâtre (z) (#30)*, le *papier crépi (#12)* et *le verre dépoli (#09)*.

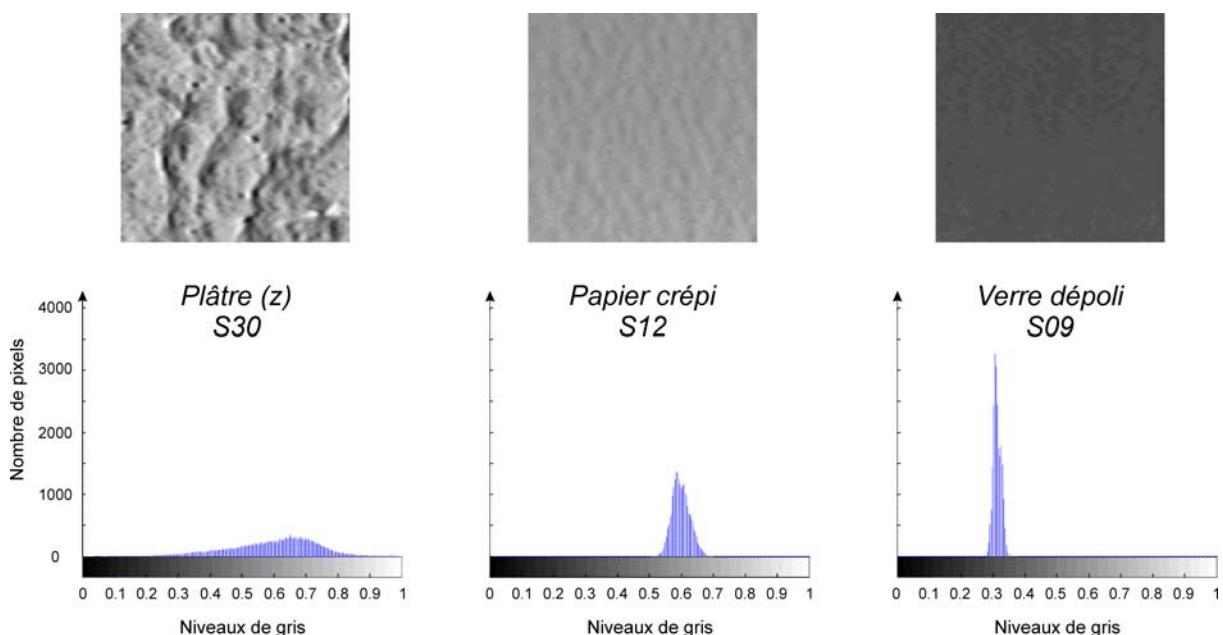


Figure 2-6 : 3 exemples de textures et leur histogramme de distribution des niveaux de gris.

Paramètres 1er ordre				
Paramètres	Formule	Plâtre(z)	Papier crépi	Verre dépoli
Luminance moyenne	$\chi = \frac{1}{N} \sum_{i,j} \alpha(i, j)$	0.60	0.59	0.31
Variance	$\mu = \frac{1}{N} \sum_{i,j} (\alpha(i, j) - \chi)^2$	0.0183	0.0005	0.0001
RMS	$RMS = \frac{\sqrt{\mu}}{\chi}$	0.23	0.040	0.038
Skewness	$Skew = \frac{1}{N} \sum_{i,j} (\alpha(i, j) - \chi)^3$	-0.55	-0.004	0.20
Kurtosis	$Kur = \frac{1}{N} \sum_{i,j} (\alpha(i, j) - \chi)^4$	3.09	3.02	2.68
SNR	$SNR = \frac{\chi}{\sqrt{\mu}}$	4.41	25.39	26.60

Paramètres 2nd ordre				
Energie	$E = \sum_{i,j} (p(i, j))^2$	0.0003	0.0044	0.0278
Entropie	$ENT = - \sum_i \sum_j p(i, j) \log p(i, j)$	8.5	5.7	3.9
Inertie	$INT = \sum_i \sum_j (i-j)^2 p(i, j)$	193.3	10.7	1.1
Homogénéité	$HOM = \sum_{i,j} \frac{1}{1 + (i-j)^2} p(i, j)$	0.11	0.32	0.66

Tableau 2-2 : Paramètres statistiques calculés sur chaque image. Ce tableau regroupe les formules des paramètres statistiques ainsi que des valeurs calculées pour trois exemples d'images de textures naturelles : le plâtre (z), le papier crépi et le verre dépoli. $\alpha(i, j)$ indique le niveau de gris du pixel (i, j) (de 0 à 1), N le nombre total de pixels. Pour les paramètres de 2nd ordre, $p(i, j)$ correspond aux éléments de la matrice de cooccurrence ; c'est à dire à la probabilité de passer d'un pixel de niveau de gris i à un pixel de niveau de gris j.

a) La moyenne χ

Cette valeur donne la valeur moyenne des niveaux de gris (ou intensité moyenne) des pixels constituant l'image. Plus l'image sera claire, plus la moyenne sera élevée. Les deux premières textures citées en exemple dans le Tableau 2-2 présentent une même moyenne de luminance alors que l'image du verre dépoli est moins lumineuse.

b) La variance μ

La variance correspond au moment d'ordre 2. Elle mesure la répartition des niveaux de gris autour de la valeur moyenne de la luminance. La texture *plâtre(z)* présente une variance plus élevée que le *papier crépi* ou le *verre dépoli*. Plus la variance est élevée, plus les variations autour la moyenne sont élevées.

c) Le skewness et le kurtosis

Le skewness (asymétrie) correspond au moment d'ordre 3 centré autour de la moyenne de la luminance. Il mesure la déviation de la distribution des niveaux de gris par rapport à une distribution symétrique. Si la déviation se fait pour des valeurs élevées par rapport à la moyenne alors le skewness sera positif (*verre dépoli*) alors que si la déviation se fait pour des valeurs inférieures, le skewness sera alors négatif (*plâtre (z)* et *papier crépi*).

Le kurtosis (aplatissement) correspond au moment d'ordre 4 centré autour de la moyenne. Il caractérise la forme du sommet de l'histogramme de distribution des niveaux de gris. Plus le kurtosis est important, plus les valeurs sont distribuées autour de la moyenne et plus le sommet de l'histogramme sera large.

d) Le signal sur bruit : SNR

Ce paramètre définit le « signal sur bruit » et équivaut au rapport de la moyenne à l'écart type. Il est particulièrement intéressant car il permet de caractériser l'hétérogénéité d'une texture. Une texture hétérogène possède un SNR faible car sa variance (et donc son écart type) est élevée. Prenons le *plâtre (z)* et le *papier crépi*, deux images avec une même moyenne. La texture *plâtre (z)* ayant une variance élevée aura un SNR plus faible que l'autre texture. Elle paraît plus hétérogène que le *papier crépi* et le *verre dépoli*. Ce paramètre est l'inverse du RMS (root mean square) ou écart quadratique moyen.

Les paramètres décrits ci-dessus sont des paramètres de 1^{er} ordre et correspondent à une description de l'histogramme en deux dimensions des niveaux de gris. Aucune information concernant le voisinage local du pixel n'est précisée dans le calcul de ces paramètres. Les méthodes d'analyse d'ordre supérieur vont permettre de décrire plus précisément des textures car elles prennent en compte les relations entre pixels voisins.

1.2. Les variables de 2nd ordre

L'ordre des méthodes est donné par le nombre de pixels mis en jeu dans le calcul des paramètres. Les méthodes de second ordre regroupent la méthode des matrices des gradients, des caractéristiques statistiques (Wu, 1992), la méthode des différences de niveaux de gris et la matrice de cooccurrence. Pour l'analyse de nos images, nous nous sommes restreints à la méthode de matrice de cooccurrence car elle est la plus communément utilisée.

La méthode de matrice de cooccurrence (ou méthode de dépendance spatiale des niveaux de gris) permet de déterminer la fréquence d'apparition d'un ensemble formé de deux pixels séparé d'une distance d dans une direction donnée. Les directions choisies sont

généralement 0° , 45° , 90° , 135° et 180° ; et 1 pour la distance d. Une fois ces matrices de cooccurrences construites, il est possible d'en extraire des paramètres qui contiennent des informations concernant la finesse et la granularité de la texture. Pour l'analyse de nos stimuli, nous avons utilisé une matrice de cooccurrence pour chacune des directions. Les paramètres extraits de cette matrice sont les suivants :

a) L'énergie ou moment angulaire d'ordre deux ou uniformité

L'énergie d'une image est un paramètre mesurant l'homogénéité de l'image. Si une image a une énergie faible alors elle contient peu de zones homogènes ; le nombre de transitions entre niveaux de gris est important. Des trois exemples, la texture *plâtre* (z) a une énergie minimale, elle est plus hétérogène que les deux autres.

b) Le contraste ou inertie

Cette variable mesure les variations locales de niveaux de gris. Si ces variations sont très fréquentes au sein de l'image, la texture présente peu de régions homogènes, son contraste sera alors assez élevé. La texture *plâtre* (z) présente plus de transitions entre niveaux de gris que la texture *papier crépi* ou *le verre dépoli*, donc moins de régions homogènes. Son contraste sera plus élevé.

c) L'entropie

L'entropie mesure la complexité de l'image. Lorsque les valeurs de la matrice de cooccurrence sont égales, l'entropie de l'image est élevée. Elle permet ainsi de caractériser le degré de granulation de l'image. Plus une image possède une entropie basse plus sa granulation sera fine, comme c'est le cas pour le *verre dépoli*.

d) L'homogénéité ou moment différentiel inverse

Ce paramètre permet de quantifier le nombre de régions homogènes présentes dans la texture. Plus la texture possède de régions homogènes, plus ce paramètre sera élevé. Le *verre dépoli* est la texture la plus homogène des trois.

Dans cette description des paramètres statistiques pour l'analyse d'une texture, les termes « homo-hétérogénéité » reviennent fréquemment car c'est une caractéristique importante de la texture pour la décrire. Plusieurs paramètres permettent de décrire cette propriété d'une texture mais ils diffèrent selon l'ordre. Par exemple, le SNR rend compte de l'homogénéité de 1^{er} ordre d'une texture (sans relation avec les pixels voisins) alors que

l'inertie ou l'homogénéité décrivent une homogénéité de 2nd ordre, elle est fonction de l'emplacement des pixels).

D'autres paramètres de 2nd ordre ont été calculés (« cluster prominence », « cluster shaded ») mais ils ne sont pas pris en compte pour les analyses de nos images. La série de paramètres de 1^{er} et de 2nd ordre décrits ci-dessus ont été calculés pour chacune des images des textures naturelles (originales et contrôles) et des formes tridimensionnelles (originales et contrôles). Les valeurs pour chaque image sont disponibles dans les annexes.

1.3. Calcul des paramètres selon le type de stimuli

a) Textures naturelles

Une analyse de variance sur les paramètres des images indique que le facteur PHASE (original et aléatoire) n'est pas significatif pour la luminance moyenne, le contraste, la variance, le SNR, le kurtosis, l'entropie et l'homogénéité pour les images du set A et celles du set C (ANOVA, $p>0,05$). Ce facteur se révèle significatif pour le skewness ($p=0,001$), l'énergie ($p=0,028$), l'inertie ($p=0,005$). L'inertie augmente pour les stimuli scrambled des textures hétérogènes comme l'aluminium (#15), la feuille de laitue (#23) ou la plante Soleirolia (#53). Pour ces textures hétérogènes, la redistribution aléatoire des phases pour les stimuli scrambled fait apparaître plus de régions homogènes que dans les images originales.

b) Formes tridimensionnelles

Des calculs similaires ont été effectués sur les images utilisées dans notre second protocole afin de comparer les images les unes par rapport aux autres. Dans un premier temps, nous avons comparé tous ces paramètres selon le type d'image (3D, 2D blob et random) à l'aide d'une analyse de variance. Tous les paramètres de 1^{er} ordre sont identiques entre ces trois types de stimuli.

Concernant les paramètres de 2nd ordre, les images 3D et 2D blob diffèrent uniquement au niveau de l'inertie ($p<0,001$). Les images des 2D blob présentent une inertie plus importante que les images originales, donc davantage de régions homogènes. En effet, ces stimuli 2D blob se caractérisent par l'apparition de patterns noirs et blancs contenus dans la forme initiale. Pour les stimuli random, la redistribution aléatoire des pixels entraîne des modifications au niveau des 4 paramètres de second ordre puisqu'elle déstructure complètement l'image.

Quant aux stimuli postérisés, ils ne respectent que la luminance moyenne des stimuli originaux. Ils ne sont réalisés qu'avec 2 niveaux de gris, les autres paramètres de 1^{er} et de 2nd ordre diffèrent par rapport aux images originales.

c) Influence de ces paramètres

Pour tester l'influence des paramètres statistiques de l'analyse des images sur les décharges de nos neurones, nous avons calculé des coefficients de corrélation de Pearson (R) entre ces paramètres et les décharges moyennes des neurones pour chaque stimulus. De plus, ces paramètres nous ont permis de mieux caractériser les dimensions présentes dans nos analyses multidimensionnelles de type MDS.

2. Contenu fréquentiel des images

2.1. Méthode de Van der Schaaf et van Hateren

Pour comparer le contenu en fréquences spatiales de nos images de formes tridimensionnelles, nous avons utilisé une méthode de calcul présenté par Van der Schaaf et Van Hateren en 1996 (van der Schaaf et van Hateren, 1996). Ces calculs ont été effectués uniquement sur les images des formes tridimensionnelles car les images contrôle des textures naturelles respectent les amplitudes des images originales (seules les phases sont randomisées). Cette méthode nous permet de caractériser le contenu fréquentiel des images par le calcul de deux coefficients (1/f)³ un pour l'orientation horizontale (1/f_{horizontal}) et l'autre pour l'orientation verticale (1/f_{verti}).

Le Tableau 2-3 résume les valeurs caractéristiques de ces coefficients calculés selon le type de l'image 3D, 2D blobs, postérisés ou random. Les résultats de ces calculs indiquent que les stimuli 2D blob présentent un spectre de fréquence spatiale similaire à celui des images tridimensionnelles originales. En effet les coefficients 1/f sont statistiquement identiques à la fois pour l'orientation verticale et horizontale (ANOVA, p>0,05). Les images des formes tridimensionnelles et leurs images 2D blob sont donc caractérisées par un même contenu fréquentiel. Par contre, le spectre de fréquences spatiales pour les stimuli random est différent de celui des images originales ; les coefficients 1/f chutent. Ces images random contiennent davantage de hautes fréquences spatiales que les images originales. Quant aux stimuli postérisés, ils sont définis par des pixels avec deux niveaux de gris. Il est logique que

³ f = Fréquence spatiale (cycle/deg)

les spectres de fréquence spatiale de ces stimuli diffèrent légèrement de celui des images originales.

3D	1/f _{horizontal}	1/f _{vertical}	2D blobs	1/f _{horizontal}	1/f _{vertical}
Moyenne	1,68	1,64		1,64	1,67
Écart-type	0,08	0,07		0,11	0,08
E. S. M	0,01	0,01		0,02	0,01
Maximum	1,79	1,77		1,80	1,82
Minimum	1,43	1,53		1,43	1,50
Intervalle	0,37	0,25		0,37	0,32
N	32	32		32	32
Médiane	1,71	1,64		1,64	1,68

Postérés	1/f _{horizontal}	1/f _{vertical}	Random	1/f _{horizontal}	1/f _{vertical}
Moyenne	1,51	1,54		0,45	0,47
Ecart-type	0,11	0,14		0,06	0,05
E. S. M	0,02	0,03		0,01	0,01
Maximum	1,82	1,81		0,61	0,61
Minimum	1,36	1,09		0,36	0,39
Intervalle	0,46	0,71		0,26	0,22
N	32	32		32	32
Médiane	1,50	1,55		0,43	0,47

Tableau 2-3 : Tableau récapitulatif des coefficients (1/f_{horizontal} et 1/f_{vertical}) caractéristiques des spectres de fréquences spatiales présentes dans les images utilisées pour le second protocole.

2.2. *Les filtres linéaires de type GABOR*

Une problématique importante et complexe de la vision par ordinateur consiste à classifier des textures de façon à créer des banques d'images aidant à l'identification d'une nouvelle texture, indépendamment, par exemple, de l'illumination. Un algorithme possible est l'emploi de banques (aussi appelées « pyramides ») de filtres. Une grande majorité des systèmes de reconnaissance à base de caractéristiques sont modélisés à partir de familles de filtres passe-bande orientés (Gabor 2D, dérivées de fonctions gaussiennes,...) selon des stratégies de codage inspirées des travaux de Hubel et Wiesel (sélectivité à l'orientation dans le cortex visuel, (Hubel et Wiesel, 1969)).

Nous avons choisi d'utiliser une pyramide de filtres GABOR 2D de plusieurs tailles (4) et orientations (8) (voir Tableau 2-4). Chaque image de texture naturelle est convoluée avec chaque filtre et on mesure l'énergie de chaque convolution. L'image filtrée est élevée au carré, puis moyennée. L'énergie de convolution est définie comme la racine carrée de cette moyenne.

Nous avons ensuite cherché à déterminer si l'analyse des textures par la population neuronale de V4 est similaire à celle effectuée par un ensemble de filtres. Les énergies ont été

utilisées pour établir des analyses multidimensionnelles par clusters (arbre de Ward) ou euclidiennes (MDS) que nous avons comparées avec les mêmes analyses multidimensionnelles obtenues sur les réponses neuronales. Les comparaisons entre MDS de filtres et des réponses neuronales ont été effectuées à l'aide du test de PROTEST proposé par Jackson en 1995 (Jackson, 1995). Une autre analyse multidimensionnelle (Factorisation Non Négative, NNF) a permis de confirmer les résultats.

Taille (Pixels)	Période (Pixels)	Sdt X (Pixels)	Sdt Y (pixels)	Fréquence (cycles/pixels)	Fréquence (cycles/degrés)
240	60	30	30	$4/240 = 0.0167$	0.835
120	30	15	15	$4/120 = 0.0333$	1.665
60	15	8	8	$4/60 = 0.0667$	3.335
30	7	4	4	$4.29/30 = 0.1430$	7.15
15	4	2	2	$3.75/15 = 0.25$	12.5

Tableau 2-4 : Filtres Gabor 2D utilisés dans la pyramide. Dans nos protocoles, une taille de 150 pixels équivaut à 3°. Chaque filtre est défini par une fréquence spatiale donnée et par une orientation (de 0° à 157,5° par pas de 22,5°).

E. ANALYSE DES DONNÉES

1. Décharges neuronales

1.1. Cellules individuelles

La réponse du neurone est quantifiée en fréquence de décharge soit le nombre de potentiels d'action par seconde (PA/s ou Hz). Pour l'analyse quantitative des modifications d'activité neuronale, nous avons distingué 2 périodes : une période de référence et une période d'activité. La période de référence a une durée de 250 ms et démarre 400 ms avant le début de la présentation des stimuli. La période d'activité dure également 250 ms et débute 50 ms après l'apparition du stimulus visuel. L'animal fixe pendant ces 2 périodes temporelles. Nous comparons la fréquence moyenne de décharge pendant cette période de référence et la fréquence moyenne de décharge pendant la période d'activité à l'aide d'un test paramétrique T de Student. Si une différence significative est détectée dans ce test, nous considérons que le neurone est activé pendant la période de présentation de stimuli (test-T, $p < 0.05$). Pour certaines analyses statistiques, nous avons normalisé les décharges moyennes du neurone en fonction du stimulus préféré. Ainsi le neurone aura pour ce stimulus préféré une décharge moyenne de 1, et les décharges moyennes pour les autres stimuli sont alors calculées en fonction de la réponse maximale du neurone pour le stimulus préféré.

Pour les neurones activés, nous avons ensuite effectué une analyse de variance (ANOVA) avec un nombre de facteurs dépendant du protocole.

1.2. Les analyses de population

Pour certaines analyses, les neurones ne sont plus considérés individuellement mais sont regroupés au sein d'une population. Les principales analyses de population sont le décours temporel des réponses moyennes, les analyses par rang et les analyses multidimensionnelles. Les deux premiers types d'analyse seront décrits dans le chapitre 'Résultats' lorsqu'une figure illustre une de ces analyses. Et la 3^{ème} dans le paragraphe suivant.

Par ailleurs, la détermination des paramètres des textures auxquels les neurones de V4 sont sélectifs est une question difficile car il n'existe pas de définition paramétrique simple d'une texture. Nous avons donc effectué deux types d'analyses multidimensionnelles de données : une analyse hiérarchique par clusters et une analyse multidimensionnelle

euclidienne (MDS). Une telle approche est originale car elle vise à appréhender les mécanismes d'action de toute une population de neurones. L'analyse hiérarchique est basée sur les distances entre les éléments (par exemple, les réponses des cellules ou des filtres) et l'analyse multidimensionnelle se base sur des coefficients de proximité calculés à partir des données brutes. Elle organise l'espace représentationnel en dimensions. Le calcul de la contrainte ou stress, définie comme une mesure de la qualité de l'ajustement de la configuration des données (D raw stress) permet de limiter le nombre de dimensions.

2. Mesures de latences

Outre les fréquences de décharge, les latences de réponse sont des données importantes pour chaque cellule enregistrée. Pour mesurer les latences de réponse, nous avons utilisé trois méthodes distinctes.

2.1. Méthode améliorée de Maunsell et Gibson (1992)

La première méthode est comparable à celle utilisée par Maunsell et Gibson (1992), légèrement modifiée par la suite par Nowak et al (1995) et Hupé et al (2002) (Maunsell et Gibson, 1992; Bullier et Nowak, 1995; Girard *et al.*, 2001). Pour chaque neurone, l'activité spontanée est calculée dans une période de 200 ms débutant 400 ms avant l'apparition des stimuli visuels. Plusieurs largeurs de classes sont utilisées (1, 2, 5, 10 et 20 ms). Les intervalles représentant l'activité spontanée se répartissent alors selon une distribution de Poisson ou une distribution gaussienne. Les tracés des courbes de ces fonctions théoriques permettent d'extraire un paramètre caractéristique des distributions : la moyenne de l'activité spontanée (λ). Si ce paramètre λ est inférieur à 1 alors la distribution théorique choisie sera celle de Poisson ; la distribution est gaussienne si la valeur de λ est supérieure à 1 (Figure 2-7).

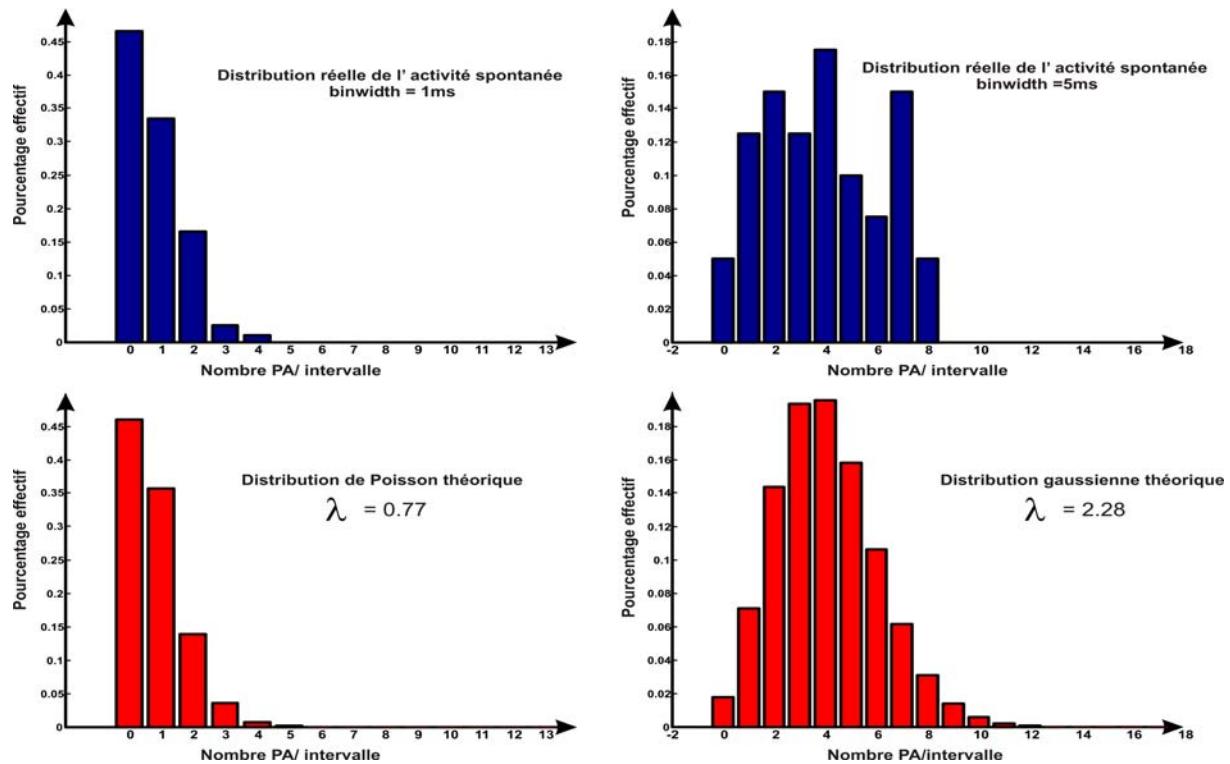


Figure 2-7 : Exemples de distributions de Poisson (à gauche) ou gaussienne (à droite) sur l'activité d'un neurone pendant la période de référence. Sur les graphes de droite, les intervalles ont une taille de 1 ms et pour les graphes de gauche de 5 ms. Les distributions en bleu correspondent aux distributions réelles de l'activité spontanée et celles en rouge aux distributions théoriques.

La Figure 2-8 présente l'exemple d'un neurone. Le calcul du paramètre λ nous permet d'extraire un seuil (p). La latence de réponse du neurone est définie comme le premier bin où $p < 0,01$; le p du bin suivant doit également être inférieur à 0,01 et le troisième bin inférieur à $p < 0,05$. Ce calcul est effectué pour les différentes tailles de bin. Si la taille du bin augmente, la distribution tend à être gaussienne et la moyenne λ augmente. Les plus grandes tailles de bin sont généralement utiles lorsque l'activité spontanée du neurone tend à être nulle.

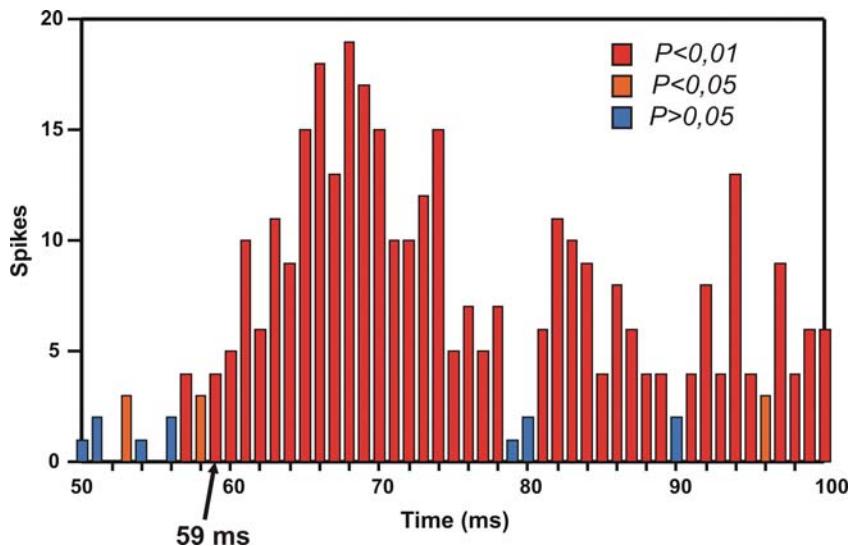


Figure 2-8 : Exemple de calcul de la latence de réponse pour un neurone de V4 (T-sfs1_11). Les intervalles sont de 1 ms. Le code couleur correspond aux valeurs P du test statistique. La période temporelle ici représentée est de 50 ms, débutant 50 ms après l'apparition des stimulations visuelles. Sur cet exemple, le neurone présente une latence de 59 ms.

2.2. Calcul des dérivés seconde

La seconde méthode utilisée consiste à étudier les variations des dérivées seconde (d'') de l'activité neuronale pendant la période d'activité afin de déterminer le moment où l'activité s'accélère. Pour effectuer ces calculs, dans un premier temps, nous calculons l'activité neuronale par intervalles de 5ms (Figure 2-9). Sur cette figure, la courbe en pointillés verts représente les dérivées seconde (d''). Le neurone ici présenté est le même que celui de la Figure 2-8. Les pics de cette courbe indiquent une augmentation soudaine de l'activité neuronale alors que les déflections négatives reflètent une diminution de l'activité. Pour chaque neurone, il suffit de relever les pics supérieurs à un seuil arbitraire (ici, $d'' > 0,1$). Le premier pic indique la latence de réponse du neurone. Toutefois, il est possible que l'activité spontanée du neurone présente d'importantes variations dues à la présence de *burst* de potentiels d'actions. Ces variations entraîneront aussi l'apparition de pics positifs sur la courbe de d'' . Afin d'exclure ces pics, pour chaque neurone, un seuil d'activité est calculé en fonction de son activité spontanée ($\lambda + 2 * \sigma$, où λ est la moyenne de l'activité spontanée en Hz et σ , l'écart-type). Les décharges des 2 intervalles suivant l'intervalle correspondant au pic positif doivent aussi être supérieures à ce seuil. Ainsi, les augmentations soudaines d'activité neuronales sont exclues pour le calcul des latences.

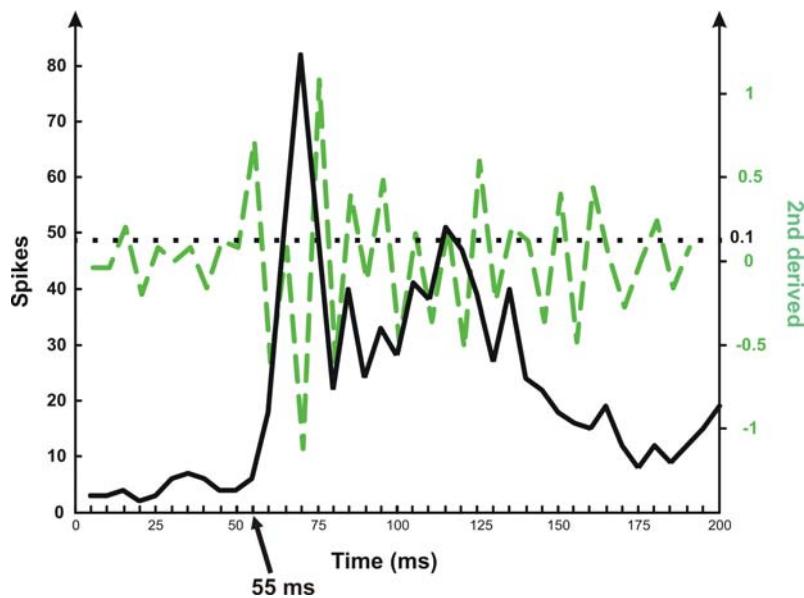


Figure 2-9 : Exemple de calcul de latence grâce à la courbe de variation des dérivés seconde (en vert). L'autre courbe (en noir) représente l'histogramme de réponse (PSTH) du neurone (intervalle de 5ms). Le temps 0 indique l'apparition du stimulus.

2.3. Spike Train Analysis (STA)

La méthode de Spike Train Analysis (STA) est communément utilisée pour calculer les latences des neurones. Elle est décrite dans les publications de Schmolesky et al. (1998), Hanes et al. (1995) et Thompson et al. (1996). (Hanes *et al.*, 1995; Leventhal *et al.*, 1995; Schmolesky *et al.*, 1998). Nous l'avons reproduit pour calculer nos latences. Contrairement aux 2 précédentes méthodes, cette méthode de calcul n'est pas basée sur les PSTHs des cellules mais consiste à déterminer pour chaque essai le début de l'augmentation significative de la décharge. Un calcul mathématique permet d'extraire une valeur de latence pour chaque cellule (voir (Hanes *et al.*, 1995; Leventhal *et al.*, 1995; Schmolesky *et al.*, 1998)).

Les valeurs calculées selon ces trois méthodes sont comparées grâce à un test non paramétrique de Mann Whitney.

Chapitre 3

Bases neuronales du codage des textures naturelles dans l'aire V4 du singe éveillé

Ce chapitre présente les résultats obtenus sur le codage des textures naturelles dans l'aire V4. Son objectif est de montrer qu'il existe une sélectivité aux textures naturelles au sein de l'aire V4. Nous caractériserons alors cette sélectivité par des paramètres statistiques des images comme la luminance moyenne ou le contraste. Nous examinerons ensuite si les réponses de la population sélective sont similaires à celles d'un ensemble de filtres GABOR, outil mathématique communément utilisé dans une tâche de classification des textures. Enfin, un intérêt particulier sera porté sur l'influence de la direction d'illumination sur la sélectivité des neurones. Dans une première partie (A), nous présenterons l'étude de la rétinotopie des aires V4 des animaux ainsi que les populations de neurones prises en compte pour chaque protocole utilisé. En effet, nous avons utilisé 3 sets différents de stimuli (set A, B et C), les résultats pour ces sets sont détaillés séparément dans différentes parties même si des analyses statistiques similaires nous permettent de les regrouper dans la partie discussion (E).

A. INTRODUCTION ET PRESENTATION DE L'ETUDE.

Les objets qui nous entourent ont à leur surface une texture qui leur est propre. Les humains peuvent utiliser cette propriété matérielle pour reconnaître ou classifier les objets entre eux (Price et Humphreys, 1989; Humphrey *et al.*, 1994; Rossion et Pourtois, 2004). La texture d'un objet se caractérise par la répétition d'un pattern de pixels plus ou moins régulier défini par le terme de texton par Bela Julesz (Julesz, 1981). C'est un indice monoculaire statique important qui permet de décrire la composition physique des objets et les propriétés de leur surface ; est-elle lisse ou rugueuse ? Régulière ou irrégulière ? De plus, la texture qui recouvre un objet présente aussi un aspect 3D plus ou moins prononcé. De par ses caractéristiques 3D, une source d'illumination modifie les patterns de micro-ombrage à la surface de la texture. Lorsque l'angle d'incidence de la source d'illumination varie, les micropatterns sont modifiés à la surface de l'objet. Malgré ces différences visuelles, il nous est possible de reconnaître des textures naturelles indépendamment des conditions d'illuminations.

Les résultats de plusieurs études neurophysiologiques sur le système visuel du primate indiquent que la texture permet une dissociation (ou segmentation) des objets du fond visuel (Lamme, 1995) ou de détecter les bordures (von der Heydt *et al.*, 1984). Toutefois, il n'est pas clairement établi comment s'effectue le codage neuronal de la reconnaissance des propriétés de surface indépendamment des conditions d'illuminations. La voie visuelle ventrale du cortex simien est spécifiquement dévolue à la reconnaissance de l'objet. L'aire V4 fait parti de cette voie ventrale et semble être impliquée dans le traitement des textures car plusieurs études ont démontré des déficits dans des processus de segmentation de textures chez des macaques lésés au niveau de l'aire V4 (De Weerd *et al.*, 1996; Merigan, 1996). Les neurones de V4 sont sélectifs à l'orientation, la largeur et la longueur de simples barres (Desimone et Schein, 1987) même si les éléments constitutifs d'une texture (les textons) sont plus complexes que de simples barres (Julesz, 1981). De plus, les neurones de V4 sont sensibles à des formes complexes (Kobatake et Tanaka, 1994), à des réseaux non-cartésiens (Gallant *et al.*, 1996) et à des contours spécifiques de la forme externe du stimulus (Pasupathy et Connor, 1999; Pasupathy et Connor, 2001). Enfin, V4 projette massivement vers le cortex inférot temporal qui contient des neurones sélectifs à des gradients de textures (Liu *et al.*, 2004). Une interrogation subsiste sur comment des éléments complexes comme des éléments constitutifs d'une texture peuvent être interprétés par le système visuel comme une propriété

matérielle de l'objet. Une récente étude a démontré que les neurones de V4 sont sélectifs au gradient d'illumination et à la densité des éléments d'une texture (Hanazawa et Komatsu, 2001).

Notre travail propose d'étudier comment l'aire V4 d'un singe macaque vigile code des textures naturelles aussi diverses que le béton ou le lin. Pour étudier cela, nous avons utilisé plusieurs sets d'images 2D représentant des textures naturelles illuminées selon différents angles d'incidence. Le but de cette étude est double : premièrement, il s'agit de déterminer si une population de neurones au sein de l'aire V4 code pour une reconnaissance ou une classification de textures ; deuxièmement, d'établir si cette classification est invariante aux conditions d'illumination. Une invariance aux conditions d'illumination corroborait les résultats de plusieurs études computationnelles qui ont démontré une classification de textures invariante à la direction d'illumination en utilisant comme outil mathématique, une banque de filtres linéaires (Leung et Malik, 2001; Varma et Zisserman, 2002; Hayman *et al.*, 2004). Grâce à cette étude, nous avons démontré qu'une population de neurones de V4 agit comme un ensemble de filtres linéaires afin de classifier des textures indépendamment des conditions d'illumination.

B. RETINOTOPIE DES AIRES V4 ET POPULATIONS NEURONALES

1. Trajectoires des électrodes

Nous avons utilisé respectivement 60 et 33 points de pénétration dans les 2 hémisphères cérébraux droits de deux animaux, Tima et Zrem. Les cartes de ces pénétrations sont représentées dans la Figure 3-1. L'emplacement des sillons chez les 2 singes est supposé à partir des empreintes présentes sur la partie d'os enlevé lors de l'opération et des observations effectuées pendant celle-ci. Des photographies numériques prises pendant l'opération du second animal ont permis de mieux positionner les sillons. Une analyse histologique post-mortem sur le second animal (Zrem) et des photographies d'un scanner à rayon X (pour Tima) confirment l'emplacement des chambres d'enregistrement et des sillons. Les résultats de ces analyses sont présentés dans le chapitre 2.

Un total de 409 neurones a été enregistré chez Tima ; 265 pour le projet texture, (WP3) et 144 pour le projet ombrage (WP2). Chez le second animal Zrem, nous avons enregistré l'activité unitaire de 104 cellules (69 pour WP3, et 35 pour WP2). Les données de WP2 et WP3 sont regroupées ici pour la cartographie.

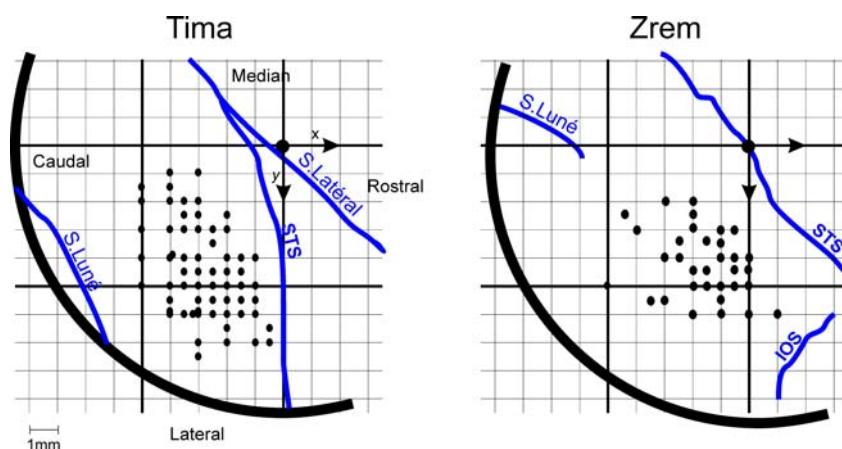


Figure 3-1 . Carte de pénétrations dans l'hémisphère droit de Tima et de Zrem (n= 60 trajectoires pour Tima et n= 33 pour Zrem). Chaque point correspond à un point d'entrée des électrodes défini par ses coordonnées rostro-caudale (x) et médiolatérale (y) lues sur le vernier du microdescendeur. La position des sillons est déduite des observations pendant la chirurgie, de photographies scanner à rayons X (Tima) et d'une analyse histologique post-mortem (Zrem). (STS : Sillon temporal supérieur, Sillon latéral, Sillon luné et IOS : sillon inféro-occipital). L'arc de cercle noir schématise la position de la chambre d'enregistrement. Les cartes montrent que nous sommes dans V4 entre les sillons STS et luné pour Tima et au carrefour des sillons STS, luné et IOS pour Zrem. La partie de V4 enregistrée sur Tima paraît plus dorsale que celle enregistrée chez Zrem.

2. Champs récepteurs et rétinotopie

La Figure 3-2 présente deux exemples de séquences enregistrées sur deux neurones individuellement afin de positionner leur champ récepteur (CR) (cf. paragraphe Méthode). Le neurone de la Figure 3-2A répond préférentiellement aux stimuli blancs ; son activité augmente lorsque les stimuli sont présentés dans une zone circulaire centrée en (-1 ; +4). Cette zone a un diamètre de 2°. Le neurone de la Figure 3-2B est davantage sélectif aux stimuli noirs et la zone du champ visuel qui s'active est plus étendue que celle du neurone A (environ 4° de diamètre). Après une telle cartographie, nous positionnons typiquement le set de stimuli (textures ou formes) au centre du CR en (-2 ; +3).

En étudiant la rétinotopie de l'aire V4 chez le macaque, Gattass et al (1998) ont démontré une relation linéaire entre l'excentricité des CR et leur taille. Nous nous sommes basés sur cette relation pour adapter la taille de nos stimuli afin que nos stimuli recouvrent une surface maximale du CR. Par exemple, pour les neurones de la Figure 3-2, nous avons utilisé des stimuli de 2° pour le neurone A et de 3° pour le neurone B (soit 100 pixels et 150 pixels respectivement).

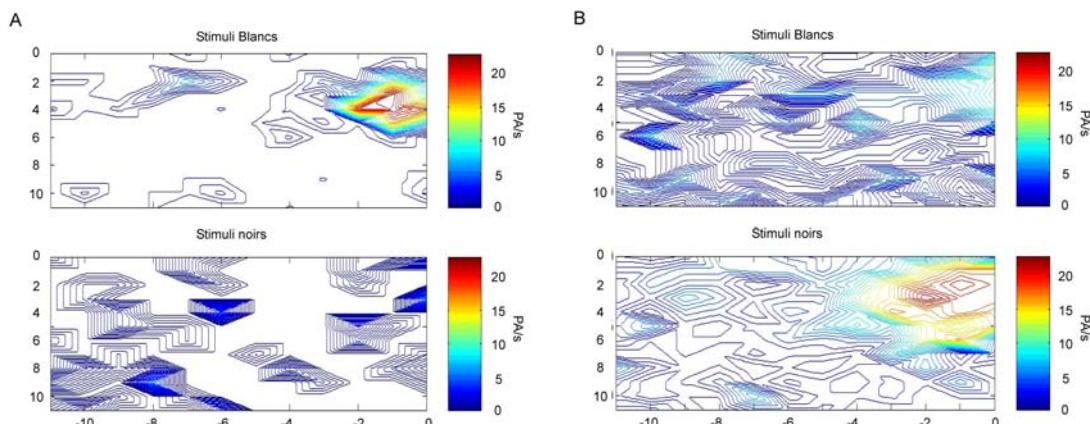


Figure 3-2: Deux exemples de séquences enregistrées sur deux neurones distincts (A : T-SFS85M.1 et B : T-SFS61M.1). Chaque quadrant de la figure représente le champ visuel inférieur gauche, la fovea étant en (0;0). Les axes sont gradués en degré d'angle visuel (DAV). Le code couleur indique le niveau de réponse du neurone (en PA/s). Les figures du haut représentent les réponses des neurones aux stimuli blancs et les celles du bas les réponses aux stimuli noirs. Le neurone de gauche a un CR de 2° de diamètre environ centré sur (-1 ; +4). Celui de droite un diamètre de 4° centré sur (-2 ; +3).

Grâce à cette méthode rapide, nous avons pu disposer nos stimuli visuels au centre des CR pour chacune des cellules. De plus, les coordonnées des trajectoires dans la chambre d'enregistrement sont connues grâce aux verniers du microdescendeur. Nous avons ainsi pu établir une carte rétinotopique de la région corticale enregistrée pour chaque animal. Pour cela, un programme développé dans Matlab associe, grâce à une couleur, la position du CR du

neurone aux coordonnées stéréotaxiques de la cellule. La Figure 3-3 présente les cartes rétinotopiques de nos deux animaux.

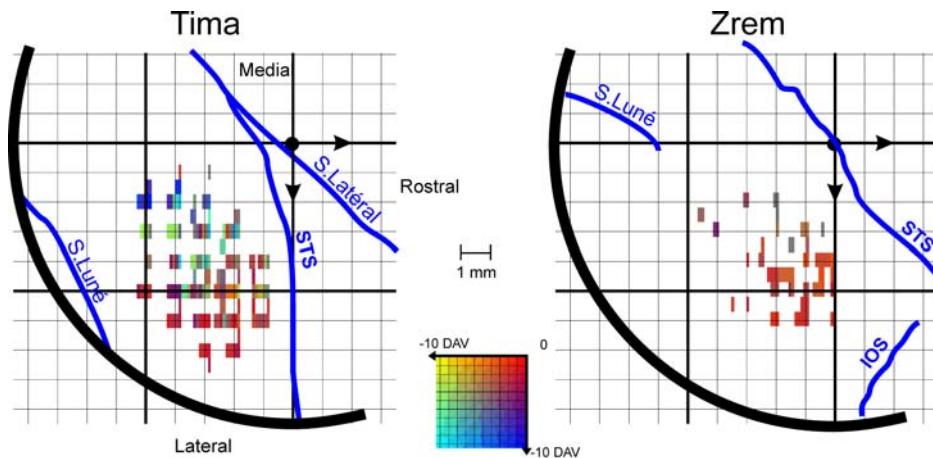


Figure 3-3: Cartes rétinotopiques de Tima et de Zrem. Chaque barre verticale correspond à une cellule. Le code couleur de la barre est fonction de l'excentricité du neurone enregistré (Tima, $n= 213$ cellules et Zrem, $n= 76$). La couleur rouge indique l'emplacement de la fovéa, les carrés bleutés, le méridien vertical et les carrés jaune orangés, le méridien horizontal.

Dans cette cartographie, j'ai pris en compte uniquement les neurones présentant une activité moyenne pendant la période d'activité différente de celle de la période de référence (test-T de Student, $p<0,05$) soit 76 neurones de Zrem et 213 cellules de Tima (pool de neurones enregistrés avec une présentation parafoviale des stimuli).

Pour Tima, les cellules enregistrées dans la partie latérale de l'aire V4 présentent les CR les plus proches de la fovéa tandis que les cellules enregistrées en position plus médiane présentent des CR plus périphériques. Toute la surface du champ visuel inférieur gauche n'a pas été explorée mais il semble que le méridien vertical soit représenté en position médiane (carrés bleus sur la Figure 3-3) et le méridien horizontal plus proche du STS (carrés jaune orangés). De plus, au cours d'une séance d'enregistrement aux coordonnées (-1 ;+7) (non indiqué sur la figure), nous avons remarqué des bruits proches d'une rythmicité cardiaque laissant supposer la proximité d'un gros vaisseau. La position du STS proche de ces coordonnées semble être vérifiée par ces observations.

Pour Zrem, les cellules enregistrées en position latérale présentent une excentricité moindre, se rapprochant de la fovéa. La chambre d'enregistrement de Zrem semble mieux positionnée sur la partie fovéale de V4 que celle de Tima. En effet, l'excentricité moyenne ($4.06^\circ \pm 1.02$) des neurones de Zrem est significativement inférieure à celle de Tima ($6.2^\circ \pm 2.8$) (T-test, $p<0.001$) (Figure 3-4A). Les excentricités rencontrées chez Tima définissent une

gamme plus étendue que celles enregistrées sur Zrem puisque aucune cellule de Zrem ne présente une excentricité supérieure à 8° (Figure 3-4B).

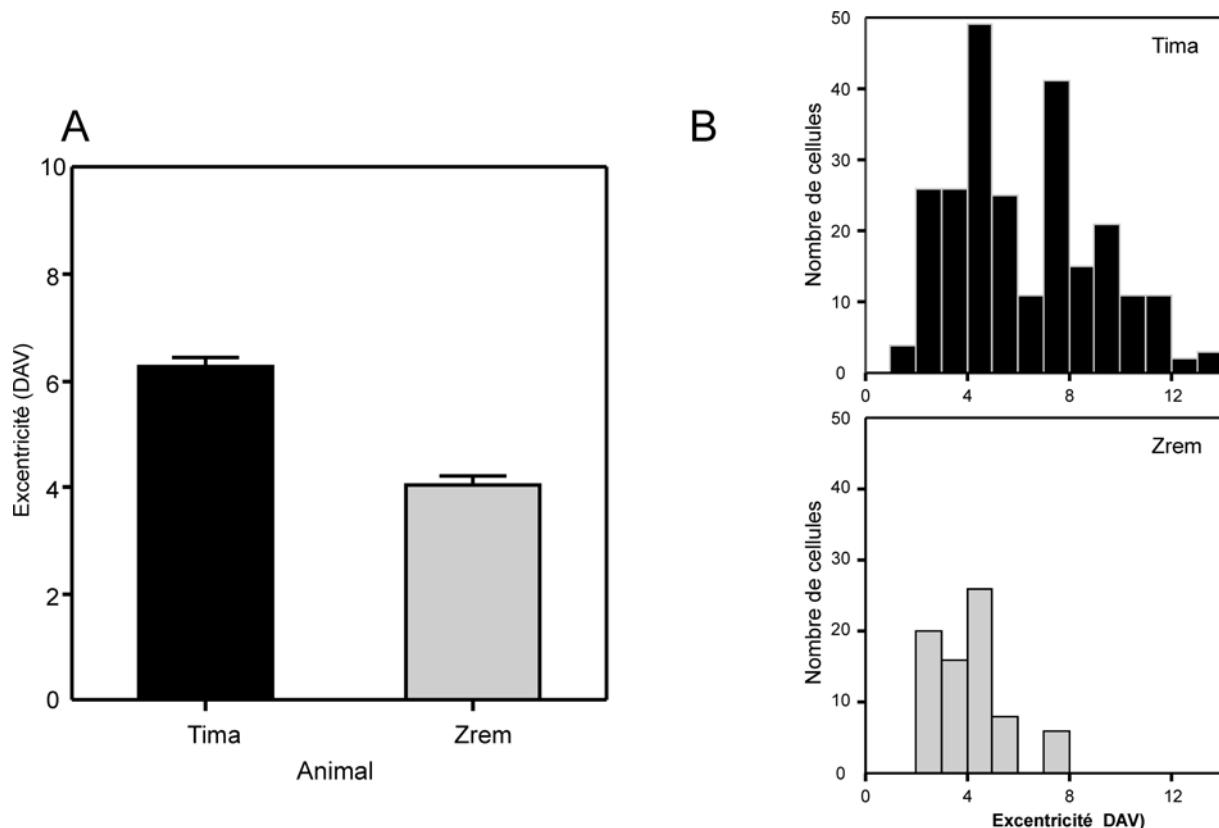


Figure 3-4 : A. Excentricité moyenne des neurones enregistrés sur les deux animaux (la barre d'erreur indique l'erreur standard de la moyenne, esm). B. Répartition des neurones suivant leur excentricité pour les 2 animaux (Timma, n= 213 et Zrem, n= 76). La largeur du pas est de 1°.

3. Activité moyenne

Pour les différentes analyses statistiques effectuées par la suite, nous avons considéré un pool commun de cellules en réunissant les cellules enregistrées sur les 2 animaux. Avant cela, nous avons vérifié qu'une trop grande différence interindividuelle n'était pas observée entre les réponses neuronales. La Figure 3-5 présente les activités moyennes des deux populations de neurones calculées pendant les deux périodes d'activité : la période de référence (-400 à -150 ms par rapport à l'apparition des stimuli) et la période de réponse (50 à 300 ms après l'apparition des stimuli).

L'activité moyenne de la population de neurones est similaire chez les deux singes au cours des deux périodes d'analyses (Anova 2 facteurs [singes, périodes], p=0.643). Ce résultat montre que nous pouvons utiliser un pool commun de neurones pour la suite des analyses si nécessaire.

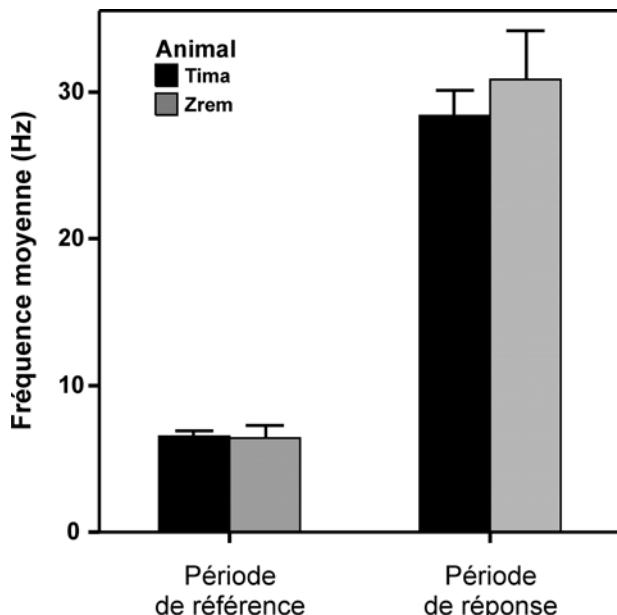


Figure 3-5: Fréquence moyenne de décharge de la population de neurones pour les deux animaux pendant deux périodes temporelles; la période de référence (-400ms à -150ms) et la période de réponse suivant l'apparition des stimuli (50ms à 300ms). Les barres d'erreur représentent l'esp. (Tim, n= 213 et Zrem, n= 76).

4. Populations d'étude

Pour ce paradigme de textures, plusieurs protocoles ont été utilisés avec un nombre variable de textures, d'angles d'incidence de la source d'illumination et de stimuli contrôles. Les protocoles sont définis dans le chapitre 2. Le Tableau 3-1 reprend le nombre de neurones enregistrés sur chaque animal par protocole.

Au total, nous avons enregistré 329 cellules pour ce paradigme « textures naturelles ». Pour 230 d'entre elles, nous avons effectué au minimum 5 présentations d'un même stimulus. Parmi ces cellules, une grande majorité des neurones (n= 210, 91%) présente une réponse différente de l'activité de référence : la plupart des cellules présente une augmentation d'activité lors de leur réponse (n= 204) et 6 cellules présentent une diminution significative d'activité. Ce pool de 210 cellules constitue notre population d'étude pour le paradigme « textures naturelles ».

Tima			
	N neurones enregistrés	N neurones (5 présentations)	N neurones (p<0.05)
Set A fovea	107	51	39
Set A parafovéal	115	96	94
Set B	31	31	26
Set C	7	4	4
Total Tima	260	182	163
Zrem			
Set C	69	48	47
Total Tima & Zrem	329	230	210

Tableau 3-1 : Nombre de neurones enregistrés par singe et par protocoles. La première colonne indique le nombre de neurones enregistrés lors des descentes d'électrodes, la deuxième colonne (bleu) le nombre de neurones enregistrés avec au moins 5 présentations pour chaque stimulus et enfin la troisième colonne (orange) signale le nombre de neurones dont le niveau de réponse pendant la période d'activité est significativement différent de celle enregistrée pendant la période de référence (T-test, $p < 0,05$). Seul le set C de stimuli a été enregistré sur les deux singes, cependant plusieurs stimuli sont communs entre les sets A, B et C.

C. SELECTIVITE AUX TEXTURES NATURELLES - SET A : 12

TEXTURES, 2 PHASES ET 3 ANGLES D'INCIDENCE

1. Base de données

Nous avons enregistré 51 cellules dans l'aire V4 de l'hémisphère droit de Timo avec le set A en position fovéale et 96 neurones en position périphérique. D'après le Tableau 3-1, une grande majorité de neurones (90%, n= 133 cellules) ont une réponse significative pendant la présentation des stimuli (test T de Student, $p<0,05$). Cet ensemble de 133 cellules constitue le pool de neurones sur lequel vont être effectuées les analyses statistiques suivantes. Les neurones enregistrés en position fovéale et périphérique ont été regroupés car les proportions de neurones sélectifs aux facteurs Texture, Type ou Angle d'incidence (d'après un test ANOVA 3 facteurs) sont égales quel que soit le positionnement des stimuli (Test CHI², $p=0,896$). Toutefois, le groupe de neurones enregistrés avec des stimuli en périphérie diffère du groupe A fovéa par une augmentation significative de leur activité moyenne (test-T, $p<0,05$):

Groupe A fovea : n= 39 cellules, Activité moyenne + esm = 12 + 1,4 Hz

Groupe A périph : n= 93 cellules, Activité moyenne + esm = 27 + 3,2 Hz.

2. Mise en évidence d'une sélectivité aux textures naturelles

La première analyse statistique sur les réponses des cellules est une analyse de type ANOVA effectuée pour chaque neurone. Les résultats nous permettent de déterminer si la texture naturelle, l'angle d'illumination ou la phase du stimulus (phase originale ou aléatoire) influent sur la réponse des neurones. Parmi ces 133 cellules, 105 cellules ont un taux de décharge significativement différent suivant la texture présentée soit 79% de la population totale (ANOVA, effet principal, $p<0,05$).

2.1.Exemples de cellules et critère de sélectivité

a) Les textures naturelles sont des stimuli efficaces pour les neurones de V4

Pour le reste des analyses, nous nous référons uniquement aux 105 cellules sélectives à la texture aussi appelées « **cellules TS** ». Pour cette première analyse, nous considérons les réponses des neurones à une texture où toutes les illuminations sont cumulées. La cellule de la Figure 3-6 est un exemple de cellule TS dans lequel seules les réponses aux stimuli originaux sont représentées. Ce neurone présente une forte sélectivité aux textures. Sa décharge augmente pour des textures aussi diverses que le *papier de verre*, le *pain* ou encore le *lin*. Au

contraire, il décharge très peu pour des textures comme le *velours* ou le *béton* même si les réponses moyennes à ces textures sont significatives par rapport à l'activité de référence (test T, $p<0,05$).

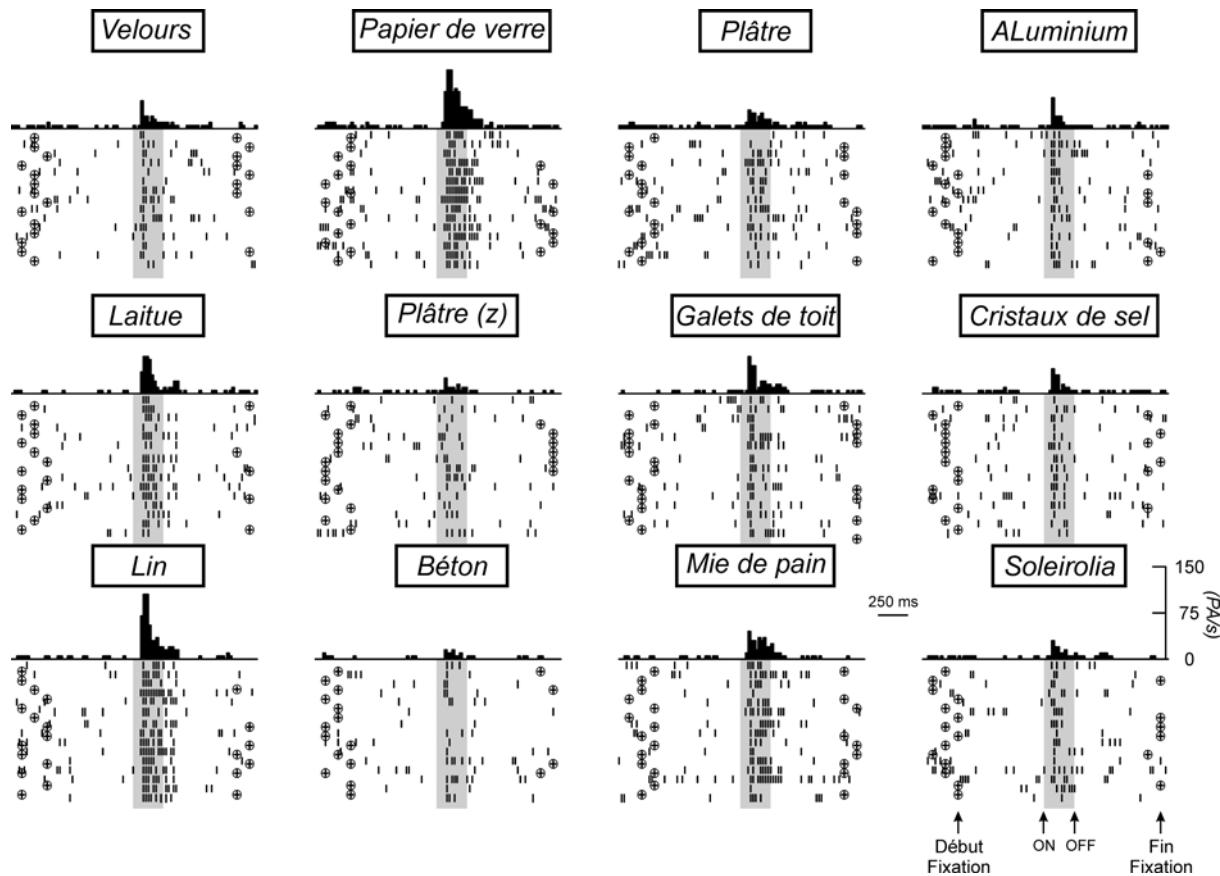


Figure 3-6 : Exemple de neurone de V4 enregistré avec le set A (T-MP2.11). Le pattern de réponse de ce neurone est organisé en fonction de la texture présentée. L'activité neuronale est représentée sous forme de diagrammes de points ou « rasters displays ». Chaque point (ou barre verticale) correspond à un potentiel d'action et chaque ligne de points au déroulement d'un essai. Les essais sont alignés, du point de vue temporel, sur le moment de présentation des stimuli visuels. Toutes les différentes illuminations sont cumulées. Au-dessus des rasters sont représentés les histogrammes de réponses ou PSTHs (peri-stimulus time histogram) avec des intervalles de 25 ms. Le rectangle gris schématise la période de présentation des stimuli. Les petites croix entourées d'un cercle indiquent le début et la fin de la fixation.

Une large majorité de cellules TS répondent significativement à au moins 6 textures par rapport à la période de référence (87%, $n=90/105$). Parmi ces cellules, 60 cellules répondent aux 12 textures, comme le neurone de la Figure 3-6 ; les 30 cellules restantes répondent à un nombre de textures entre 5 et 11 ($n=13$ cellules pour 11 textures naturelles). Seules 3 cellules répondent à 2 textures ; les 11 autres cellules répondent entre 3 et 5 textures. D'après cette analyse, les textures naturelles sont des stimuli efficaces pour activer les cellules de V4 puisque la majorité des cellules voient leur taux de décharge augmenter par rapport à une période de référence.

b) Calcul d'un index de sélectivité I_{max}

Nous avons vu dans le paragraphe précédent que la population de cellules TS répond sélectivement à un ensemble de textures par rapport à la période de référence. Par exemple, la cellule présentée dans la Figure 3-6 voit son activité neuronale augmenter pendant la présentation des stimuli visuels pour les 12 textures. Toutefois, au niveau de sa décharge neuronale, on peut noter une certaine variation de son activité selon la texture naturelle présentée. Sa décharge est maximale pour *le papier de verre* (46,4 + 2.5 Hz) et décroît fortement pour le *béton* (8.14 PA/s +/- 1.2). Afin de mieux caractériser cette variabilité de réponse, nous avons calculé un index de sélectivité I_{max} . Il s'agit, dans un premier temps, de déterminer pour chaque cellule, la texture naturelle parmi les 12 textures testées qui engendre une réponse moyenne maximale (R_{max}). Cette texture est appelée texture maximale. L'histogramme de la Figure 3-7 présente les textures maximales de la population TS. Ainsi, 3 textures se retrouvent en majorité codées par les cellules TS ; la *feuille de laitue* ($n=20$ cellules), le *béton* ($n=18$) et la *feuille d'aluminium* ($n=17$). Les autres textures naturelles sont codées par un plus petit nombre de cellules TS. Ce constat nous permet de calculer l'index I_{max} tel que :

$$I_{max} = (0.75 * R_{max}) + \sigma_{max}$$

R_{max} = Taux de décharge moyen du neurone à la texture maximale (PA/s).

σ_{max} = erreur standard de la moyenne de R_{max} .

La réponse du neurone pour une texture préférée est supérieure à I_{max} .

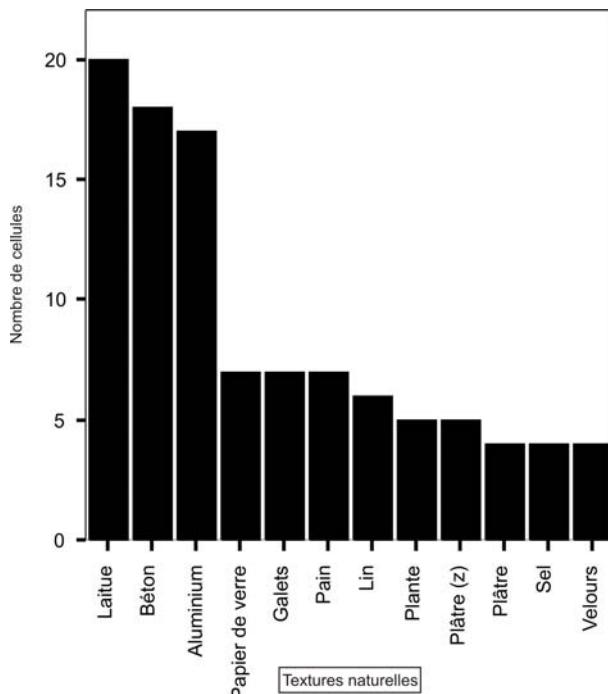


Figure 3-7: Textures naturelles maximales. Le nombre de cellules ($n=105$ cellules TS) est représenté en fonction des textures naturelles que la cellule code de façon maximale. 20 cellules TS préfèrent ainsi la *laitue*.

La Figure 3-8 présente les décharges moyennes de 3 cellules en fonction des textures naturelles. L'histogramme en A montre que cette cellule répond préférentiellement au *lin* et au *papier de verre*. Le neurone en B ne répond qu'à une seule texture, la *feuille de laitue* et le neurone en C répond à un ensemble de quatre textures : l'*aluminium*, la *feuille de laitue*, le *pain* et la *plante*.

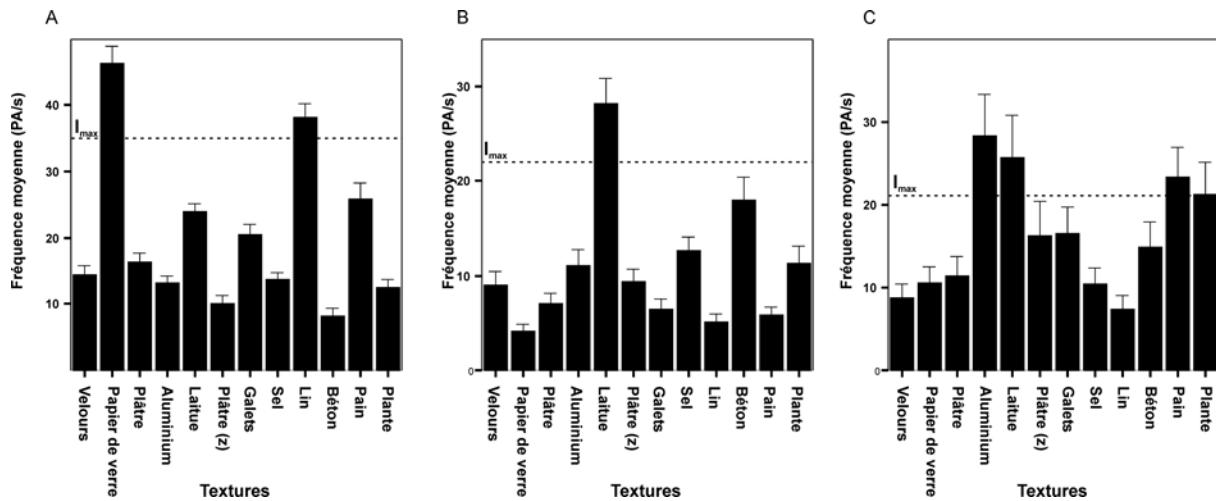


Figure 3-8 : 3 exemples de cellules de V4 (A. TMP2.11. B.TMP46.11. C.TMP112.11). La décharge moyenne pour chaque texture est représentée (+ esm). La ligne en pointillée indique le seuil I_{max} pour chaque cellule. A. Cet exemple est celui représenté dans la Figure 3-6. Ce neurone code pour 2 textures préférées (le papier de verre et le lin). B. Ce neurone ne présente qu'une seule texture préférée (la laitue). C. Ce neurone répond à 4 textures ; l'aluminium, la laitue, le pain et la plante.

En se basant sur ce critère (I_{max}), on observe que les cellules TS répondent en moyenne à un ensemble de 5 textures mais présentent des patterns de sélectivité différents. La majorité des cellules TS répondent au moins à cinq textures ($n = 64, 61\%$). Au contraire, 17 cellules TS (16%) présentent un pattern de sélectivité plus fin en ne répondant qu'à une seule texture (comme l'exemple B). Dix autres cellules répondent significativement à seulement 2 textures naturelles.

Grâce au calcul de l'index I_{max} , nous pouvons aussi savoir à quelle fréquence les 12 textures sont codées au sein de notre population. Nous relevons ainsi pour chaque cellule TS, les textures pour lesquelles la réponse moyenne est supérieure à I_{max} et nous faisons le total sur la population TS. Les résultats sont représentés dans la Figure 3-9.

D'après la Figure 3-9, nous pouvons dégager arbitrairement trois groupes de textures en fonction du nombre de cellules TS qui y répondent significativement. Le premier groupe (1) est constitué par les textures auxquelles les neurones répondent le plus fréquemment : il s'agit de la *feuille de laitue*, l'*aluminium*, le *béton* et la *plante Soleirolia* avec respectivement, 63, 63, 59 et 55 % de cellules TS qui y répondent. Nous retrouvons les textures préférées des

cellules de la Figure 3-7. Ces textures naturelles présentent un aspect visuel particulier ; ce sont les textures les plus hétérogènes parmi les 12 textures testées. Le second groupe (2) comprend le *plâtre* (*z*), les *cristaux de sel*, le *pain* et les *galets* (avec un nombre de cellules TS sélectives entre 47% et 40%) et enfin le troisième groupe (3) est constitué du *lin*, du *velours*, du *plâtre* et du *papier de verre* auxquelles moins de cellules TS répondent significativement (28 à 34%). Ces trois groupes de textures (1 - *Laitue, aluminium, béton et plante* ; 2 - *Pain, plâtre(z), galets et sel* ; et 3 - *Lin, plâtre, velours et papier de verre*) sont représentés dans la Figure 3-9.

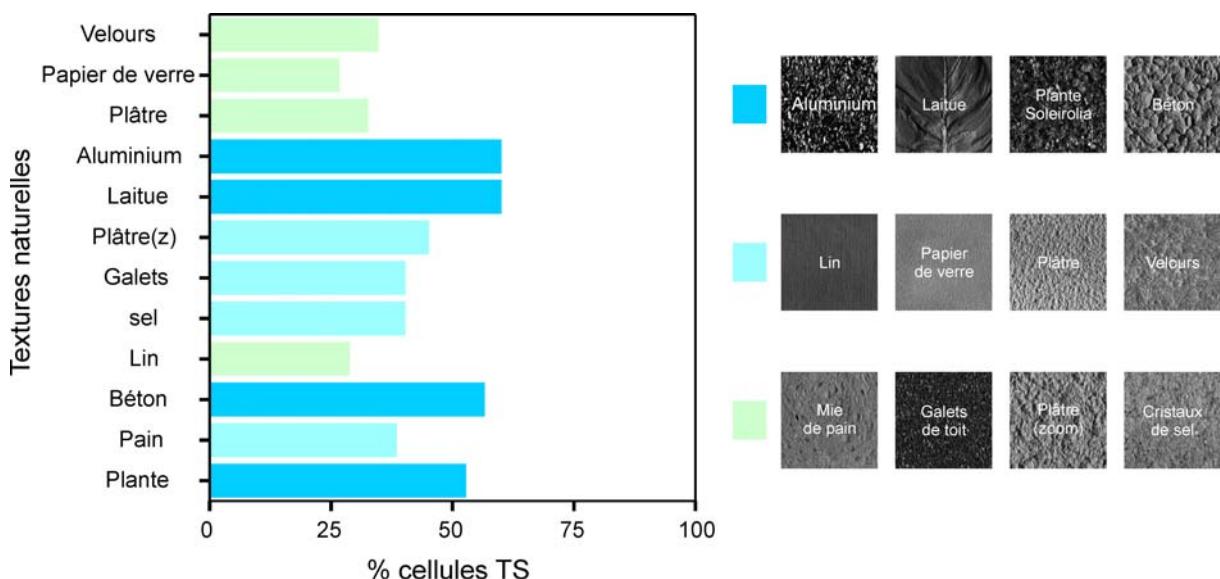


Figure 3-9: Pourcentage de cellules TS codant les textures naturelles d'après le critère I_{max} . Les différentes couleurs correspondent à trois groupes de textures naturelles. Les groupes de textures ont été fait arbitrairement en fonction de la fréquence de codage de la population TS. Les photographies à droite de graphique illustrent les 3 groupes de textures établis.

2.2. La sélectivité est préservée si on déplace le stimulus dans le CR du neurone

Pour un nombre restreint de cellules ($n=10$), nous avons effectué des enregistrements en déplaçant de 1° le stimulus dans le CR du neurone par rapport à sa position initiale (au centre du CR). Ce test a permis de vérifier si la sélectivité aux textures naturelles n'est pas due à des indices locaux du stimulus dans le CR, comme par exemple les bordures des stimuli ou la nervure verticale de la *feuille de laitue*. La Figure 3-10 présente un exemple de neurone enregistré avec ce type de contrôle. Ce neurone répond préférentiellement au *béton* (#45) lorsque les images sont présentées au centre de son CR et décharge peu pour la *plante Soleirolia* (#53). Cette sélectivité est préservée lorsque les stimuli sont déplacés. Pour 8 cellules, la texture préférée en position reste celle qui engendre un taux de décharge maximal

lorsque les stimuli sont déplacés. Pour les 2 autres cellules, le taux de décharge pour la texture préférée reste important.

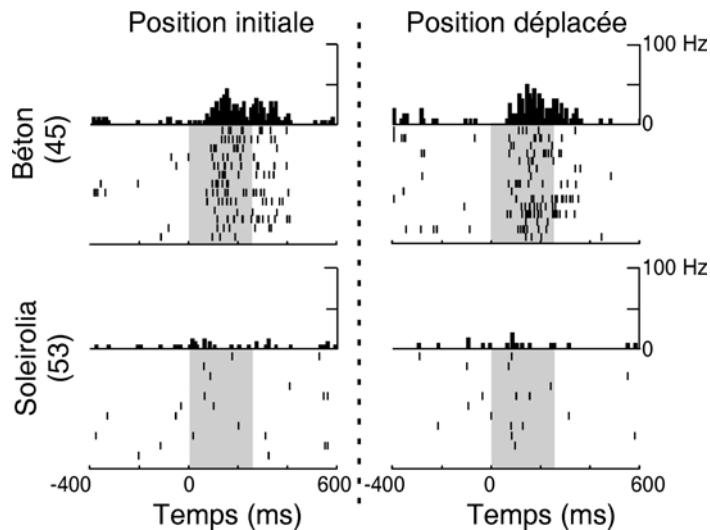


Figure 3-10 : Exemple de cellule testée avec 2 positions des stimuli dans son CR, la position initiale (au centre du CR) et une position dite *déplacée* de (+1 ; +1)° dans le CR (T-MP53.12). La sélectivité du neurone pour le béton (#45) persiste malgré la position déplacée. Le neurone ne décharge pas pour la plante Soleirolia (#53).

2.3. Synthèse sur la sélectivité des neurones TS aux textures naturelles

Dans cette première partie, nous nous sommes intéressés à la sélectivité des neurones de V4 pour les textures naturelles. Les principaux résultats à retenir sont :

- Les textures naturelles sont des stimuli efficaces pour les neurones de V4.
- La majorité des cellules de V4 présentent une modulation de leur réponse en fonction de la texture naturelle testée (ANOVA, $p < 0,05$). Nous avons appelé ces cellules ‘TS’.
- Les cellules TS ne répondent pas qu’à une seule texture mais plutôt à un ensemble de textures, en moyenne 5 d’après I_{max} .
- Les textures naturelles préférées sont la *laitue*, le *béton* et l’*aluminium* ; soient 3 textures particulièrement hétérogènes.
- Une cellule TS peut coder à la fois une texture homogène et une texture hétérogène.
- La sélectivité des neurones TS est conservée lorsque les stimuli sont déplacés à l’intérieur du champ récepteur du neurone.

Ces résultats mettent en évidence la complexité des réponses de la population TS à l’ensemble des textures naturelles. Les analyses suivantes ont pour objectif de déterminer quels paramètres présents dans les images des textures naturelles permettent d’expliquer une réponse identique d’un neurone pour des textures aussi différentes que la *laitue* ou le *béton*.

3. Influence des paramètres statistiques des images

3.1. Paramètres de 1^{er} ordre : moyenne, variance, skewness et kurtosis de la luminance, RMS, et SNR

Chaque image de texture naturelle est caractérisée par un ensemble de paramètres statistiques calculés à partir des histogrammes de distribution des niveaux de gris (voir méthodes). Une image présente donc des paramètres de 1^{er} ordre qui lui sont propres. Afin de caractériser l'influence que pourraient avoir ces paramètres sur la décharge des neurones, nous avons calculé des coefficients de corrélation de Pearson entre les réponses moyennes des cellules et les différents paramètres de 1^{er} ordre ; moyenne, variance, skewness et kurtosis de la luminance, RMS, et SNR (cf. Méthodes pour définitions).

Dans un premier temps, nous nous sommes intéressés à la luminance moyenne. La Figure 3-11 présente deux exemples de cellules TS corrélées à la luminance moyenne des images. La cellule en A voit son activité neuronale augmenter pour des images à forte luminance (i.e. des images claires) et diminuer pour des images à basse luminance (i.e. des images foncées). La décharge de cette cellule est donc corrélée positivement à la luminance moyenne des images ($R_{Pearson} = 0,734, p < 0,001$). De manière inverse, la cellule B montre une corrélation négative ($R_{Pearson} = -0,834, p < 0,001$) à la luminance moyenne.

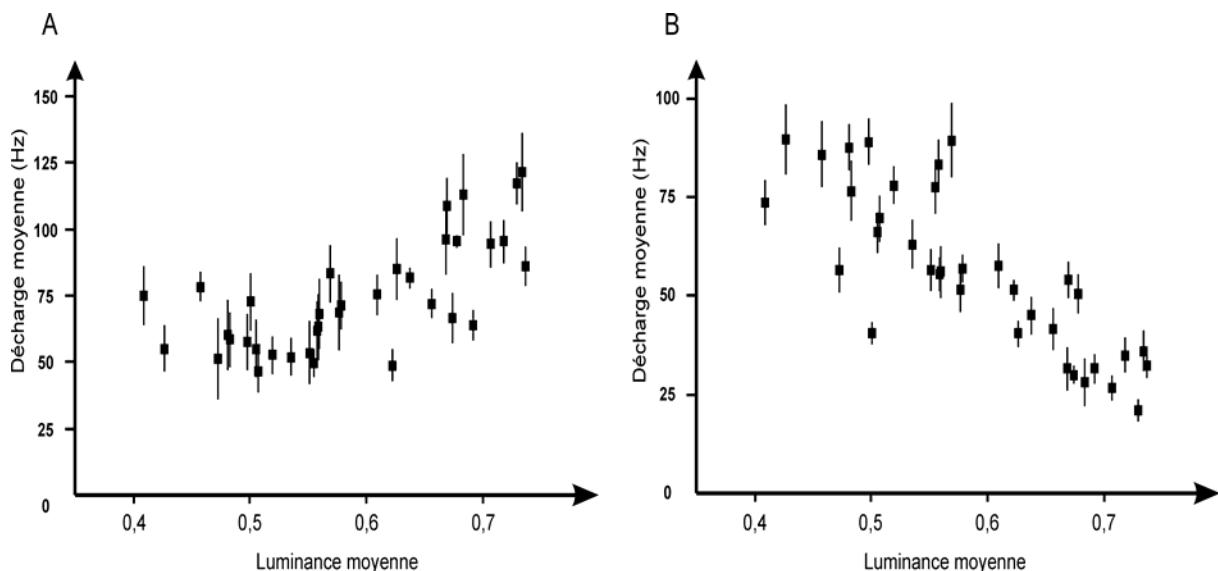


Figure 3-11 : Exemples de cellules TS corrélées à la luminance moyenne des images (A : TMP93.21 et B : TMP80.31). Chaque carré représente la décharge moyenne pour une image donnée (+/- esm). La cellule présentée en A montre une corrélation positive par rapport à la luminance moyenne alors que la cellule en B montre une corrélation négative.

Nous nous sommes aussi intéressés à l'influence des 5 autres paramètres statistiques de 1^{er} ordre. Pour chaque neurone, nous avons calculé un coefficient de corrélation (R

Pearson) pour chacun des paramètres. Le calcul de ces coefficients nous a permis d'établir des histogrammes de distribution pour la population TS présentés dans la Figure 3-12.

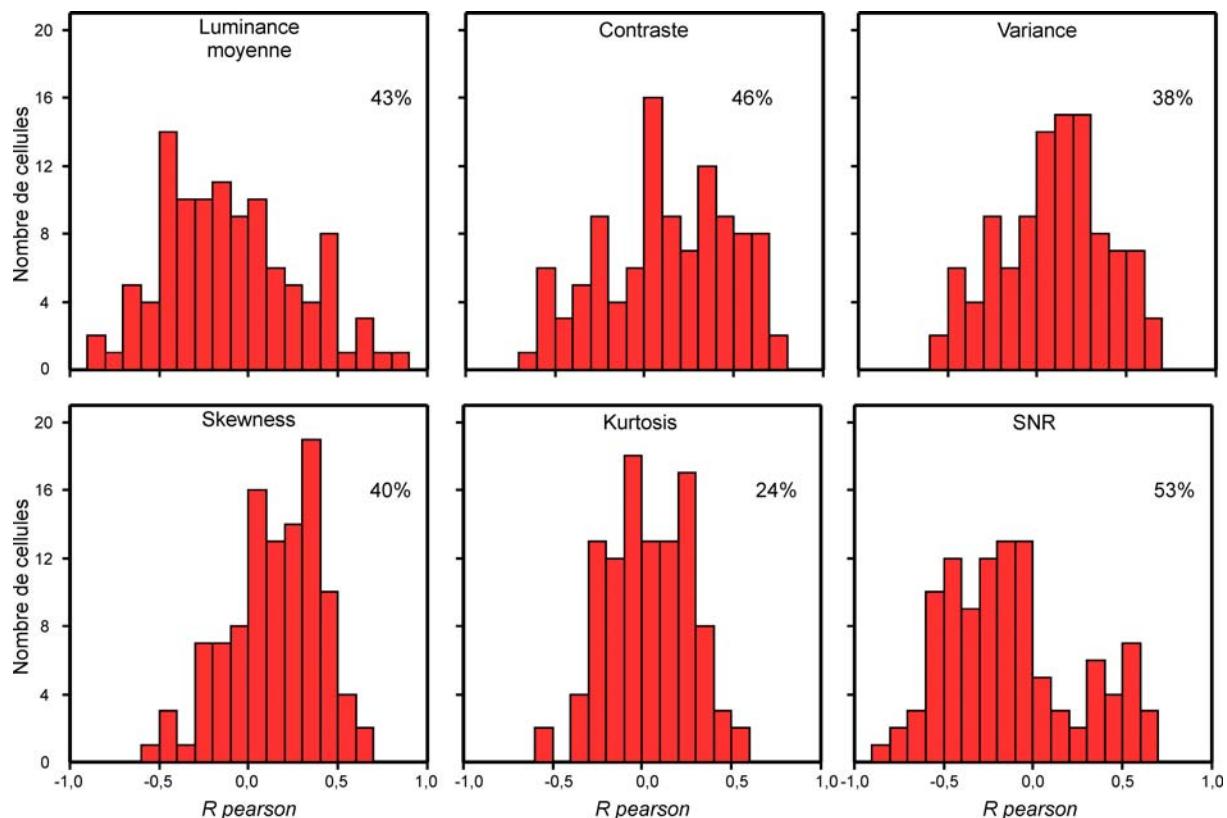


Figure 3-12: Histogrammes de distribution des coefficients de corrélation (*R Pearson*) pour chacun des paramètres statistiques de 1^{er} ordre. Le pourcentage indique le nombre de cellules TS corrélées significativement au paramètre ($p<0.05$). Taille du pas =0,1.

Ces histogrammes de distribution des coefficients de corrélations ont des médianes proches de 0 (med= -0,12 ; 0,17 ; 0,15 ; 0,18 ; 0,04 et -0,21 pour la luminance moyenne, le contraste, la variance, le skewness, le kurtosis et le SNR respectivement). Ce résultat indique que les neurones TS tendent à ne pas être corrélés à chacun de ces paramètres. Les histogrammes de distribution relatifs à la luminance moyenne, au contraste et au SNR ont des bases plus élargies que ceux établis pour le skewness (ou asymétrie), le kurtosis (aplatissement) et la variance. Ce résultat signifie que davantage de cellules sont corrélées aux paramètres : luminance moyenne, contraste et SNR. En effet, 43, 46 et 53% des cellules TS sont corrélées significativement à la luminance moyenne, au contraste moyen ou au SNR des images ($p<0.05$)¹. Ces 3 paramètres de 1^{er} ordre influencent particulièrement les réponses des cellules TS.

¹ Pour rappel, le SNR est l'inverse du contraste moyen. Ces deux paramètres décrivent une même caractéristique d'une texture.

3.2. Paramètres de 2nd ordre : Energie, entropie, inertie et homogénéité

De manière similaire, nous avons calculé les coefficients de corrélation (R Pearson) entre les paramètres de 2nd ordre et les réponses de cellules TS aux 36 stimuli originaux. La Figure 3-13 représente les histogrammes de distribution des coefficients de corrélation calculés pour les 4 paramètres de 2nd ordre.

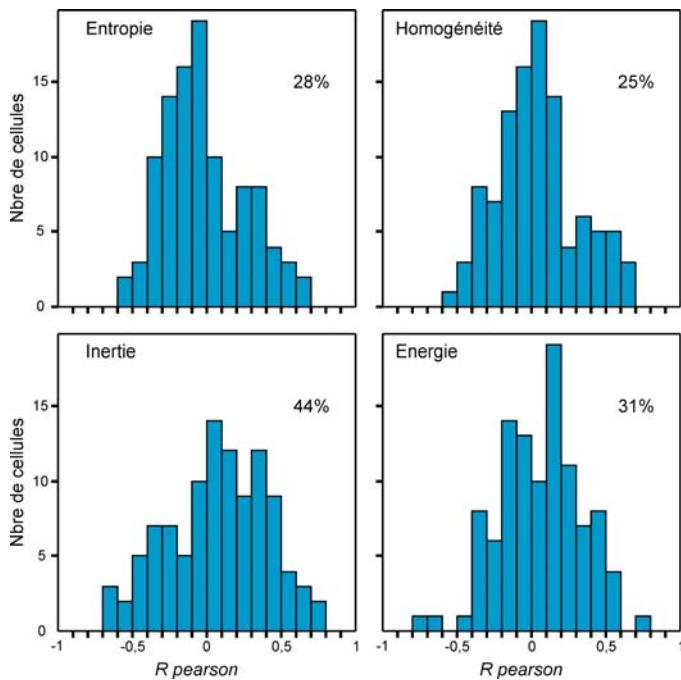


Figure 3-13 : Histogrammes de distributions des coefficients de corrélation (R Pearson) pour les 4 paramètres de 2nd ordre. Les pas ont une largeur de 0.1.

Le paramètre de 2nd ordre qui entraîne une plus forte corrélation au sein de la population TS est l'inertie ; 44% des cellules TS y sont corrélées significativement ($p < 0.05$). Pour les autres paramètres, les bases des histogrammes sont plus réduites, moins de cellules TS y sont corrélées (31, 28 et 25% pour l'énergie, l'entropie et l'homogénéité respectivement). Les paramètres de 2nd ordre ici pris en compte sont calculés pour une direction horizontale (0°). Des coefficients de corrélation ont également été calculés avec des paramètres pour des directions de 45° , 90° et 135° . Les pourcentages de cellules corrélées n'excèdent pas 30%.

Aucun des paramètres statistiques des images ne semble expliquer à lui seul les différentes réponses de la population TS à notre set d'images de textures naturelles même si la luminance moyenne et le SNR sont des paramètres influents sur le taux de décharges des neurones TS. De plus, suite aux précédentes analyses, nous avons déterminé que les cellules TS répondent plutôt à une gamme de textures qu'à une texture seule et que les textures qui engendrent une réponse maximale sont majoritairement des textures hétérogènes. Il serait intéressant de connaître les particularités du groupe de textures auquel la population TS va

répondre de façon similaire. De par la complexité de ces résultats et vu le nombre important d’interactions possibles entre les textures naturelles relevées au sein du codage de la population TS, nous avons effectué des analyses multidimensionnelles (analyse hiérarchisée par clusters et MDS) afin de mieux caractériser les réponses de notre population TS.

4. Analyses multidimensionnelles sur les réponses de la population TS

4.1. Analyse hiérarchisée par clusters

Dans un premier temps, pour chaque neurone, les réponses moyennes aux 36 stimuli originaux sont standardisées. La méthode d’analyse utilisée est la méthode de Ward. La Figure 3-14 présente le résultat de l’analyse par cluster effectuée sur les réponses moyennes des neurones TS aux 36 stimuli originaux. Ce résultat se présente sous la forme d’un arbre aussi appelé dendrogramme.

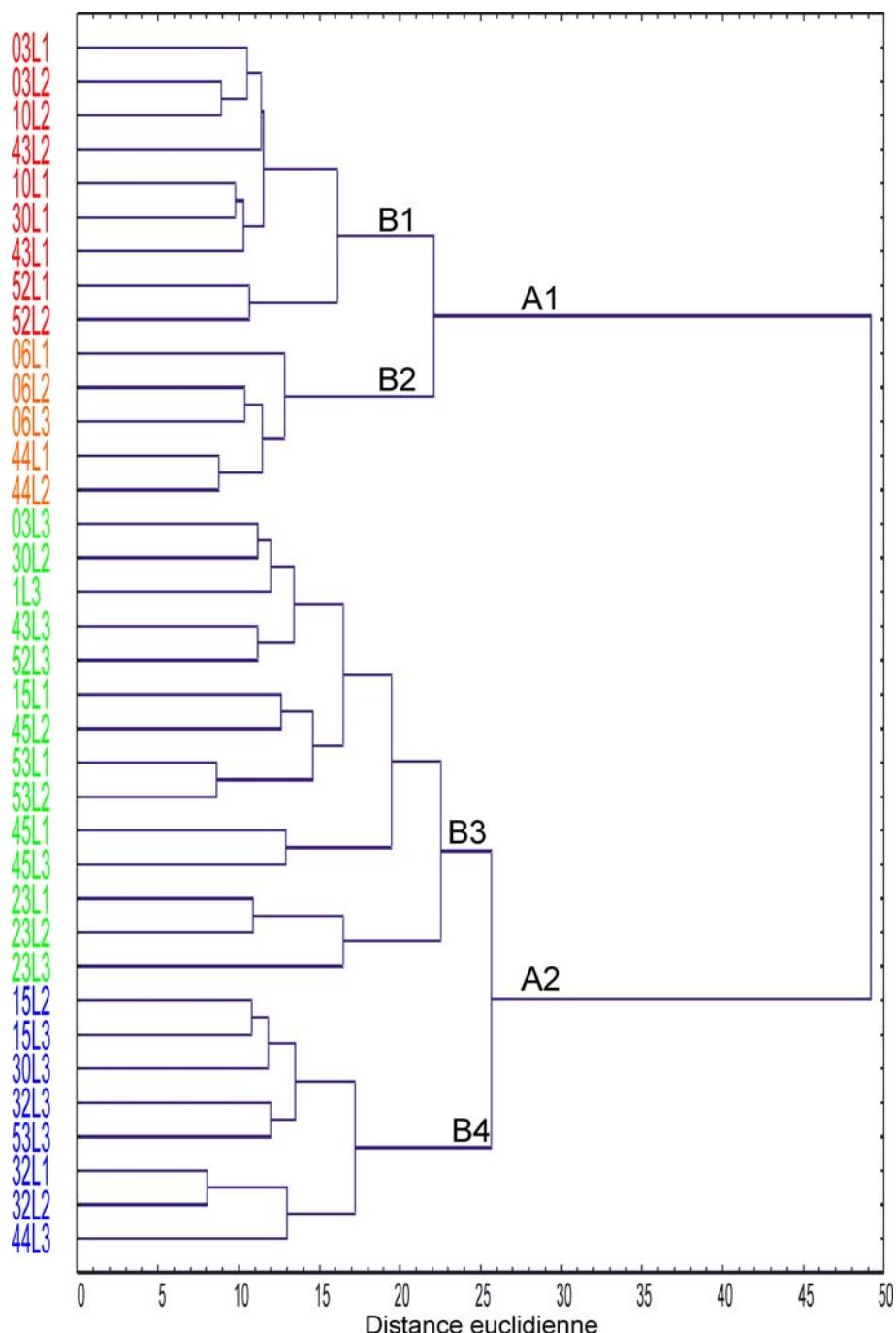


Figure 3-14 : Analyse hiérarchisée par clusters sur les réponses moyennes de population TS aux 36 stimuli originaux (n=105 cellules). Chaque branche horizontale à gauche correspond à un stimulus visuel. Les deux premiers chiffres indiquent le numéro de la texture suivi de l'angle d'éclairage (L1 =22.5°, L2 =45° et L3 =67.5°). Les couleurs schématisent la présence d'un cluster (B1, B2, B3 et B4).

Cette représentation graphique permet de caractériser les réponses moyennes de la population TS. L'arbre se dissocie en plusieurs niveaux, nous nous intéresserons qu'aux deux premiers niveaux de dissociation. Au premier niveau (distance $d=49$), l'arbre se dissocie en deux groupes (ou clusters) ; le premier, A1, est constitué des clusters B1 et B2 et le second, A2, des clusters B3 et B4. Les réponses à ces deux ensembles de textures naturelles sont très

différentes et les neurones TS vont répondre de façon similaire aux textures regroupées à l'intérieur d'un même cluster. La Figure 3-15 reprend les 4 clusters en regroupant les images des textures naturelles suivant le cluster auquel elles appartiennent. Afin de différencier les clusters les uns par rapport aux autres, nous nous sommes basés sur les paramètres statistiques des images. Une analyse de variance de type ANOVA sur les 8 paramètres, avec comme seul facteur le cluster (A ou B) auquel appartient le stimulus visuel est effectuée. Les résultats indiquent que le regroupement selon A1 ou A2 est influent pour les 8 paramètres (Anova, $p<0.05$). Il en est de même pour la séparation en 4 clusters B1, B2, B3 et B4. Pour décrire les images contenues dans les clusters, nous nous sommes intéressés en particulier à 2 paramètres : la luminance moyenne et le SNR. En effet, d'après l'analyse des coefficients de corrélation (R Pearson, paragraphe C.3), ces deux paramètres semblent les mieux appropriés pour décrire les décharges des cellules. Les valeurs moyennes des luminances et SNR des clusters sont détaillées dans le graphe de la Figure 3-15.

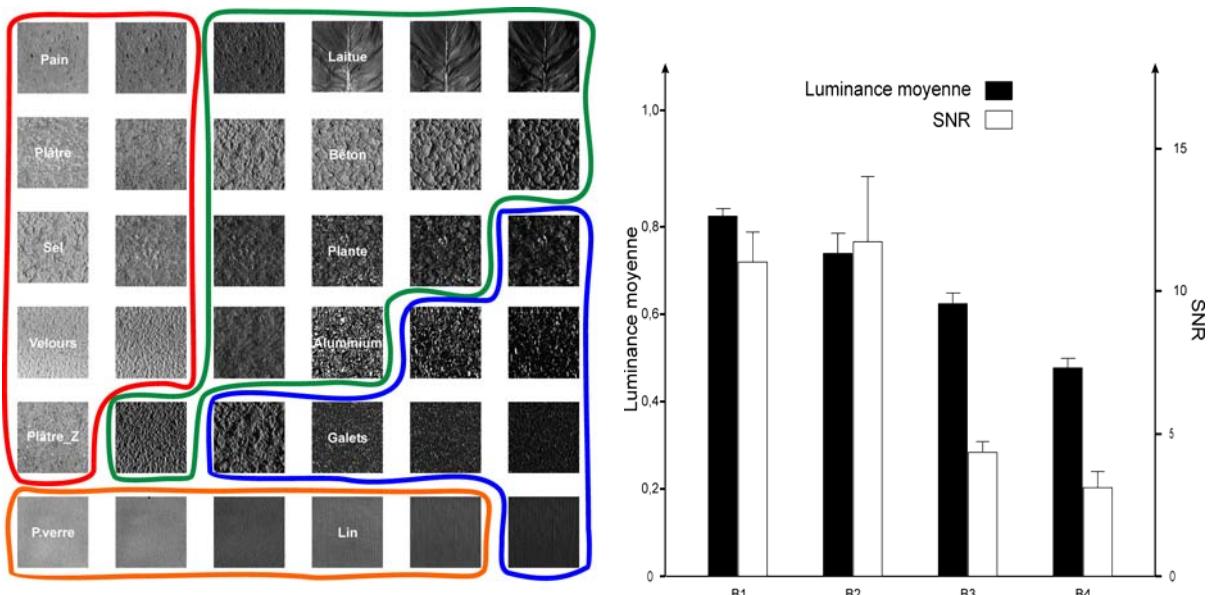


Figure 3-15 : A gauche : Images des textures naturelles regroupées selon les 4 clusters identifiées après analyses par cluster des réponses moyennes de la population TS. Sur une ligne sont représentées les images de deux textures. En colonne sont regroupées les images illuminées par un même angle ; l'image de gauche pour une illumination de 22.5° , l'image du milieu pour une illumination de 45° et l'image la plus à droite pour une illumination de 67.5° . Les images sont disposées arbitrairement et les traits de couleurs schématisent les clusters identifiés dans la Figure 3-14. A droite : Luminance moyenne et SNR des clusters B1, B2, B3 et B4. Chaque bâton indique la valeur moyenne et la barre d'erreur, l'esm.

Les textures appartenant aux clusters A1 et A2 sont respectivement hétérogènes et homogènes car elles diffèrent significativement par leur SNR moyen ($\text{SNR}_{\text{moy}} = 11,3$ et $3,9$ respectivement, ANOVA, $p<0,001$). Les textures dans le cluster A1 présentent également des luminances moyennes importantes alors que les textures de A2 des luminances moyennes

basses ($\text{Lum}_{\text{moy}} = 0,74$ et $0,56$ respectivement, ANOVA, $p<0,001$). Au second niveau de dissociation (B), les clusters B1 et B2 présentent des mêmes valeurs de SNR_{moy} et de Lum_{moy} . Le cluster B3 diffère du cluster B4 par la présence de textures plus claires ($\text{Lum}_{\text{moy}} = 0,59$ et $0,45$ respectivement, TUKEY, $p=0,039$). Aucun des paramètres de 1^{er} et de 2nd ordre que nous avons calculés ne nous permet de différencier les clusters B1 et B2 entre eux. Pour les clusters B3 et B4, hormis la luminance moyenne, seul le contraste de 2nd ordre (l'inertie pour une direction de 0°) est significativement différent (Test TUKEY, $p=0,019$).

A l'aide d'une analyse par cluster, nous avons donc identifié 2 types de réponses de la population TS correspondant aux 2 clusters A1 et A2. Les images regroupées dans ces clusters sont caractérisées par des paramètres semblables. Le cluster A1 comprend des images de textures plutôt homogènes et à forte luminance alors que le cluster A2 comprend des images de textures plutôt hétérogènes et à faible luminance. Cette séparation au sein des réponses moyennes de la population TS selon le SNR ou la luminance moyenne n'est pas exclusive ; par exemple, l'image du *plâtre* (L2) et celle du *béton* (L1) ont des valeurs de paramètres identiques mais se retrouvent dans des clusters différents (B1 et B3 respectivement).

4.2. Analyse multidimensionnelle de type MDS

L'analyse suivante est une analyse multidimensionnelle de type MDS sur les réponses moyennes de la population TS. Le test de stress (voir Figure 3-16) permet de restreindre l'analyse à 3 dimensions. En effet, le stress (*D star raw stress*) augmente de manière exponentielle à partir de 3 dimensions faisant apparaître un « coude » (Kruskal et Wish, 1978). La Figure 3-16 reprend les 3 représentations bidimensionnelles de la MDS sur la population TS.

Les clusters identifiés suite à la précédente analyse ne sont pas totalement distincts dans les 3 dimensions mais les textures d'un même cluster, identifiées par un code couleur, tendent à être proches. La première représentation bidimensionnelle (dimensions 1 vs 2) permet la meilleure séparation visuelle. Les stimuli compris dans le cluster A1 se retrouvent dans la partie gauche du graphe (valeurs de dimension 1 négatives) et les stimuli du cluster A2 dans la partie droite (valeurs majoritairement positives). La dimension 2 permet une séparation visuelle entre les clusters B1 et B3 (valeurs majoritairement positives) et les clusters B2 et B4 (valeurs négatives). Quant à la troisième dimension, elle ne permet pas de mettre en évidence une autre séparation entre les clusters. La Figure 3-17 dispose les images des textures sur les représentations 2D.

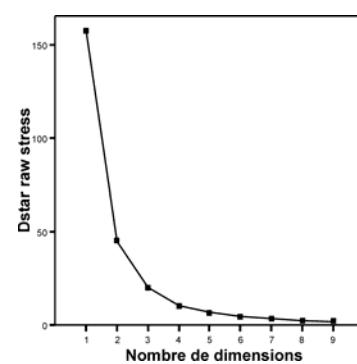
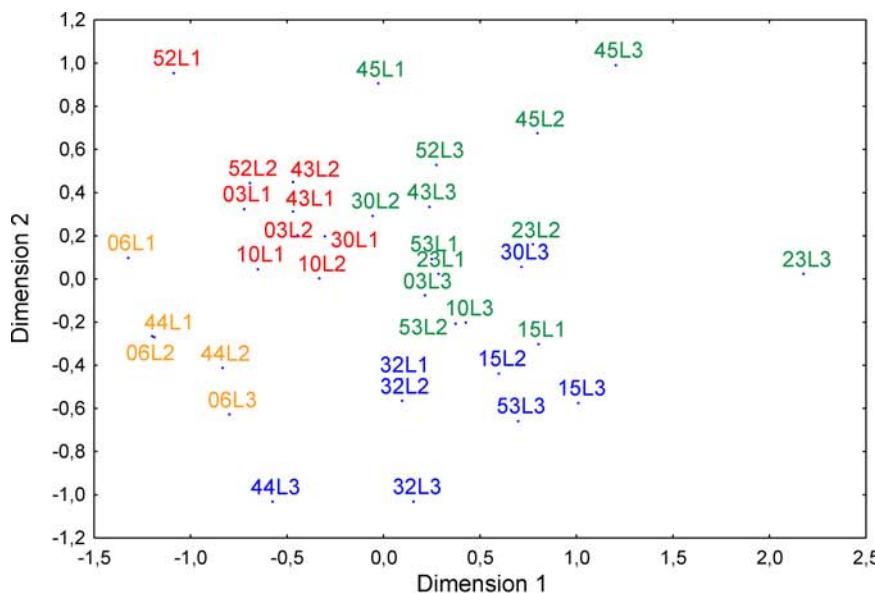
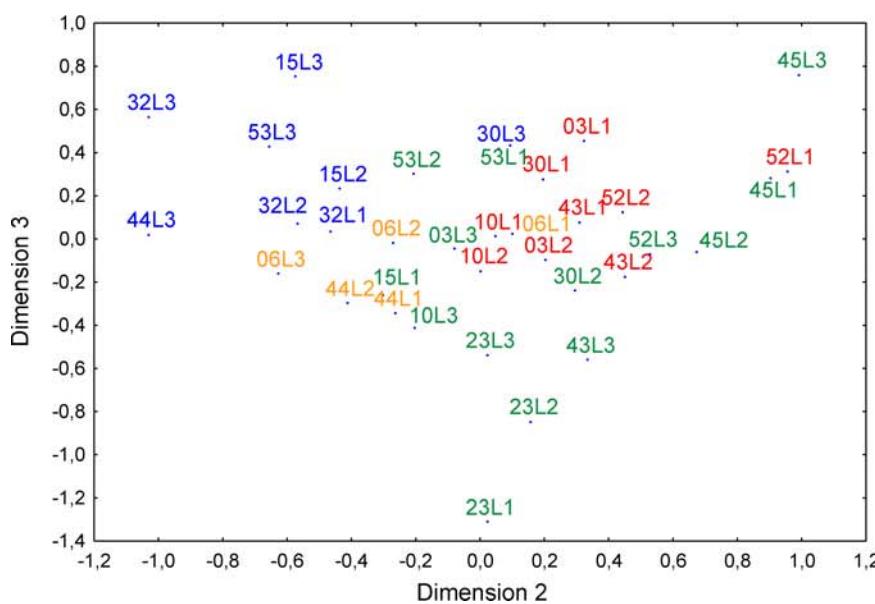
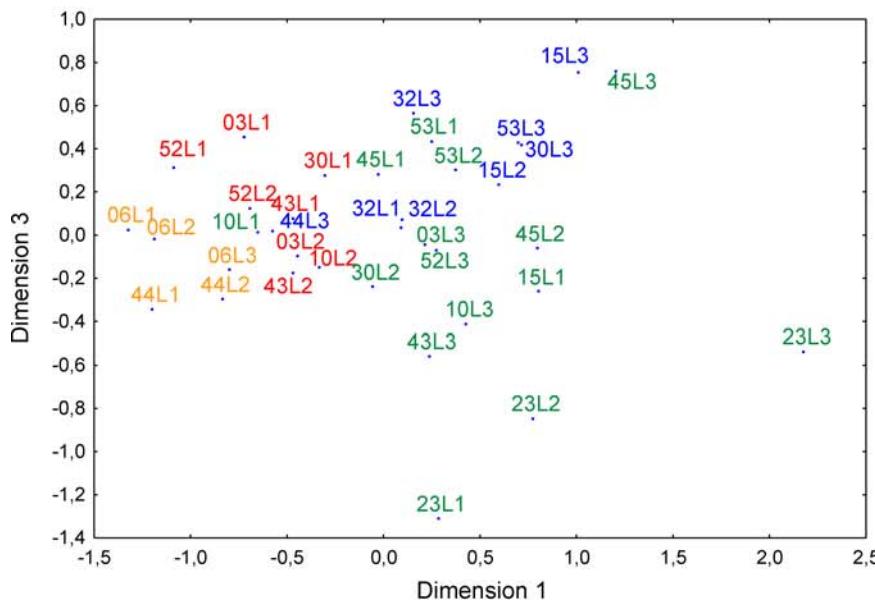


Figure 3-16 : MDS sur les réponses moyennes de la population TS aux 36 stimuli originaux. Représentations bidimensionnelles (dimension 2 VS dimension 1, dim 3 VS dim 1 et dim 3 VS dim 2). Les distances sont des distances euclidiennes. Le code couleur respecte les clusters identifiés (B1, B2, B3 et B4) avec la précédente analyse. En haut à droite : Valeur de stress (D_{star} raw stress) en fonction du nombre de dimensions.



Afin de caractériser les dimensions de cette MDS, nous avons calculé des coefficients de corrélation de Spearman (*R Spearman*) entre les coordonnées et les paramètres statistiques de chaque image. Nous ne relevons que le coefficient de corrélation le plus important ; Celui-ci permettra de caractériser les dimensions de notre MDS. Ainsi, la dimension 1 est fortement corrélée au SNR des images (*R Spearman* = -0.870, $p < 0.001$). Les textures homogènes (SNR fort) ont des valeurs négatives dans cette dimension. C'est le cas du cluster A1. Les textures hétérogènes (SNR faible) ont des valeurs positives, comme le cluster A2. La dimension 2 est corrélée significativement avec la luminance moyenne des images (*R Spearman* = 0.520, $p < 0.001$). Les images à forte luminance auront donc des valeurs de la dimension 2 majoritairement positives. C'est ce que nous pouvons observer dans la troisième représentation 2D de la Figure 3-17. Contrairement aux dimensions 1 et 2 corrélées à des paramètres de 1^{er} ordre, la dimension 3 est uniquement corrélée à des paramètres de 2nd ordre moyennés à travers toutes les matrices de cooccurrence (Energie_{moy}, *R Spearman* = 0,41, $p = 0,19$: Inertie_{moy}, *R Spearman* = 0,44, $p = 0,07$: Homogénéité_{moy}, *R Spearman* = 0,49, $p = 0,02$).

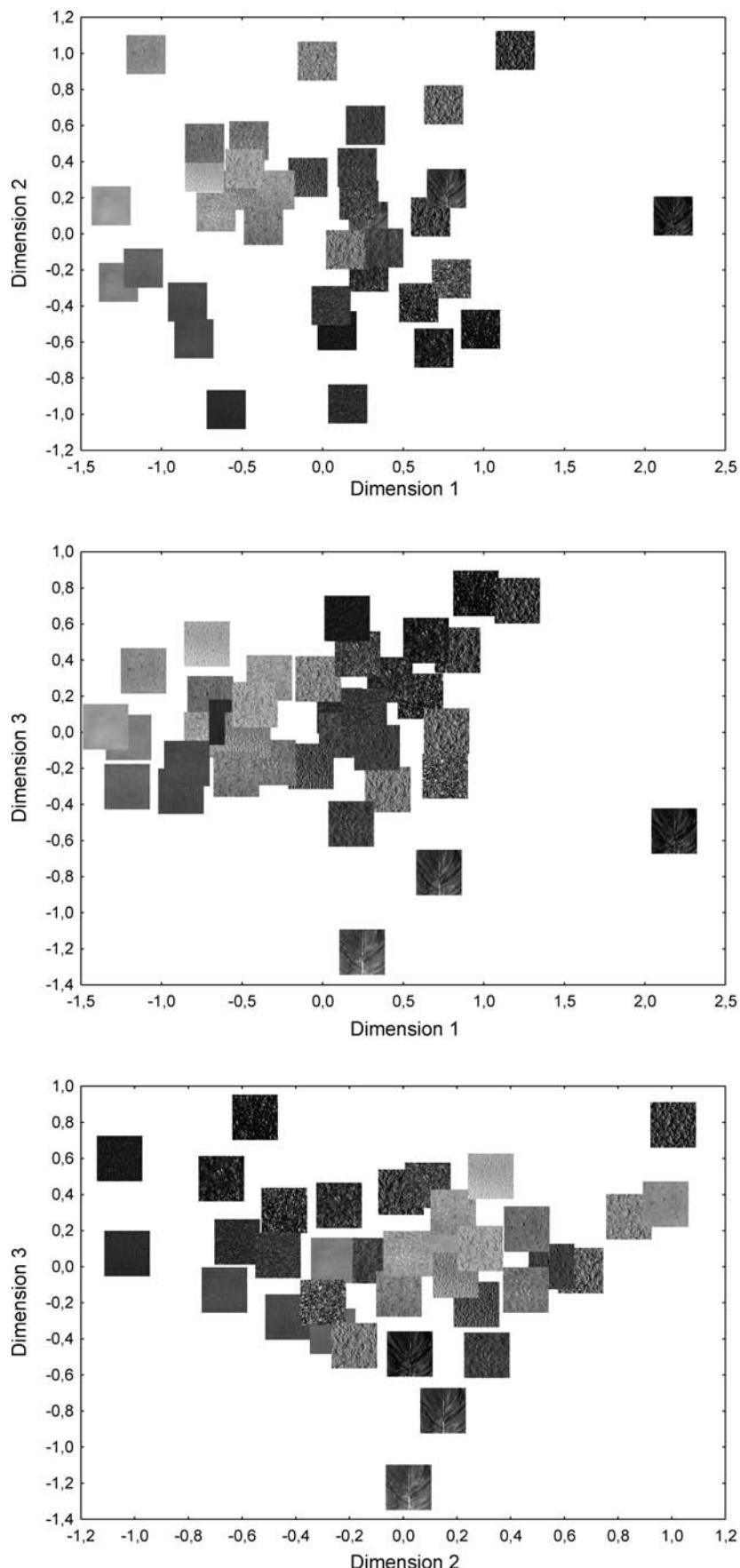


Figure 3-17 : Représentations 2D de la MDS sur les réponses de la population TS. Les photographies des textures remplacent les noms des stimuli de la Figure 3-16.

5. Influence de la phase de l'image

Pour tester l'influence de la phase de l'image sur la réponse des neurones de V4, nous disposons de stimuli contrôle : les stimuli *scrambled* (voir matériel et méthodes). La structure tridimensionnelle due au pattern de micro-ombrages présente dans les stimuli originaux disparaît dans les stimuli *scrambled*. Ainsi, l'aspect visuel des stimuli originaux diffère des stimuli où les phases sont mélangées même si les stimuli *scrambled* et originaux contiennent les mêmes amplitudes du spectre de Fourier.

5.1. Préférence des cellules TS pour les stimuli originaux

a) Résultats du test ANOVA

Les résultats du test ANOVA indiquent que 51 cellules TS montrent un effet principal significatif de la phase dans leur réponse, soit 49% de la population TS. Pour la majorité de ces cellules ($n= 47$), la réponse moyenne aux stimuli originaux est supérieure à la réponse moyenne aux stimuli contrôles. Autres que ces cellules montrant un effet principal de la phase sur leurs taux de réponse, 23 cellules ont un effet de la phase uniquement pour certaines textures naturelles (interaction PHASE X TEXTURE significative). Parmi la population TS, nous avons donc catégorisé un pool de 74 cellules influencées par la phase des stimuli soit 73 % de la population TS totale (ANOVA effet principal + interaction, $p<0.05$).

b) Sélectivité de la population TS

La Figure 3-18 représente la réponse moyenne de la population TS en fonction des textures naturelles et de la phase des stimuli. La réponse moyenne de la population TS aux stimuli originaux est supérieure à celle aux stimuli contrôle (T-test, $p<0,001$). Toutefois, la différence de réponses entre stimuli originaux et contrôles varient selon la texture naturelle considérée. Par exemple, la population TS montre une réponse similaire pour les deux types de stimuli pour 5 textures: le *papier de verre*, le *lin*, le *velours*, le *pain* et le *plâtre* (Test TUKEY, $p=0.764$, $p=0.971$, $p=0.451$, $p=0.174$ et $p=0.099$ respectivement). Ce sont des textures homogènes ($SNR_{moyen} >8$). Les textures naturelles telles que l'*aluminium*, la *laitue* et la *plante* montrent une augmentation fortement significative de leur taux de réponse ($p<0.001$). Ce sont des textures très hétérogènes ($SNR_{moyen} <4$). Quant aux autres textures (le *plâtre(z)*, les *galets*, le *sel* et le *béton*), elles sont faiblement dissociées ($p<0.05$). Ce sont des textures hétérogènes ($SNR_{moyen} <4$) hormis le *sel* qui peut être considéré comme une texture homogène ($SNR_{moyen} >9$). Ce résultat est à considérer pour la population TS totale mais qu'en est-il pour chaque cellule ? Au niveau individuel, un test TUKEY nous indique pour chaque texture naturelle si le neurone augmente son taux décharge pour les stimuli originaux.

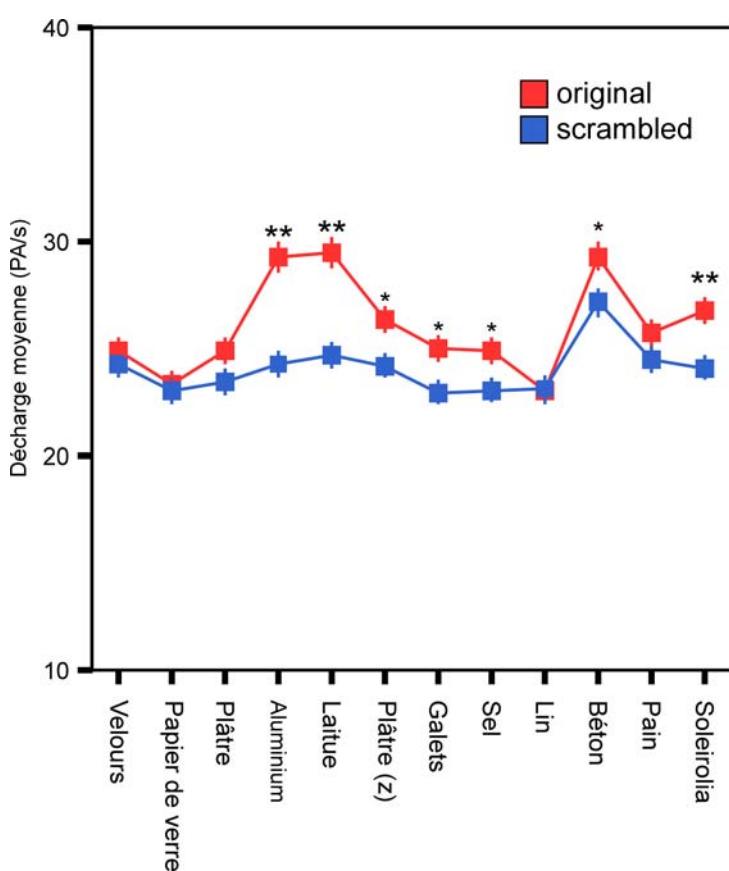


Figure 3-18 : Réponses moyennes de la population TS en fonction des textures naturelles et de la phase des stimuli. Les astérisques schématisent les résultats du test TUKEY sur la phase. * : $p<0.05$; ** : $p<0.001$.

c) Préférence individuelle

Au sein de la population TS influencée pour la phase des stimuli ($n=74$), les cellules ont une réponse supérieure pour, en moyenne, 2 textures originales parmi les 12 testées (test TUKEY, $p<0.05$). Les textures originales les plus différenciées par cette sous-population de cellules TS sont la *laitue*, l'*aluminium* et la *plante*, soient les textures hétérogènes codées préférentiellement par les cellules TS dans leur globalité (voir paragraphe C.2.1.b). Nous nous intéressons dorénavant qu'aux textures préférées des neurones. De manière générale, l'analyse de la population TS montre que 22 % des textures préférées engendrent une réponse neuronale supérieure pour les stimuli originaux par rapport aux stimuli contrôles. La grande majorité des textures (77%) présentent un même niveau de réponse pour les stimuli originaux et contrôles. Le reste des textures (1%) regroupe les textures naturelles pour lesquelles les cellules TS ont un taux de décharge significativement plus important pour les stimuli contrôle. La Figure 3-19 décrit pour chaque texture, le pourcentage de ces cellules qui diffèrent le stimulus original du stimulus contrôle.

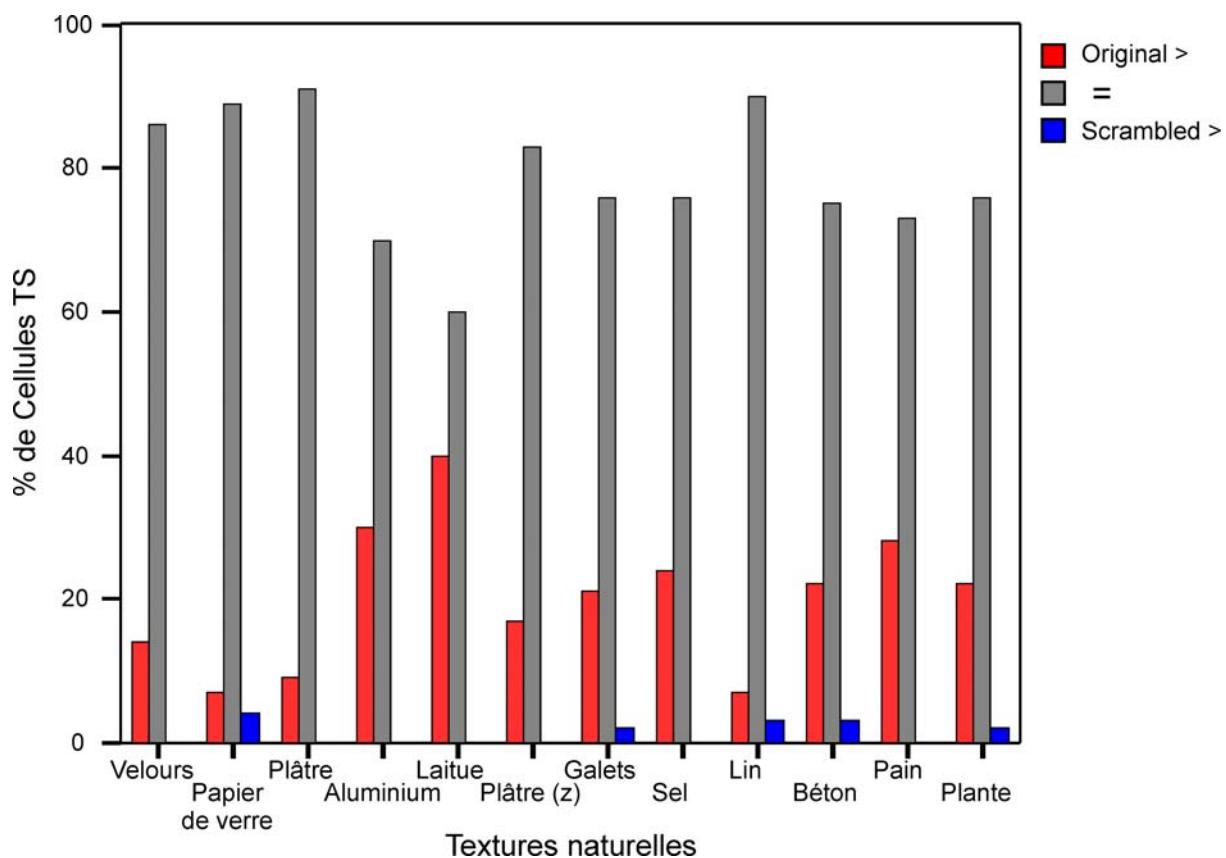


Figure 3-19 : Pourcentage de cellules TS qui diffèrent les stimuli originaux des stimuli contrôles pour chaque texture naturelle. Les bâtons rouges indiquent les pourcentages de cellules qui diffèrent les stimuli originaux des stimuli contrôles avec un codage supérieur pour les stimuli originaux (Original >); les bâtons gris, les pourcentages de cellules qui ne diffèrent pas les stimuli (=) et enfin les bâtons bleus, les pourcentages de ces cellules qui diffèrent le stimulus original du stimulus contrôle mais avec un codage inférieur (Scrambled >).

Pour décrire la différence de codage entre stimuli originaux et contrôle pour chaque texture naturelle, je conserve la séparation en trois groupes des textures naturelles observée suite à l'analyse de la Figure 3-9). Dans un premier temps, nous nous intéressons aux neurones répondant sélectivement aux textures homogènes (*papier de verre*, *plâtre*, *pain*, *lin* et *velours*). Le neurone de la Figure 3-20 est un exemple de cette catégorie de cellule. Il répond sélectivement au *papier de verre* et au *lin* (d'après I_{max} , colonne de gauche).

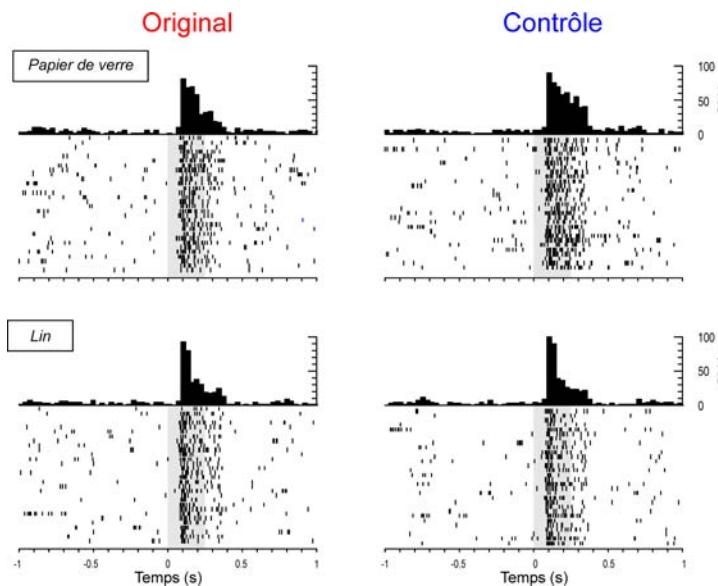


Figure 3-20 : Réponses d'un neurone (TMP2_11) aux textures naturelles *papier de verre* et *lin* en fonction de la phase des stimuli. Les conventions sont identiques à celles de la Figure 3-6. Les différentes illuminations sont cumulées. Les réponses moyennes aux stimuli originaux sont représentées à gauche de la figure, en haut pour le *papier de verre* et en bas pour le *lin*. Sur la colonne de droite sont représentées les réponses moyennes aux stimuli contrôle correspondant. Les réponses aux stimuli originaux et scrambled sont équivalentes.

La cellule de la Figure 3-20 a une même taux de décharge pour les stimuli originaux et scrambled des deux textures homogènes (TUKEY, $p=0.314$ et $p=0.212$ pour le *papier de verre* et le *lin* respectivement). Au niveau de la population TS, les textures homogènes, qu'elles soient présentées en phase originale ou aléatoire, engendrent une réponse similaire des neurones. Par exemple, 86% des cellules TS ont une réponse identique pour la texture *velours* quelque soit la phase des stimuli ; 89% pour le *papier de verre*, 91% pour le *plâtre*, 73% pour le pain et 90% pour le *lin*. Pour résumer, peu de cellules TS différencient les stimuli originaux des stimuli contrôle de ces textures homogènes (Textures du groupe 3, cf. Figure 3-9).

De la même manière, nous nous intéressons aux neurones qui répondent aux textures hétérogènes (*béton*, *sel*, *galets* et *plâtre (z)*). Les pourcentages de réponses homologues baissent légèrement puisque 78%, 76%, 76% et 83% des cellules TS ont un même niveau de réponse (par ordre respectif).

Enfin, pour les neurones qui répondent aux textures très hétérogènes (*aluminium*, *laitue* et *plante*), les taux d'homologie chutent. Le neurone de la Figure 3-21 en est un exemple.

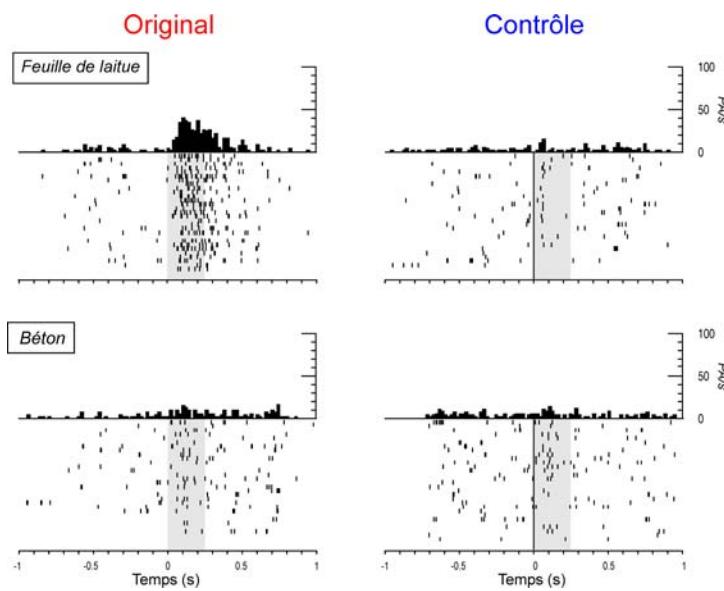


Figure 3-21 : Exemple de cellule sélective (TMP46_11) à une texture hétérogène. Mêmes conventions que la Figure 3-20.

Cette cellule répond sélectivement à la *feuille de laitue* mais présente une faible modulation de sa décharge pour le *béton* (colonne de gauche). Pour les stimuli contrôle, sa décharge neuronale chute significativement pour la *feuille de laitue* ($p= 0.001$), et reste faible pour les stimuli contrôle du *béton*. Au sein de la population TS, la texture *laitue* est la texture préférée; 61 % des cellules TS y répondent sélectivement ($n= 63$). Parmi ces cellules TS, 40% ont une réponse pour les stimuli originaux significativement supérieure ($n= 25$, comme ici présentée dans l'exemple de la Figure 3-21). Cette texture naturelle engendre le plus fort pourcentage de cellules TS sélectives qui différencient les stimuli originaux des stimuli contrôles soit 40%. Pour les autres textures hétérogènes préférées (*aluminium* et *plante*), les pourcentages de cellules sélectives à la phase des stimuli chutent (30% et 22% respectivement).

En résumé, pour la majorité des textures (77%), les neurones TS ne différencient pas les stimuli originaux des stimuli contrôles. Toutefois, si on observe individuellement les neurones, les cellules différencient mieux les stimuli originaux des textures hétérogènes (*laitue*, *aluminium* et *plante*) de leurs stimuli contrôles que les textures homogènes (*lin*, *papier de verre*, *plâtre* et *velours*).

5.2. Histogramme de population et analyse par rang

La Figure 3-22A présente le décours temporel de la réponse moyenne aux stimuli originaux et aux stimuli contrôle de la population TS. Les réponses pour chaque type de stimuli sont alignées sur la présentation des stimuli visuels. Le décours temporel de la réponse moyenne de la population TS aux stimuli originaux ou contrôle est caractérisé par l'apparition

de 2 phases distinctes correspondantes à 2 périodes temporelles. Dans la première phase, un pic de réponse apparaît après une latence de 50 ms ; ce pic n'est pas sélectif au type de stimuli présenté. En effet, les réponses moyennes aux stimuli originaux et aux stimuli contrôles sont identiques pendant la période temporelle de ce premier pic d'activité soit de 50 ms à 90 ms (ANOVA type, $p=0.081$). La sélectivité pour les stimuli originaux apparaît dans une période plus tardive dans la seconde phase après une latence de 90 ms. Cette sélectivité se signale par une augmentation significative de la fréquence moyenne pour les stimuli originaux par rapport aux stimuli contrôle pendant une période de 110ms (de 90 à 300 ms, ANOVA $p<0.001$).

L'analyse suivante compare les réponses pour tous les stimuli originaux à celles pour les stimuli *scrambled* correspondants, sans prendre en compte le critère I_{max} et sans cumuler les angles d'incidences. Pour cela, nous avons classé pour chaque neurone les réponses aux stimuli originaux dans un ordre décroissant (le premier stimulus étant la texture préférée du neurone). Ce même ordre est utilisé pour représenter la réponse au stimulus contrôle (*scrambled*) correspondant. Pour cette analyse, les illuminations sont considérées séparément, car chaque stimulus original a un stimulus contrôle. Afin d'écartier une variabilité interneurones au niveau du taux des décharges des neurones, les réponses des cellules sont normalisées d'après la réponse maximale au stimulus préféré. Les réponses normalisées pour chaque rang sont ensuite moyennées au sein de la population TS.

La Figure 3-22B représente les réponses moyennes normalisées de la population TS en fonction du rang du stimulus et de la phase des stimuli. Cette figure montre la présence d'une sélectivité pour les stimuli originaux qui se traduit par une augmentation de la décharge moyenne de la population TS pour les stimuli préférés (ceux qui ont un rang petit) et une diminution en fonction du rang du stimulus. La réponse moyenne de la population pour le stimulus original préféré (rang 1) est proche de 1 (0.97). Ceci indique que pour la majorité des cellules, le stimulus qui engendre une décharge maximale du neurone est un stimulus original et non un stimulus *scrambled*. Une sélectivité apparaît également au niveau de la réponse moyenne pour les stimuli contrôles. En effet, le coefficient des droites de régression linéaire (non illustrés) pour les stimuli originaux et les stimuli contrôle sont identiques ($R^2 = 0,94$) même si la différence entre les réponses moyennes pour le stimulus préféré (rang 1) et celle pour le stimulus le moins efficace (rang 36) est plus importante pour les stimuli originaux que pour les stimuli contrôle (0.78, 0.21 respectivement). Un autre résultat intéressant est à noter ; les réponses moyennes aux stimuli contrôles ne sont pas complètement indépendantes des réponses originales. Autrement dit, si un stimulus est préféré par le neurone, son contrôle aura

également tendance à être celui qui donne la plus forte réponse parmi les stimuli *scrambled*. Ce résultat est confirmé par le calcul du coefficient de corrélation de Spearman qui montre une forte corrélation entre les deux séries de réponses moyennes (courbe rouge vs. courbe bleue, $R_{Spearman} = 0,98$, $p < 0,001$). Cette corrélation significative pourrait s'expliquer par le fait que les amplitudes du spectre de Fourier sont conservées entre les stimuli originaux et contrôles.

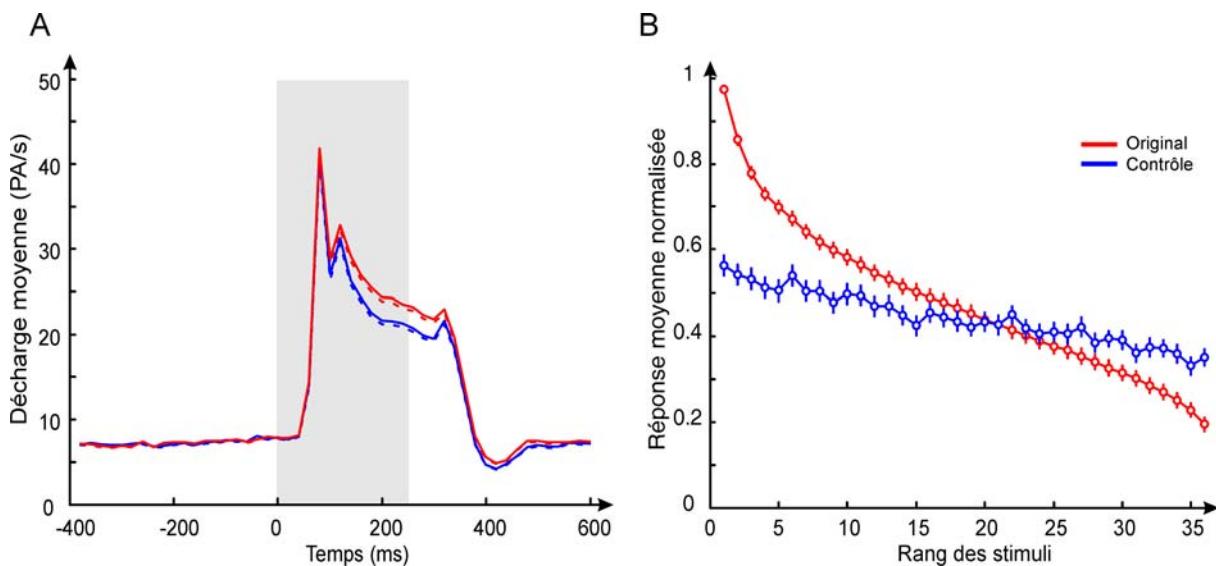


Figure 3-22: A. Histogramme de réponses de la population TS selon le type des stimuli (+ esm : ligne continue, - esm : ligne pointillée). Au sein de la population TS, les réponses sont moyennées pour chaque type de stimuli par intervalle de 20 ms et sont alignées sur la présentation des stimulations visuelles (t0). B. Décharges moyennes normalisées de la population TS en fonction du rang du stimulus (+/- esm). La couleur des courbes correspond au type des stimuli, en rouge pour les stimuli originaux et en bleu pour les stimuli contrôles ou *scrambled*. Une sélectivité apparaît au niveau des réponses aux stimuli originaux. Les réponses aux stimuli originaux et contrôles sont corrélées significativement ($R_{Spearman}=0,98$, $p < 0,001$).

Pour résumer, au sein de la population TS, l'analyse par rang nous démontre la présence d'une sélectivité pour les stimuli originaux générant une augmentation de la décharge neuronale pour les stimuli originaux préférés par rapport aux stimuli contrôle correspondant même si les réponses aux stimuli contrôle correspondants ne sont pas complètement indépendantes. Cette sélectivité apparaît dans une période tardive de réponse de la population ICS.

5.3. Synthèse

En comparant les réponses des cellules aux stimuli originaux à celles pour les stimuli *scrambled*, nous avons mis en évidence plusieurs particularités :

- La population TS voit son taux de décharge moyen augmenter pour les stimuli originaux en comparaison avec les stimuli *scrambled*.

- 77% des cellules présentent une modulation de leur réponse en fonction de la phase des stimuli.
- Les neurones TS différencient mieux les stimuli originaux de leurs stimuli *scrambled* pour les textures hétérogènes.
- Il existe une corrélation significative entre les réponses moyennes aux stimuli originaux et celles aux stimuli *scrambled* correspondant.

6. Influence de l'angle d'incidence de la source d'illumination

Dans les analyses précédentes, nous avons caractérisé une population de neurones sélectives aux textures naturelles. Cette population TS répond significativement mieux aux stimuli originaux qu'aux stimuli *scrambled*. Cependant, dans toutes ces précédentes analyses, nous avons considéré les réponses aux textures naturelles indépendamment des angles d'incidence (sauf pour l'analyse par rang) ; nous cumulions les réponses aux trois illuminations (22,5°, 45° et 67,5°). Dans cette partie, nous allons étudier l'influence de ce facteur sur les réponses des neurones TS dans le codage des textures afin de démontrer la présence ou pas d'une invariance du taux de décharge à l'angle d'incidence.

6.1. Sélectivité à l'angle d'incidence

Les résultats du test ANOVA indiquent que 66% de cellules de V4 (n=69) présentent une sélectivité à l'angle d'incidence (effet principal, $p<0,05$). Parmi ces 69 cellules, 65 sont des cellules TS. La Figure 3-23 présente un exemple de cellule TS dont la décharge neuronale est influencée par l'angle d'incidence. Cette sélectivité n'est pas présente pour toutes les textures. Par exemple, cette cellule montre un même taux de réponse pour les 3 images du *lin* et a des taux de réponses différents pour le *béton*. De plus, la même sélectivité pour les angles d'illumination n'est pas observée pour toutes les textures naturelles ; pour le *velours*, l'angle préféré est de 22,5° alors que pour le *béton*, la réponse maximale est pour l'angle de 67,5°. Parmi les cellules TS influencées par l'angle d'incidence, la majorité présente une interaction significative entre la texture naturelle testée et l'angle d'incidence (63%, soit 41 cellules TS).

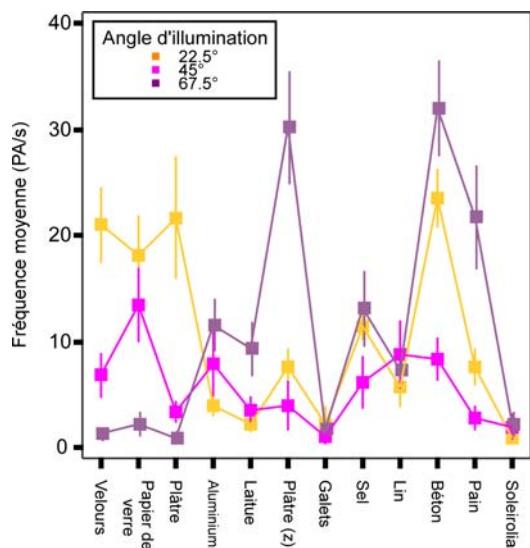


Figure 3-23: Exemple de cellule TS sélective à l'angle d'incidence aux stimuli originaux (TMP53.12). Seules les réponses aux stimuli originaux sont considérées.

6.2. Index de sélectivité à l'angle d'incidence (DMI)

D'après l'analyse précédente, la majorité des cellules TS influencées par l'angle d'incidence présente des patterns de sélectivité différents selon la texture naturelle considérée. Nous nous sommes donc restreints à déterminer si une invariance est présente pour les textures préférées des neurones TS. Pour cela, nous avons calculé un index de sélectivité à l'angle d'incidence (Direction Modulation Index ou DMI) décrit dans l'étude de Vogels et Biederman (Vogels et Biederman, 2002). Pour chaque neurone, nous choisissons parmi les 12 textures naturelles, la texture naturelle préférée, c'est à dire celle qui engendre une réponse moyenne maximale du neurone. La texture naturelle préférée se présente sous la forme de 3 stimuli selon l'angle d'incidence. Nous calculons la différence entre la réponse moyenne au stimulus préféré (celui qui engendre une réponse neuronale maximale parmi les trois stimuli originaux) et la réponse moyenne minimale. Cette différence est ensuite divisée par la moyenne des erreurs standard de ces deux conditions. Une cellule qui code avec un même niveau de réponse les 3 stimuli a un index de sélectivité DMI égal à 1. Cette cellule est dite invariante à l'angle d'incidence pour la texture naturelle préférée. Si l'index augmente, alors la cellule présente une variabilité selon l'angle d'incidence.

Le neurone de la Figure 3-24A est un exemple de cellule invariante à l'angle d'incidence alors que la cellule de la Figure 3-24B voit ses réponses modulées. La cellule de la Figure 3-24A répond préférentiellement à la *feuille de laitue*. Cet exemple de cellule présente un DMI de 0,84. A l'inverse, la cellule de la Figure 3-24B est sélective à l'angle d'incidence. En effet, cette cellule voit sa décharge neuronale varier pour la *mie de pain* selon l'angle d'incidence. La réponse de ce neurone est maximale pour un angle d'incidence de 22,5°. Elle est presque nulle pour l'angle de 67,5°. Cette cellule présente un DMI de 6,31.

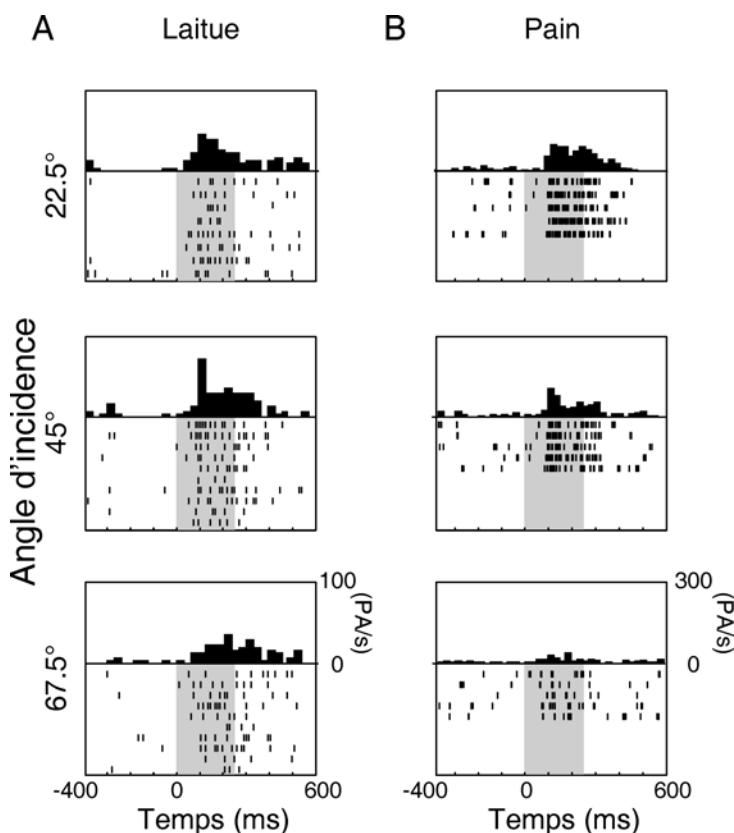


Figure 3-24 . A : Exemple de cellule TS non sélective à l'angle d'incidence pour sa texture maximale, la laitue (TMP46.11). Chaque quadrant correspond à un angle d'incidence 22,5°, 45° et 67,5°. B : Exemple de cellule TS sélective à l'angle d'incidence (TMP74.31). Cette cellule répond aux stimuli originaux de la *mie de pain* et présente une réponse neuronale maximale pour un angle d'incidence de 22,5°. Les intervalles ont une largeur de 25 ms.

La Figure 3-25 présente l'histogramme de distribution des DMI de la population TS ($n=105$). La médiane de cet histogramme est de 0,97. La modulation de la réponse due à l'angle d'incidence est donc proche de la déviation standard de la réponse du neurone. Ceci indique un résultat important ; la majorité des cellules TS présentent une relative invariance au niveau de leur réponse neuronale pour la texture naturelle préférée quelle que soit les conditions d'illumination (comme l'exemple de la cellule de la Figure 3-24A). Au sein de la population TS, seules 14 cellules montrent un index de sélectivité supérieur à 2 ; la cellule de la Figure 3-24B représente l'extrême, avec un index de 6,31.

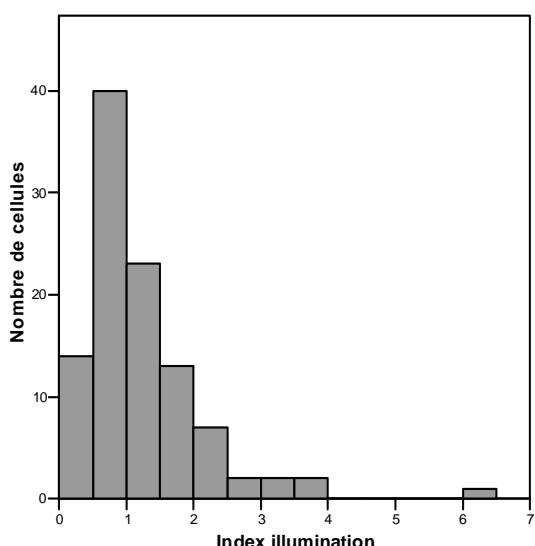


Figure 3-25 : Histogramme de l'index d'illumination DMI pour la population TS. La largeur du pas est de 0,5.

6.3. Synthèse sur l'influence de la direction d'illumination

- 66% des cellules de V4 (n=69) présentent une modulation de leurs réponses en fonction de l'angle d'incidence (ANOVA, effet principal, $p<0.05$).
- Parmi ces cellules, la majorité sont des cellules TS (96%, n=65).
- Le pattern de sélectivité diffère selon la texture naturelle considérée.
- Les cellules TS présentent une relative invariance à l'angle d'incidence pour leur texture naturelle maximale.

7. Approche computationnelle pour une classification de textures à l'aide d'une pyramide de filtres GABOR

7.1. Energie de convolution des filtres

Le calcul d'outils mathématiques permettant de segmenter une image texturée ou de classer des textures entre elles a fait l'objet de nombreuses études dans le domaine de la vision artificielle (Rao et Lohse, 1996; Varma et Zisserman, 2002; Hayman *et al.*, 2004), voir chapitre 1, p21). Dans ces différentes études, l'utilisation d'un ensemble de filtres semble biologiquement plausible et apporte d'intéressants résultats (Turner, 1986; Leung et Malik, 2001; Hayman *et al.*, 2004). Les analyses suivantes montrent qu'une population de neurones de V4 se comporte comme un ensemble de filtres linéaires pour analyser les textures naturelles. Nous utilisons un ensemble ou « pyramide » de 12 filtres GABOR. Chaque image texturée est filtrée par chacun des filtres. Ce filtrage consiste en une convolution. L'énergie de la convolution est utilisée comme paramètre de sorties des filtres.

7.2. Analyses multidimensionnelles de l'énergie de la pyramide de filtres

a) Analyse hiérarchisée par clusters

Chaque stimulus est filtré par la pyramide des filtres et fournit une énergie de convolution pour chacun des filtres GABOR. Une analyse par cluster sur ces énergies de convolutions nous permet de représenter les stimuli sous la forme d'un arbre (Figure 3-26).

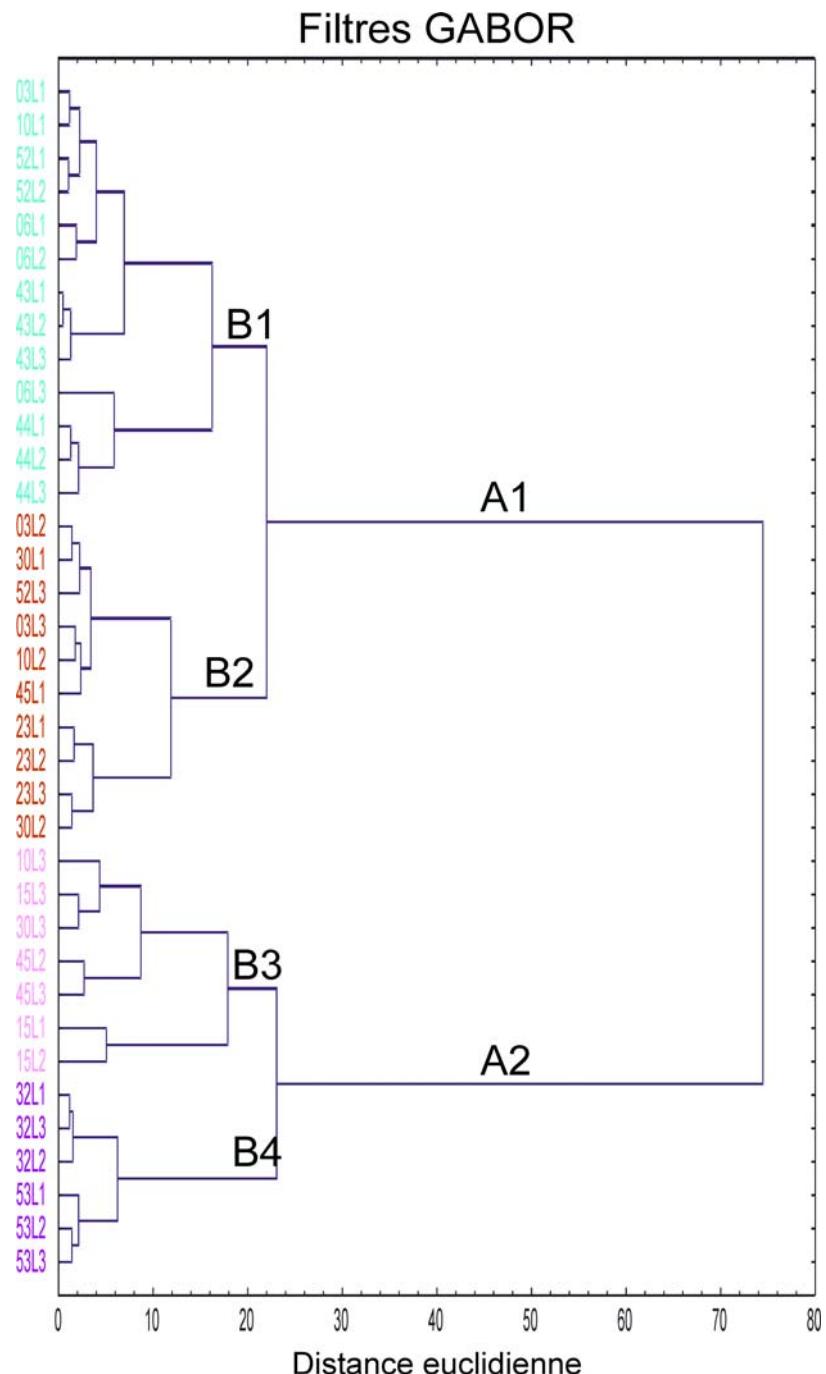


Figure 3-26 : Analyse hiérarchisée par clusters des énergies de convolutions des filtres GABOR pour les 36 stimuli originaux. Les énergies de convolution sont standardisées pour chaque filtre, et la méthode d'analyse est la méthode de Ward. Le code couleur est arbitraire et correspond à 4 clusters identifiés (B1, B2, B3 et B4)

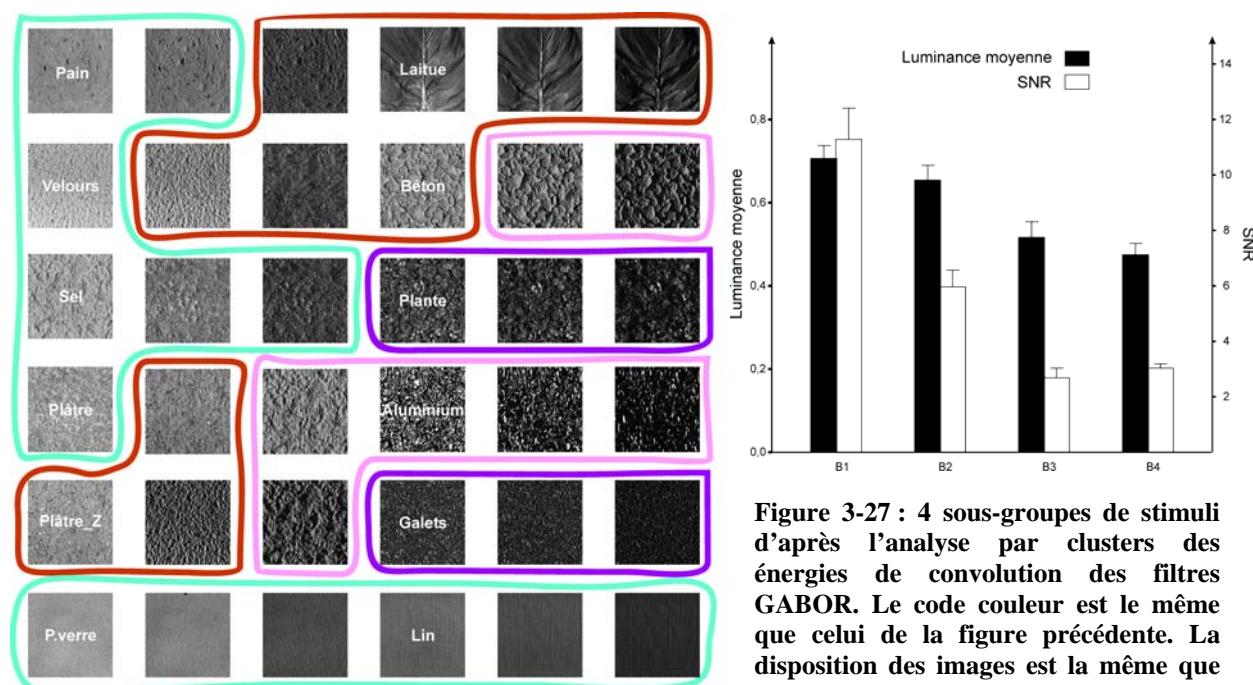
Les stimuli peuvent être séparés, d'après les énergies de convolution des filtres GABOR, en 2 groupes distincts (A1 et A2). Chacun de ces groupes comprend lui-même 2 clusters (B1, B2 et B3, B4 respectivement). La Figure 3-27 reprend les 4 sous-groupes de stimuli en y regroupant les images des textures naturelles selon le cluster. Le graphe de droite présente les luminances et les SNR moyens des clusters.

Le cluster B1 comprend 13 stimuli dont les 3 stimuli du *papier de verre*, du *lin* et du *sel*. Ce sont principalement des textures homogènes plutôt à fortes luminances (SNR_{moyen}= 11.29 et lum_{moyen}= 0.70). Le cluster B2 comprend des images plus hétérogènes (SNR_{moyen}= 5.98) que celles de B1 mais la luminance reste relativement élevée (lum_{moyen}= 0.65).

Le second groupe de textures naturelles (A2) est constitué de 2 sous-ensembles B3 et B4 chacun comprenant respectivement 7 et 6 images de textures. Ils sont caractérisés par la présence de textures hétérogènes et à faible luminance (B3 ; SNR_{moyen}= 2.70 et lum_{moyen}= 0.51 et B4 ; SNR_{moyen}= 3.03 et lum_{moyen}= 0.48).

Une analyse ANOVA avec comme unique facteur le cluster auquel l'image appartient nous indique que les clusters sont influencés par la luminance et le SNR (ANOVA, effet principal, $p < 0.001$). De plus, un test post-hoc TUKEY nous indique que les clusters B1 et B2 ont des luminances moyennes similaires, c'est le cas également pour les clusters B3 et B4.

Au niveau du SNR, seul le cluster B1 diffère des trois autres clusters avec un SNR_{moyen} élevé (SNR_{moyen}= 11.29). Ces deux paramètres ne permettent pas de séparer les clusters B3 et B4 et l'analyse des autres paramètres nous indique que seule la variance moyenne du cluster B3 diffère significativement de celle du cluster B4 (B3 ; var_{moyen} = 0.040 et B4 ; var_{moyen} = 0.024).



droite : Luminance et SNR moyens selon les clusters identifiés d'après l'analyse par clusters des énergies de convolution des filtres GABOR. Le code couleur est le même que celui de la figure précédente. La disposition des images est la même que celle utilisée pour la Figure 3-15. A

droite : Luminance et SNR moyens selon les clusters identifiés d'après l'analyse par clusters des énergies de convolution des filtres GABOR. Le code couleur est le même que celui de la figure précédente. La disposition des images est la même que celle utilisée pour la Figure 3-15. A

La pyramide de filtres GABOR semble donc se baser sur plusieurs paramètres (SNR et luminance) afin de différencier et de regrouper en clusters les images des textures naturelles. De par la contribution de multiples paramètres, une analyse multidimensionnelle (MDS) sur les énergies de convolution des filtres GABOR nous a paru nécessaire afin de représenter les relations entre les différentes images.

b) Analyse de type MDS sur l'énergie de la pyramide de filtres

La Figure 3-29 reprend les représentations bidimensionnelles de cette analyse multidimensionnelle. Les valeurs de stress (non représentées ici) permettent de restreindre l'analyse à 3 dimensions. La représentation A (dim 1 vs dim 2) permet une meilleure séparation visuelle entre les clusters. Les clusters B1 et B2 ont des valeurs majoritairement négatives alors que les clusters B3 et B4 des valeurs exclusivement positives.

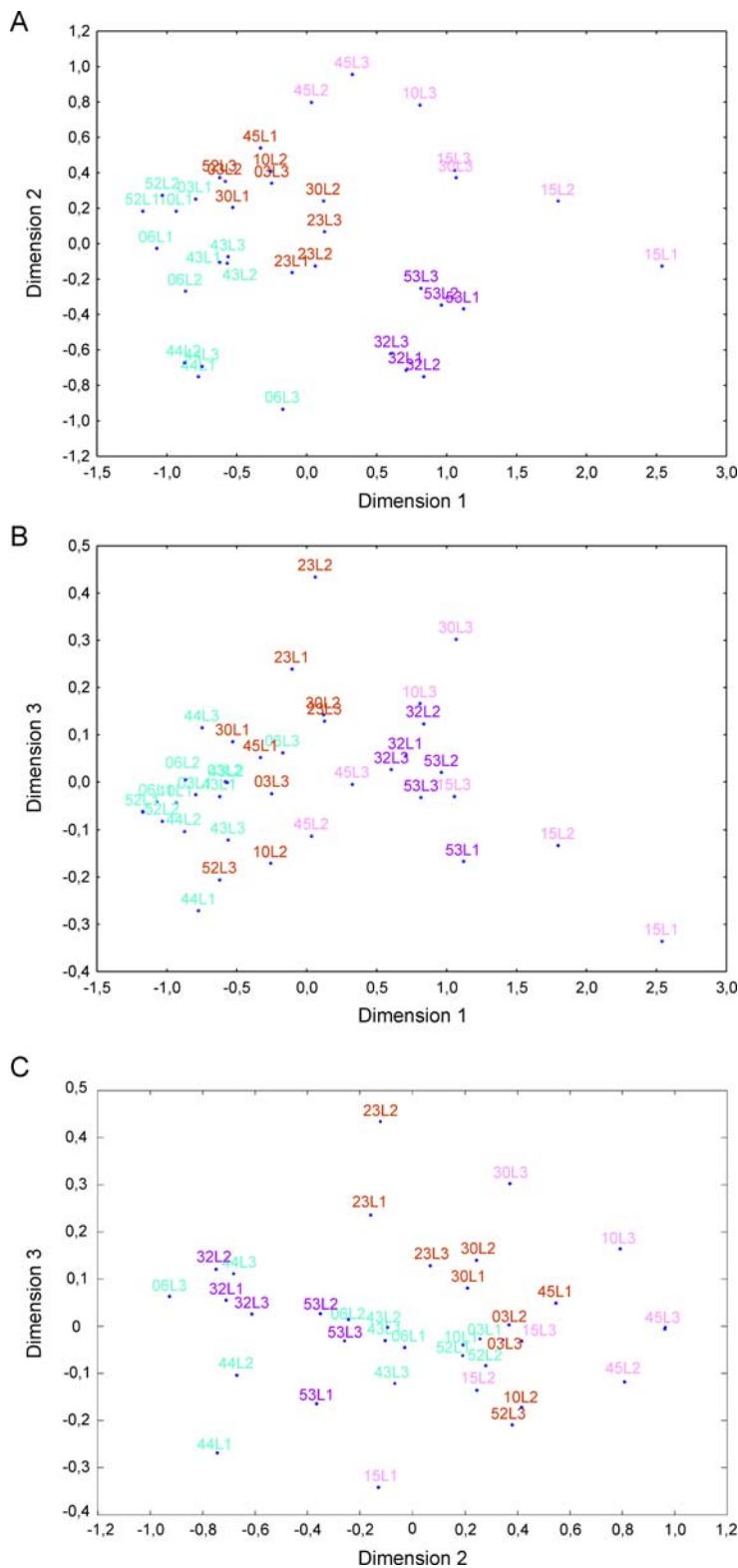


Figure 3-28: Analyse multidimensionnelle (MDS) des énergies de convolution de la pyramide des filtres GABOR pour les 36 stimuli originaux sous la forme de 3 représentations bidimensionnelles (A : Dim 1 vs Dim 2, B : Dim 1 vs Dim 3 et C : Dim 2 vs Dim 3). Chaque couleur respecte le cluster auquel appartient l'image.

7.3. Comparaison entre les réponses de la population TS et l'énergie de la pyramide de filtres GABOR

Afin de comparer visuellement les clusters identifiés suite à la méthode de Ward sur la population TS et les MDS sur les réponses neuronales et des filtres, nous nous référons aux couleurs utilisées pour caractériser les clusters dans l'analyse hiérarchisée sur la population TS (cf. Figure 3-14). La Figure 3-29 reprend les représentations 2D (dim 1 vs. dim 2) des MDS sur les réponses de la population TS (A) et sur les énergies de convolution des filtres GABOR (B). Les couleurs laissent apparaître une similitude entre les réponses de la population TS et les filtres GABOR. Afin de tester quantitativement cette similitude, nous avons effectué un test de rotation de PROCRUSTES² sur les matrices de distances euclidiennes. Ce test de permutation aléatoire permet de vérifier si la somme des déviations des résidus est inférieure à celle qui serait obtenue avec une matrice aléatoire.

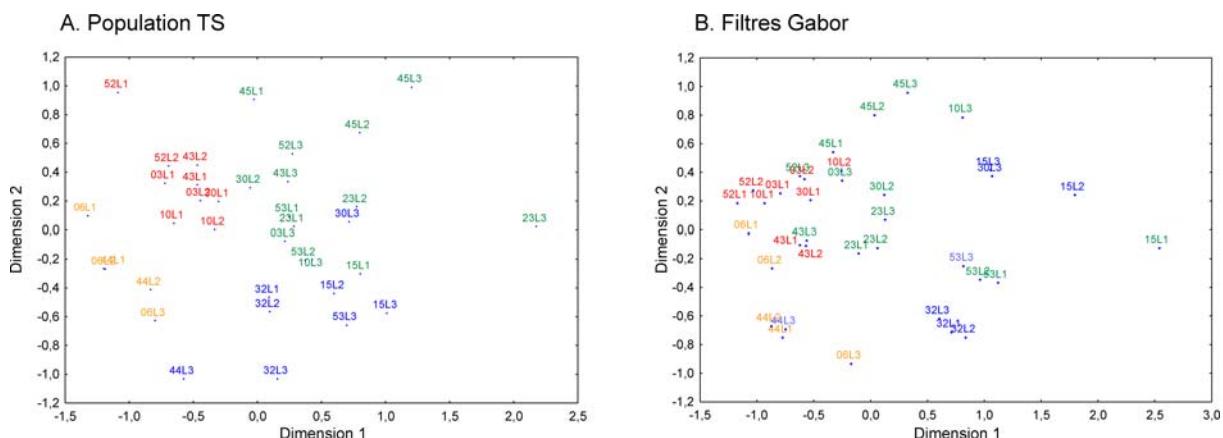


Figure 3-29 : Représentations 2D (Dim 1 vs Dim 2) pour les réponses moyennes des cellules TS (A) et pour les filtres GABOR (B). Les couleurs correspondent aux clusters identifiés suite à l'analyse sur les réponses des filtres.

Les résultats montrent une importante concordance entre les réponses des cellules TS et les énergies de convolution des filtres GABOR (p rejection=0,0001). Cette concordance entre les réponses des cellules et les réponses des filtres est confirmée par une seconde analyse statistique, une factorisation non négative (NNF). Cette analyse calcule une corrélation entre les réponses des filtres GABOR et celle des cellules TS dans un espace à n dimensions. La Figure 3-30 montre la représentation graphique de cette corrélation. Les réponses des cellules et des filtres sont fortement corrélées ($p1 = 0,876$ et $p2 = 0,867$). Ceci

² Ce test de signification est un test de permutation aléatoire basé sur une analyse des « Procrustes ». La matrice de données est soumise à plusieurs transformations (rotation, translation, dilatation ...) afin de minimiser la somme des déviations des résidus au carré entre les points pour chacune des observations. Ce test a été présenté par D.A Jackson en 1995 pour une étude parue dans *Ecoscience* sur la position géographique et la composition chimique de lacs.

indique que la population TS se comporte comme un ensemble de filtres GABOR pour le codage des textures naturelles.

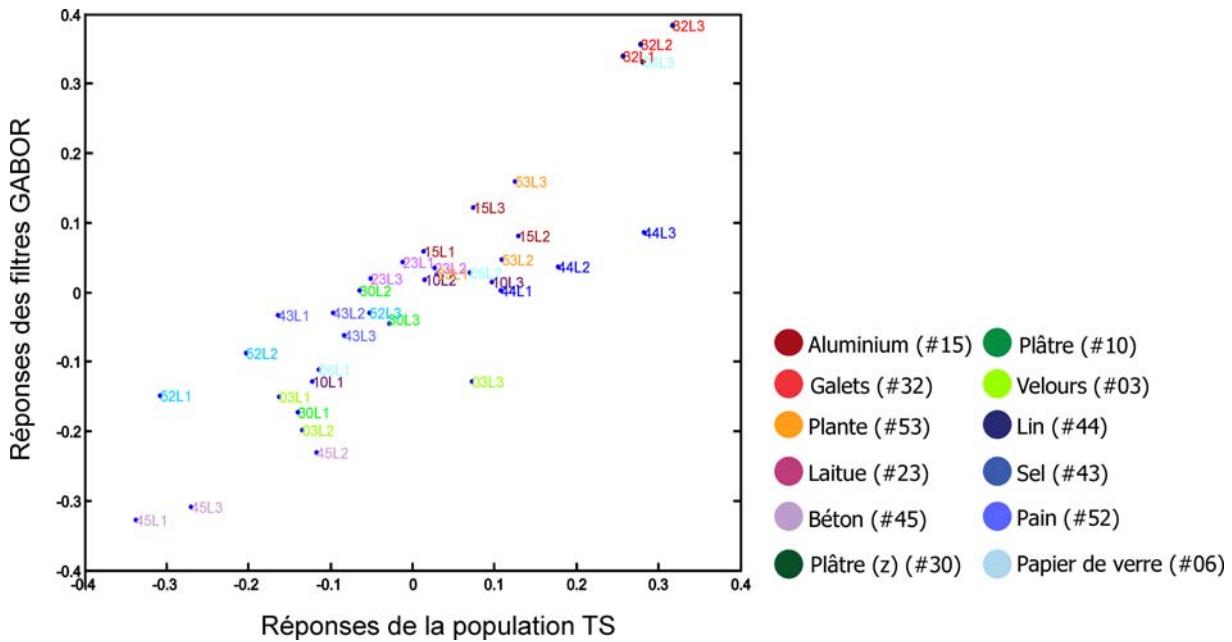


Figure 3-30 : représentation graphique du NNF. Chaque couleur schématise une texture.

7.4. Synthèse

- La pyramide de filtres GABOR classe les textures naturelles en clusters selon leurs SNR et leurs luminances moyennes.
- Les réponses neuronales de la population TS et les sorties des filtres sont fortement corrélées (Test de procrustes et NNF).
- La population TS se comporte comme une pyramide de filtres GABOR pour le codage des textures naturelles.

8. Conclusion générale

Dans le premier protocole, nous avons utilisé 12 textures naturelles illuminées selon trois angles d'incidence. A chacun de ces stimuli naturels correspond un stimulus contrôle dont les phases du spectre de Fourier sont randomisées mais les amplitudes conservées.

Les analyses statistiques ont démontré qu'une large majorité des cellules de V4 ($n=105$) présente des niveaux de réponses variables en fonction de la texture naturelle présentée. Ce sont des cellules « Texture Selective » ou TS. Parmi les textures préférées, nous retrouvons des textures hétérogènes telles que l'*aluminium*, la *feuille de laitue* ou le *béton*.

Enfin les analyses multidimensionnelles signalent que la luminance moyenne et le SNR seraient des paramètres importants pour le codage des textures naturelles au sein de l'aire V4. Toutefois, dans ce premier groupe de stimuli (set A), le SNR des images est fortement corrélé à la luminance moyenne (cf. Figure 3-32). Ceci est dû au choix arbitraire des 12 textures naturelles qui ne représentent pas un panel assez large des textures naturelles du point de vue de ces paramètres. Afin de mieux caractériser l'influence de ces paramètres sur le codage des neurones de V4, nous avons sélectionné de nouvelles textures naturelles parmi la base de données CURET; le set B.

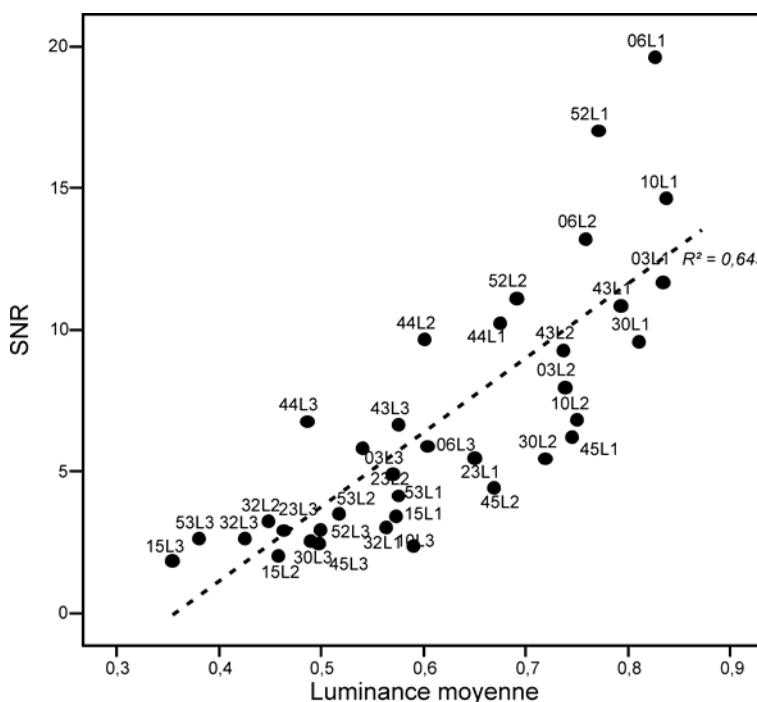


Figure 3-31 : SNR des images du set A en fonction de leurs luminances moyennes. Ces 2 paramètres sont corrélés linéairement ($R^2=0,645$). Chaque point représente un stimulus.

D. ETUDE SYSTEMATIQUE DE L’INFLUENCE DU SNR ET DE LA LUMINANCE MOYENNE - SET B : 65 TEXTURES NATURELLES

1. Population d’étude

Nous avons enregistré 31 cellules chez Tima avec le set B de 65 textures naturelles illuminées uniquement selon un angle L3 (67.5°). Parmi ces 31 cellules, 26 montrent une augmentation significative de leur réponse pendant la période de présentation des stimuli en comparaison avec la période de référence, soit 84 % des cellules (Test-T, $p<0,05$). Ce pool de 26 cellules constitue le groupe de neurones sur lequel est effectué un test statistique de type ANOVA avec comme unique facteur la texture naturelle présentée.

2. Cellules TS

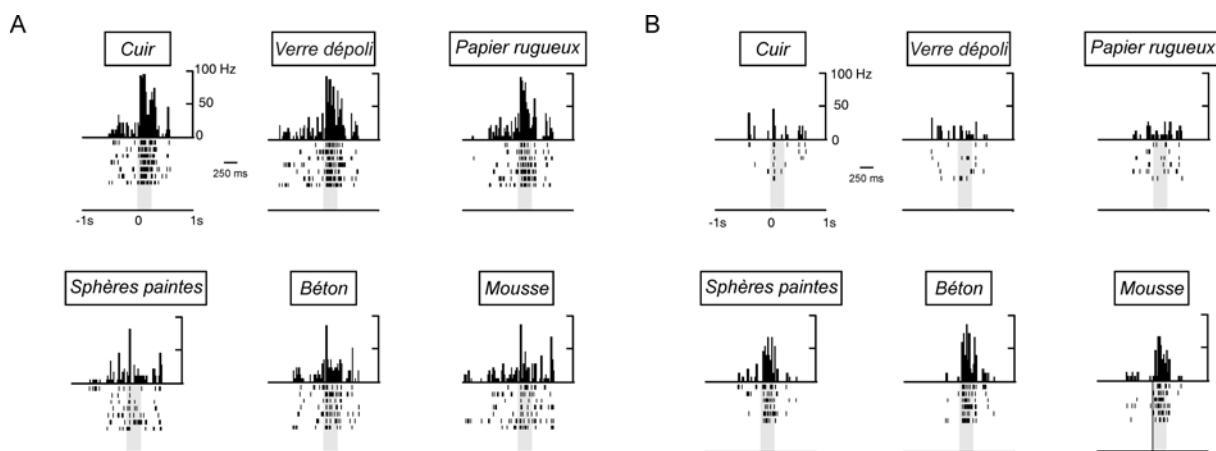


Figure 3-32: Deux exemples de cellules sélectives aux textures naturelles (A : TMP-124.11 ; B, TMP-128.11). Mêmes conventions que la Figure 3-6. Les intervalles des histogrammes ont une largeur de 25 ms.

Parmi ces 26 cellules, 18 neurones ont une décharge neuronale significativement différente selon la texture naturelle présentée (ANOVA, $p<0,05$), soit 69 % des cellules. Ces 18 cellules sont des cellules TS. La Figure 3-32 présente 2 exemples de neurones TS et leur taux de réponses pour 6 textures du set B. Le choix de ces 6 textures n'est pas arbitraire puisqu'elles diffèrent par leur hétérogénéité. En effet, les 3 textures de la première ligne sont des textures homogènes ($SNR > 15$) alors que les trois autres textures sont plutôt hétérogènes ($SNR < 3.5$). La cellule A répond préférentiellement à des textures naturelles homogènes. Son activité neuronale augmente pour le *cuir*, le *verre dépoli* et le *papier rugueux* et répond peu aux 3 autres textures, plutôt hétérogènes. La cellule B présente un pattern de réponse inverse ;

sa décharge neuronale augmente pour les textures hétérogènes (*sphères peintes, béton et mousse*) et reste minime pour les 3 textures homogènes. Ces deux cellules semblent donc fortement corrélées aux SNR des images. La cellule A est corrélée positivement ($R\ Pearson = 0.531$, $p<0.001$). Sa décharge augmente pour les textures homogènes à fort SNR et diminue pour les textures hétérogènes. A l'inverse, la cellule B est corrélée négativement ($R\ Pearson = -0.714$, $p<0.001$); son taux de décharge augmente pour les textures hétérogènes. Au sein de la population TS, 13 cellules sont corrélées significativement au SNR des images comme ces deux exemples de cellules ($p<0.05$).

Nous avons également caractérisé la sélectivité de ces cellules TS pour les textures naturelles en calculant un index de sélectivité I_{max} (même calcul que pour la précédente population de neurones). Ces cellules TS répondent en moyenne à 5 textures naturelles parmi les 65. Parmi les textures préférées, nous retrouvons le *plastique rugueux* (#04, SNR = 1,38), l'*éponge* (#21, SNR = 2,45) et le *béton* (#49, SNR = 1,12) soit des textures hétérogènes.

3. Particularités de ce nouveau set de stimuli

Le set B présente un nombre plus important de textures naturelles ($n=65$) que le set A tout en conservant les 12 textures initialement testées. Ce set est particulièrement intéressant car pour plusieurs textures naturelles, les paramètres SNR et luminance moyenne ne sont pas corrélés linéairement comme nous l'observions pour le set A. La Figure 3-33 présente deux graphes résumant les relations entre ces deux paramètres (A) et entre les coefficients de corrélation calculés entre les réponses moyennes aux stimuli et les valeurs des luminances moyennes ou des SNR (B).

Sur le graphique A, les 12 textures naturelles choisies dans le set A présentent une relation linéaire entre les luminances moyennes et les SNR ($R\ Pearson = 0,676$, $p=0,016$). De ce fait, le set A ne présente pas de variabilité importante au niveau du SNR des textures naturelles ; aucune des 12 textures ne présente un SNR supérieur à 10 alors que les luminances moyennes des images varient. Pour les nouvelles textures du set B ($n=53$), le SNR et la luminance moyenne des images ne sont pas corrélées significativement ($R\ Pearson = 0,130$, $p=0,353$). La relation linéaire entre les 2 paramètres pour les 12 textures originales explique l'importante corrélation entre les coefficients de Pearson [r] calculés pour le SNR et la luminance moyenne, ($R=0.798$, $p<0.001$). Dans le set élargi de textures naturelles, ces deux paramètres ne sont pas corrélés ($R=0.140$, $p>0.05$). Au sein de la population TS, les décharges neuronales de 9 cellules sont corrélées significativement à la luminance moyenne des images. Certaines cellules sont significativement corrélées aux SNR des textures mais pas à la luminance moyenne ($n=7$)

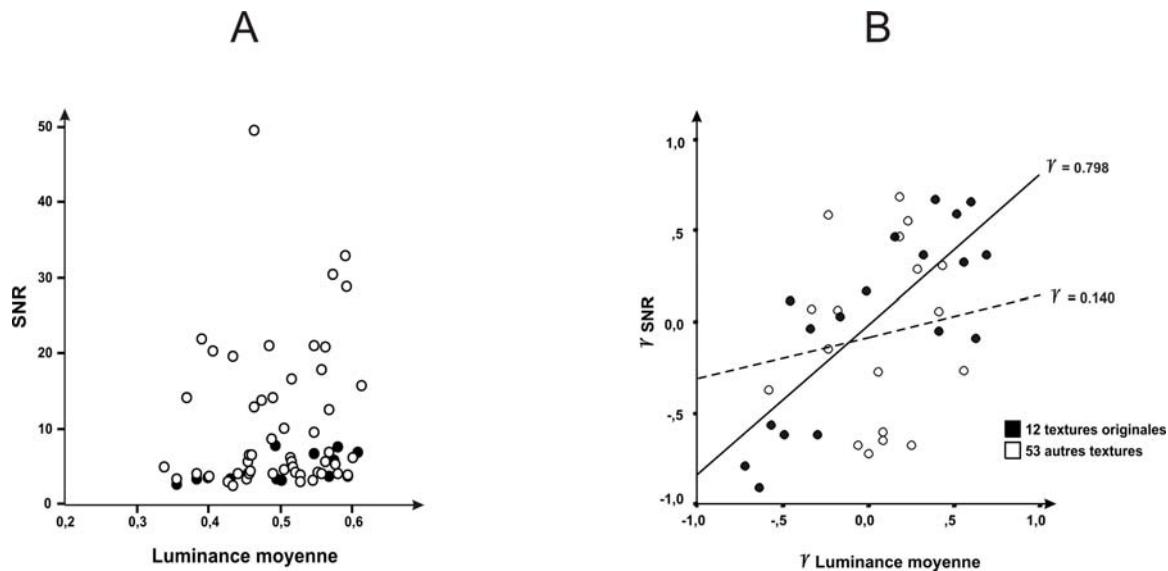


Figure 3-33 : A. Relation entre les paramètres SNR et luminance moyenne des textures naturelles du set B. les points correspondent aux textures naturelles initialement choisies dans le set A (noir) ou complètement nouvelles (blanc). B. Relation entre les coefficients de corrélations de Pearson [r] calculés pour les 18 cellules TS pour ces deux paramètres selon les textures naturelles.

4. Analyses de population

4.1. Analyse hiérarchisée par cluster

Comme nous l'avons effectué pour la population TS enregistrée avec le set A de stimuli, nous effectuons des analyses multidimensionnelles (analyse hiérarchisée par cluster et MDS) sur cette nouvelle population TS ($n=18$ cellules). La Figure 3-35 présente le résultat de l'analyse hiérarchisée par cluster pour la population TS enregistrée avec le set B. Les réponses aux 65 stimuli peuvent se diviser en deux grands groupes A1 et A2. Le cluster A1 comprend 2 sous-ensembles (B1, B2) de taille relativement similaire (13 et 10 stimuli respectivement). Le second groupe A2 est constitué de 2 sous-ensembles (B3 et B4) et comprend 39 stimuli (34 et 8 respectivement).

Afin de décrire les images contenues dans chaque cluster, nous nous sommes intéressées aux valeurs moyennes de la luminance et du SNR de chaque cluster (Figure 3-36). Ces deux paramètres semblaient le mieux décrire les réponses de la population TS pour le premier set A. Une analyse ANOVA indique que la luminance et le SNR des images varient selon les clusters ($p<0,001$). Selon ces critères, le cluster B1 comprend des textures plutôt sombres et très homogènes ($Lum_{moyen} = 0,55$ et $SNR_{moyen} = 16,52$) et le cluster B2 des images plus claires mais également homogènes ($Lum_{moyen} = 0,45$ et $SNR_{moyen} = 19,09$). Le cluster B1 diffère du cluster B2 par sa luminance moyenne (Test TUKEY, $p<0,001$) et non par son SNR moyen (Test TUKEY, $p=0,755$). Le groupe A1 est donc majoritairement constitué de textures

naturelles relativement homogènes. A l'inverse, le cluster A2 comprend des textures plutôt hétérogènes. Le cluster B3 comprend un nombre important d'images (n=34). Les textures y sont plutôt hétérogènes et à faiblement illuminées ($\text{Lum}_{\text{moyen}} = 0,51$ et $\text{SNR}_{\text{moyen}} = 3,74$). Le cluster B4 regroupe 8 images ($\text{Lum}_{\text{moyen}} = 0,43$ et $\text{SNR}_{\text{moyen}} = 4,69$). Celui-ci diffère de B3 au niveau de sa luminance et non de son SNR (Test TUKEY, $p= 0,012$, et $p=979$ respectivement).

En résumé, le cluster A1 comprend des textures homogènes ($\text{SNR}_{\text{moyen}}$ élevé) alors qu'A2 est principalement constitué de textures hétérogènes ($\text{SNR}_{\text{moyen}}$ bas). Chacun de ces clusters comprend deux sous-ensembles séparés selon leur luminance moyenne (B1 et B3, $\text{Lum}_{\text{moyen}}$ élevé soit des images claires et B2 et B4, $\text{Lum}_{\text{moyen}}$ basse soit des images foncées). La Figure 3-37 dispose les images selon leur luminance et SNR, le code couleur regroupe les images en fonction de leur appartenance au cluster.

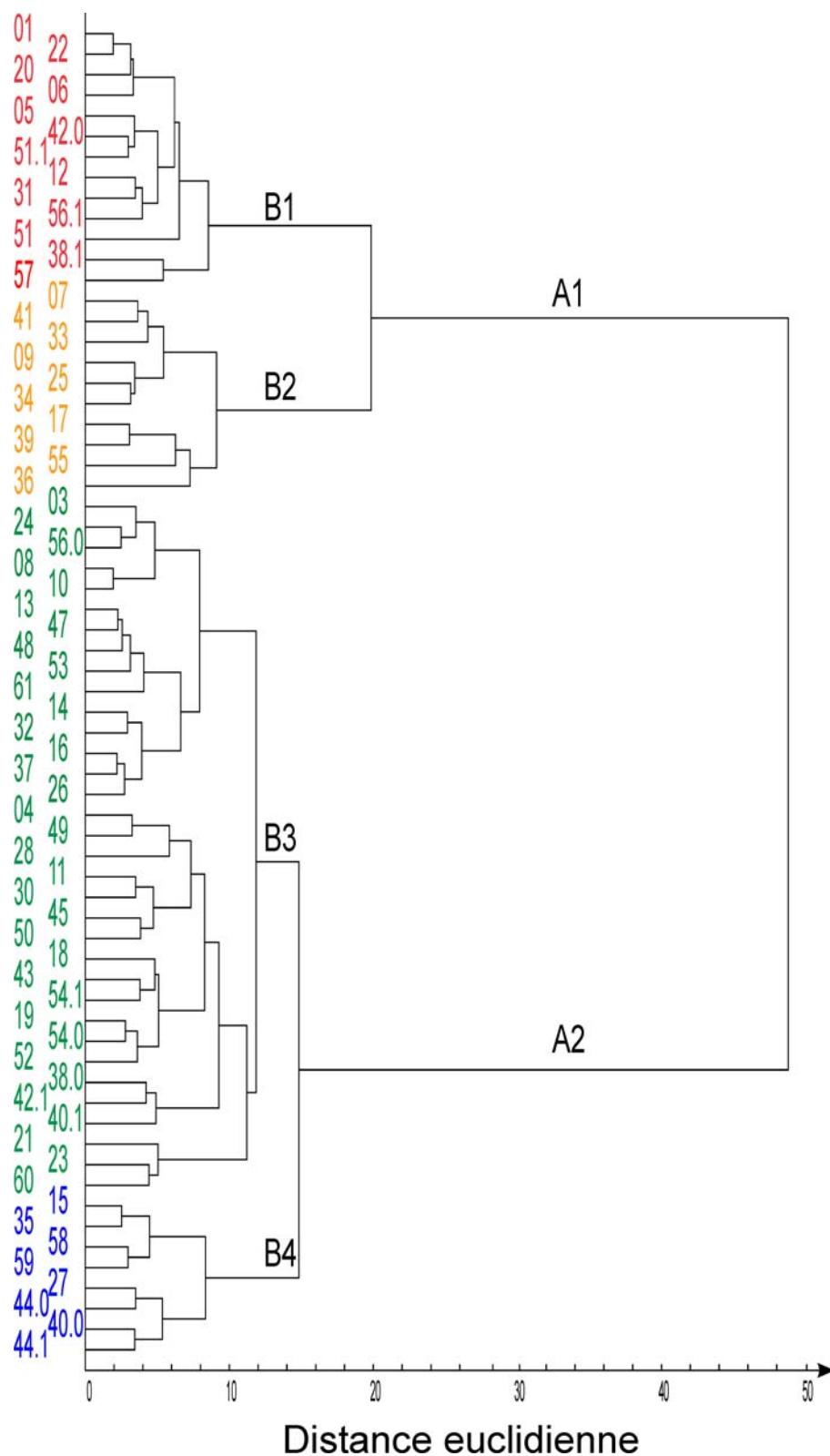


Figure 3-34 : Analyse hiérarchisée par clusters sur les réponses cellulaires de la population TS. Les couleurs correspondent aux 4 clusters identifiés.

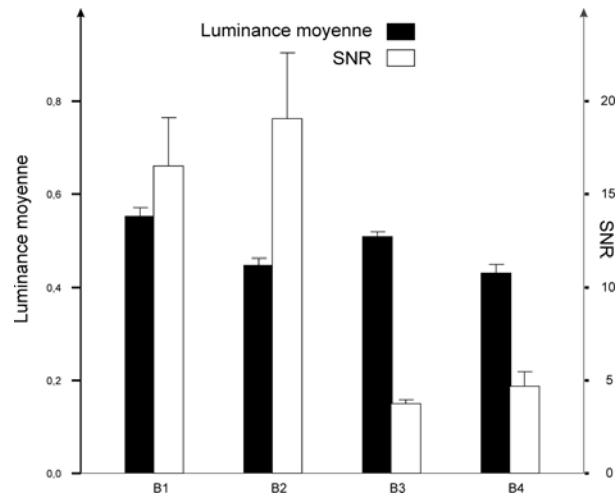


Figure 3-35 : Luminances moyennes et SNR moyens des clusters pour la population TS enregistrée avec le set B.

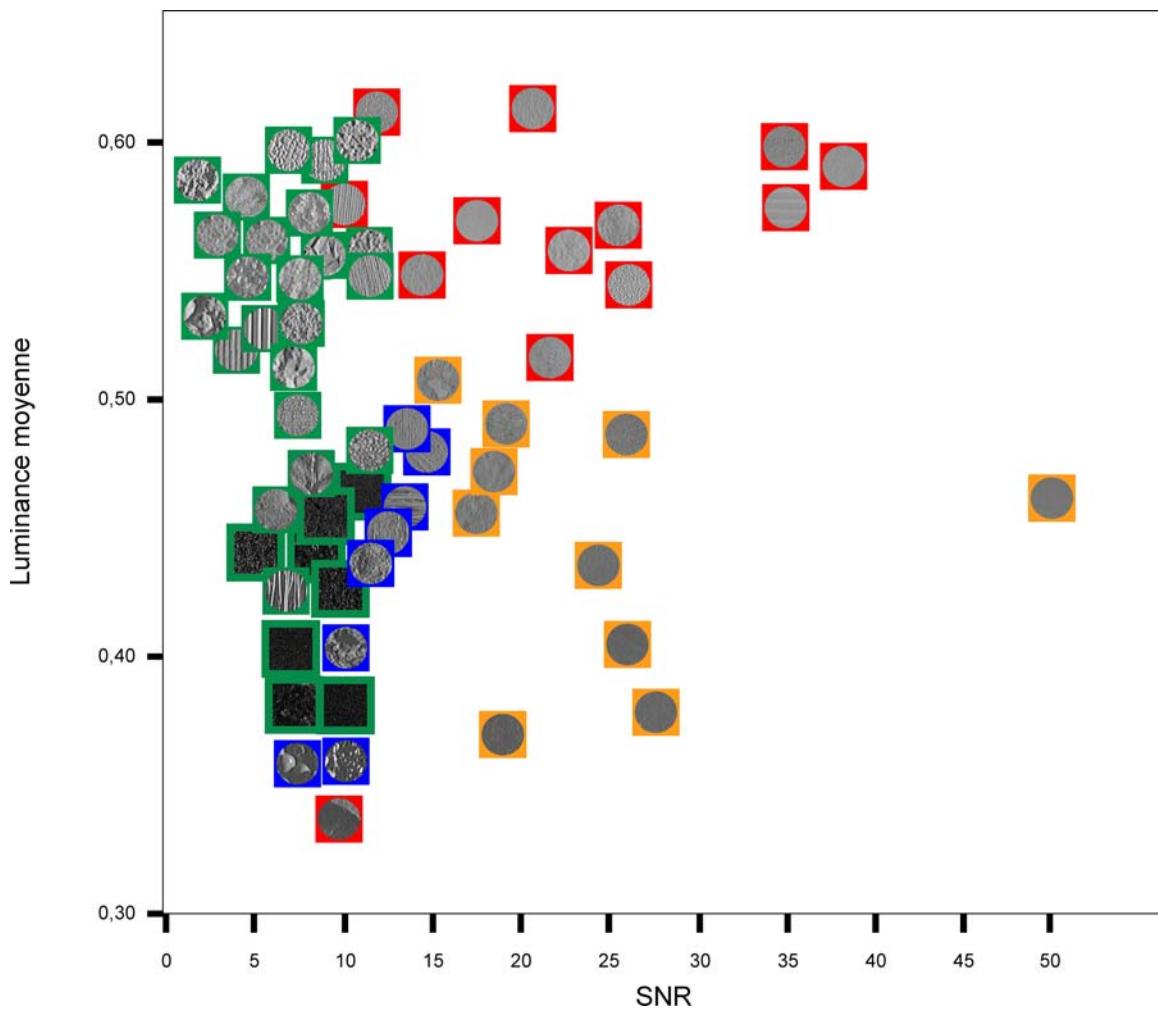


Figure 3-36 : 4 clusters d'images de textures naturelles en fonction du SNR et de la luminance moyenne des images.

4.2. Analyse multidimensionnelle (MDS)

Une analyse multidimensionnelle (MDS) des réponses de la population TS au set B nous permet de caractériser les paramètres importants dans le codage des textures naturelles. La Figure 3-37 représente des représentations 2D de la MDS sur les réponses moyennes des cellules TS. Les valeurs de stress permettent de restreindre l'analyse MDS à 3 dimensions.

La dimension 1 dissocie le cluster A1 du cluster A2. Les réponses aux stimuli compris dans les clusters B1 et B2 présentent des valeurs de la dimension 1 exclusivement positives alors que celles pour les clusters B3 et B4 des valeurs majoritairement négatives. D'après les précédentes analyses, le cluster A1 diffère du cluster A2 par une augmentation significative du SNR (ANOVA, $p<0.001$). Une analyse de corrélation de Spearman nous indique que les valeurs de la dimension 1 sont corrélées significativement avec le SNR des images ($R_{Spearman} = 0.898$, $p<0.001$) et non avec la luminance moyenne des images ($R_{Spearman} = 0.152$, $p= 0.228$). Au contraire, la dimension 2 est corrélée significativement avec les luminances moyennes des images ($R_{Spearman} = -0.760$, $p<0.001$) et pas avec le SNR des images ($R_{Spearman} = -0.004$, $p=0.975$). Les clusters B2 et B4 présentent des luminances moyennes assez basses et ont des valeurs positives dans la dimension 2. Pour les clusters B1 et B3, l'inverse se produit. Quant à la dimension 3, nous n'avons pas trouvé de paramètre permettant de la caractériser ; aucune corrélation ne se révélant significative pour cette dimension.

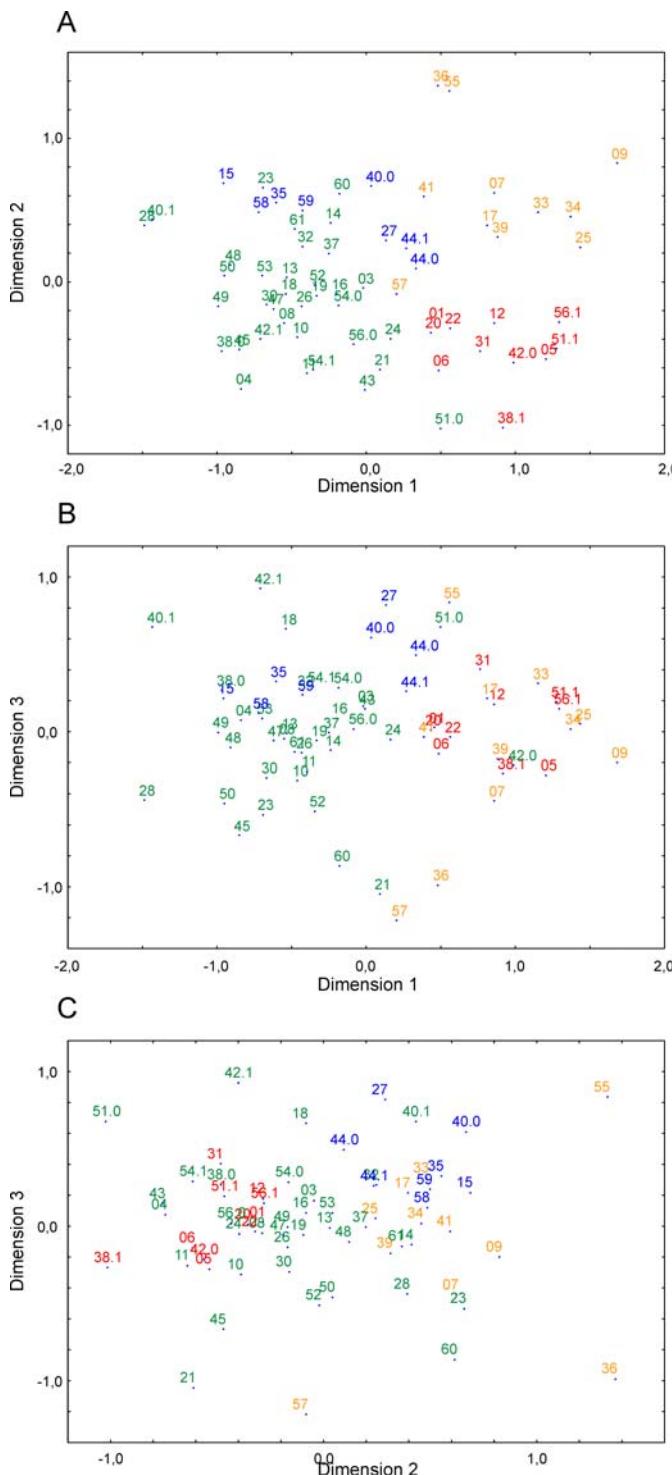


Figure 3-37 : Représentations 2D de la MDS sur les réponses de la population TS. Chaque point correspond à la réponse moyenne de la population TS à un stimulus (représenté par le numéro de la texture). La couleur de point correspond au cluster auquel il appartient d'après l'analyse hiérarchisée.

Pour résumer, nous avons réussi à caractériser 2 dimensions permettant d'expliquer la réponse de notre population TS au set B des textures naturelles, le SNR et la luminance moyenne.

5. Comparaisons entre les réponses de la population TS et les énergies de convolution des filtres Gabor

5.1. Test de Procrustes

Comme nous l'avons effectué sur le set A, nous comparons les réponses moyennes de la population TS et les énergies de convolution des filtres Gabor pour le set B. L'ensemble de filtres utilisés pour ces analyses est identique à celui décrit dans la première partie des résultats. Les couleurs employées respectent les clusters identifiés suite à l'analyse hiérarchisée (cf. Figure 3-35). La Figure 3-38 montre 2 représentations à 2 dimensions des MDS; la première (A) reprend la représentation des réponses moyennes des cellules TS aux 65 textures naturelles (cf. Figure 3-37A), la seconde (B), les réponses des filtres aux 65 textures naturelles.

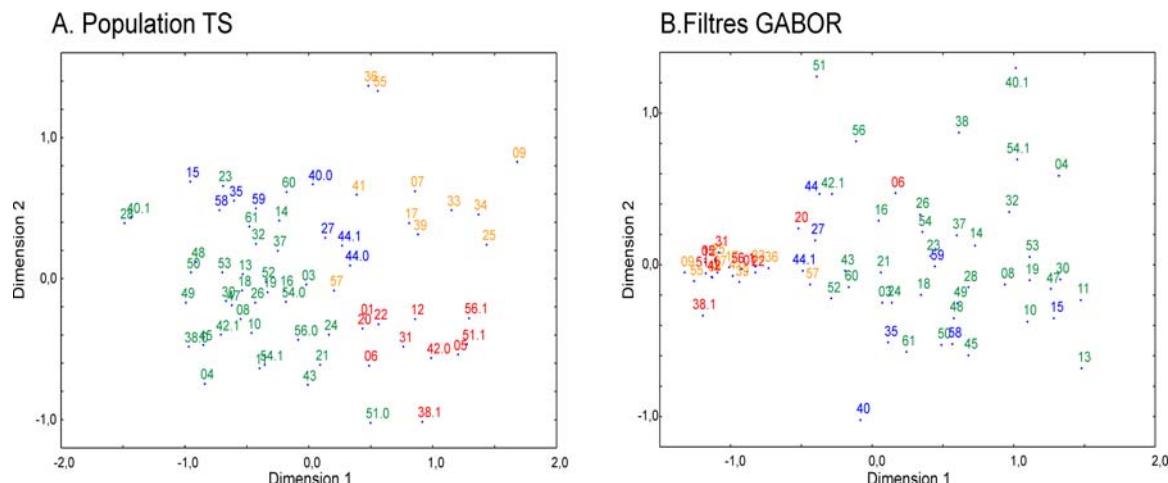


Figure 3-38 : Représentations 2D (dim 1 vs dim 2) MDS des réponses de la population TS (A) et des filtres (B) au set B. Les couleurs sont les mêmes que celles utilisées dans l'analyse hiérarchisée par cluster de la Figure 3-35.

Le test statistique (PROTEST) nous indique que les matrices de distances euclidiennes de la population TS et des filtres Gabor sont corrélées (p rejection=0,001). Les réponses des cellules TS à ce set de stimuli sont donc comparables aux réponses d'un ensemble de filtres Gabor. Ce résultat confirme notre première analyse effectuée sur le set A.

5.2. Factorisation non négative (NNF)

La Figure 3-39 confirme la forte corrélation observée entre la réponse de la population TS et les réponses des filtres. Cette représentation est une factorisation non négative (NNF) où les réponses aux filtres GABOR sont disposées en fonction des réponses moyennes des cellules TS. Cette représentation statistique nous fournit 2 coefficients ($p1= 0,837$ et $p2=0,854$) indiquant qu'il existe une corrélation importante entre les réponses des cellules TS et les réponses des filtres.

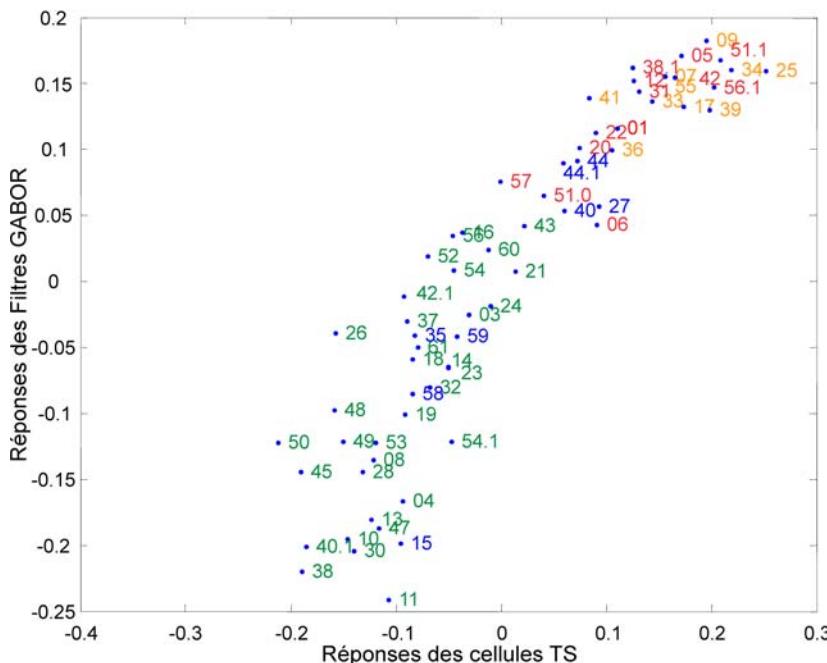


Figure 3-39 : Représentation graphique de la Factorisation non négative entre les réponses des filtres GABOR et les réponses moyennes des cellules TS. Le code couleur correspond aux clusters.

6. Synthèse des analyses sur le set B

Ce nouveau set de stimuli a la particularité de contenir davantage d'images de textures toutes illuminées selon un même angle d'incidence (L3). Ce set d'images est assez varié pour que la luminance et les paramètres d'homogénéité (SNR, variance) ne soient pas corrélés. Suite aux analyses multidimensionnelles (Analyse hiérarchisée par cluster et MDS), nous avons démontré que la population TS se base sur ces paramètres pour différentier les textures naturelles entre elles. Elle se comporte comme une pyramide de filtres GABOR pour regrouper les textures naturelles en clusters. Ces résultats confirment donc les observations effectuées sur le set A avec un set d'images plus élargies en textures naturelles. Afin de valider les résultats chez un second animal (Zrem), nous avons utilisé un set étendu rassemblant les caractéristiques des sets A et B (large panel de textures naturelles illuminées avec 2 angles différents avec leurs images 'contrôle').

E. ETUDE SYSTEMATIQUE DE LA SELECTIVITE AUX TEXTURES NATURELLES : SET C : 41 TEXTURES NATURELLES, 2 ANGLES D'INCIDENCE, 2 STIMULI CONTROLES

1. Population d'étude

Nous avons enregistré 51 cellules avec le set C d'images de textures naturelles (4 chez Tima et 47 chez Zrem). Parmi ces 51 cellules, 36 ont été testées avec les stimuli contrôles *scrambled*. Comme définis dans le set A, ces stimuli contrôles correspondent à des redistributions aléatoires des phases des images de textures naturelles illuminées par un angle de 22.5°. Les 15 autres cellules ont été testées avec des réseaux orientés (voir paragraphe E.5 pour les détails).

Les résultats du test statistique ANOVA indiquent que 97 % (n= 49) des cellules ont une décharge neuronale différente selon les textures naturelles présentées (effet principal, $p < 0,05$). Ce sont des cellules TS. Pour cette sous population de neurones TS enregistrée avec 41 textures naturelles différentes, nous avons calculé de la même manière que pour le set A, un critère de sélectivité I_{max} afin de quantifier leur sélectivité aux textures naturelles.

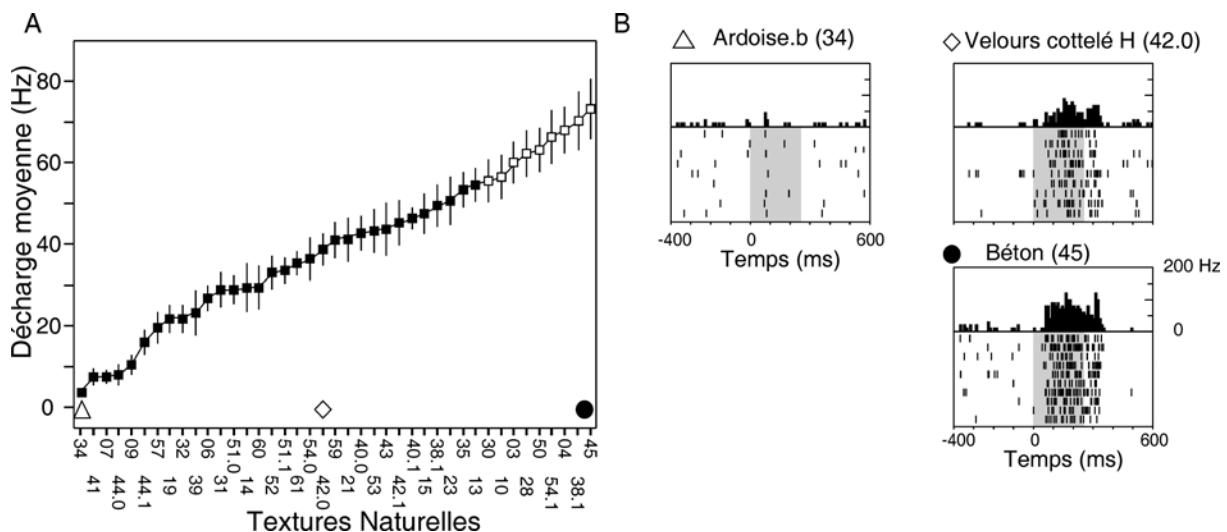


Figure 3-40: Exemple de cellule TS. A : Courbe obtenue à partir des réponses moyennes pour chaque texture (les réponses aux angles d'incidence L1 et L2 sont cumulées). Les réponses moyennes sont ordonnées selon un ordre croissant. Les symboles sur l'axe des abscisses correspondent aux réponses du neurones représentées en B. Les carrés blancs indiquent les réponses moyennes au-dessus de I_{max} , les barres d'erreurs, l'esm. B : Chaque PSTH et raster représentent les réponses moyennes aux textures Ardoise_b, Velours cotelé_H et Béton. Les réponses aux deux illuminations sont cumulées. La taille des intervalles est de 10ms.

La Figure 3-40 montre un exemple de cellules TS enregistrée avec le set C. D'après le critère I_{max} , cette cellule TS répond sélectivement à 9 textures. En moyenne, les neurones TS

répondent à 6 textures naturelles parmi les 41 présentées. La majorité des cellules TS montrent une forte sélectivité aux textures naturelles. En effet, 26 cellules TS répondent à moins de 3 textures naturelles, soit 53 % de la population TS. Pour les autres cellules TS, 12 d'entre elles répondent entre 4 et 10 textures naturelles et les 11 cellules restantes répondent à plus de 10 textures naturelles. Ces résultats sont résumés dans la Figure 3-41. Cette figure montre également que 12 cellules TS sont très sélectives aux textures naturelles ; elles ne répondent qu'à une texture parmi les 41 présentées. A l'inverse, une cellule est peu sélective aux textures naturelles et répond à 31 textures.

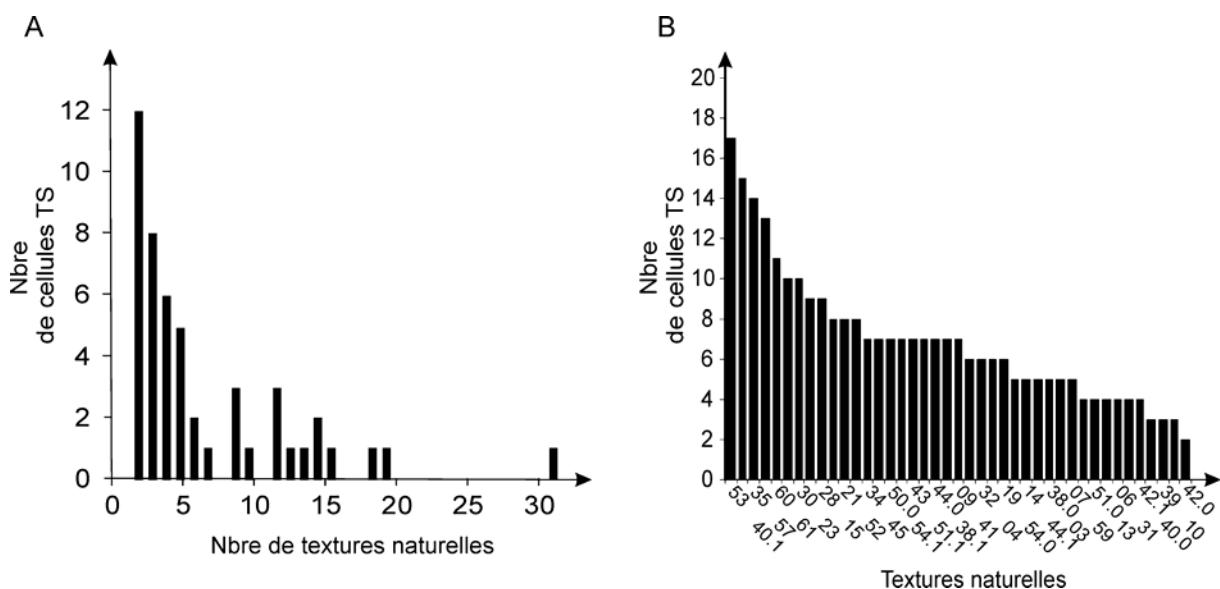


Figure 3-41: A : Nombre de textures naturelles codées par la population TS d'après le critère I_{max} . 12 cellules sont très sélectives et ne répondent qu'à une seule texture d'après I_{max} . B : Nombre de cellules TS en fonction de la texture naturelle présentée. Les textures naturelles sont classées selon un ordre décroissant : de la texture préférée par la population TS à la texture la moins codée par cette population.

La Figure 3-41B présente les textures les plus fréquemment codées au niveau de la population TS. Parmi les textures préférées, nous retrouvons la *plante Soleirolia* (#53), la *paille V* (#41.1), les *sphères peintes* (#35) et la *queue de paon* (#57) codées respectivement par 17, 15, 14 et 13 cellules TS.

2. Analyses de population

2.1. Analyse hiérarchisée par clusters

Comme nous l'avons effectué pour les populations TS testées avec les sets A et B d'images, l'étape suivante d'analyse est une analyse hiérarchisée par clusters. Les réponses moyennes pour les 82 stimuli originaux sont standardisées pour chaque cellule. La méthode de Ward nous permet d'établir un arbre (Figure 3-42).

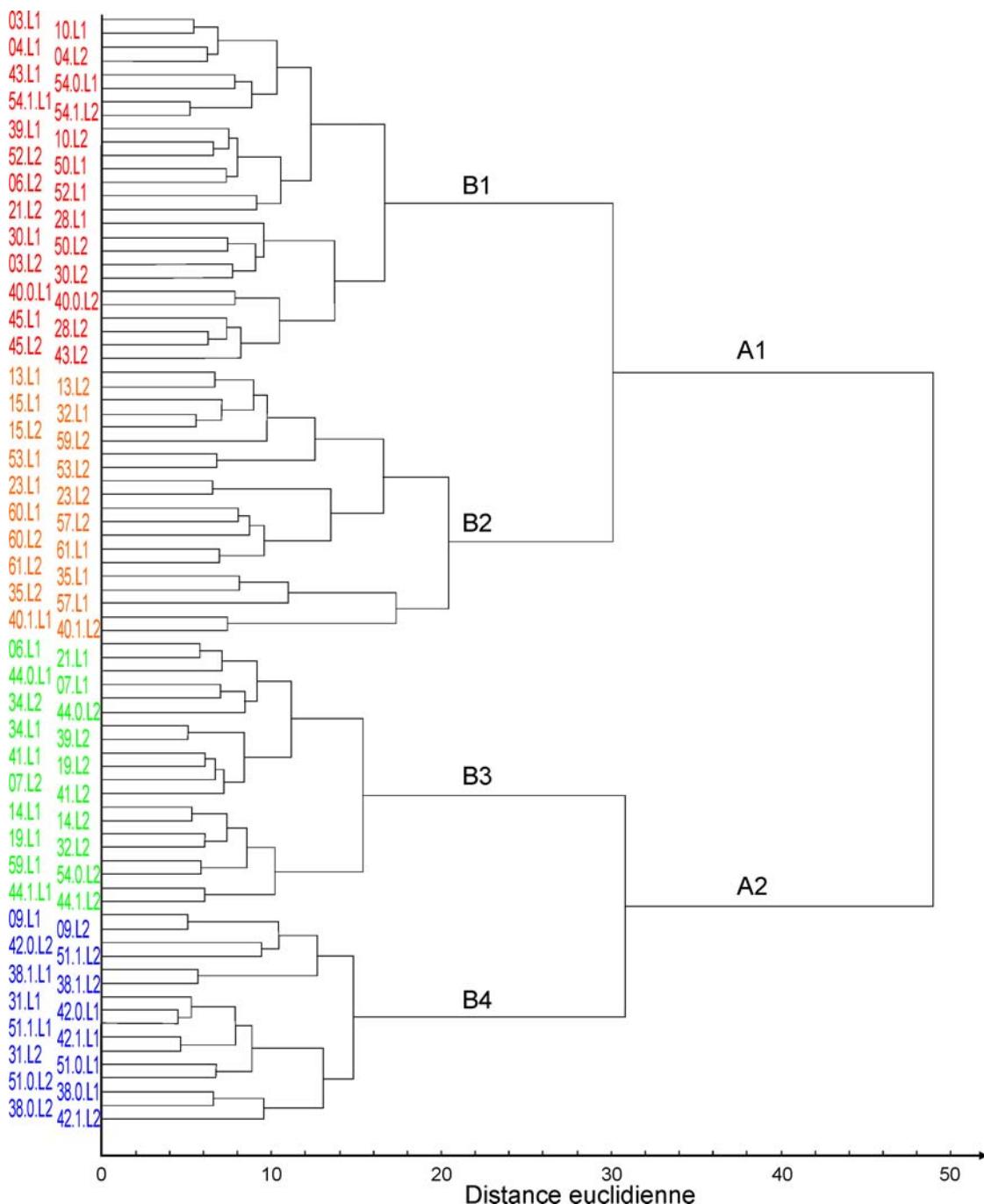


Figure 3-42 : Analyse hiérarchisée par cluster. Chaque branche représente un des 82 stimuli originaux. Les couleurs permettent de distinguer 4 clusters : B1 en rouge, B2 en orange, B3 en vert et B4 en bleu.

Premièrement, l'arbre se dissocie en deux clusters distincts (A1 et A2). La branche A1 contient 2 clusters (B1 et B2) regroupant respectivement 26 et 20 stimuli. La branche A2 contient 2 autres clusters ; B3 constitué de 2N stimuli et B4, de 16 stimuli. Les textures appartenant aux clusters A1 et A2 sont respectivement hétérogènes et homogènes car elles diffèrent par leur variance de luminance ($Var_{moyen} = 0,027$ et $0,008$ respectivement, ANOVA, $p < 0,001$). Le niveau B de dissociation sépare des textures selon leur luminance moyenne. Les textures dans B1 et B4 ont des valeurs de luminances moyennes élevées ($Lum_{moyen} = 0,49$ et

0,48 respectivement, ANOVA, $p=0,986$) tandis que les textures contenues dans B2 et B3 ont des valeurs de luminances moyennes assez basses (Lum_{moyen} = 0,23 et 0,28 respectivement, ANOVA, $p=0,986$). De plus, les textures dans B1 et B4 diffèrent significativement de celles contenues dans B2 et B3 par leur luminance moyenne ($p<0,001$). Par ailleurs, les textures de B4 diffèrent de B1, B2 et B3 par des valeurs du SNR élevées (SNR_{moyen} = 3.87, 1.66, 4.49 et 12.05 respectivement, $p<0.001$). Les textures de B4 sont donc particulièrement homogènes (cf. Figure 3-43). Nous n'avons pas observé de différences significative entre les clusters avec les paramètres de 2nd ordre ($p>0,05$). Enfin, il existe une tendance à ce que les mêmes textures illuminées différemment soient appariées au sein d'un même cluster puisque 35 paires (L1 & L2) parmi les 41 se retrouvent appariées dans un même cluster B. Les pourcentages de textures appariées sont de 85 % dans B1, 90% dans B2, 70% dans B3 et 100% dans B4. La Figure 3-43B reprend les valeurs moyennes des SNR et luminances selon le cluster.

La Figure 3-43A illustre les 4 clusters identifiés suite à cette analyse hiérarchisée en regroupant avec le code couleur les images appartenant à un même cluster.

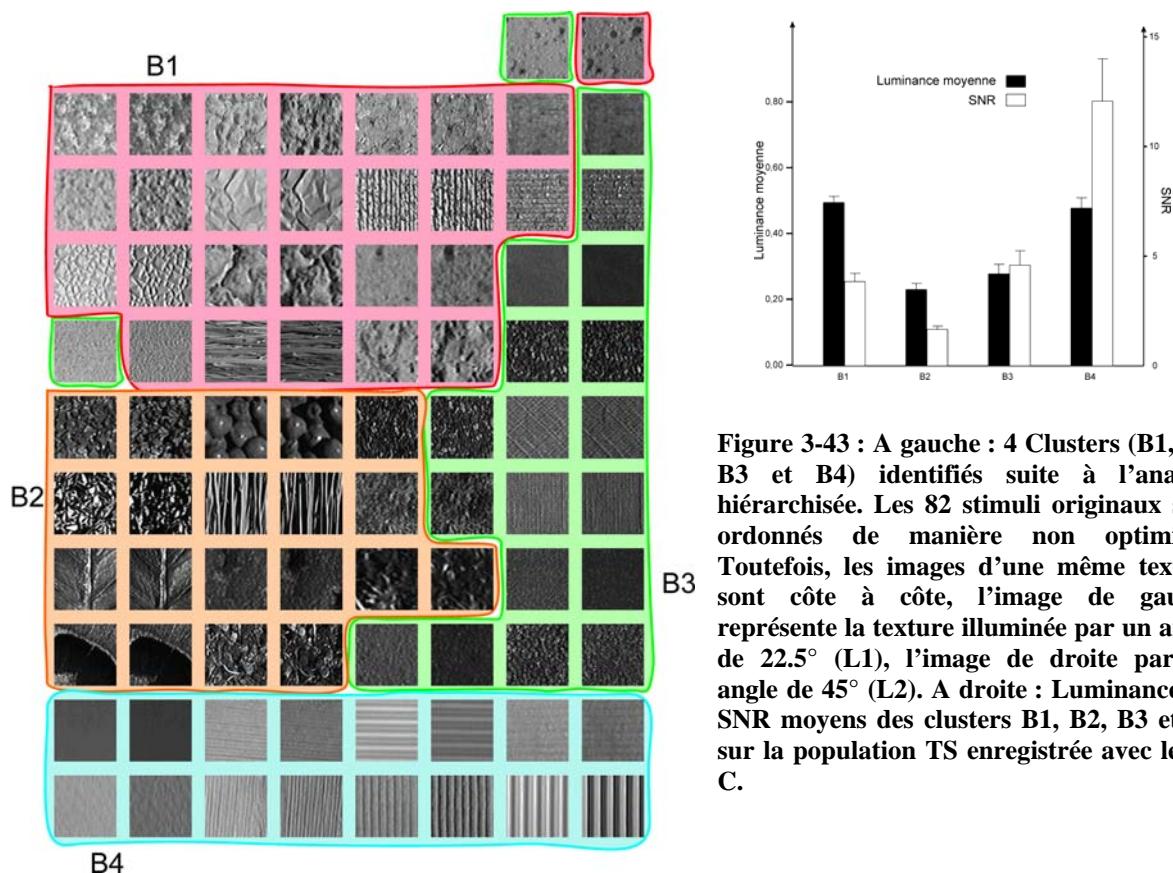


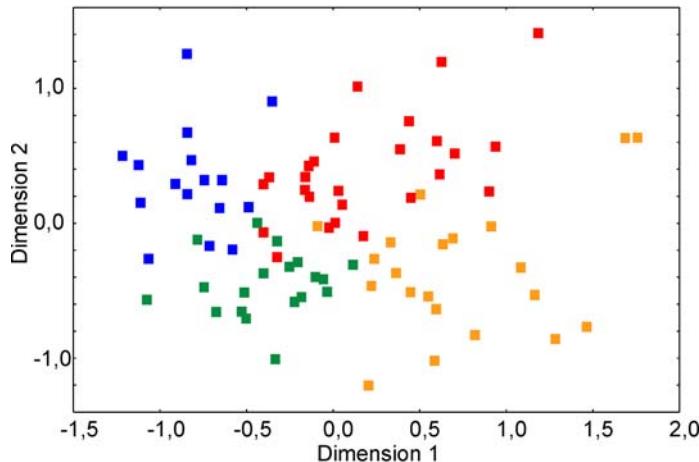
Figure 3-43 : A gauche : 4 Clusters (B1, B2, B3 et B4) identifiés suite à l'analyse hiérarchisée. Les 82 stimuli originaux sont ordonnés de manière non optimisée. Toutefois, les images d'une même texture sont côte à côte, l'image de gauche représente la texture illuminée par un angle de 22.5° (L1), l'image de droite par un angle de 45° (L2). A droite : Luminances et SNR moyens des clusters B1, B2, B3 et B4 sur la population TS enregistrée avec le set C.

2.2. Analyse multidimensionnelle de type MDS

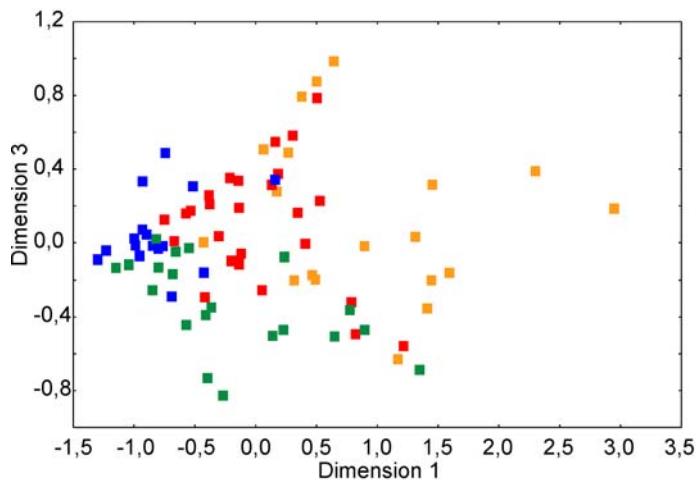
Afin de définir combien de dimensions permettent de représenter les réponses de la population TS, nous avons effectué une analyse multidimensionnelle de type MDS basé sur

les distances euclidiennes. Le calcul de valeurs de stress permet de restreindre cette analyse à 3 dimensions (Kruskal et Wish, 1978). Le résultat de cette analyse est présenté dans la Figure 3-44.

A. Dimension 1 vs. Dimension 2



B. Dimension 1 vs. dimension 3



C. Dimension 2 vs. dimension 3

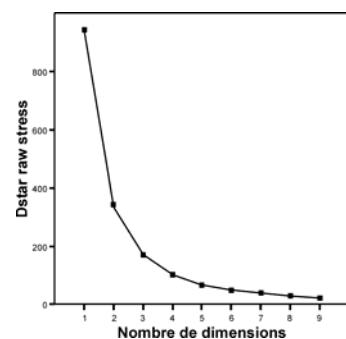
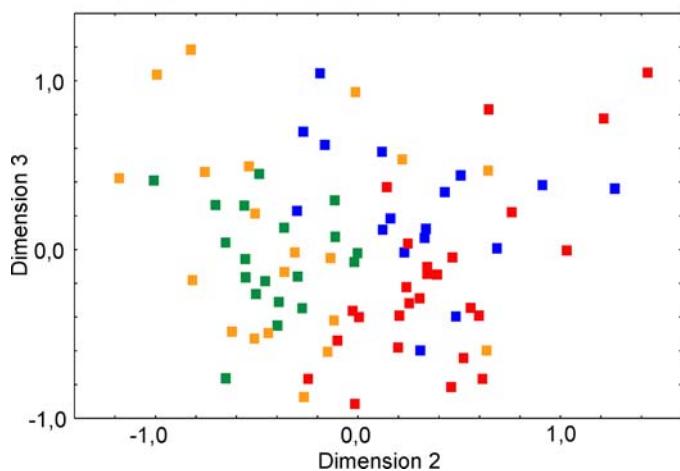


Figure 3-44 : Représentations bidimensionnelles de la MDS effectuée selon les réponses de la population TS. Le code couleur correspond aux clusters identifiés dans la figure précédente. Chaque point correspond à un stimulus. En haut à droite : valeur de stress (D_{star} raw stress) en fonction du nombre de dimensions. Un coude apparaît à partir de 3 dimensions. Pour l'analyse de type MDS, nous nous restreignons donc à 3 dimensions.

L'impression visuelle (cf. Figure 3-46) indiquant que les textures se regroupent selon leur hétérogénéité et leur luminance est confirmée par le calcul de coefficient de corrélation ($R_{Spearman}$). La dimension 1 de la MDS est corrélée significativement avec le SNR et la variance des images ($R_{Spearman} = 0.738$, $p < 0.001$ et $R_{Spearman} = 0.639$, $p < 0.001$ respectivement. La dimension 2 est corrélée significativement avec la luminance moyenne ($R_{Spearman} = 0.670$, $p < 0.001$). Contrairement aux dimensions 1 et 2 corrélées à des paramètres de 1^{er} ordre, la dimension 3 est uniquement corrélée avec des paramètres de 2nd ordre moyennés sur les différentes matrices de cooccurrence (Energie : $R = 0.527$, $p < 0.001$; Entropie : $R = 0.498$, $p < 0.001$; Homogénéité : $R = 0.486$, $p < 0.001$ et Inertie : $R = 0.301$, $p = 0.006$). Les couleurs des clusters de l'analyse hiérarchisée (cf. Figure 3-42) permettent de mieux comparer visuellement les 2 analyses multidimensionnelles. La Figure 3-44A nous montre que les 4 principaux clusters (B1, B2, B3 et B4) identifiés par l'analyse hiérarchisée sont dissociés dans cette représentation 2D.

Pour résumer, nous retrouvons les caractéristiques des clusters que j'ai détaillées dans la partie précédente. Pour la dimension 1, les images de textures hétérogènes (SNR bas) se retrouvent dans la partie droite du graphique (clusters B1 et B2) alors que les textures plutôt homogènes sont situées dans la partie gauche du graphique (clusters B3 et B4). Pour la dimension 2, les images à forte luminance moyenne se retrouvent majoritairement dans la partie supérieure du graphique (clusters B1 et B4), les images à faible luminance moyenne dans la partie inférieure. La Figure 3-46 reprend la représentation bidimensionnelle de la Figure 3-44A en y disposant les images des textures. Les textures sombres se retrouvent le plus souvent la partie inférieure du graphique, les textures claires dans la partie supérieure. Les textures homogènes sont disposées dans la partie gauche du graphique et les textures hétérogènes dans la partie droite.

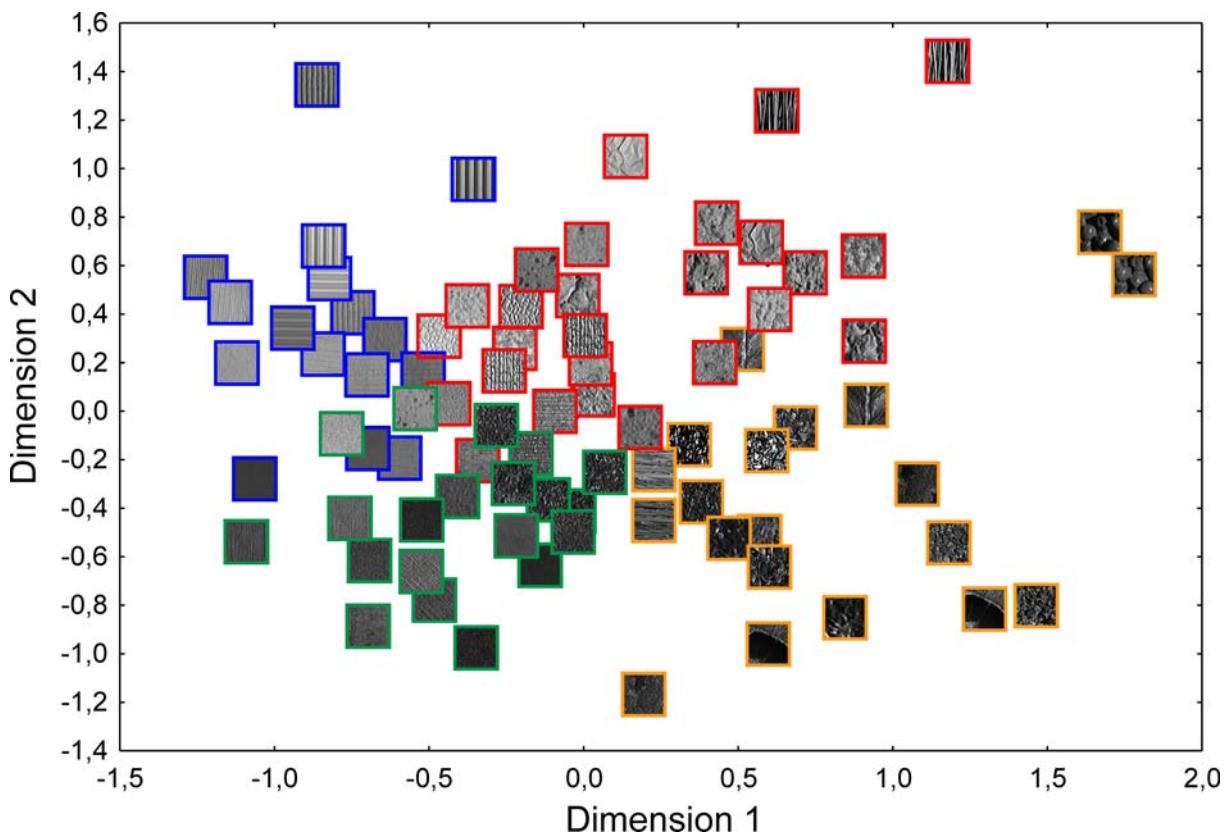


Figure 3-46 : MDS (dimension 2 en fonction de la dimension 1). Les couleurs qui encadrent les images respectent les clusters définis grâce à l’analyse hiérarchisée.

3. Comparaison entre les réponses de la population TS et les sorties des filtres Gabor

Pour comparer les réponses de la population TS aux énergies de convolution des filtres Gabor, la Figure 3-46B reprend les dimensions 1 et 3 de la MDS ‘spectrale’ obtenue avec les réponses des filtres et la Figure 3-46A la MDS ‘neuronale’ avec les mêmes dimensions 1 et 3. Cette MDS permet une meilleure comparaison visuelle entre les réponses des neurones TS et celles des filtres. Le test de rotation PROCRUSTES indique qu’il existe une corrélation significative entre les réponses neuronales et spectrales ($m^2=0.68$, $p<0,0001$).

Les réponses de la population TS sont donc comparables à celles d’une pyramide de filtres pour une classification de textures. Pour quantifier cette similarité, nous recherchons pour chaque image de texture quelle est l’image qui entraîne une réponse similaire de la population TS ou de la pyramide des filtres. Concrètement, nous recherchons, dans chaque représentation 3D, les points les plus proches deux à deux. Ainsi, dans la MDS neuronale, la plus petite distance entre des images d’une même texture mais illuminée sous L1 et L2 est observée pour 29 images (parmi 82). Ceci est confirmé dans la MDS de la Figure 3-46 où les images d’une même texture tendent à être relativement proches. Cet effet est plus prononcé

dans la MDS spectrale car 51 des images sont voisines de leur homologue qu'elle soit illuminée sous L1 ou L2. Les filtres sont donc plus efficaces que les cellules TS pour identifier deux images d'une même texture mais illuminées sous différents angles. Ainsi, les filtres regroupent 26 des 29 paires identifiées par les cellules Ces 26 paires sont à la fois hétérogènes ($n = 13$, cluster A1) ou homogènes ($n = 13$, cluster A2).

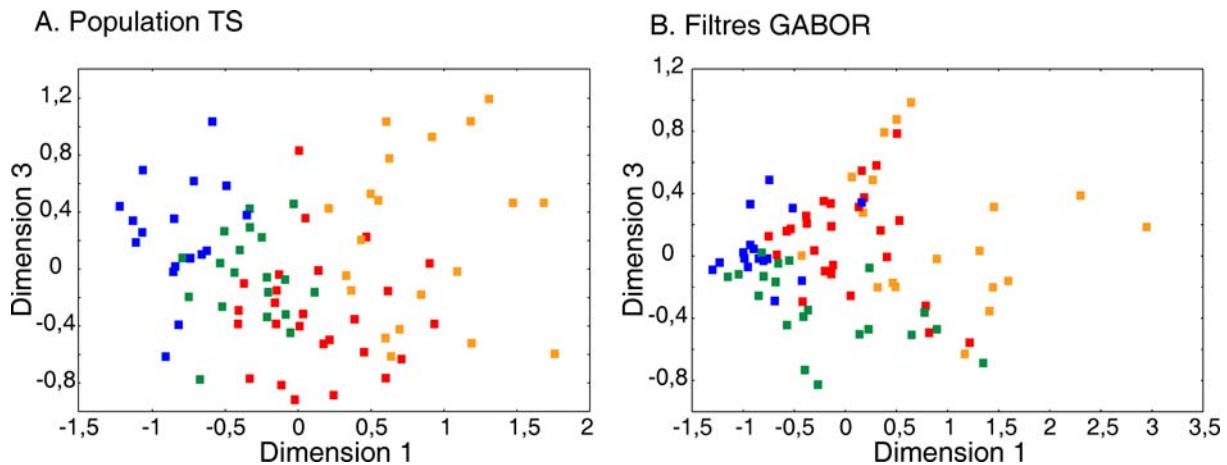


Figure 3-46 : Représentations bidimensionnelles des réponses de la population TS (A) et de la pyramide de filtres GABOR (B). Le code couleur correspond aux clusters identifiés dans l'analyse hiérarchisée effectuée sur les réponses de la population TS.

4. Influence de la phase

4.1. Résultats du test ANOVA

Parmi les 51 cellules testées avec le set C, 36 ont été enregistrées avec les stimuli contrôle de type *scrambled*. Les cellules restantes ont été testées avec des réseaux orientés (voir paragraphe E.5). Parmi ces 36 cellules, 35 cellules sont des cellules TS. Au sein de cette population TS, 16 cellules montrent un effet principal significatif de la phase des stimuli sur leurs réponses neuronales soit 46 % (ANOVA 3 facteurs, $p < 0,05$). La totalité des 16 cellules code en moyenne de manière supérieure les stimuli originaux par rapport aux stimuli *scrambled*. Parmi les cellules restantes, 13 autres cellules TS n'ont pas d'effet principal de la phase mais montrent une interaction significative entre la phase des stimuli et les textures naturelles. Au total, 29 cellules TS sont influencées à la phase des stimuli sont 83% (ANOVA, effet principal + interaction, $p < 0,05$).

Pour les 35 cellules TS, nous vérifions uniquement pour les textures codées (d'après I_{max}) si la réponse aux stimuli *scrambled* est significativement différente (Test post hoc TUKEY, $p < 0,05$). Sur la totalité des textures codées spécifiquement par la population TS, seules 17% des textures montrent un taux de réponse significativement supérieur pour les stimuli originaux par rapport aux stimuli *scrambled*. Parmi les 41 textures, l'*aluminium* (#15)

engendre des réponses différentes selon le type des stimuli pour 4/5 cellules. C'est la seule texture qui engendre un pourcentage de réponses différentes aussi fort. Les autres textures montrent des pourcentages faibles.

4.2. Analyse par rang

Afin de comparer les réponses moyennes aux stimuli originaux et celles aux stimuli contrôles, j'ai effectué une analyse par rang. La méthode utilisée est identique à celle effectuée pour le set A. Cependant, seuls les stimuli illuminés par un angle de 22.5° sont pris en compte pour cette analyse (soit 41 stimuli) car ils ont des stimuli contrôles correspondants. La Figure 3-47 montre le résultat de cette analyse.

Les réponses moyennes normalisées pour les stimuli originaux se caractérisent par la présence d'une importante sélectivité. En effet, la décharge moyenne de la population augmente pour les stimuli originaux préférés (rangs petits) et décroît en fonction du rang du stimulus. Le coefficient de la droite de régression linéaire ($r = 0.908$) indique la présence d'une importante sélectivité. Une sélectivité est également présente pour les décharges moyennes aux stimuli contrôles même si elle est moins prononcée ($r = 0.724$). Enfin, comme nous l'observions pour le set A, les réponses aux stimuli contrôles ne sont pas indépendantes des réponses aux stimuli originaux. En effet, il existe une importante corrélation entre ces deux séries de réponses normalisées ($R_{Spearman} = 0.879$, $p < 0.001$).

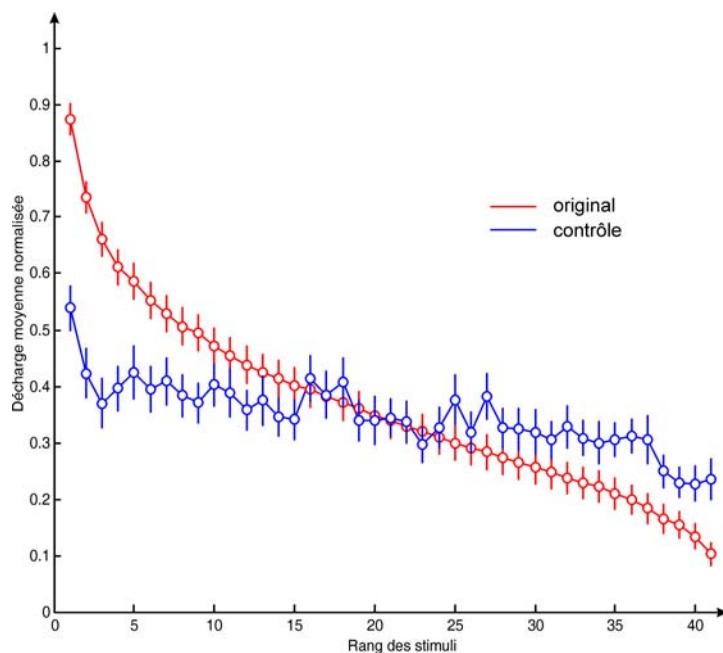


Figure 3-47: Analyse par rang. Les conventions sont les mêmes que celles utilisées dans la Figure 3-22. Les réponses aux stimuli originaux et scrambled sont corrélées significativement ($R_{Spearman} = 0.879$, $p < 0.001$).

5. Un nouveau contrôle : les réseaux orientés

5.1. Population d'étude

Pour une population assez restreinte de neurones ($n = 15$ cellules, 11 chez Zrem et 4 chez Tima), nous avons testé des réseaux orientés comme stimuli contrôles en remplaçant les stimuli *scrambled*. Les réseaux présentent 6 fréquences spatiales et 8 orientations différentes (voir annexes pour illustration). Deux gammes de 5 fréquences spatiales ont été testées au sein de cette population de cellules ; 8 cellules de 0,42 à 7,15 cycles/deg et 7 cellules de 0,85 à 12,5 cycles/deg. Parmi les 15 cellules, 14 sont des cellules TS. La totalité des analyses est effectuée sur cette population TS.

Afin de tester la sélectivité à l'orientation ou à la fréquence spatiale du réseau, un test ANOVA deux facteurs (orientation X fréquence spatiale) est effectué pour chacune des cellules. Parmi les 14 cellules, 86% des cellules montrent une réponse significativement différente selon la fréquence spatiale, 66% selon l'orientation des réseaux (Anova, deux facteurs, effet principal, $p < 0,05$). Parmi cette population TS, 53% des cellules présentent une interaction significative entre ces 2 facteurs. La Figure 3-48 présente deux cellules dont les réponses neuronales varient en fonction de la fréquence spatiale des réseaux. La cellule en A répond préférentiellement à des réseaux de basse fréquence (entre 0,85 et 1,66 cycles/deg) alors que pour la cellule B, la réponse moyenne normalisée augmente pour des fréquences spatiales hautes (7,15 cycles/deg). Si nous cumulons les réponses moyennes à ces réseaux des 14 cellules, il n'y a pas de fréquence spatiale ou d'orientation particulière qui semble être préférentiellement codée (Anova deux facteurs, $p_{freq\ spatial} = 0,518$ et $p_{orientation} = 0,625$).

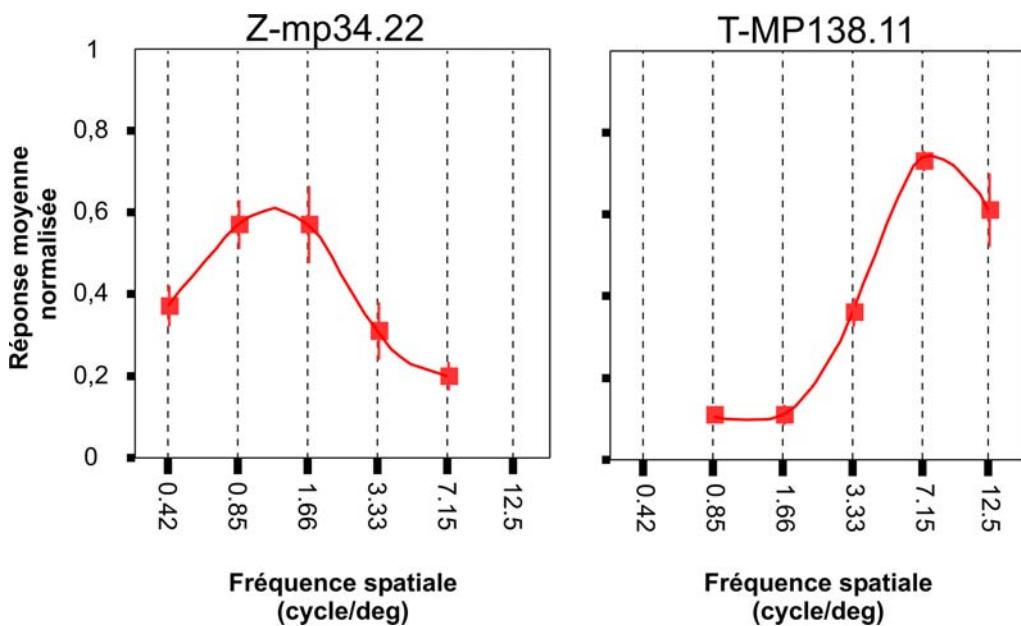


Figure 3-48 : Exemples de cellules sensibles à la fréquence spatiale. Les décharges de ces cellules (Z-mp 34.22 et T-mp 138.11) sont représentées sous la forme de réponses moyennes normalisées. Chaque réponse moyenne par stimuli est normalisée selon la décharge maximale pour le stimulus préféré (textures naturelles ou réseaux) et les réponses sont moyennées par fréquence spatiale en cumulant les 8 orientations. Les barres d'erreur représentent les erreurs standard de la moyenne.

L'étape suivante des analyses consiste à comparer les réponses moyennes aux textures et celles aux réseaux orientés pour chaque cellule. Pour cela, nous relevons pour chaque cellule, les textures codées d'après le critère I_{max} ainsi que la fréquence spatiale des réseaux qui engendre une réponse maximale du neurone. Par exemple, la cellule T-MP138.11 voit son taux de décharge augmenter pour une fréquence spatiale de 7.15 cycles/deg ; ses textures préférées sont le lin (H et V, #44), le plastique rugueux (#04) et la feuille de maïs (H, #51_1) ; soient des textures visuellement homogènes. Parmi la population TS, seules 2 cellules (dont TMP138.11) montrent des taux de décharge maximaux pour des fréquences spatiales supérieures à 7.15 cycles/deg. A l'inverse, la cellule Z-MP34.22 préfère des fréquences spatiales plutôt basses (0.85 à 1.66 cycles/deg) et décharge pour la paille (V et H, S40), la feuille d'aluminium (#15), le papier froissé (#28) et les sphères peintes (#35), textures visuellement hétérogènes.

5.2. Analyse hiérarchisée par cluster

Les réponses moyennes aux textures et aux réseaux orientés de la population de neurones sont traitées à l'aide d'une analyse hiérarchisée par clusters. Seules les réponses moyennes aux textures illuminées par un angle incident de 22.5° sont ici prises en compte pour restreindre l'analyse à 81 observations (41 images de textures + 40 images de réseaux orientés). Les réseaux orientés ont une fréquence spatiale s'échelonnant de 0.8 et 7.5 cycles/deg. La Figure 3-49 présente l'arbre de l'analyse hiérarchisée par clusters.

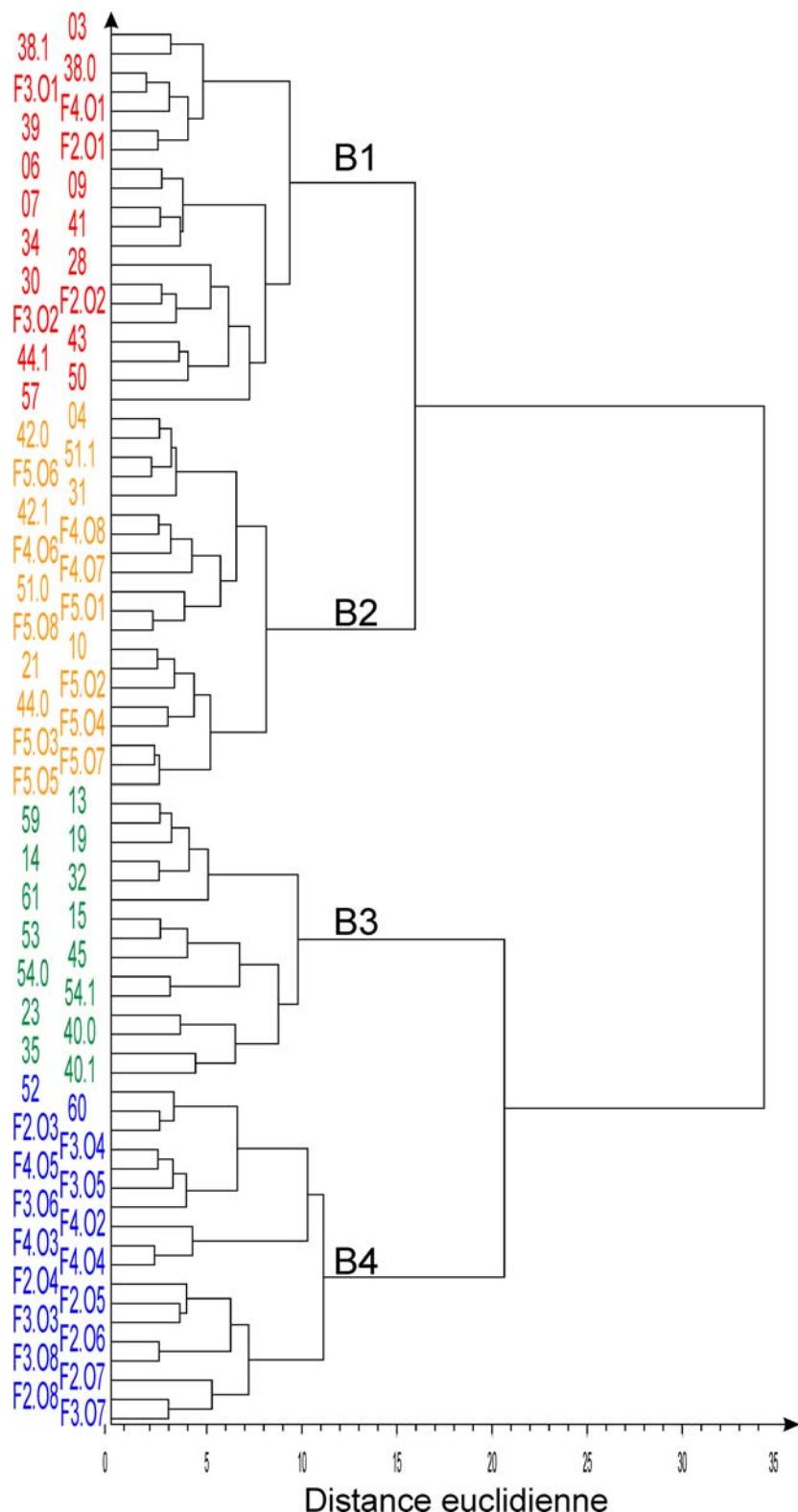


Figure 3-49 : Dendrogramme effectué sur la population de neurones TS testés avec les réseaux orientés (n=15). L'illumination pour les images des textures est de 22.5°. Les réseaux orientés ont des fréquences spatiales allant de 0.8 cycles/deg (F2) à 7.5 cycles/deg (F5) et des orientations qui varient de 0° (O1, orientation horizontale) à 157.5° (O8) par pas de 22.5°.

Les réponses moyennes peuvent être regroupées en 4 clusters (B1, B2, B3 et B4). La figure suivante détaille les 4 groupes de stimuli. Le groupe B1 regroupe 20 images dont 5

images de réseaux orientés. Les fréquences spatiales F2, F3 et F4 y sont représentées pour une même orientation horizontale. Les textures y sont relativement homogènes (voir détail des clusters, Figure 3-50). Le cluster B2 comprend également 20 images dont 11 de réseaux orientés (les 8 orientations pour les réseaux avec une fréquence spatiale de 7.15 cycles/deg, et 3 réseaux à 3.33 cycles/deg). Le cluster B3 comprend exclusivement des images de textures naturelles visuellement hétérogènes ($n=15$). Le cluster B4 comprend 18 images dans 15 représentantes des réseaux orientés. Les fréquences spatiales représentées varient entre 0.8 et 3.33 cycles/deg, les orientations de 22.5 à 157.5°.

En analysant les textures appartenant à chaque cluster, on peut remarquer que les textures du cluster A1 (B1 + B2, $n=24$) sont principalement homogènes ($SNR_{moyen} = 8,4$). A l'inverse, les textures du cluster A2 (B3 + B4, $n= 17$) sont hétérogènes ($SNR_{moyen} = 2,4$). Cette différence de SNR entre les clusters A1 et A2 est significative (ANOVA, cluster, $p<0,001$). D'après l'analyse hiérarchisée par cluster, la population TS présente à un premier niveau deux patterns de réponses suivant l'hétérogénéité de la texture. Les textures homogènes entraînent un taux de réponse de la population identique à celui pour des réseaux orientés définis majoritairement par des fréquences spatiales hautes. Dans B2, on retrouve, par exemple, tous les réseaux définis par une fréquence spatiale de 7,15 cycles/deg. Au contraire, les textures hétérogènes, principalement regroupées dans le cluster A2, se retrouvent avec des réseaux orientés définis par des fréquences spatiales faibles.

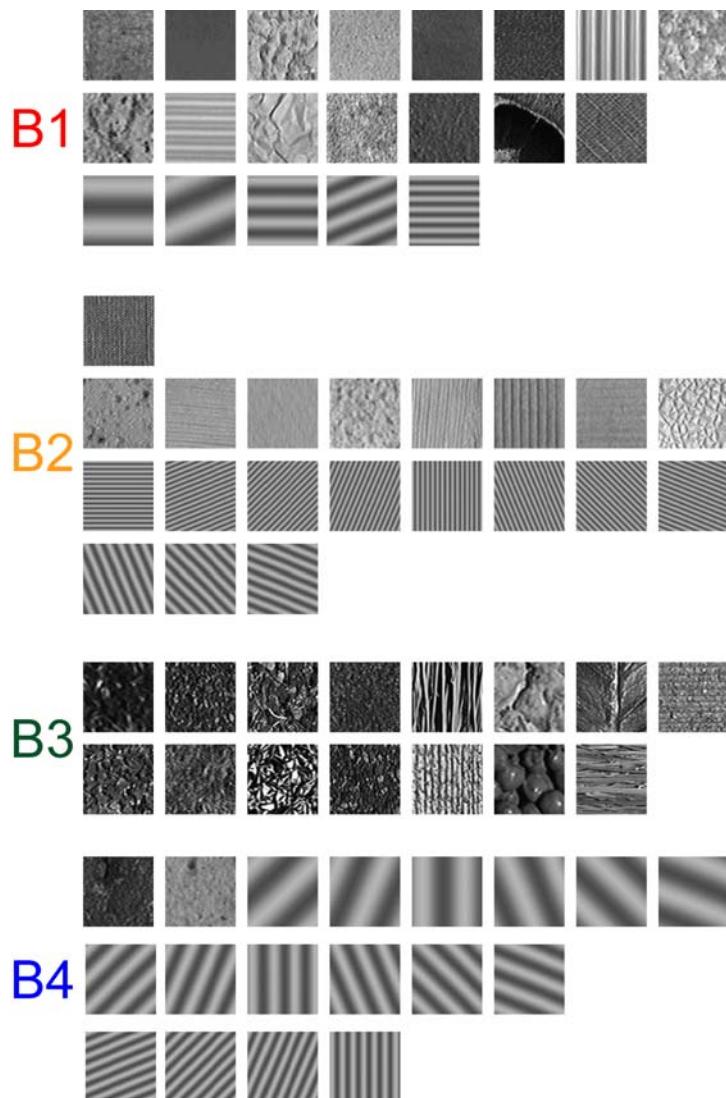


Figure 3-50 : 4 clusters identifiés grâce à l’analyse hiérarchisée sur la population TS testé avec les réseaux orientés.

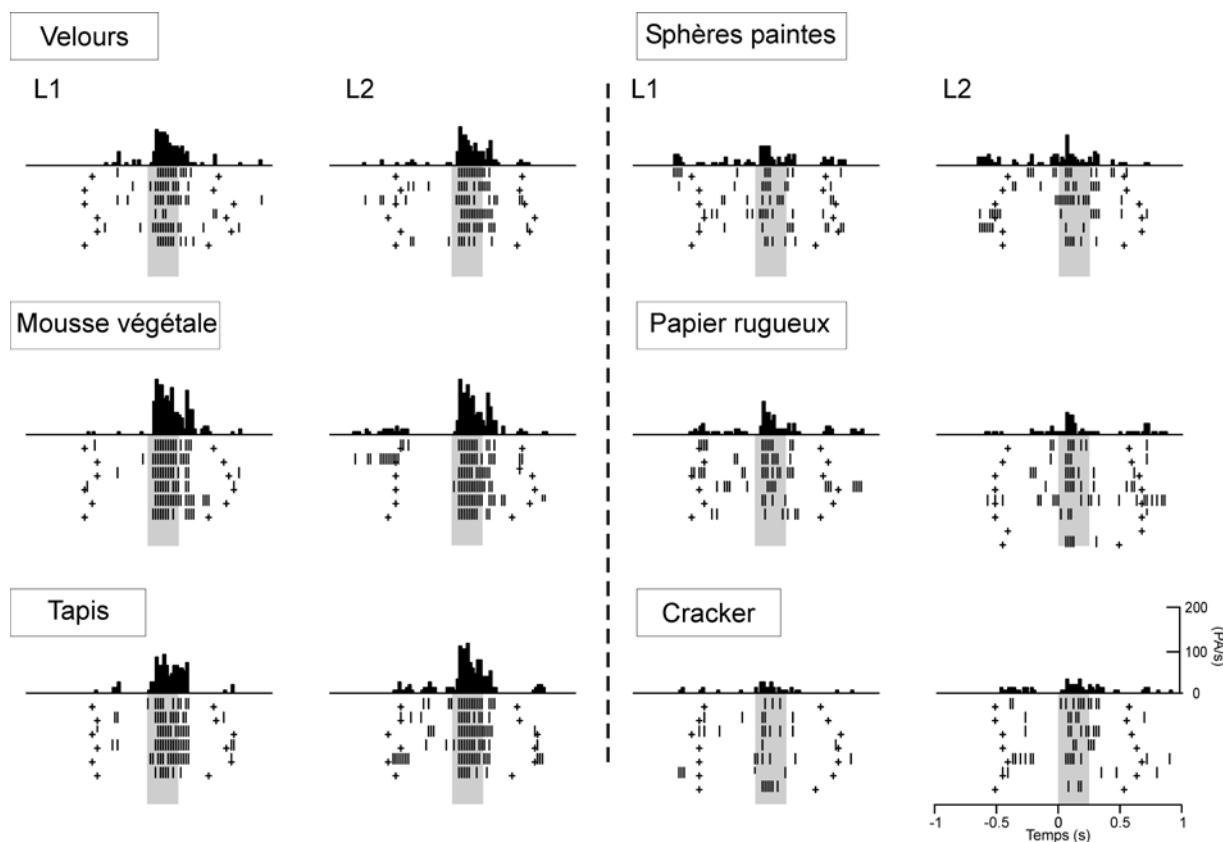


Figure 3-51 : Exemple de cellule TS invariante à l’angle d’incidence. Son taux de décharge ne varie pas selon l’illumination pour ses 3 textures préférées (Velours, Mousse végétale et Tapis) mais aussi pour 3 textures pour lesquelles cette cellule décharge peu (Sphères peintes, Papier rugueux et Cracker). Mêmes conventions que la Figure 3-6.

Une analyse par rang nous permet de mettre en évidence cette invariance au niveau de la population TS. Les séries des réponses moyennes normalisées de la population TS en fonction de l’angle d’incidence sont corrélées significativement ($R_{Spearman} = 0,972$, $p < 0,001$). Un stimulus préféré pour un angle d’incidence aura tendance à être également préféré pour l’autre angle d’incidence.

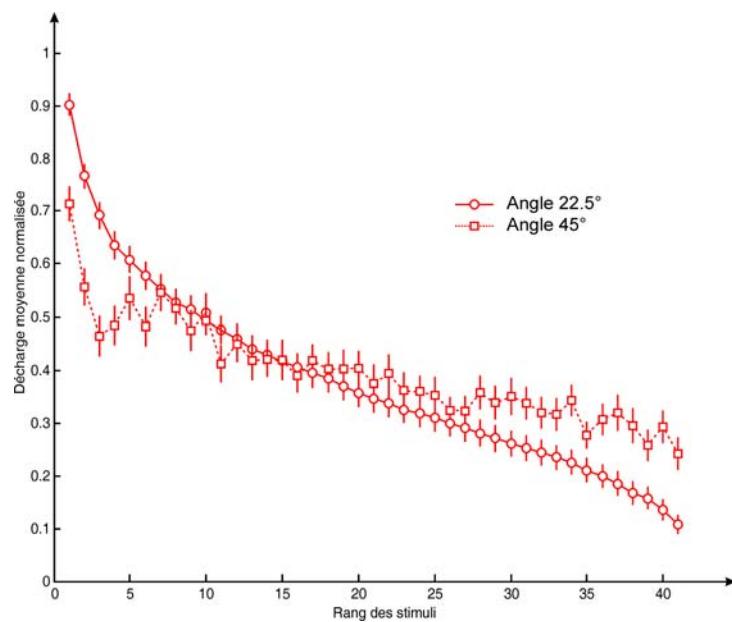


Figure 3-52 : Décharges moyennes normalisées en fonction du rang des stimuli. Selon l'angle d'incidence d'illumination L1.

F. DISCUSSION

1. Synthèse des principaux résultats

Cette étude a permis de démontrer l'existence d'une sélectivité aux textures naturelles dans l'aire V4 du singe. Cette sélectivité a été quantifiée et comparée avec les réponses d'une pyramide de filtre GABOR, outil mathématique communément utilisée dans une classification de textures. De cette étude, plusieurs résultats importants sont à retenir :

- Les textures naturelles sont des stimuli efficaces pour les cellules de V4. Le taux de décharge des cellules augmente lors de la présentation de textures naturelles en comparaison avec une période de référence sans présentation de stimuli. **La majorité des cellules de V4 est sélective aux textures naturelles** qui leur sont présentées (79%, 69% et 97% de cellules TS respectivement pour les sets A, B et C). Les cellules TS répondent en moyenne à une gamme de 5 textures et codent en majorité pour un ensemble de textures hétérogènes même si certaines répondent à la fois pour une texture hétérogène et pour une texture homogène.
- La sélectivité du neurone est préservée lorsque les stimuli sont déplacés dans le CR. **Les neurones TS ne codent donc pas les textures naturelles sur la base d'informations locales** mais comme une entité globale.
- Aucun des paramètres statistiques des images ne semble expliquer à lui seul les différences de réponses de la population TS à notre set d'images de textures naturelles. Toutefois, **la luminance moyenne et le SNR sont deux paramètres déterminants** pour le codage de la population TS. Les analyses multidimensionnelles (Analyse hiérarchisée par clusters et MDS) permettent de mettre en évidence des ensembles ou clusters de textures auxquels la population TS répond de façon similaire. La luminance moyenne et le SNR permettent de caractériser les clusters établis suite à l'analyse hiérarchisée ainsi que 2 dimensions des MDS. Contrairement aux dimensions 1 et 2 corrélées à des paramètres de 1^{er} ordre, la dimension 3 est corrélée à des paramètres décrivant l'hétérogénéité à un 2nd ordre.
- **Les réponses des cellules TS sont influencées par la phase des stimuli** (77% du set A, 83% pour le set C). La presque totalité des neurones TS qui montrent un effet principal de la phase sur leur taux de réponse ont un taux de décharge plus important pour les stimuli originaux que pour les stimuli *scrambled* (96% pour le set A ; 100% pour le set C). Toutefois, la population TS présente des taux de réponses différents selon la phase et la

texture naturelle présentée. Les textures hétérogènes sont bien différenciées de leurs stimuli contrôles alors que les textures homogènes (*lin, papier de verre*) montrent un même taux de réponse de la population TS. Pour la majorité des textures codées préférentiellement d'après I_{max} , les neurones TS ne différencient pas les stimuli originaux des stimuli contrôle (soit 77%). Toutefois, si on observe individuellement les neurones, les cellules différencient mieux les stimuli originaux des textures hétérogènes (*laitue, aluminium, plante et béton*) de leurs stimuli contrôles que les textures homogènes (*lin, papier de verre, plâtre et velours*). Pour les textures homogènes les originaux et les contrôles sont visuellement très proches. D'après l'analyse par rang, les réponses moyennes de la population TS aux stimuli contrôles ne sont pas complètement indépendantes des réponses originales. Autrement dit, si un stimulus est préféré par le neurone, son contrôle aura également tendance à être celui qui donne la plus forte réponse parmi les stimuli contrôles.

- La majorité des cellules TS est influencée par **l'angle d'incidence de la source d'illumination**. Toutefois, le pattern de sélectivité varie selon la texture naturelle. D'après notre index DMI, **la population TS est invariante à l'angle d'incidence** pour la texture naturelle préférée de chaque neurone.

Une approche computationnelle à l'aide d'une pyramide de filtres GABOR permet de définir plusieurs clusters de textures naturelles sur la base des énergies de convolution des filtres. Les réponses neuronales de la population TS et celles des filtres GABOR aux stimuli originaux sont corrélées. **Notre population TS se comporte comme une pyramide de filtres GABOR pour ébaucher une classification de textures naturelles.**

2. Localisation des sites d'enregistrements

Comme des études comportementales continuent avec Tima, aucune confirmation histologique des sites d'enregistrement n'a été effectuée sur cet animal. Il nous faut donc prouver par d'autres moyens que nous avons bien enregistré dans l'aire V4. Pour cela, nous nous basons sur les cartes rétinotopiques des aires V4. Nous retrouvons les positions respectives du méridien vertical (proche du sillon luné) et du méridien horizontal (dans la partie antérieure de la chambre d'enregistrement) décrites dans d'autres études (Gattass *et al.*, 1988; Fize *et al.*, 2003). De plus, comme cela été décrit dans l'étude de Gattass et al (1988), les champs récepteurs fovéaux sont en position latéro-antérieure tandis que les champs récepteurs en périphérie sont dorso-postérieurs. En accord avec ces mêmes résultats, la distance au niveau de la surface corticale entre les neurones avec un CR parafovéal (2° d'excentricité) et ceux avec 14° d'excentricité est d'environ 6 mm pour le singe Tima. La

possibilité que certaines cellules soient enregistrées au niveau de l'aire V2 est exclue pour plusieurs raisons :

(1) les CR dorso-postérieurs ont des excentricités pouvant atteindre 14 °, ce qui exclut des enregistrements dans la portion latérale de V2 au fond du sillon luné.

(2) Latéralement, les représentations fovéales de V2 et V4 dans le sillon préluné sont côte à côte mais nos CR les plus centraux ont 2° d'excentricité et sont donc trop périphériques pour être dans V2.

(3) Les tailles des CR aux différentes excentricités sont en accord avec celles décrites dans l'étude de Gattass et al (1988) avec une augmentation approximative de la taille de un degré² par degré d'excentricité.

Pour le singe Zrem, les traces des sillons, les réponses neuronales et les tailles des CR sont très similaires à ceux rencontrés chez Tim, même si les CR sont plus fovéaux.

Ces cartes rétinotopiques établies, l'histologie post-mortem sur Zrem et une radiographie scanner démontrent que nos enregistrements cellulaires s'effectuent au sein même de l'aire V4 entre les sillons luné, IOS et STS.

3. Sélectivité aux textures naturelles

3.1. Sélectivité générale

L'étude de Kobatake et Tanaka a démontré que le long de la voie visuelle ventrale, la sélectivité des neurones à des formes devient de plus en plus complexe (Kobatake et Tanaka, 1994). Les neurones de V1 ne répondent qu'à des barres orientées tandis que les neurones du cortex inférot temporal répondent davantage à des formes complexes. Au niveau de V4, des études d'électrophysiologie ont montré que les neurones de V4 sont sélectifs à des attributs simples d'une forme comme l'orientation, la longueur ou la largeur d'une barre mais également à l'orientation et à la fréquence spatiale de réseaux (Schein et Desimone, 1990). Au niveau de la sélectivité à la forme externe des stimuli, les études de Pasupathy ont montré que les neurones de V4 codent la forme complexe d'un stimulus en le décomposant en plusieurs fragments (Pasupathy et Connor, 1999; Pasupathy et Connor, 2001; Pasupathy et Connor, 2002). D'autres études ont révélé que les neurones de V4 sont sélectifs à des réseaux non-cartésiens (hyperboliques, polaires ou concentriques) (Gallant *et al.*, 1993; Gallant *et al.*, 1996). Les neurones de V4 pourraient donc coder pour des éléments d'un alphabet permettant de décrire une forme complexe ou le contenu d'une forme. Ainsi, ils coderaient le contour externe en le décomposant en plusieurs fragments et le contenu interne sous la forme de textons. Par ailleurs, une étude en neuroimagerie (IRMf) chez les humains révèle que l'aire

V4 réagit à des textures naturelles (Puce *et al.*, 1996). Toutefois, il n'est pas clairement établi que l'aire V4 de cette étude correspond à l'aire homologue de V4 décrite dans l'étude de Denys *et al.* (2004) (Denys *et al.*, 2004).

Nos résultats sont en accord avec ces études puisqu'ils indiquent que la majorité de nos cellules est sélective à la texture naturelle. De plus, comme il a été remarqué dans les études de Pasupathy *et al.* (1999) ou Gallant *et al.* (1996), le pattern de sélectivité est assez large ; nos neurones TS déchargent pour en moyenne 5 textures (Gallant *et al.*, 1996; Pasupathy et Connor, 1999). Un premier argument plausible pour expliquer cette sélectivité est que les neurones de V4 sont sensibles aux contours externes des stimuli. Le fait que la sélectivité des neurones persiste lorsque les stimuli sont déplacés de 1° dans le CR des cellules permet d'exclure cette hypothèse. Ce résultat corrobore ceux de Gallant *et al.* (1996) qui avaient utilisés des réseaux cartésiens ou non-cartésiens et qui observaient des réponses invariantes des cellules de V4 à la position des stimuli dans les CR.

3.2. Paramètres des textures

Nous avons remarqué que, typiquement, une cellule TS de V4 répond à un ensemble de textures naturelles et que les textures préférées au sein de la population TS sont plus généralement des textures hétérogènes comme la feuille de laitue (#23) ou l'aluminium (#15) pour le set A, et la plante Soleirolia (#53) ou la paille (#40.0) pour le set C. A l'inverse, des textures homogènes telles que le Lin (#44) ou le papier de verre (#06) apparaissent moins souvent comme étant des textures préférées de la population TS. Les préférences de réponses de la population TS pour les textures hétérogènes pourraient être la conséquence d'un contenu riche du spectre de fréquences spatiales dans ce type d'images. En effet, les phases et l'amplitude du spectre de Fourier sont essentielles pour déterminer l'aspect visuel d'une image (Tadmor et Tolhurst, 1993). Afin de tester cela, nous avons utilisé des stimuli contrôles définis par une redistribution aléatoire des phases de Fourier (*scrambled*). Cette redistribution aléatoire modifie l'aspect visuel des images originales, en particulier pour les textures hétérogènes. Une majorité des cellules TS (77%) présentent un taux de décharge moyen plus important pour les textures originales que pour les stimuli contrôles, et ceci est particulièrement vrai pour les neurones qui codent préférentiellement les textures hétérogènes comme la feuille de laitue (#23). Les textures hétérogènes en comparaison avec les textures homogènes contiennent de forts indices 3D dus à leurs patterns de micro-ombrages. Les neurones TS se basent sur ces indices 3D pour les dissocier de leurs stimuli *scrambled*. A l'inverse, les textures homogènes contenant peu d'indices de profondeur, sont peu dissociées par les neurones TS.

Les neurones sensibles à la profondeur dans une texture ne devraient pas répondre seulement de manière différentielle mais de manière supérieure à des images contenant de la profondeur. Le fait que les cellules de V4 préfèrent les stimuli originaux plutôt que leurs stimuli *scrambled* particulièrement pour les textures hétérogènes suggère que les neurones de V4 encodent l'aspect 3D des textures. Ce résultat indique qu'il existe un traitement de la profondeur au niveau de la population des neurones au sein de l'aire V4. Toutefois, comme les stimuli sont présentés en situation passive aux animaux, nous ne pouvons conclure réellement sur la perception de la profondeur des animaux.

3.3. Dimensions de la classification des textures

Les textures naturelles sont des stimuli complexes qui ne peuvent être décrits par un simple paramètre (Haralick, 1979). Par exemple, nous, humains, qualifions et classifions les textures visuellement par des adjectifs dont Bhushan a établi une liste exhaustive (Bhushan *et al.*, 1997). Nous pouvons également reconnaître une texture sur la base d'informations tactiles (Picard *et al.*, 2003). Toutes ces études ont utilisé comme méthode d'analyse des méthodes multidimensionnelles qui révèlent que plusieurs dimensions sont nécessaires pour obtenir une classification correcte de textures. Nous avons effectué ce même type d'analyse sur les réponses de notre population TS (Analyse hiérarchisée par clusters et MDS). Ces analyses nous ont permis de mettre en évidence que 3 dimensions permettent d'expliquer les réponses des neurones de V4. Nous avons caractérisé 2 dimensions en termes de luminance et d'hétérogénéité de 1^{er} ordre. Quant à la 3^{ème} dimension, elle est corrélée à des paramètres quantifiant l'hétérogénéité à un 2nd niveau. Or les paramètres de 2nd ordre sont issus de calculs effectués à partir de matrices de cooccurrences orientées et nous avons montré que les réponses de la population TS sont corrélées avec celles issues d'une pyramide de filtres GABOR orientés (test de PROCRUSTES et NNF). La 3^{ème} dimension pourrait alors être dépendante de l'orientation locale d'ensembles de pixels. Fonctionnellement, ces dimensions pourraient être interprétées comme des dimensions perceptives requises pour classifier des textures entre elles. Une bonne classification des textures tend à regrouper des textures similaires entre elles. La population de V4 classifie-t-elle réellement les textures ? Nos analyses par clusters nous indiquent que les images d'une même texture mais illuminées par différents angles (L1 et L2, cf. Figure 3-43) ont tendance à appartenir à un même cluster. A l'aide de ces résultats, nous pouvons donc suggérer que les neurones au sein de V4 participent à une classification de textures en se basant sur la luminosité et sur l'hétérogénéité de la texture.

3.4. Comparaison avec les études computationnelles

Les modèles théoriques sur la classification des textures cherchent à mimer de la manière la plus parfaite possible notre perception visuelle en condition réelle. Un problème intéressant concerne le fait que les humains puissent reconnaître une texture naturelle indépendamment de la direction d'illumination. Pourtant, des changements de la direction d'illumination modifient fortement les patterns de micro-ombrages des textures naturelles. Ces modèles théoriques utilisent communément comme outil mathématique une pyramide de filtres linéaires de type GABOR comme celle que nous avons utilisée lors de cette étude. Dans leur étude où Leung et Malik (2001) définissent les *textons 3D* comme vocabulaire permettant de classer les textures, ils soulignent que leur algorithme montre un taux de réussite important (87%) dans une tâche de classification de textures pour des textures illuminées selon un nouvel angle (Leung et Malik, 2001). Les textures présentées ont été préalablement *apprises* par l'algorithme mais illuminées selon un angle différent. Par ailleurs, d'autres études ont développé des algorithmes plus élaborés (tous basés sur des réponses de filtres) qui sont plus efficaces dans une tâche de classification de textures. En effet, ces algorithmes peuvent classifier des nouvelles textures même si l'angle d'illumination ou le point de vue varient (Varma et Zisserman, 2002; Hayman *et al.*, 2004).

Au cours de notre étude, nous avons démontré que les cellules TS se comportent comme un ensemble de filtres GABOR pour classifier les textures naturelles. D'après Jones et Palmer (1978), une pyramide de filtres GABOR mime plutôt les propriétés des CR des cellules simples de V1 que celles de V4 (Jones et Palmer, 1987). Toutefois, la théorie de Beck (1982) propose que le traitement de la texture s'effectue selon plusieurs niveaux de filtrages comme Chubb et Landy (1991) le proposent dans leur modèle de « filter-rectify-filter » (Beck, 1982; Chubb et Landy, 1991). Ces modèles soulignent que la segmentation des textures est effectuée lors d'une seconde étape de filtrage. La première étape de filtrage concerne les cellules simples de V1 alors que le second niveau de filtrage permet la détection des bordures définies par la texture. Selon le modèle de Thielscher et Neumann (2005), cette étape serait effectuée au sein de l'aire V4 permettant ainsi d'encoder l'orientation de la surface, ceci ne serait pas possible dans les aires inférieures à V4 d'après Lennie (Lennie, 1998; Thielscher et Neumann, 2005). Ce rôle fonctionnel élaboré de l'aire V4 dans la segmentation des textures a été confirmée dans des études de neuroimagerie chez l'Homme. En effet, l'étude de Kastner *et al.* (2000) démontre que les bordures définies par des textures entraînent une meilleure activation dans V4 que dans des aires plus bas-niveau comme V1 et V2 (Kastner *et al.*, 2000). De plus, des études de lésions de V4 chez le macaque démontrent

que les animaux lésés montrent des déficits dans des tâches de segmentation de textures (De Weerd *et al.*, 1996; Merigan, 2000). Enfin, nous avons démontré que les réponses de la population TS et celles des filtres sont corrélées significativement. Cette analogie est due principalement aux paramètres statistiques des images caractérisant leur hétérogénéité (SNR ou variance) et leur luminance. De manière surprenante, une dimension de la MDS obtenue avec le set C est corrélée significativement avec les moyennes de paramètres statistiques de 2nd ordre, ce qui ne pourrait pas être présent au sein de V1.

Toutefois, une étude de lésion de V2 montre que les animaux montrent d'importants déficits dans une tâche de segmentation de textures (Merigan *et al.*, 1993). De ce fait, il n'est pas clairement établi comment s'effectuent les différentes étapes du traitement perceptif de la texture de V1 à V4. Existe-t-il simplement une différence d'échelle entre les aires ? Elles opéreraient des processus de traitement identiques ; les CR de V4 regrouperaient alors plusieurs CR de V1 ou de V2. De récentes études démontrent que ce processus de traitement ne peut pas s'expliquer par un modèle linéaire de filtres. Les neurones de V1 peuvent détecter des relations spatiales particulières entre les pixels au sein d'une texture visuelle (Purpura *et al.*, 1994). Lamme (1995) démontre que la segmentation de textures dans V1 permet de distinguer une figure du fond visuel (Lamme, 1995). Plus récemment, Hegde et Van Essen (2006) se sont interrogés sur l'encodage de stimuli complexes dans V1, V2 et V4. Ils utilisent pour cela des stimuli similaires à ceux utilisés par Gallant dans ses études ; des réseaux cartésiens et non-cartésiens et des arcs ou des courbes. Leurs résultats démontrent qu'il n'existe pas de différence majeure au niveau de traitement de la forme entre les aires V1, V2 et V4. Les modulations des CR des neurones des aires inférieures seraient dues aux connections feedback de V4 vers V2 ou V1 (Hegde et Van Essen, 2006).

4. Invariance à la direction d'illumination

Les humains peuvent facilement reconnaître une texture même si elle est illuminée sous différentes conditions. Nous pouvons admettre qu'une telle faculté de reconnaissance est également présente chez le singe. En effet, dans son milieu naturel, les conditions d'illuminations varient tout au long de la journée et n'altèrent pas sa reconnaissance des objets. Une interrogation intéressante concerne les bases neuronales de cette invariance. Nos stimuli sont illuminés selon 3 angles d'incidence différents produisant en conséquence 3 patterns distincts de micro-ombrages. Ces patterns diffèrent particulièrement si la texture est hétérogène et si elle contient des aspérités (des crevasses ou des crêtes).

Les modèles structuraux énoncent qu'une reconnaissance d'objets n'est pas influencée par une variation d'illumination car la reconnaissance de l'objet se fait grâce à sa

décomposition en éléments structuraux qui ne contiennent pas ce type d’information (Marr et Nishihara, 1978; Biederman, 1987). Des résultats d’études psychophysiques corroborent ces théories en démontrant que les sujets peuvent reconnaître facilement des visages ou des objets lorsque les conditions d’illumination varient (Nederhouser *et al.*, 2001). Ce mécanisme d’invariance à la direction d’illumination est présent au sein du cortex IT d’après l’étude électrophysiologique de Vogels et Biederman (2002). Les neurones de IT sont invariants pour différentes directions d’illumination pour leur forme préférée (Vogels et Biederman, 2002).

Les modèles d’alignement proposent que la direction d’illumination soit encodée dans la représentation de l’objet ou du visage (Ullman, 1989; Poggio et Edelman, 1990). Ces théories s’appuient sur des résultats psychophysiques qui démontrent que la reconnaissance des visages ou d’objets est influencée par la direction d’illumination (Gauthier et Tarr, 1997; Troje et Bulthoff, 1998; Braje, 2003). A l’aide d’enregistrements électrophysiologiques, Hanazawa et Komatsu (2001) démontrent que les cellules de V4 sont influencées par la direction du gradient d’illumination. Ils remarquent également la présence d’un biais pour des directions d’illumination le long des l’axe vertical ; les neurones ont tendance à être plus sélectifs pour ces directions d’illumination (Hanazawa et Komatsu, 2001). Contrairement à leur résultat, notre étude démontre que les réponses des cellules TS présentent une relative invariance aux angles d’incidence pour la texture préférée. Toutefois, notre protocole est différent car nous utilisons 3 angles d’incidence mais seulement une seule direction d’illumination ; la source lumineuse est positionnée sur la droite des images. Nous avons donc mis en évidence une relative invariance pour cette direction d’illumination. Il pourrait exister au sein de V4 des sous populations de neurones sélectives à une direction d’illumination mais invariantes aux angles d’incidence dans cette direction. De tels mécanismes pourraient être démontrés en enregistrant les réponses des cellules à des textures naturelles illuminées sous d’autres directions ; par la gauche, par-dessous ou par-dessus. De plus, il serait intéressant de démontrer l’invariance à plusieurs angles d’illumination dans une tâche de conditionnement plus cognitive dans laquelle l’animal a été entraîné à reconnaître des textures sous différentes conditions d’illumination. Cette généralisation de la reconnaissance des textures sous différentes illuminations est envisageable au sein de l’aire V4 puisqu’il a été démontré que cette aire participe à des phénomènes de *perceptual learning* (Rainer *et al.*, 2004).

5. Conclusion

Nos résultats ont démontré que l’aire V4 joue un rôle déterminant dans la perception visuelle des textures indépendamment des conditions d’illumination. Les neurones de V4 se basent sur des paramètres tels que la luminance moyenne ou l’hétérogénéité pour classifier les

textures entre elles. Ces cellules se comportent comme un ensemble de filtres GABOR afin d'ébaucher une classification des textures naturelles. Le traitement de cet attribut intrinsèque de l'objet nous permettrait de le reconnaître et de le segmenter du fond visuel. Outre la forme et la couleur, l'aire V4 participe également à l'encodage de la texture, propriété matérielle de l'objet. De même, puisque les réponses de V4 sont influencées par la congruence visuo-somesthésique des textures (Haenny et Shiller, 1988), il est possible que les informations visuo-tactiles, essentielles pour décrire une texture convergent vers cette aire corticale. En conséquence, il serait intéressant de tester les réponses des neurones de V4 lorsque des singes effectuent des tâches actives de reconnaissance de textures basées sur des informations visuelles ou tactiles congruente ou incongruente.

En termes de conclusion, je reprendrai une phrase publiée dans la revue de Landy et Graham (2004) sur la perception des textures. Les auteurs écrivent: “a complete theory of texture perception might involve recognition that natural textures area associated with real world materials, and the appearance of texture may well relate to perception of the particular material from which the image derived (wood, plastic, water, grassland, etc.) or properties of the real-world material that might relate to actions the observer might wish to take. This is the concept of an “affordance” (Gibson, 1979). Is the material sticky? Will it crumble in my hand? Will I be able to walk on it in bare feet?” (Landy et Graham, 2002).

Cette étude nous a permis de mettre en évidence une sélectivité de l'aire V4 pour des patterns de micro-ombrages qui définissent la texture naturelle de l'objet. Or, le gradient d'illumination génère également à la surface des patterns de macro-ombrages qui nous permettent d'extraire la forme tridimensionnelle de l'objet. Une sélectivité à ces patterns de macro-ombrages de l'aire V4 sera testée dans l'étude suivante.

Chapitre 4

Codage des ombrages dans l'aire V4 du singe vigile

Ce chapitre présente les résultats obtenus sur l'encodage des gradients d'illumination au sein de l'aire V4. Dans le chapitre 1, nous avons vu que les ombrages à la surface d'un objet sont considérés comme un indice pictural de la profondeur. Cette étude comporte plusieurs problématiques. Dans la première, il s'agit de déterminer si un codage de la profondeur sur la base d'ombrages est présent au sein de l'aire V4. Si un tel codage est présent, nous caractériserons la sélectivité pour les formes tridimensionnelles présentées et enfin nous testerons si le codage pour une forme tridimensionnelle est invariant selon la direction d'illumination.

A. MISE EN EVIDENCE D'UN CODAGE DE LA PROFONDEUR SUR LA BASE D'OMBRAGES AU SEIN DE L'AIRE V4

1. Base de données

Au cours de cette étude, nous avons utilisé plusieurs protocoles sur les deux animaux avec des stimuli contrôles différents et un nombre variable de stimuli. Le tableau suivant décrit le nombre de neurones enregistrés selon le protocole utilisé sur chaque animal.

Tima			
	N neurones enregistrés	N neurones (5 présentations)	N neurones (p<0.05)
Set S1	60	49	46
Set S2	40	31	31
Set S3	15	13	13
Zrem			
Set S3	35	31	29
Total Tima & Zrem	150	124	119

Tableau 4-1: Nombre de neurones enregistrés par singe et par protocoles. Les colonnes sont identiques à celles du tableau 1 du chapitre 3. Set S1 : 8 formes 3D, 2D blobs et Random ; Set S2 : 8 formes 3D, 2D blobs et postérésés ; Set S3 : 4 formes 3D, 2D blobs et postérésés..

Nous avons enregistré la réponse de 124 cellules de V4 (n=93 pour Tima, et n=31 pour Zrem) au cours des différents paradigmes. Parmi ces 124 neurones, une grande majorité (96%) montrent une activité neuronale significativement différente pendant la présentation des stimuli par rapport à la période de référence soit 119 neurones (90 cellules sur Tima et 29 cellules pour Zrem). Ce pool de 119 cellules constitue l'ensemble des cellules sur lesquelles sont effectuées les analyses statistiques.

2. Etude de la sélectivité 3D

Afin de mettre en évidence les sélectivités des neurones de V4, nous opérons un test ANOVA 3 facteurs (type, direction d'illumination et forme) sur les réponses neuronales de chaque cellule de V4. Le type des stimuli indique s'il s'agit d'un stimulus 3D, d'un 2D blob, d'un random ou d'un stimulus postéréisé (indépendamment de la forme du stimulus).

2.1. *Résultats du test ANOVA.*

Parmi les 119 cellules, 93 cellules (soit 78% de la population) présentent un effet principal du type de stimuli au niveau de leur réponse neuronale (ANOVA, p<0,05). Ces cellules sont dites « cellule ICS » (Inner content selective). Un test post-hoc TUKEY permet

de comparer deux à deux les réponses moyennes aux stimuli 3D, aux stimuli 2D blobs, aux stimuli random ou postérés afin de déterminer si chaque cellule voit son taux de réponse augmenter pour un type particulier de stimuli. Ainsi, ceci nous permettra de savoir si les cellules ICS codent préférentiellement les stimuli 3D par rapport aux stimuli contrôles.

Dans un premier temps, nous intéresserons uniquement aux réponses aux stimuli 3D et 2D blobs. Parmi les 93 cellules ICS, 60 présentent des réponses moyennes significativement différentes entre ces deux types de stimuli (ANOVA, TUKEY, $p<0.05$). Parmi elles, 23 cellules ICS montrent une réponse moyenne aux stimuli 3D plus importante alors que les 37 autres cellules ont un codage plus important pour les stimuli 2D blobs. Les 33 cellules ICS restantes de la population globale ne présentent pas de différence significative au niveau du taux de réponse pour ces deux types de stimuli; leurs réponses moyennes aux stimuli 3D et 2D blobs sont similaires. La première catégorie de cellules est aussi appelée « cellules 3D », la seconde « cellule 2D blobs » et la troisième « cellule 2D ». La Figure 4-1 présente des exemples de cellules pour ces 3 catégories de neurones.

Le neurone de la Figure 4-1A présente un niveau moyen de réponse pour les stimuli 3D supérieur à celui pour les stimuli 2D blobs. La différence de réponses entre les deux types de stimuli varie en fonction de la forme présentée ; les formes 3D n°1, n°3, n°4, n°5 et n°7 sont dissociées de leurs contrôles respectifs alors que ce neurone montre un même taux de réponse pour les formes n°2, n°6 et n°8 quel que soit le type de stimuli. Cette variation de réponse sera analysée dans le paragraphe suivant avec le calcul d'un index de dissociation. La cellule de la Figure 4-1B est un exemple de cellule répondant significativement aux stimuli 2D blobs. Les réponses moyennes aux stimuli 2D blobs sont supérieures à celles pour les stimuli 3D. Enfin, le troisième exemple de neurone (Figure 4-1C) présente une cellule qui ne dissocie pas les deux types de stimuli. Cette cellule présente un même taux de réponse quelle que soit la forme présentée.

En résumé, parmi la population ICS, nous avons caractérisé trois catégories de cellules ayant des patterns différents de réponses ; 25% sont des cellules 3D, 40% des cellules 2D blobs et 35% des cellules 2D. Intéressons-nous particulièrement aux cellules 3D. La Figure 4-2 reprend deux exemples de cellules 3D. Les réponses pour les différentes formes sont dissociées afin de mettre en évidence la présence d'une interaction TYPE X FORME. La cellule de la Figure 4-2B a une interaction significative entre ces deux facteurs (ANOVA, interaction, $p<0.05$). Parmi les 23 neurones, 13 cellules montrent une interaction significative comme pour cet exemple. Les différences de réponses entre les stimuli 3D et 2D blobs varient selon la forme présentée. Cette cellule voit sa décharge neuronale augmenter

significativement pour les formes n°1, n°2, n°3 et n°8 pour les stimuli 3D par rapport aux stimuli 2D blobs correspondants alors que les décharges moyennes pour les formes n°4, n°5 et n°7 sont identiques quel que soit le type du stimulus (Test TUKEY, $p>0.05$). Le niveau de réponse pour la forme n°6 est significativement supérieur pour les contrôles 2D blobs. Les 10 autres cellules ne montrent aucune interaction significative. La Figure 4-2A montre un exemple de ce type de cellule ; les réponses moyennes aux formes 3D n°2, n°3, et n°4 sont significativement supérieures aux réponses moyennes des stimuli contrôles correspondants (Test TUKEY, $p<0,05$). Pour résumer, il est donc important de noter que les cellules 3D n'ont pas de réponse significativement supérieure pour toutes les formes 3D et ne vont dissocier que certaines formes tridimensionnelles de leurs contrôles respectifs au niveau de leur décharge neuronale.

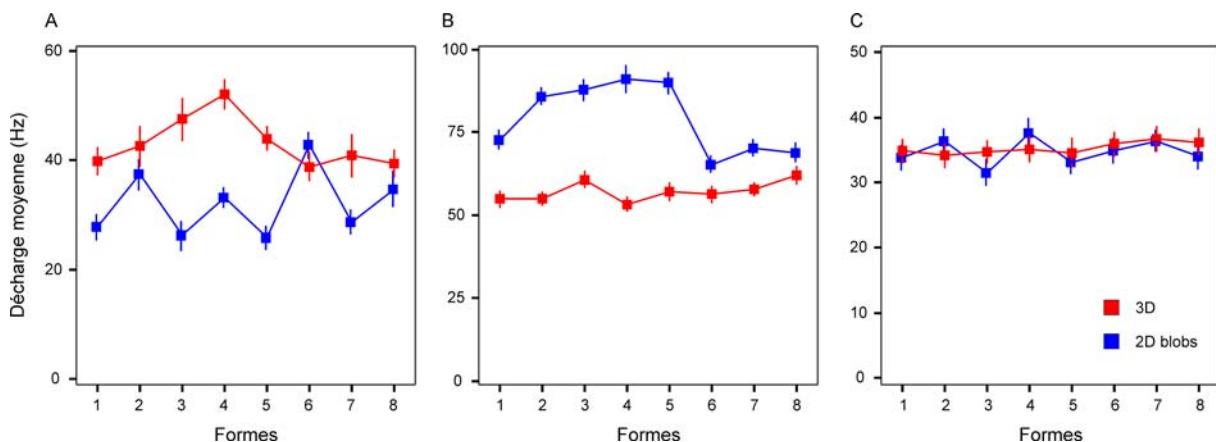


Figure 4-1 : Trois exemples de cellules de V4. A. Cellule 3D sélective, T-SFS30.31 (25% des cellules influencées par le type des stimuli) B. Cellule 2D blobs sélective, T-SFS20.11 (40%). C. Cellule non sélective, T-SFS10.31 (35%). Toutes les illuminations et formes utilisées sont cumulées. Les couleurs correspondent aux types des stimuli ; rouge pour les stimuli 3D et bleu pour les stimuli 2D blobs. Les barres d'erreurs représentent l'erreur standard de la moyenne (esm).

Afin de mettre en évidence quelles formes 3D sont codées préférentiellement par rapport à leur contrôle 2D, le test post-hoc TUKEY nous indique si la réponse pour une forme 3D donnée est significativement supérieure à son contrôle 2D blobs. Au niveau de la population 3D, les cellules dissocient en moyenne 2,5 formes 3D de leurs contrôles 2D blobs respectifs. Le nombre maximum de formes 3D dissociées est de 5. Parmi les formes 3D les plus dissociées, nous retrouvons la forme n°3 pour 61% des cellules 3D et la forme n°1 pour 45% des cellules 3D. Inversement, les formes n°5 et n°7 sont très peu dissociées par les cellules 3D, car seules 14% et 20% des cellules 3D les dissocient respectivement.

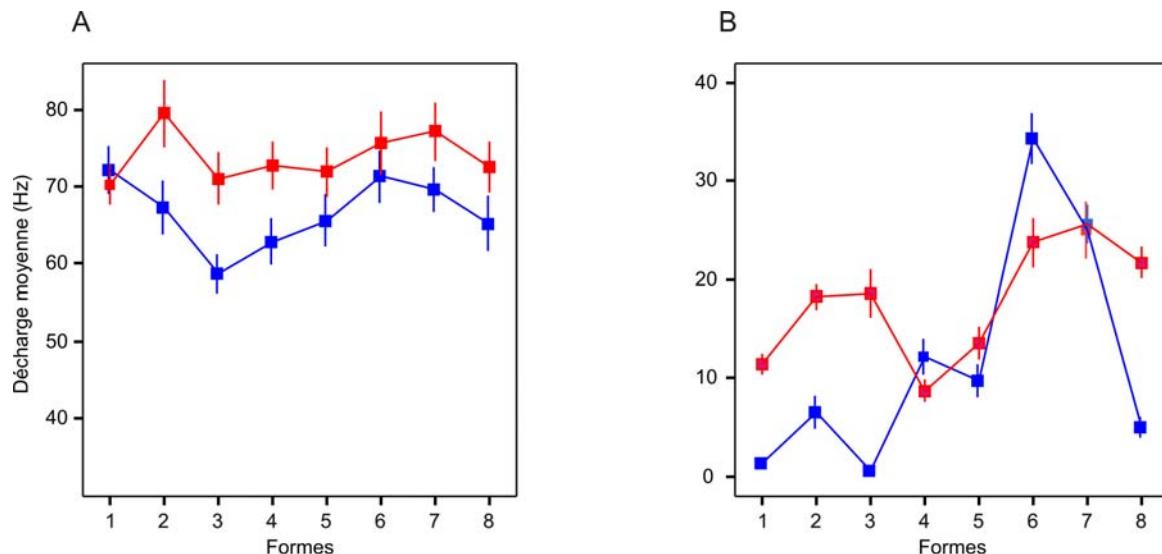


Figure 4-2 : Décharges moyennes de 2 cellules 3D en fonction des formes et des types des stimuli. Le code couleur est le même que la Figure 4-1. Les barres d'erreurs indiquent l'esm. La cellule A (T-SFS22.11) présente pour la majorité des formes un taux de réponse supérieur pour les stimuli 3D alors que la cellule B (T-SFS5.21) présente une interaction significative entre la forme et le contenu interne des stimuli.

Des cellules autres que les cellules 3D dissocient également les formes 3D de leur contrôles. Par exemple, 14 cellules 2D blobs dissocient au moins une forme 3D de son contrôle respectif mais cette forme 3D ne fait pas parti des stimuli qui engendrent une réponse maximale du neurone.

2.2. Index de dissociation.

Les deux exemples de cellules de la Figure 4-2 montrent que la différence de réponses entre les stimuli 3D et leurs stimuli 2D blobs varie en fonction de la forme présentée. De plus, les réponses pour les stimuli 2D blobs, même si elles sont inférieures en moyenne aux réponses pour les stimuli 3D, atteignent un niveau de réponse important. Afin de caractériser cette différence de codage pour chaque cellule, nous avons calculé un index de réponse (I_{3D}) pour chacune des 8 formes.

$$I_{3D} = (R_{3D} - R_{2D}) / SDm$$

R_{3D} = Réponse moyenne pour une forme 3D

R_{2D} = Réponse moyenne pour les stimuli 2D blobs correspondant

SDm = Moyenne des déviations standard de R_{3D} et R_{2D} .

Un index (I_{3D}) entre +1 et -1 indique que la réponse moyenne aux stimuli 3D et la réponse moyenne aux stimuli 2D sont équivalentes. Au-delà de +1.2 la réponse moyenne aux stimuli 3D est significativement supérieure à celle pour les stimuli contrôles correspondant (test post-hoc TUKEY). Le Tableau 4-2 résume le calcul de ces index (I_{3D}) pour les 8 formes pour les deux cellules de la Figure 4-2.

Formes	n°1	n°2	n°3	n°4	n°5	n°6	n°7	N°8
Cellule A	-0,40	1,25	1,82	1,74	1,19	0,60	0,72	0,86
Cellule B	7,37	2,54	3,75	-0,61	0,75	-1,85	0,11	4,83

Tableau 4-2 : Valeurs de I_{3D} pour chaque forme (de 1 à 8) pour les deux cellules 3D de la Figure 4-2, (T-SFS22.11 et T-SFS5.21). Chaque valeur indique le I_{3D} pour une forme donnée. La forme 3D optimale (I_{3D} maximum) est la forme n°4 et n°1 respectivement pour les cellules A et B.

La cellule de la Figure 4-2A présente un index I_{3D} maximal pour la forme n°3. Les 2 autres formes dissociées d'après les résultats du test post-hoc TUKEY, ont des index de 1.25 et de 1.74 pour les formes n°2 et n°4 respectivement. Pour la cellule Figure 4-2B, la plus forte dissociation s'observe pour la forme n°1 ($I_{3D} = 7,37$). Les trois autres formes dissociées (n°2, n°3 et n°8) ont toutes des index I_{3D} supérieurs à 2,5. Toutes les formes dissociées ont un I_{3D} supérieur à 1.2. Nous utiliserons ce critère pour déterminer si une forme 3D est dissociée ou pas de son contrôle 2D correspondant.

L'histogramme de la Figure 4-3 présente la distribution des index I_{3D} pour la population 3D. Cet histogramme présente une médiane de 0.67 indiquant que la majorité des formes 3D présentent un même taux de décharge que leurs stimuli 2D blobs correspondant. Toutefois d'après ce critère ($I_{3D} > 1.2$), 48 formes 3D sont dissociées soit en moyenne 2,2 pour les neurones 3D enregistrés avec 8 formes et 1,75 pour ceux enregistrés avec 4 formes.

A l'aide des calculs de ces index I_{3D} , nous avons pour chaque cellule la forme 3D optimale, c'est à dire celle qui engendre une dissociation maximale (I_{3D} maximum). Pour les cellules 3D, la forme 3D n°3 est la plus fréquemment dissociée de son contrôle 2D blobs car 14 cellules 3D la dissocient. La forme 3D optimale ne correspond pas forcément à la forme 3D préférée (celle qui engendre une réponse moyenne maximale). Pour 10 cellules 3D, la forme 3D optimale correspond à la forme 3D préférée.

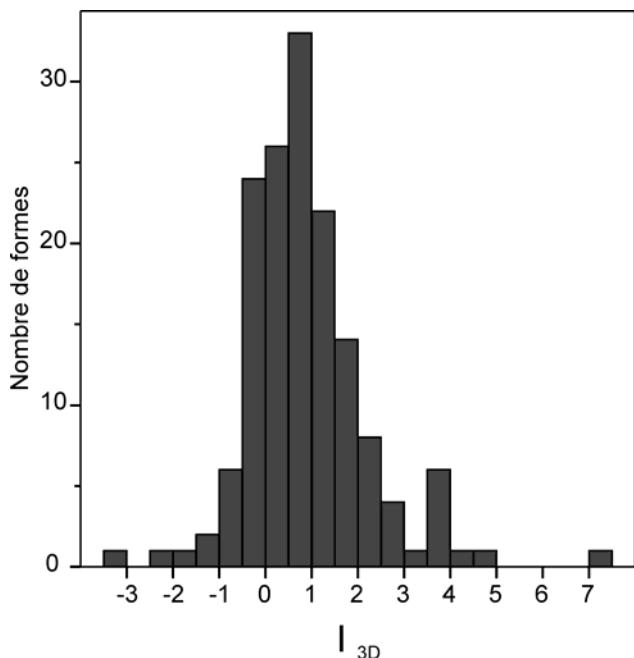


Figure 4-3 : histogramme de distribution des I_{3D} pour la population 3D (n=152 formes). Cet histogramme présente une médiane de 0.67 indiquant que la majorité des formes 3D ne sont pas dissociées de leur contrôle 2D blobs respectifs.

3. Réponses des cellules ICS aux stimuli random.

Au total, nous avons testé 46 cellules avec les stimuli random. Parmi ces 46 cellules, une importante majorité (n=40) montre une modulation de leur réponse selon le type de stimuli présenté soit 40 cellules ICS (ANOVA 3 facteurs, $p<0,05$). Les analyses présentées dans cette partie concernent ce pool de 40 cellules ICS.

3.1. Analyse par rang

Une analyse par rang entre les réponses aux stimuli 3D, 2D blobs et random nous permet de représenter les relations de codage entre les types de stimuli. Pour uniformiser ces écarts au sein de la population, nous avons normalisé les réponses pour chaque cellule. La Figure 4-4 représente l'analyse par rang effectuée sur la population ICS testée avec les stimuli random.

Les réponses moyennes aux stimuli 3D montrent une sélectivité dans la mesure où elles décroissent en fonction du rang des stimuli. Une telle sélectivité n'est pas présente pour les stimuli 2D blobs et random. En effet, les coefficients des droites de régression linéaire (R^2) pour les stimuli 3D sont supérieurs à ceux calculés pour les stimuli 2D blobs et random (les droites de régression ne sont pas représentées sur la figure mais $R^2_{3D}=-0,97$; $R^2_{2Dblobs}=-0,66$; $R^2_{random}=-0,23$). Toutefois, les réponses moyennes aux stimuli 3D et 2D blobs sont corrélées ; le coefficient de corrélation de Spearman indique que les réponses moyennes aux stimuli 2D blobs contrôles ne sont pas indépendantes du stimulus 3D présenté : si un stimulus 3D est préféré par la cellule alors la décharge augmente pour son contrôle 2D ($R_{Spearman}=0.82$, $p<0,001$). Une corrélation significative est également présente entre les réponses moyennes

aux stimuli 3D et random même si elle n'est pas aussi importante que celle observée entre les réponses moyennes aux stimuli 3D et 2D blobs ($R_{Spearman}=0.45$, $p=0,009$). En conséquence, lorsqu'un stimulus 3D est codé préférentiellement par la cellule alors son stimulus 2D blobs ou random aura tendance à être codé de manière préférentielle parmi la série de stimuli contrôles. Toutefois, il existe une différence du taux de décharge entre ces stimuli sinon les 3 courbes seraient parfaitement parallèles et se juxtaposeraient. Ce résultat intéressant peut s'expliquer par le fait que les stimuli 3D, 2D blobs et random ont des paramètres statistiques de 1^{er} ordre identiques, ainsi le neurone parmi une série de stimuli contrôle codera le stimulus ayant une combinaison de paramètre similaire au stimuli 3D correspondant. De plus, les stimuli diffèrent par leur contenu mais non par leur contour externe. Cette tendance à préférer le stimulus contrôle correspondant peut être due à la sélectivité du neurone pour le contour externe des stimuli.

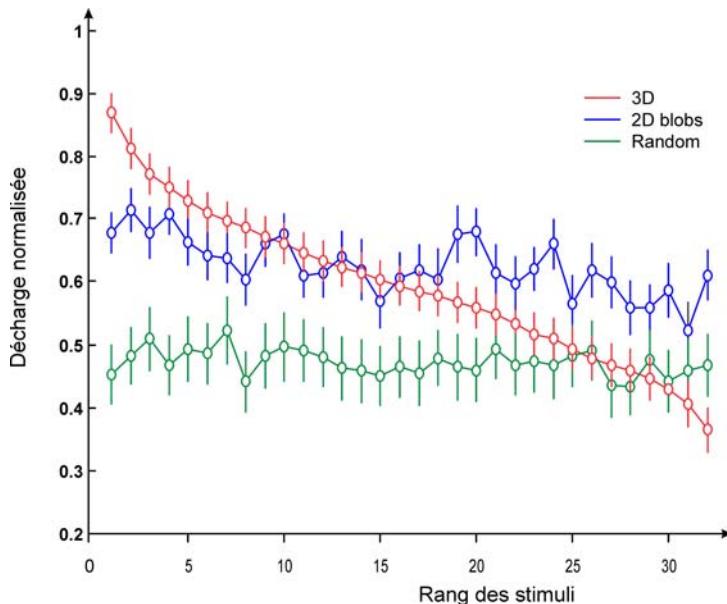


Figure 4-4 : Analyse par rang de la population ICS testée avec les stimuli random. Les réponses moyennes aux stimuli 3D et 2D blobs sont corrélées significativement ($R_{Spearman}=0,82$, $p<0,001$). Il en est de même pour les stimuli 3D et random ($R_{Spearman}=0,45$, $p=0,009$).

A l'aide de cette analyse par rang, nous avons démontré que la population ICS présente une préférence pour les stimuli 3D et que les réponses aux stimuli contrôles ne sont complètement indépendantes des réponses aux stimuli 3D. Afin de caractériser les différences de réponses au sein de la population ICS pour les stimuli 3D, 2D blobs et random, nous avons effectué des analyses multidimensionnelles (Analyse hiérarchisée par clusters et MDS).

3.2. Analyse hiérarchisée par clusters

Dans un premier temps, les réponses de chaque neurone de la population ICS aux stimuli 3D, 2D blobs et random sont standardisées afin de limiter toute variation interneurones. Un dendrogramme est ensuite réalisé à partir des réponses standardisées en utilisant la méthode de WARD (Figure 4-5). Sur cette figure, nous observons que les réponses moyennes de la population ICS se séparent à une distance euclidienne importante (distance, $d=170$) en deux clusters distincts (A1 et A2). Le cluster A2 comprend la totalité des stimuli random et A1, les stimuli 3D et 2D blobs. Cette séparation indique que la population ICS répond très différemment aux stimuli random par rapport aux deux autres types de stimuli, la distance euclidienne étant relativement importante. A une distance euclidienne moindre ($d=38$), le cluster A1 se divise en deux sous-groupes B1 et B2. Le cluster B1 regroupe la totalité des stimuli 3D et 5 stimuli 2D blobs alors que le cluster B2 est uniquement constitué de stimuli 2D blobs. L'analyse par cluster indique donc que la population des cellules ICS dissocie les stimuli selon leur contenu.

3.3. MDS

D'après l'analyse par clusters, les réponses des neurones ICS aux stimuli random semblent relativement différentes de celles aux stimuli 3D et 2D blobs. Une analyse multidimensionnelle de type MDS nous permet de visualiser les similarités de réponses entre les stimuli. La Figure 4-6 présente les représentations 2D d'une analyse multidimensionnelle effectuée sur les réponses moyennes de la population ICS. L'analyse des valeurs de stress permet de restreindre l'analyse à 3 dimensions.

Une meilleure dissociation entre les types de stimuli s'observe dans la représentation 2D de la Figure 4-6A. Les réponses aux stimuli random sont regroupées vers la droite du graphique avec des valeurs de la dimension 1 supérieures à 0.8. Quant à la dissociation entre les réponses aux stimuli 2D blobs et 3D, elle s'effectue dans la dimension 2 ; les réponses aux stimuli 3D présentent des valeurs dans cette dimension positive alors que les réponses aux stimuli 2D blobs ont des valeurs en majorité négatives. Cette dissociation est confirmée dans la Figure 4-6C (dim 3 vs dim 2). La dimension 3 ne nous permet pas de dissocier les réponses moyennes entre elles.

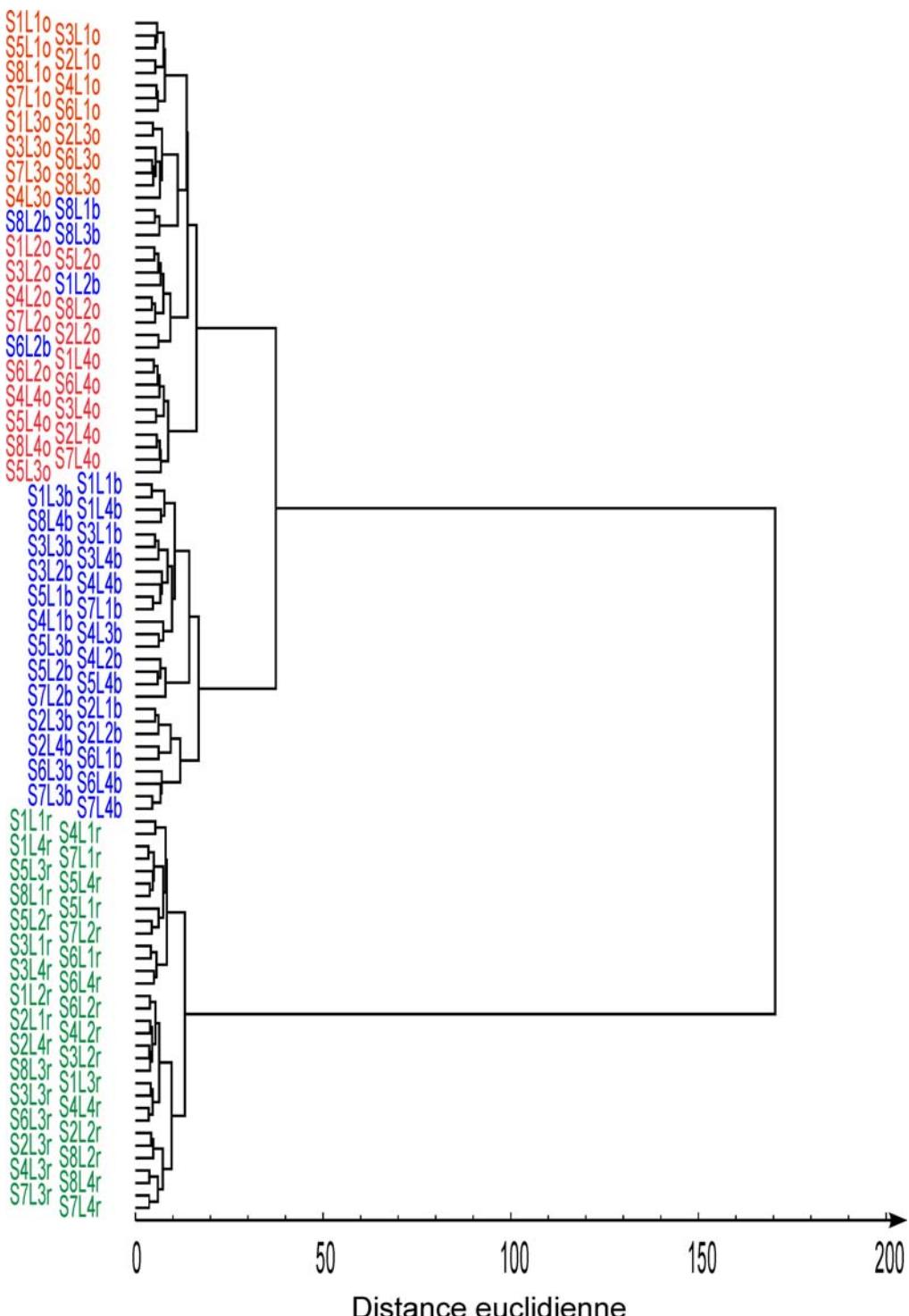


Figure 4-5 : Dendrogramme effectué sur les réponses standardisées de la population ICS testée avec les stimuli 3D, 2D blob et random (n=40 cellules). Chaque ligne horizontale représente un stimulus. Le code couleur correspond à chaque type de stimulus (en rouge pour les stimuli 3D, en bleu pour les stimuli 2D blob et en vert pour les stimuli random). Nom des stimuli : ex : S1L4r : “S” pour le numéro de la forme, ici Forme n°1 (de S1 à S8), “L” pour la direction d’illumination (L1 : haut, L2 : gauche, L3 : droite et L4 : bas), “r/o/b/p” pour le type des stimuli (o : 3D, b : 2D blob, r : random et p : postérisé). Les réponses moyennes se regroupent selon le contenu 3D, 2D blobs ou random des stimuli.

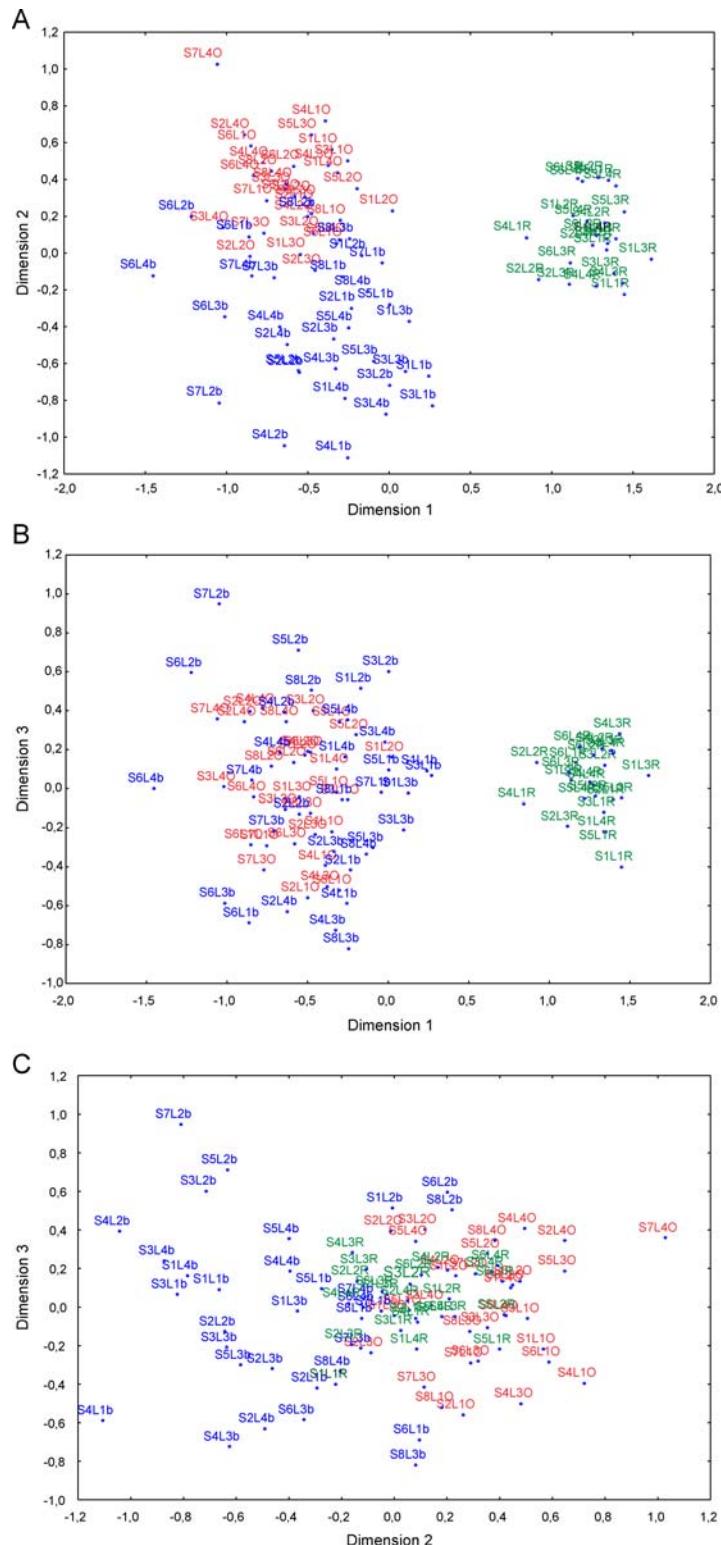


Figure 4-6 : Représentations 2D des réponses moyennes de la population ICS (n=40) testée avec les stimuli random (A : dim 1 vs dim 2, B : dim 1 vs dim 3 et C : dim 2 vs dim 3). Chaque point correspond à un stimulus ; le code couleur correspond au type des stimuli comme celui utilisé dans la Figure 4-5. Ce type d'analyse est limitée à 88 observations or notre set d'images comprend 96 stimuli différents. D'après l'analyse par clusters les réponses aux stimuli random sont relativement proches entre elles et quels que soient les 24 stimuli random pris en compte, le résultat des MDS est identique ; nous avons donc retiré 8 stimuli random pour réaliser ces MDS.

3.4. Réponses individuelles des cellules 3D aux stimuli random.

Les 40 cellules ICS testées avec les stimuli random regroupent trois groupes distincts de neurones, chacun ayant un pattern de sélectivité différent; 12 cellules 3D, 13 cellules 2D blobs et 15 cellules 2D. Sur la totalité de la population ICS, une large majorité ($n=36$) montrent une modulation significative de leur réponse entre les stimuli random et 3D; 28 cellules montrent une diminution de leur réponse neuronale pour les stimuli random et 8 cellules montrent une augmentation significative de leur réponse (TUKEY, $p<0,05$). La totalité des cellules 3D testées avec les stimuli random présentent un niveau de réponse significativement inférieur pour les stimuli random (TUKEY, $p<0,05$). La Figure 4-7 présente un exemple de cellule 3D testé avec les stimuli random. Il s'agit de la même cellule présentée dans la Figure 4-1A.

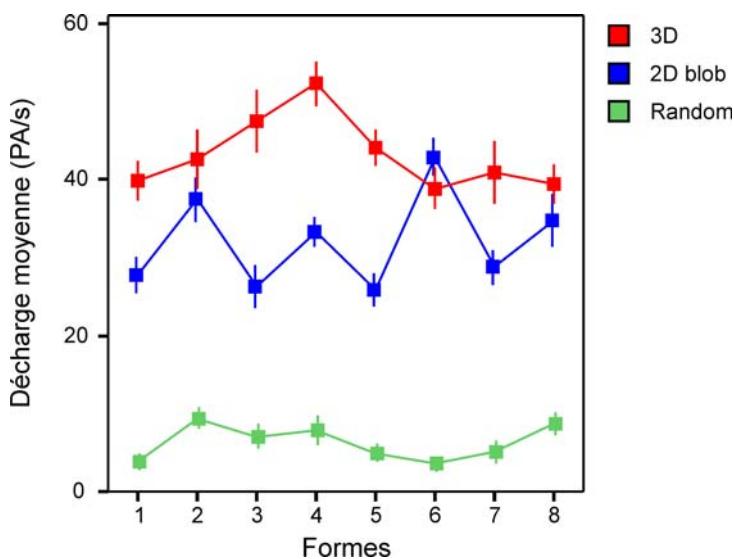


Figure 4-7 : Exemple de cellule 3D (T-SFS30.31). Ce neurone montre un faible taux de décharge pour les stimuli random.

De la même manière qu'avec les stimuli 2D blobs, nous avons calculé un index de dissociation (I_{rand}) entre les stimuli 3D et les stimuli random. Au total, 12 cellules 3D ont été testées avec les stimuli random soit 96 formes 3D (pour chaque cellule, 8 formes ont été testées). La Figure 4-8 présente les index de dissociation pour les stimuli random et les stimuli 2D blobs au sein de cette population 3D. Cette figure nous indique qu'une grande majorité des formes 3D sont dissociées de leur forme random ($n= 79/96$ formes soit 82%, $p<0,05$). Au niveau des formes 2D blobs, la dissociation est moins importante puisque seules 28 formes sont dissociées soit 29% ($p<0,05$, $I_{3D} >0$). Pour une importante majorité des formes, les cellules 3D ne différencient pas les stimuli 3D des 2D blobs. Ce graphique nous indique également que 26 formes sont doublement dissociées ($I_{3D} >1.2$ & $I_{rand} >1.2$, $p<0,05$). Les 12

cellules 3D ont au moins une forme 3D doublement dissociée. Le nombre maximum de formes doublement dissociées est de 4 pour une cellule 3D.

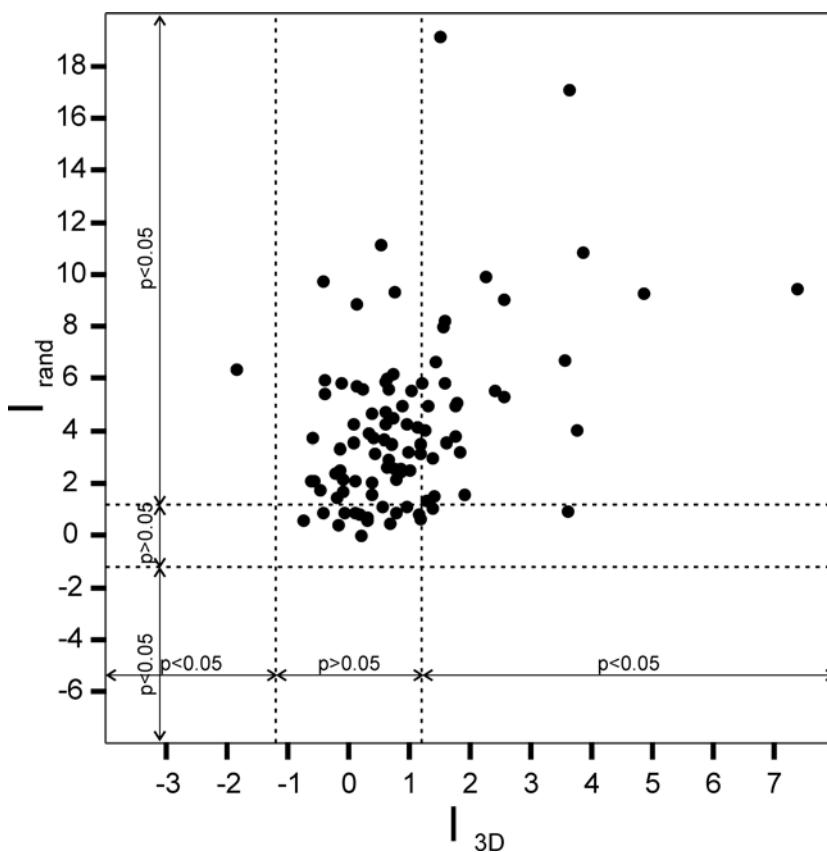


Figure 4-8 : Index de dissociation I_{rand} en fonction de I_{3D} pour la population 3D testée avec les stimuli random (n=12 cellules et 8 formes soit 96 points). Chaque point représente une forme 3D. Les lignes pointillées indiquent le seuil significatif du test TUKEY ($p<0,05$ et $p>0,05$) entre les formes 3D et 2D blobs, et entre les formes 3D et random.

4. Réponses des cellules ICS aux stimuli postérisés

4.1. Population de base

Au cours des enregistrements, 73 cellules parmi les 119 ont été testées avec des stimuli postérisés; 44 cellules ont été enregistrées sur Tima et 29 sur Zrem. Les régions d'enregistrements des cellules chez Tima correspondent à celles enregistrées avec le précédent protocole (stimuli random) et pour Zrem, la région corticale coïncide avec celle explorée pendant la première étude sur les textures naturelles.

Un test ANOVA à trois facteurs permet de vérifier l'importance du facteur TYPE sur la réponse du neurone ; 73 % (53 cellules) de la population totale testée avec les stimuli postérisés sont des cellules ICS (ANOVA, effet principal, $p<0,05$). Ces 53 cellules ICS regroupent 11 cellules 3D, 26 cellules 2D blobs et 18 cellules 2D.

4.2. Analyse par rang.

Afin de constituer un pool commun de neurones entre celles enregistrées avec le set S2 et S3, nous avons ciblé nos analyses de population sur les 4 formes communes à ces 2 sets ; les formes n°3, 4, 5 et 6. Le pool de stimuli sur lequel sont effectuées les analyses de

population comprend donc 48 images différentes ; 16 stimuli 3D, 16 stimuli 2D blobs et 16 stimuli postérisés. La Figure 4-9 présente l’analyse par rang effectuée sur les réponses de la population ICS.

Comme nous l’observions avec la population ICS testée avec les stimuli random, une sélectivité au niveau des réponses pour les stimuli 3D est présente et se traduit par une augmentation de la réponse moyenne pour les stimuli préférés. La décharge moyenne décroît en fonction du rang des stimuli et les réponses moyennes aux stimuli 3D et 2D blobs sont corrélées significativement ($R_{Spearman} = 0.78$, $p < 0,001$). Les réponses moyennes aux stimuli postérisés sont corrélées significativement avec celles correspondant aux stimuli 3D ($R_{Spearman} = 0.72$, $p = 0.002$). Il existe donc une correspondance entre les réponses aux stimuli 3D et postérisés indiquant qu’une réponse pour le stimulus postérisé a une tendance à être plus importante si le stimulus 3D est préféré par le neurone.

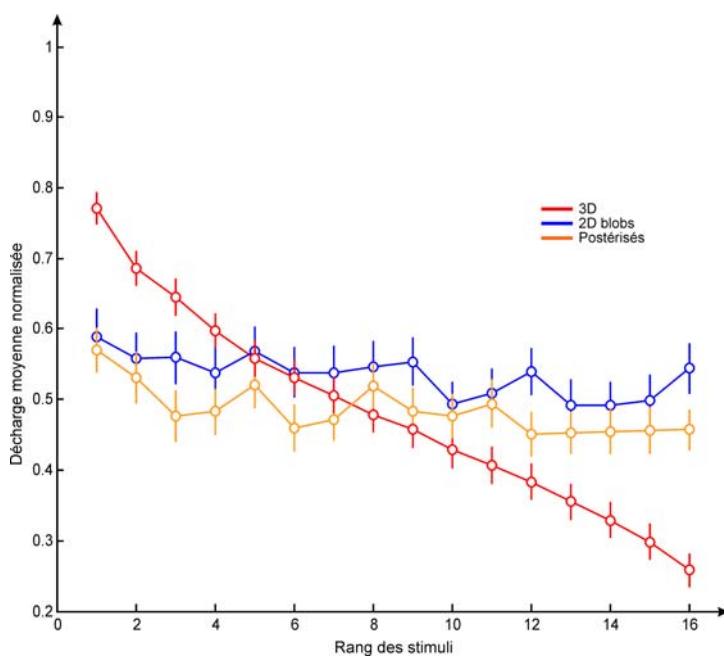


Figure 4-9 : Analyse par rang sur les réponses normalisées de la population ICS aux stimuli 3D, 2D blobs et postérisés. Les barres d’erreurs indiquent les écarts-types. Les réponses aux stimuli 3D et 2D blobs sont corrélées significativement ($R_{Spearman} = 0.78$, $p < 0,001$), idem pour les réponses aux stimuli 3D et postérisés ($R_{Spearman} = 0.72$, $p = 0.002$).

4.3. Analyse hiérarchisée par cluster

Afin de tester les similarités de réponses entre les stimuli 3D, 2D blobs et postérisés, nous avons effectué deux analyses multidimensionnelles (Analyse hiérarchisée par cluster et MDS). La Figure 4-10 représente l’arbre obtenu suite à cette analyse hiérarchisée par cluster sur les réponses de la population ICS. Cette analyse hiérarchisée révèle plusieurs niveaux de séparation. Nous nous intéresserons qu’au deux premiers niveaux (A et B). A une distance euclidienne de 43, l’arbre se divise en 2 clusters A1 et A2. Le cluster A2 est uniquement composé de stimuli 2D blobs. Seuls les stimuli 2D blobs de la forme n°6 sont exclus de ce

cluster et se retrouvent dans le cluster A1. Le cluster A1 se divise ensuite ($d=28$) en 2 sous-groupes B1 et B2. Le cluster B1 est caractérisé par la présence de la totalité des stimuli postérisés et de 2 stimuli 3D (S5L2 et S5L4). Le cluster A1 comprend, quant à lui, une majorité de stimuli 3D et les 4 stimuli 2D blobs de la forme n°6. Hormis quelques exceptions, l’analyse hiérarchisée par clusters nous indique que la population ICS dissocie les différents types de stimuli. Cette analyse nous indique également que les réponses moyennes aux stimuli 3D et postérisés ($d=28$) sont relativement proches en comparaison avec les réponses moyennes aux stimuli 2D blobs ($d=43$).

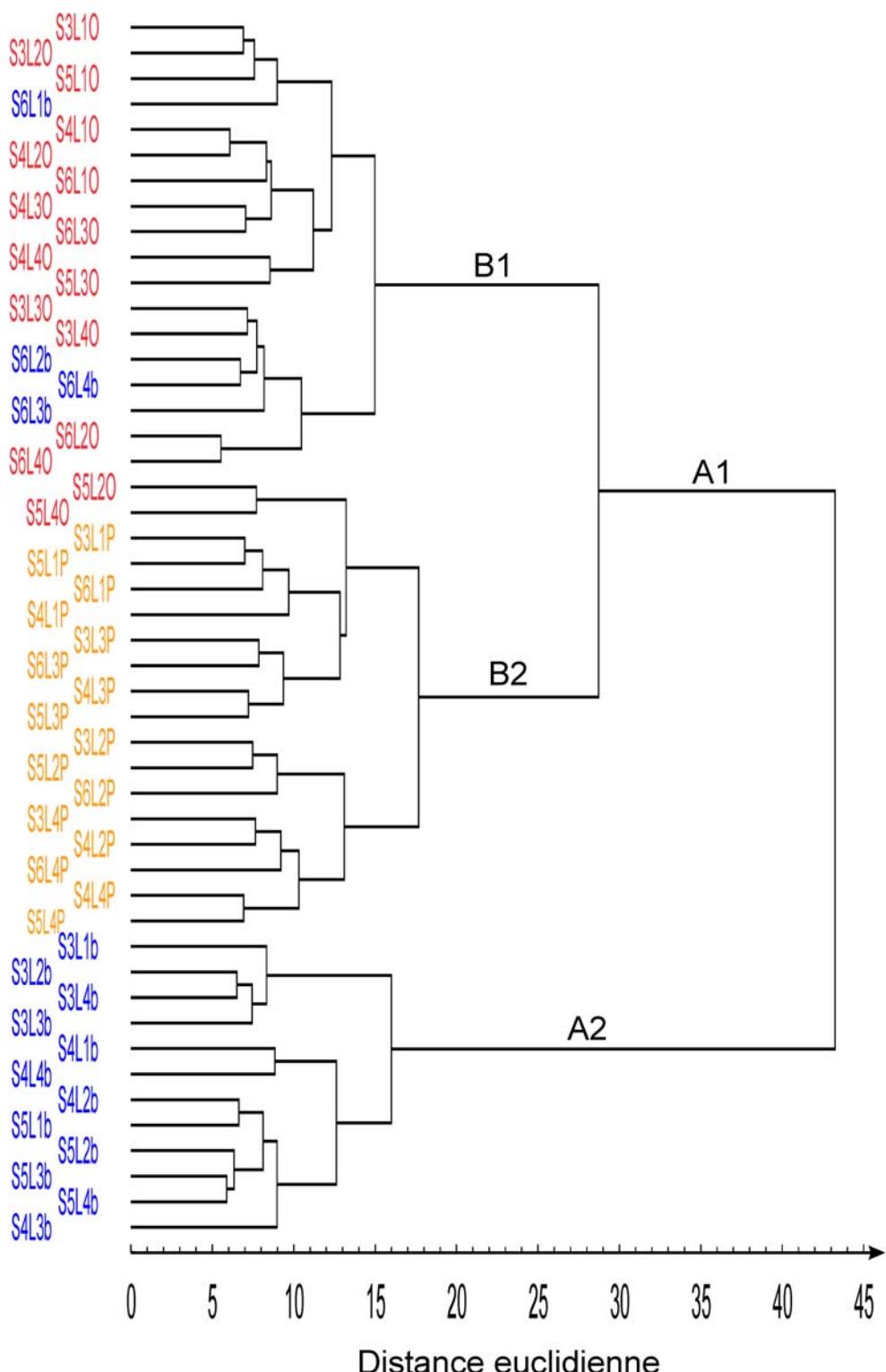


Figure 4-10 : Dendrogramme effectué sur les réponses standardisées de la population ICS testée avec les stimuli 3D, 2D blob et postérisés (n=53 cellules). Chaque ligne horizontale représente un stimulus. Le code couleur correspond à chaque type de stimulus (en rouge pour les stimuli 3D, en bleu pour les stimuli 2D blob et en orange pour les stimuli postérisés). Les réponses moyennes se regroupent selon le contenu 3D, 2d blobs ou postérisés des stimuli.

4.4.MDS

La Figure 4-11 contient les représentations bidimensionnelles de la MDS effectuée sur la population ICS testée avec les stimuli postérisés ($n=53$ cellules). Elle permet de comparer les réponses aux stimuli 3D et celles aux stimuli 2D blobs et postérisés correspondants. La représentation de la Figure 4-11A dissocie le mieux les stimuli selon leur contenu ou type. La dimension 1 permet de dissocier les stimuli 3D et postérisés des stimuli 2D blobs, soit le cluster A1 du cluster A2. Les stimuli présents dans le cluster A1 ont des valeurs majoritairement négatives alors que les stimuli 2D blobs ont des valeurs exclusivement positives dans cette dimension. La dimension 2 permet de dissocier les stimuli 3D (valeurs négatives en majorité) des stimuli postérisés (valeurs positives sauf pour le stimulus S3L1P). La dimension 3 ne permet pas d'observer une meilleure dissociation entre les différents types de stimuli (cf. Figure 4-11B et C).

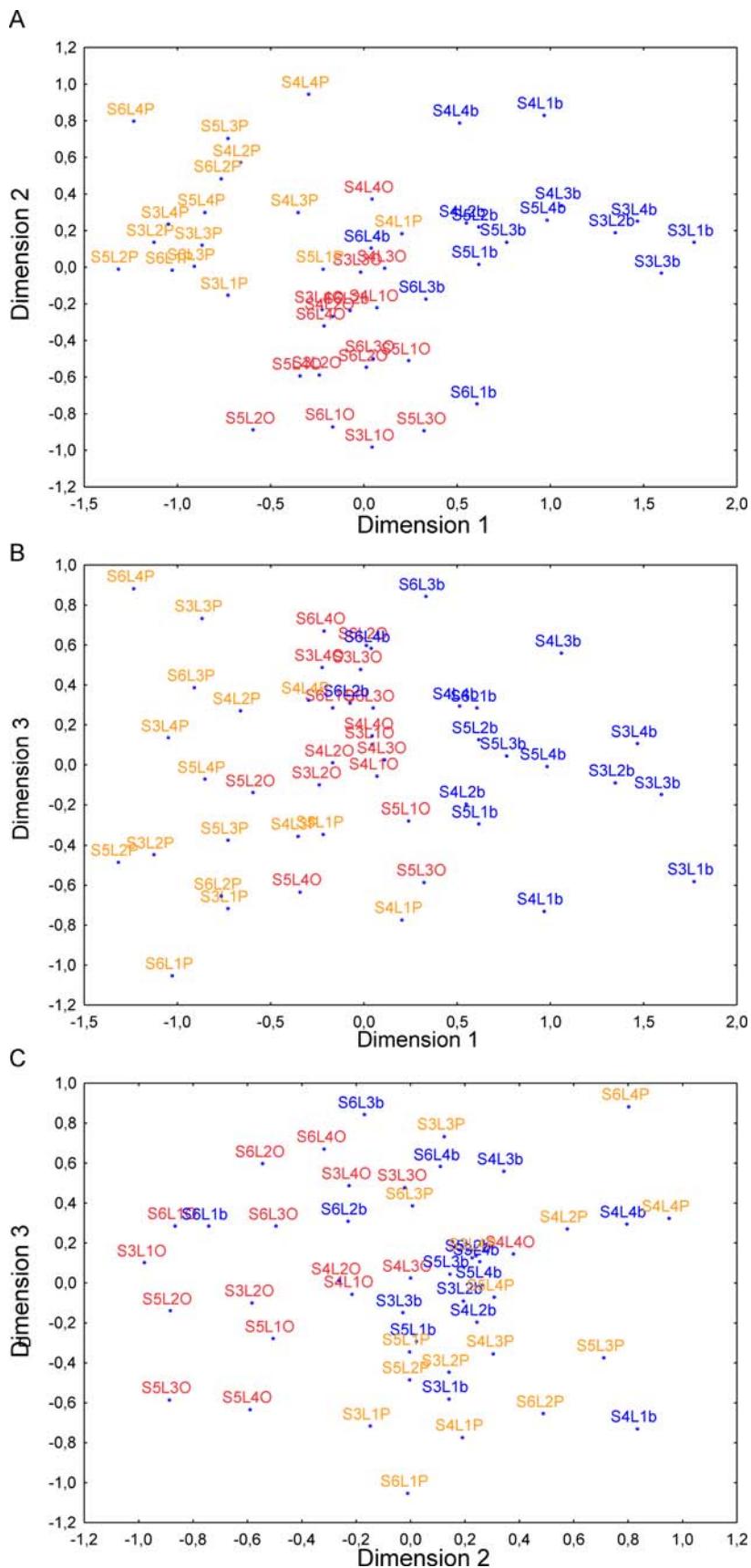


Figure 4-11 : Représentations 2D des réponses moyennes de la population ICS ($n=40$) testée avec les stimuli random (A : dim 1 vs dim 2, B : dim 1 vs dim 3 et C : dim 2 vs dim 3). Chaque point correspond à un stimulus ; le code couleur correspond au type des stimuli comme celui utilisé dans la Figure 4-5.

4.5. Réponses individuelles des cellules 3D aux stimuli postérésés.

Suite aux résultats du test ANOVA, nous avons noté que 11 cellules 3D ont été testées avec les stimuli postérésés. Ce paragraphe a pour but d'étudier plus précisément les réponses de ces neurones 3D aux stimuli postérésés. Parmi les 11 neurones 3D, 2 cellules montrent une augmentation significative de leur réponse pour les stimuli 3D par rapport aux stimuli postérésés (TUKEY, $p<0,05$), 7 cellules montrent une réponse moyenne identique pour les deux types de stimuli et 2 cellules ont une réponse moyenne supérieure pour les stimuli postérésés. Seules deux cellules montrent donc une réponse moyenne supérieure pour les stimuli 3D par rapport aux stimuli 2D blobs et aux stimuli postérésés. Les cellules de la Figure 4-12 représentent chacune un exemple de ces catégories. La cellule A a une réponse moyenne supérieure pour les stimuli postérésés par rapport à celle pour les stimuli 3D pour les formes n°4 et n°6. la décharge de ce neurone est identique pour les deux autres formes. La cellule B a un taux de décharge plus important pour les stimuli 3D des formes n°3 et n°6 par rapport aux stimuli 2D blobs ou postérésés. Quant à la cellule C, elle présente un même niveau de réponse pour les stimuli 3D et postérésés pour les formes n°4, n°5 et n°6.

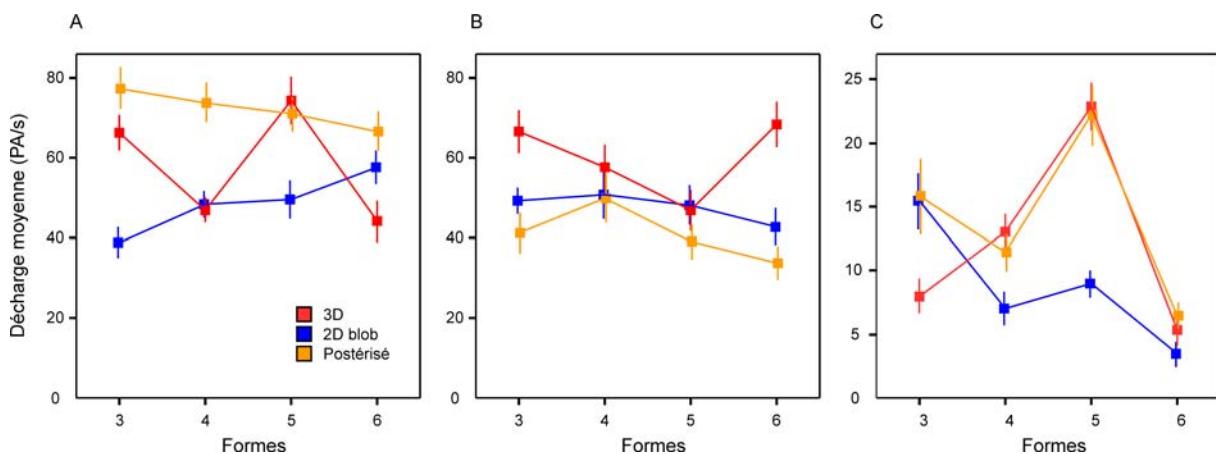


Figure 4-12 : Trois exemples de cellules 3D testées avec les stimuli postérésés (A :Z-SFS2.11, B :Z-SFS4.11, C :Z-SFS10.11). La décharge moyenne pour chacune des formes est représentée en fonction du type des stimuli. Les 4 illuminations sont cumulées. La cellule A présente une augmentation significative de son taux de décharge pour les stimuli postérésés, la cellule B pour les stimuli 3D et la cellule C présente un même niveau de réponse entre les stimuli postérésés et 3D.

Comme le montrent ces trois exemples de cellules, la majorité des cellules 3D présentent une interaction significative entre le type des stimuli et la forme testée ($n=9/11$). Suites aux analyses entre les réponses aux stimuli 3D et 2D blobs, nous avons caractérisé pour chaque neurone les formes 3D qui engendrent une différence de réponse par rapport aux stimuli 2D blobs. De la même manière, nous pouvons vérifier pour ces formes 3D dissociées si la réponse aux stimuli postérésés diffère.

Pour chaque cellule 3D, nous considérons uniquement la forme 3D optimale et nous comparons son niveau de réponse à celui pour le contrôle postérisé à l'aide d'un test post-hoc TUKEY. Le calcul d'un rapport I_{post} permet également de quantifier la différence de réponse entre les 2 types de stimuli. La cellule A de la Figure 4-12 a comme forme 3D optimale la forme n°3 et elle n'est pas différenciée lorsqu'elle se présente sous la forme des stimuli postérisés ($I_{\text{post}} = -0.80$). Pour la cellule B, la forme n°6 est la forme 3D optimale ($I_{3D} = 2.38$). Cette forme se dissocie également des stimuli postérisés ($I_{\text{post}} = 3.23$). Toutefois, cette différence n'est présente uniquement que pour les formes n°3 et n°6. Pour la cellule C, sa forme 3D optimale est la forme n° 5. Pour cette forme, la cellule C présente un niveau de réponse similaire pour les stimuli contrôles postérisés ($I_{\text{post}} = 0.12$).

La Figure 4-13 reprend les index de dissociation I_{3D} et I_{post} pour chaque forme testée pour la population 3D. Des lignes de préférence indiquent le seuil de significativité du test post-hoc TUKEY. Ainsi, 56 formes 3D ont été testées à la fois avec les stimuli 2D blobs et postérisés ; 3 neurones testés avec 8 formes et 8 neurones avec 4 formes. Parmi ces 56 formes 3D, 19 formes 3D se dissocient de leurs contrôles 2D blobs et 6 de leurs contrôles postérisés. La majorité des formes n'est pas dissociée puisque 28 d'entre elles présentent un même niveau de réponse quel que soit le type des stimuli. D'après ce graphique, seules 4 formes 3D sont doublement dissociées de leurs contrôles 2D blobs et postérisés respectifs. Trois cellules 3D opèrent à ces doubles dissociations.

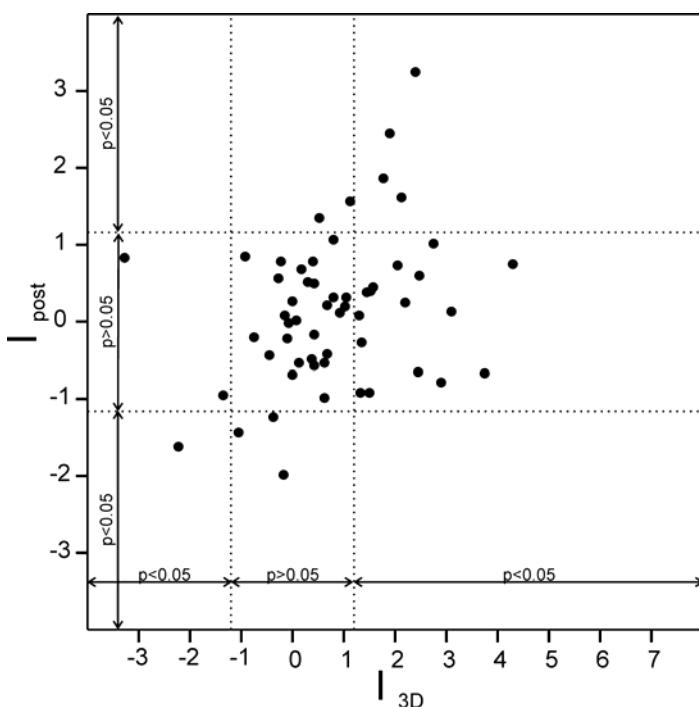


Figure 4-13 : I_{post} en fonction de I_{3D} pour 56 formes 3D testées avec les stimuli postérisés. Les lignes pointillées schématisent les seuils de significativité des I_{post} et I_{3D} . Seules 4 formes sont doublement dissociées.

5. Synthèse des principaux résultats

Au cours de cette partie, nous nous sommes attachés à caractériser le codage de la 3D créé par les ombrages au sein de l'aire V4. Pour cela nous disposons de différents types de stimuli contrôles. De manière générale, le codage de la 3D s'avère complexe et dépend de la forme de l'objet considéré. Si l'on considère les cellules individuelles, peu préfèrent systématiquement la 3D ou la 2D. Néanmoins, une dissociation des indices 3D et 2D semble se dégager au niveau de la population.

En détail, plusieurs résultats sont à retenir :

- Une majorité des cellules de V4 enregistrées présente une modulation de leur taux de réponse selon le type de stimuli présenté soit 78% de la population totale. Ces cellules sont dites ICS (n=93)
- Parmi les cellules ICS, 60 cellules montrent un taux de décharge moyen différent entre les stimuli 3D et 2D blobs : 23 cellules montrent un niveau moyen de réponse plus important pour les stimuli 3D, 37 pour les stimuli 2D blobs (TUKEY, $p<0.05$). La première catégorie de cellules est appelée « cellule 3D », la seconde « cellule 2D blob ».
- La différence de réponses entre les stimuli 3D et les stimuli 2D blobs varient selon la forme présentée. Une majorité des cellules 3D montrent une interaction significative entre le type du stimulus et la forme présentée (13/23 cellules). Les cellules 3D dissocient en moyenne 2,5 formes de leurs contrôles 2D blobs respectifs lorsqu'elles sont enregistrées avec 8 formes et 1,75 avec seulement 4 formes.
- Les analyses multidimensionnelles (analyse hiérarchisée par clusters et MDS) nous indiquent que les réponses de la population ICS se “clusterisent” en fonction du type des stimuli. Les réponses aux stimuli random sont nettement dissociables des réponses aux autres types de stimuli.
- Au niveau des réponses individuelles, les cellules 3D dissocient bien leur stimulus 3D de leur stimulus random respectif mais peu de leur stimulus postérisé.

Nous avons mis en évidence dans cette partie qu'une population de neurones de V4, la population ICS, dissocie les stimuli 3D de leurs contrôles 2D respectifs (2D blobs, random et postérisé). Au cours du paragraphe suivant, nous allons tester si une quelconque sélectivité aux formes tridimensionnelles est présente au sein de cette population ICS et notamment au de la sous-population de neurones 3D.

B. SELECTIVITE AUX FORMES TRIDIMENSIONNELLES

1. Résultats du test ANOVA

Afin de tester la sélectivité des neurones ICS aux formes qui leur sont présentées, nous reprenons les résultats du test ANOVA à 3 facteurs (forme testée, type et direction d'illumination) sur les réponses des neurones ICS. Au total, 64 cellules ICS présentent un taux de décharge variable selon la forme présentée, soit 69% de la population ICS (ANOVA effet principal, $p<0,05$). Parmi cette population sélective, une grande majorité des neurones voit leur pattern de sélectivité se modifier en fonction de la forme et du type des stimuli puisque 77% des cellules présentent une interaction significative entre la forme et le type des stimuli ($n=49$).

Dans ce paragraphe, nous nous intéresserons uniquement à la sélectivité aux formes tridimensionnelles de la population ICS. Pour cela, nous avons effectué un test ANOVA 2 facteurs sur les réponses de la population ICS aux stimuli 3D (ANOVA Forme X Direction d'illumination). Au total, 45 cellules ICS ont une réponse neuronale sélective à la forme tridimensionnelle présentée soit 48% de la population ICS totale (ANOVA effet principal, $p<0,05$). Toutefois, la majorité de ces cellules présentent également une interaction significative entre la forme et la direction d'illumination ($n=26$). Parmi cette population de cellule sélective à la forme 3D présentée, nous dénombrons 12 cellules 3D. La Figure 4-14 montre un neurone 3D sélectif aux formes tridimensionnelles. Ce neurone présente une augmentation de son taux de décharge pour les formes n°6 et n°7 et une diminution significative pour les formes n°1 et n°4 (TUKEY, $p<0,05$). La Figure 4-15 présente un exemple de cellule invariante à la forme 3D présentée.

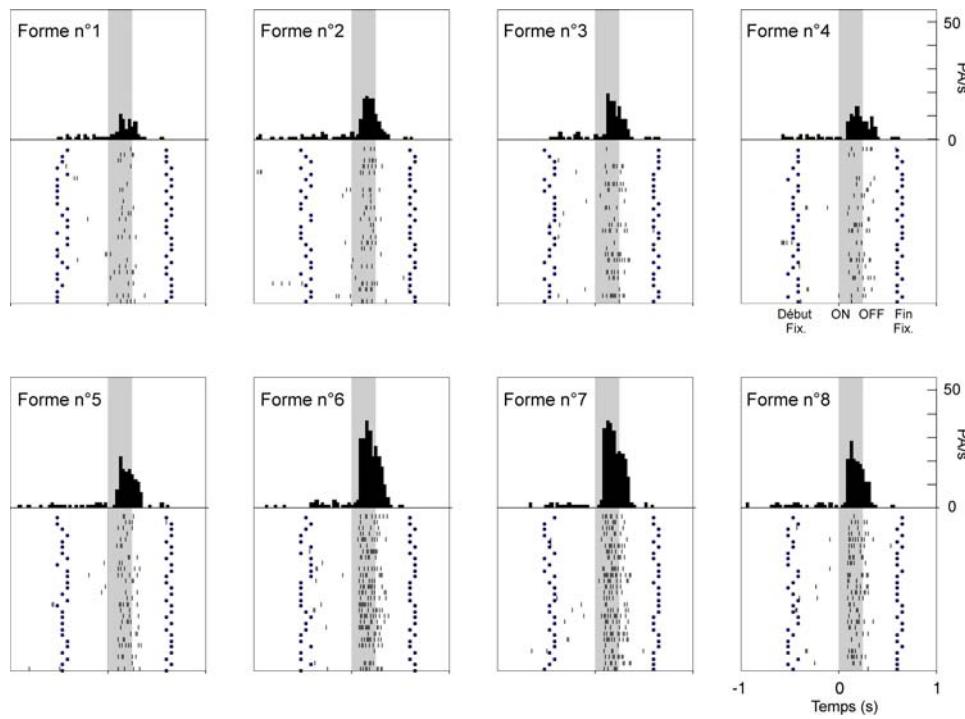


Figure 4-14 : Exemple de neurone sélectif aux formes 3D (T-sfs5.21). Les 4 illuminations sont cumulées. Les carrés bleus indiquent le début et la fin de la fixation (Début fix et Fin fix). Le rectangle gris schématise la durée de présentation des stimuli. Tous les essais sont alignés sur le temps 0. Cette cellule code préférentiellement les formes 3D n° 6 et 7. Les intervalles des histogrammes ont une largeur de 25 ms.

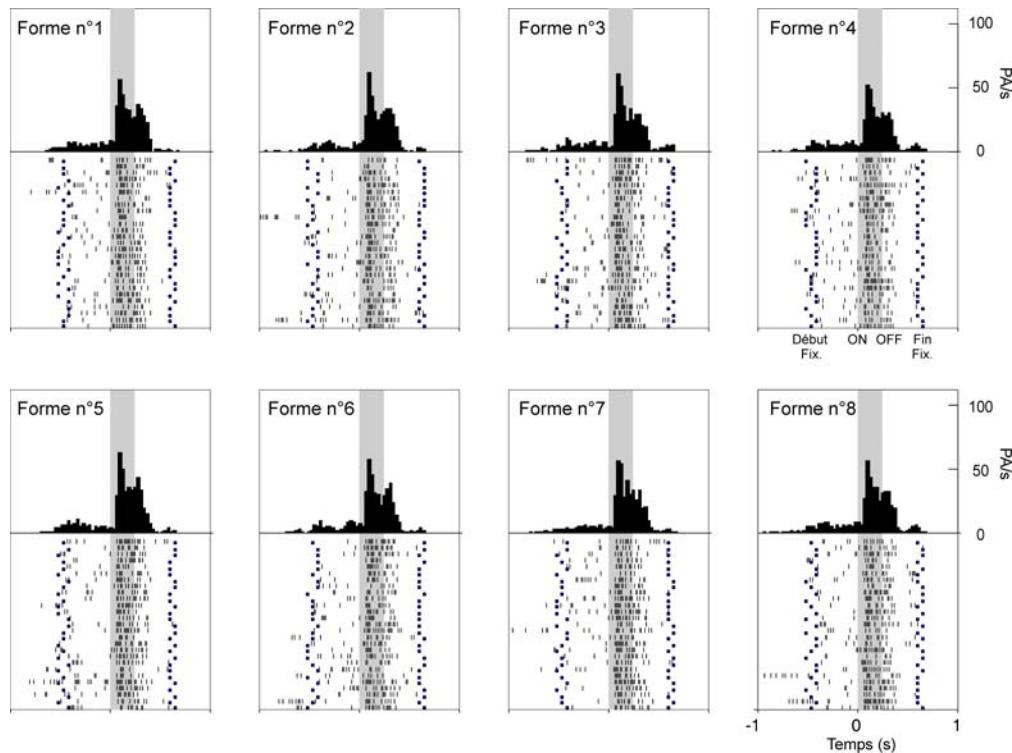


Figure 4-15 : Cellule invariante à la forme tridimensionnelle présentée (T-SFS13.31). Mêmes conventions que la Figure 4-14. Cette cellule présente un taux de réponse similaire pour les 8 formes 3D testées.

2. Index de sélectivité (SMI)

Afin mieux caractériser une éventuelle sélectivité au niveau des réponses neuronales aux différentes formes 3D au sein de la population ICS, nous avons calculé un index de sélectivité comme Vogels et Biederman le présentent dans leur étude de 2002. Pour une direction d'illumination donnée (ici lorsque les formes sont illuminées par le haut), nous choisissons la forme qui engendre une réponse moyenne maximale (R_{maxi}) et celle qui engendre une réponse moyenne minimale (R_{mini}). L'index de modulation à la forme (SMI) se définit comme le quotient entre la différence des réponses moyennes ($R_{\text{maxi}} - R_{\text{mini}}$) et la moyenne des déviations standard de ces deux réponses. Si l'index a une valeur égale à 1 alors le neurone est peu sélectif à la forme 3D et aura une réponse similaire pour toutes les formes 3D qui lui sont présentées. Les neurones de la Figure 4-14 et de la Figure 4-15 ont un SMI de 0,90 et de 3,91 respectivement. La distribution des SMI calculés sur la population ICS est représentée sous forme d'histogramme (Figure 4-16).

Cet histogramme présente une valeur médiane de 1,57 indiquant que la population ICS montre une faible modulation de leur taux de décharge selon la forme tridimensionnelle présentée. Parmi cette population ICS, nous avons auparavant catégorisé une sous population de 23 cellules 3D. Les exemples des Figures 4-14/15 font parti de cette catégorie de cellules. Si nous ne prenons en compte que cette sous-population 3D alors l'histogramme des valeurs a une médiane de 1,75 indiquant qu'une faible modulation est également présente au sein cette sous-population.

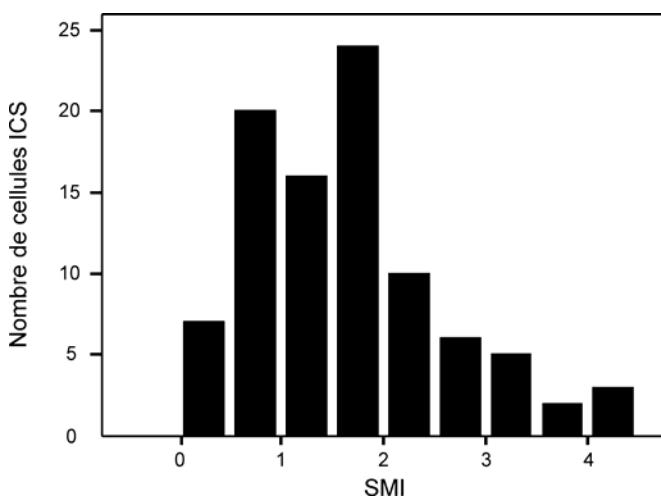


Figure 4-16 : Distribution des index de sélectivité à la forme (SMI) pour la population ICS ($n=93$). Les intervalles ont une largeur de 0,5.

3. Synthèse

Au cours de cette partie sur la sélectivité aux formes tridimensionnelle, nous avons remarqué que la population ICS présente une faible modulation de leur taux de décharge en

fonction de la forme présentée. Les formes utilisées (cf. Méthodes), de par leur pourtour, sont très proches entre elles et cette faible modulation est donc peut-être due à une faible variabilité entre les formes tridimensionnelles utilisées.

C. SELECTIVITE A LA DIRECTION D'ILLUMINATION.

1. Population d'étude

Un paramètre important au niveau de notre set d'image est la direction d'illumination. En effet, la direction d'illumination induit un gradient d'illumination sur l'objet éclairé et définit, grâce aux ombrages, la forme tridimensionnelle de l'objet. Ce gradient d'illumination est présent au sein des stimuli 3D mais également au sein des stimuli 2D blobs. Les analyses statistiques présentées ici ne concernent donc que les réponses aux stimuli de ces deux catégories.

2. Résultats ANOVA

Au niveau de la population totale, 66 cellules montrent une modulation de leur réponse en fonction de la direction d'illumination (ANOVA effet principal, $p<0,05$) soit 55% de la population totale ; 59 sont des cellules ICS. La Figure 4-17 présente deux exemples de cellules ICS dont les décharges neuronales sont influencées par la direction d'illumination.

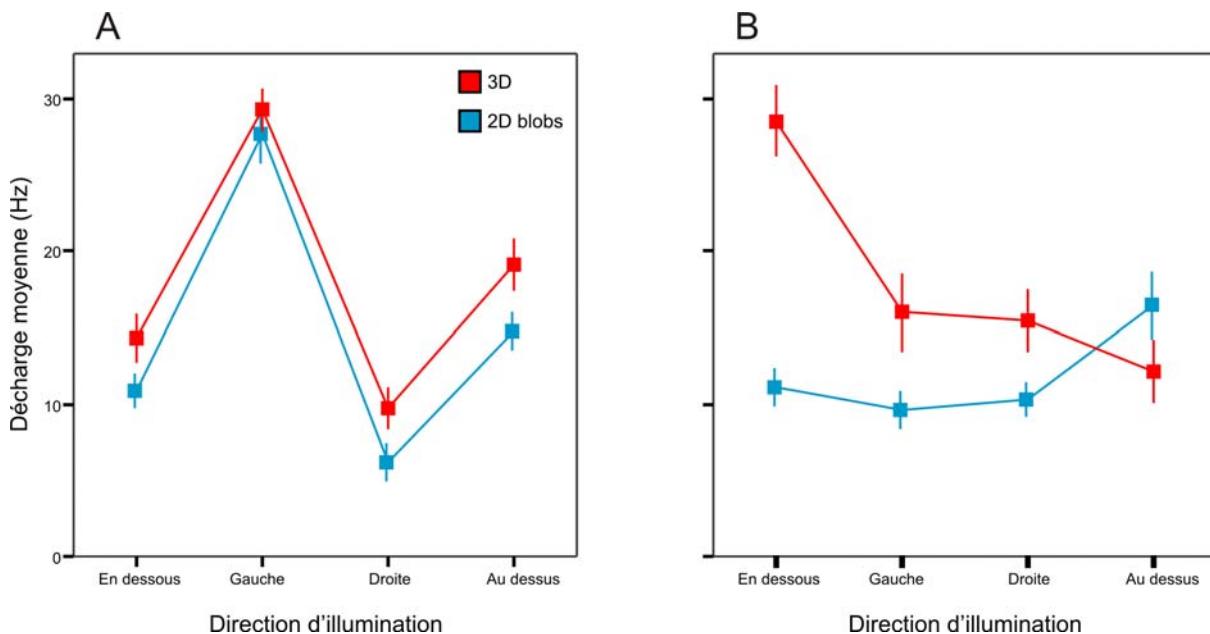


Figure 4-17 : A et B. Cellules 3D influencées par la direction d'illumination. Les réponses moyennes par direction d'illumination sont calculées en cumulant les réponses du neurone pour toutes les formes. Les barres d'erreurs représentent les esm. Le neurone A (T-SFS23.31) montre un même pattern de sélectivité que les stimuli soit en 3D ou 2D blobs ; sa décharge moyenne augmente significativement lorsque que l'illumination vient de la gauche. A l'inverse, le neurone B (Z-SFS2.11) montre des patterns de sélectivité différents selon le type 3D ou 2D blob des stimuli.

Pour la cellule A, les réponses moyennes pour les stimuli 3D et 2D blobs augmentent significativement lorsque que l'illumination provient de la gauche. A l'inverse, la réponse

neuronale décroît pour des illuminations provenant d'au-dessous ou de la droite. Ce pattern de sélectivité se retrouve à la fois pour les stimuli 3D et pour les stimuli 2D blobs. Ce neurone a une direction d'illumination préférée, lorsque l'illumination provient de la gauche (TUKEY, $p<0.05$). Parmi les cellules ICS sélectives à la direction d'illumination, 32 neurones présentent un pattern de sélectivité similaire à ce neurone (ANOVA, interaction TYPE X Direction d'illumination, $p>0,05$) soit 54% de la population. La cellule de l'exemple B est également influencée par la direction de l'illumination mais montre un pattern de sélectivité différent selon le type 3D ou 2D blobs des stimuli. La décharge neuronale augmente lorsque les stimuli 3D sont éclairés par en dessous alors qu'elle se différencie des autres directions d'illumination pour les stimuli 2D blobs lorsqu'elle provient d'au-dessus. Comme cette cellule, 27/59 cellules ICS présentent une interaction significative entre le type et la direction d'illumination (ANOVA, interaction TYPE X Direction d'illumination, $p<0,05$).

Nous nous restreignons dorénavant à l'influence de la direction d'illumination sur le codage des stimuli 3D, notamment au niveau des 23 cellules de la sous-population de cellules 3D (cf. paragraphe A.1). Parmi ces cellules 3D, 14 cellules montrent un effet principal significatif de la direction d'illumination sur leurs réponses neuronales (ANOVA, effet principal, $p<0,05$). Parmi ces 14 cellules 3D, seuls 5 neurones montrent une interaction significative entre la forme et la direction d'illumination. La Figure 4-18 présentent 2 exemples de neurones 3D. Le neurone de la Figure 4-18B voit son taux de décharge augmenter lorsque l'illumination vient de la gauche. A l'inverse, le neurone de la Figure 4-18A présente un même taux de décharge quelle que soit la direction d'illumination. Ces neurones présentent ce même pattern de sélectivité quelle que soit la forme considérée (ANOVA, interaction, $p>0,05$). Parmi les 23 cellules (Tukey test), seules 4 cellules ont une préférence significative pour une direction d'illumination ; 1 pour la gauche (Figure 4-17A), 2 pour en dessous et 1 pour au-dessus. De plus, 2 autres cellules montrent une diminution significative de leur taux de réponse lorsque l'illumination vient d'en dessous.

Afin de caractériser la sélectivité de ces cellules à une direction d'illumination, nous avons calculé un index de modulation de réponses en fonction de la direction d'illumination (DMI). Cet index a été dans un premier temps calculé sur la forme 3D préférée. Ce calcul a également été effectué sur la forme 3D optimale du neurone.

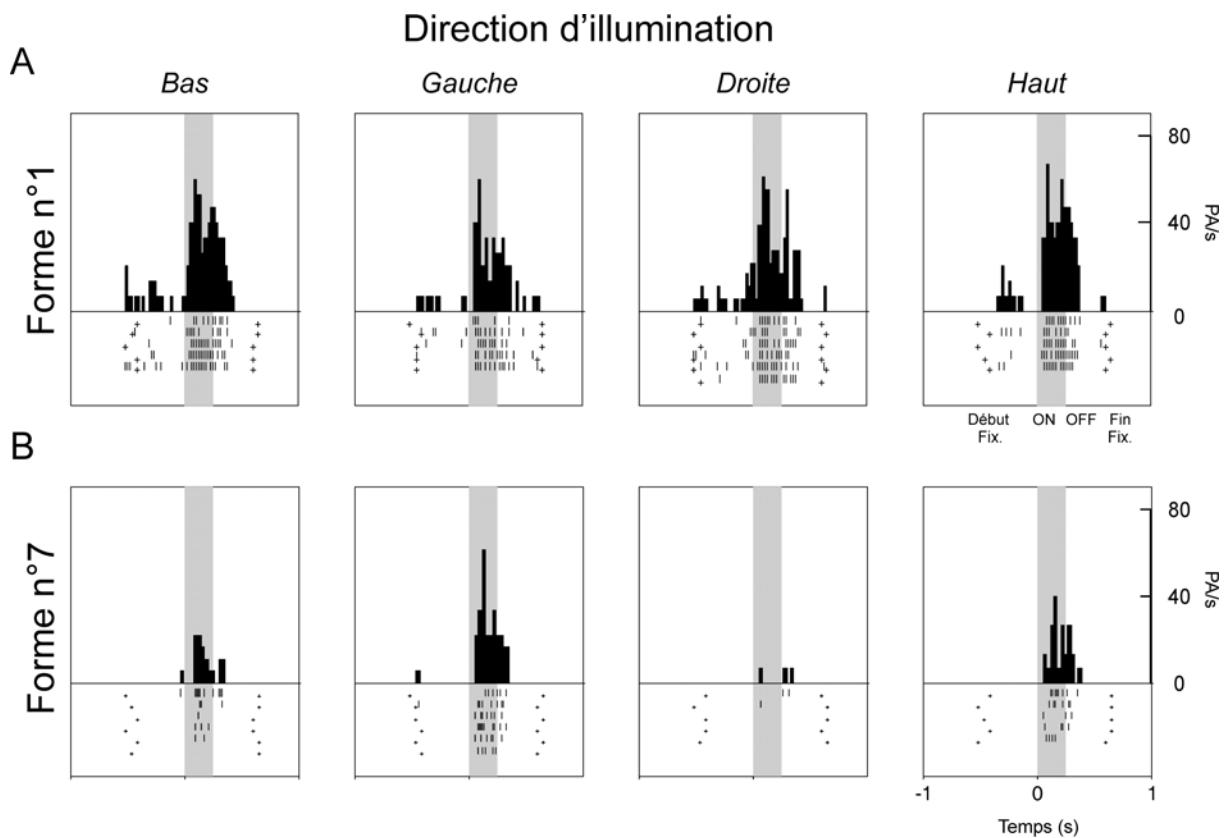


Figure 4-18 : Réponses de 2 exemples de cellules 3D à leur forme 3D préférée. Les intervalles ont une largeur de 30 ms. La cellule A (T-SFS13.31) montre un même taux de réponse quelle que soit la direction d'illumination alors que la cellule B (T-SFS23.21) voit son taux de décharge augmenter lorsque l'illumination vient de la gauche. Mêmes conventions que la Figure 4-15.

Chaque forme tridimensionnelle est éclairée selon 4 directions ; ce qui donne 4 images distinctes d'une même forme (soit 4 conditions). Le calcul détaillé ci-dessous est effectué pour chacune des cellules 3D. Au niveau de la forme 3D préférée (celle qui engendre une réponse moyenne maximale), on détermine le stimulus 3D préféré ; celui pour lequel la décharge neuronale est maximale (R_{max}). On effectue la différence entre R_{max} et la décharge minimale parmi les 3 réponses aux autres conditions d'illumination du stimulus 3D préféré. Cette différence est ensuite rapportée sur la moyenne des deux déviations standard correspondant aux 2 conditions. Un neurone ayant un index DMI proche de 1 aura un niveau de réponse similaire pour les 4 conditions d'éclairage : la modulation de réponse entre les différentes directions d'illumination n'est pas supérieure à la déviation standard. Les neurones des Figures 4-18 ont des index DMI 1.05 et 2.48 respectivement. La Figure 4-19 présente l'histogramme de distribution de ces index calculés au sein de la population 3D. Cette distribution a une médiane de 1.24, indiquant que la majorité des cellules présente une faible modulation de leur réponse en fonction de la direction d'illumination pour leur forme préférée. Pour les formes 3D optimales (rappel: séparant le mieux 3D et 2D), la médiane de

l'histogramme est très proche (1.34) ; l'influence de la direction d'illumination est également minime.

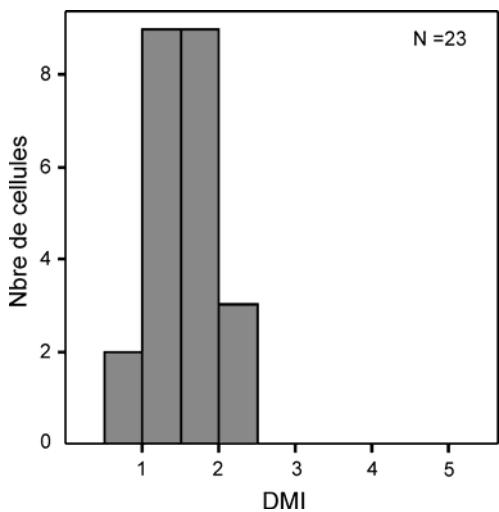


Figure 4-19 : Histogramme de distribution de DMI : index de sélectivité à la direction d'illumination. (n= 23 cellules 3D). La médiane de cet histogramme est de 1,24 indiquant que la population 3D présente une faible modulation de leur réponse en fonction de la direction d'illumination.

3. Analyse de population.

Afin de visualiser les sélectivités de la population 3D aux stimuli 3D, nous avons effectué une analyse hiérarchique par clusters. Seules les réponses aux 4 formes communes à tous les neurones enregistrés sont ici prises en compte. La Figure 4-20 reprend l'arbre établi avec les réponses moyennes standardisées aux 16 stimuli 3D. Les distances euclidiennes entre les clusters sont minimes en comparaison avec celles rencontrées dans les 2 autres arbres. Ce résultat indique que les réponses à ces stimuli 3D sont relativement proches. Toutefois, les réponses moyennes aux stimuli illuminés par en dessous (L1, en mauve, sur le graphique) se dissocient des autres réponses moyennes pour former un cluster distinct (qui contient aussi S5L3o).

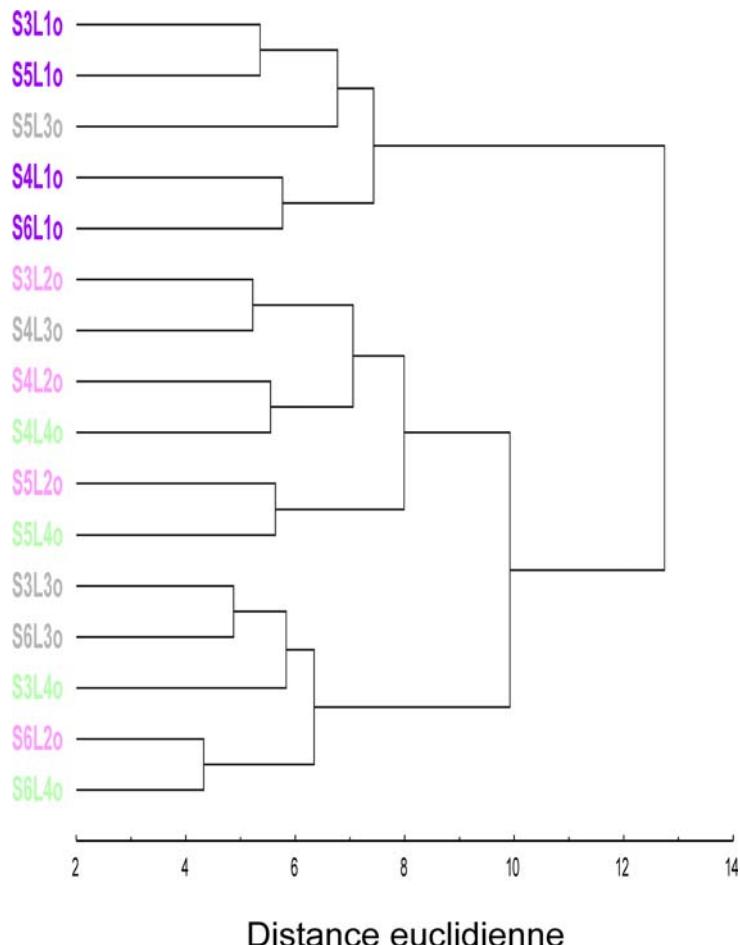


Figure 4-20 : Dendrogramme effectué sur la population 3D (n=23). Chaque réponse moyenne aux 16 stimuli 3D est représentée le long d'une branche horizontale sur la gauche du graphique. Les couleurs regroupent les stimuli en fonction de leur direction d'illumination (L1, en bas, L2, à gauche, L3 ; à droite et L4 ; en haut). Les réponses moyennes sont standardisées et la méthode utilisée est la méthode de Ward.

D. DISCUSSION

1. Synthèse des principaux résultats

Cette étude présentait 3 problématiques importantes. La première, était de tester si l'aire V4 code la 3D provenant de la distribution des ombrages sur une surface. Les deux autres objectifs étaient de tester la sélectivité aux formes tridimensionnelles et l'influence de la direction d'illumination sur les réponses des cellules aux formes tridimensionnelles. Après enregistrement d'une population de 119 cellules chez 2 animaux, plusieurs résultats sont à retenir :

- **Une majorité des cellules enregistrées est influencée par le contenu 3D, 2D blobs, random ou postérisé des stimuli (78%).** Ce sont des cellules dites ICS.
- Parmi les cellules ICS, nous avons caractérisé une **sous population de 23 cellules** qui présentent un taux de réponse supérieur pour les stimuli 3D. Ces cellules sont dites « **cellules 3D** ». C'est une faible proportion mais rapportée au nombre de neurones enregistrés dans V4, l'effet n'est pas négligeable.
- Au niveau de la sous-population de cellules 3D, les cellules ne dissocient pas toutes les formes tridimensionnelles de leurs contrôles 2D blobs respectifs car une grande partie des cellules 3D montrent une interaction significative entre le type du stimulus et la forme présentée (13/23 cellules). En moyenne, si les cellules ont été enregistrées avec 8 formes, les cellules 3D dissocient 2,5 formes de leurs contrôles 2D blobs et 1,75 si elles ont été enregistrées avec seulement 4 formes.
- A l'aide des analyses multidimensionnelles (Analyse hiérarchisée par clusters et MDS), nous avons démontré que **les réponses de la population ICS se regroupent en clusters** selon un niveau de réponse moyen similaire. A l'intérieur des clusters identifiés, nous retrouvons une majorité de stimuli définis par **un même type de pattern interne** (3D, 2D blobs, random et postérisé). Les réponses moyennes de la population ICS aux stimuli random sont nettement dissociables des réponses aux autres types de stimuli. Les réponses moyennes aux stimuli 3D sont relativement proches des réponses moyennes aux stimuli postérisés. Une représentation bidimensionnelle (dim 1

vs dim 2) permet de visualiser les clusters distinctement selon leur contenu interne.

- Les analyses par rang nous indiquent que **les réponses aux stimuli contrôles** (2D blobs, random ou postérisés) **ne sont pas complètement indépendantes** des réponses aux stimuli 3D correspondants. Autrement dit, lorsqu'un stimulus 3D a tendance à être préféré par la cellule, alors parmi tous les stimuli contrôles, la réponse du neurone au stimulus contrôle correspondant sera privilégiée.
- Les résultats du test ANOVA nous indiquent qu'une majorité de la population ICS (69%) est influencée par la forme du stimulus et de son contenu interne. Pour rappel, les stimuli 3D, 2D blobs et random sont définis par des paramètres de 1^{er} ordre identiques. Ainsi, les neurones ICS sont sélectifs à une gamme de paramètres.
- Avec le calcul d'un index de sélectivité à la forme 3D (SMI), nous avons démontré que la population ICS présente une faible modulation de leur taux de réponse suivant la forme 3D présentée. Il en est de même pour la population 3D.
- Il n'y a pas d'invariance des réponses à la direction d'illumination pour la sous population de neurones 3D pour leur forme tridimensionnelle préférée ou optimale.

2. Mise en évidence d'un codage différentiel pour les stimuli 3D

2.1. Sélectivité générale

Grâce à cette étude, nous avons démontré que les neurones de V4 sont influencés par les ombrages et codent de manière différentielle les stimuli 3D des stimuli 2D. Or, au niveau de la littérature, il a été clairement démontré que la sélectivité des neurones le long de la voie visuelle ventrale devient de plus en plus complexe de V1 au cortex inférot temporal (Kobatake et Tanaka, 1994). Les neurones de V1 sont sélectifs à l'orientation d'un stimulus simple comme une barre alors qu'au niveau de IT, les neurones codent davantage pour des formes complexes (Tanaka, 2002). Suite aux résultats d'études électrophysiologiques au niveau de V4, nous savons que les neurones de V4 sont influencés par l'orientation, la longueur ou la largeur d'une forme simple (ex ; une barre) ou d'un réseau orienté (Schein et Desimone, 1990). De plus, les études de groupe de C. Connor ont démontré que les neurones de V4 sont sélectifs à un élément particulier de la forme externe du stimulus et qu'une population

pourrait ainsi coder la forme complète d'un stimulus en le décomposant en plusieurs fragments (Pasupathy et Connor, 1999; Pasupathy et Connor, 2001; Pasupathy et Connor, 2002). D'autres études ont démontré que les neurones de V4 codent pour le contenu interne des stimuli (Gallant *et al.*, 1993; Gallant *et al.*, 1996). Les stimuli utilisés par le groupe de Gallant sont des réseaux cartésiens ou non-cartésiens (hyperpolaires, polaires ou concentriques). Les résultats de ces études ont mis en évidence que les neurones de V4 sont sélectifs au pattern interne des stimuli. De même, à l'aide d'analyses multidimensionnelles identiques à celles que nous avons utilisé (Analyse hiérarchisée par clusters et MDS), les auteurs notent qu'une dissociation en clusters des réponses moyennes des cellules selon le pattern interne des stimuli (cartésiens, hyperpolaires, polaires ou concentriques) apparaît (Gallant *et al.*, 1996). Ainsi, la population de neurones de V4 coderait tous les éléments d'un alphabet de primitives permettant l'analyse de n'importe quelle forme complexe en encodant à la fois son contour externe et son contenu interne. Les neurones de V4 coderait le contour externe en le décomposant en plusieurs fragments et le contenu interne sous la forme de textons.

Nos résultats sont en accord avec les conclusions de ces études puisque nous avons relevé une importante sélectivité de nos neurones au type des stimuli (3D, 2D blobs, random et postérisé). Ces neurones ICS présentent différents patterns de sélectivité parmi lesquels nous avons mis en évidence une sous population de neurones 3D qui codent préférentiellement en moyenne les stimuli 3D par rapport à leurs stimuli 2D blobs. Cette population représente $\frac{1}{4}$ de la population ICS. Toutefois, nous avons remarqué que les cellules 3D prises individuellement ne dissocient pas toutes les formes tridimensionnelles de leurs contrôles respectifs. Les cellules 3D dissocient en moyenne 2 formes. Cependant, les analyses multidimensionnelles nous permettent de mettre en évidence une meilleure dissociation selon le pattern interne des stimuli au niveau des réponses moyennes de la population ICS totale. En effet, nos analyses hiérarchisées par cluster mettent en évidence, pour une majorité de stimuli, un regroupement selon le pattern interne des stimuli. Les stimuli random sont tous regroupés ensemble et leur cluster est nettement dissocié des 2 autres (3D et 2D blobs). Les stimuli 3D et postérisés sont regroupés dans des clusters distincts, il en est de même pour les stimuli 2D blobs même si quelques stimuli 2D blobs sont associés aux stimuli 3D. Il est possible que ces stimuli contrôles laissent apparaître une perception tridimensionnelle notamment pour les stimuli contrôles de la forme n°6 et de la forme n°8 (cf. annexe). La Figure 4-21 reprend les illustrations de stimuli 3D et 2D blobs non dissociés

(Forme n°6) et dissociés (Forme n°7). Le stimulus contrôle de la forme n°6 est peut-être perçu comme tridimensionnel.

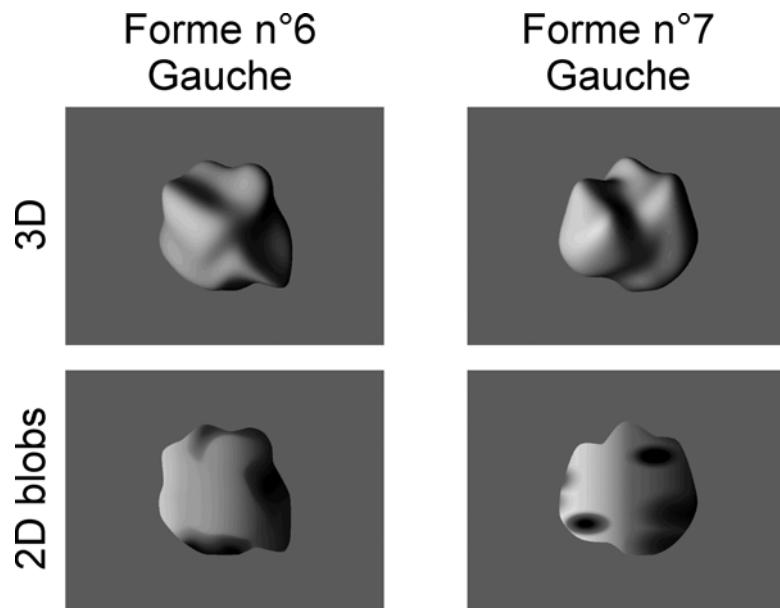


Figure 4-21 : Formes 3D et 2D blobs. Cette figure reprend les illustrations des formes 3D n°6 et 7 illuminées par la gauche et de leurs contrôles 2D blobs respectifs. La forme n°6 n'est dissociée de son contrôle dans l'analyse par clusters sur la population ICS (cf.). Son contrôle 2D blobs présente un percept tridimensionnel alors que le contrôle de forme n°7 se retrouve dans le cluster dédié aux 2D blobs.

Nous avons également démontré que les neurones ICS, pris individuellement, dissocient mal les formes tridimensionnelles de leurs stimuli contrôle et que cette faible dissociation se retrouve entre les clusters contenant les stimuli 3D et 2D blobs dans les analyses hiérarchisées en comparaison avec la dissociation des réponses pour les stimuli random ($d=170$ pour les stimuli random, $d=38$ pour les stimuli 2D blobs). Ce résultat peut s'expliquer par le fait que les stimuli 2D blobs ont été réalisés afin de produire une texture aléatoire définie par des paramètres statistiques de 1^{er} et de 2nd ordre identiques à ceux des stimuli 3D. Or nous avons démontré dans l'étude précédente que les neurones de V4 sont influencés par des paramètres statistiques similaires. A l'inverse, les stimuli 3D diffèrent de leurs stimuli random par leurs paramètres statistiques de 2nd ordre mais aussi par leur spectre de fréquence spatiale (cf. méthodes) ; ce qui entraînerait une dissociation plus importante en terme de réponse neuronale de la population ICS. L'apport des stimuli postérés nous a permis de mettre en évidence que le codage des ombrages n'est pas simplement due à la disposition de patchs noirs et blancs à l'intérieur d'une même forme externe. Toutefois, d'après l'analyse hiérarchisée, les réponses des neurones ICS aux stimuli postérés sont plus proches de celles pour les stimuli 3D que des réponses pour les 2D blobs ($d=28$ pour les

stimuli postérisés et $d=43$ pour les 2D blobs). Ce résultat indique que la population ICS code plus la 3D en se basant sur la disposition des patchs noirs et blancs que sur les paramètres statistiques de 2nd ordre. Toutefois, les stimuli postérisés ne sont pas définis en 3D selon nos tests de psychophysique chez les humains mais qu'en est-il des animaux ? Une réponse comportementale de l'animal pourrait confirmer sa perception tridimensionnelle.

Grâce aux analyses multidimensionnelles, nous avons démontré qu'une population de neurones de V4 regroupe les stimuli en fonction de leur pattern interne (3D, 2D blobs, random ou postérisé). Ces résultats rejoignent les résultats de l'équipe de Gallant sur l'encodage des réseaux non-cartésiens qui avait noté une séparation franche selon le type de réseau à l'aide des analyses multidimensionnelles similaires (Gallant *et al.*, 1996).

2.2. Codage des ombrages comme indice tridimensionnel

Grâce à cette étude, nous avons mis en évidence un codage particulier des neurones de V4 pour des stimuli tridimensionnels définis par des ombrages. D'autres études ont démontré que l'aire V4 participe au traitement d'indices tridimensionnels tel que la disparité ou la texture. La sélectivité des neurones de V4 à la disparité a été démontrée dans plusieurs travaux (Hinkle et Connor, 2001; Hinkle et Connor, 2002; Watanabe *et al.*, 2002; Hegde et Van Essen, 2005; Hinkle et Connor, 2005; Tanabe *et al.*, 2005). Toutes ces études s'accordent sur le fait que des neurones de V4 sont sensibles à la disparité des stimuli. Par exemple, Hinkle et Connor démontrent que les cellules de V4 sont sélectifs à l'orientation tridimensionnelle d'une barre (Hinkle et Connor, 2002). Watanabe et de ses collaborateurs (2002) rapportent qu'une population de neurones de V4 modulent leurs réponses en fonction des changements de disparité des stimuli (Watanabe *et al.*, 2002). Enfin, les études du groupe de Tanabe et celle de Hegde et Van essen (Hegde et Van Essen, 2005; Tanabe *et al.*, 2005) ont démontré que la sensibilité à la disparité des neurones de V4 dépend de la nature du stimulus (SFS ou RDS). Ces études concluent que les informations concernant la forme bidimensionnelle et la disparité binoculaire pourraient être traitées de manière interdépendante.

Au niveau du codage de gradient de textures, le groupe de Hanazawa (2001) a enregistré la réponse de neurones de V4 à des textures artificielles composées de plusieurs éléments circulaires où la densité et la taille varient (Hanazawa et Komatsu, 2001). Une importante majorité des cellules enregistrées est sensible à la taille et/ou à la densité des éléments qui constituent la texture. D'après les auteurs, ces cellules participeraient à la perception de textures définies en 3D sur la base d'ombrages.

Nos résultats indiquent qu'il existe un traitement de la profondeur au sein de l'aire V4 sur la base d'ombrages. Toutefois, comme les stimuli sont présentés en situation passive aux animaux, nous ne pouvons pas conclure réellement sur la perception de la profondeur des animaux.

Au niveau du codage des ombrages comme indice tridimensionnel, de nombreuses études indiquent que le cortex inférot temporal est influencé par les ombrages. Par exemple, en neuroimagerie chez l'Homme, les études de Moore et Engel (2001) et celle de Kourtzi et Kanwisher démontrent que l'aire LOC s'active lorsque que les objets sont définis en 3D à partir d'ombrages (Kourtzi et Kanwisher, 2000; Moore et Engel, 2001). Moore et Engel (2001) ont enregistré, en IRMf, une activation significative de l'aire LOC lors de la présentation d'objets 3D définis par des ombrages en comparaison avec des présentations d'objets 2D (Moore et Engel, 2001). De plus, les sujets pouvaient indiquer leur perception tridimensionnelle de l'objet par une réponse comportementale. Les activations du LOC s'accordent avec les réponses comportementales des sujets ; elles sont significatives si les sujets perçoivent de la profondeur au niveau de l'objet et décroissent si les sujets jugent l'objet en 2D. Cette activation significative de l'aire LOC lorsque que les objets 3D sont définis par des ombrages avait été préalablement décrite dans l'étude de Kourtzi et Kanwisher. L'étude de Sereno en IRMf chez des singes anesthésiés a également remarqué des foyers d'activation le long du STS pouvant correspondre à l'aire TE (Sereno *et al.*, 2002). Ces études de neuroimagerie sont en accord avec les résultats de l'étude de Vogels et Bierderman (2002) qui démontre à l'aide d'enregistrements électrophysiologiques unitaires qu'une sélectivité sur la base des ombrages est présente au sein du cortex inférot temporal chez le macaque (Vogels et Biederman, 2002).

Outre les ombrages, le cortex inférot temporal participe également au traitement de la profondeur grâce à la perspective des gradients de texture. En effet, l'étude de Liu *et al.* (2004) démontre la présence de neurones sélectifs au gradient de textures. Les conclusions de cette étude corroborent celles des études de Janssen qui démontrent que des neurones sélectifs au gradient de disparité sont présents au sein de l'aire TE (Janssen *et al.*, 1999; Janssen *et al.*, 2000).

Ainsi l'aire V4 pourrait être le siège de traitement d'indices tridimensionnels tels que la disparité, les gradients de texture et les ombrages (Hanazawa et Komatsu, 2001; Hinkle et Connor, 2001; Hinkle et Connor, 2002) et véhiculerait ces informations vers le cortex inférot temporal vers lequel l'aire V4 projette massivement (Boussaoud *et al.*, 1991).

3. Sélectivité à la forme 3D présentée

Au cours de cette étude, nous avons démontré que la population ICS présente une faible modulation de leur taux de décharge pour les différentes formes tridimensionnelles qui leur sont présentées. Cependant, une forte sélectivité à la forme tridimensionnelle a été observée au sein du cortex inférotemporal (Vogels et Biederman, 2002). Les stimuli tridimensionnels utilisés dans leur étude sont constitués de geons (cylindres, cônes ou pavés rectangulaires) et ont des formes tridimensionnelles facilement dissociables visuellement. Nos stimuli tridimensionnels représentent toutes des formes patatoïdes où les concavités et les convexités sont propres à chaque forme. Ces formes sont visuellement très proches. Afin de réellement tester la sélectivité à une forme tridimensionnelle, il faudrait utiliser des stimuli identiques à ceux utilisés par Vogels et Bierderman (2002).

4. Reconnaissance des objets et direction d'illumination.

Les modèles théoriques sur la reconnaissance des objets ne s'accordent pas sur l'influence de la direction d'illumination sur celle-ci. Les modèles structuraux proposent que la reconnaissance d'un objet ne soit pas influencée par une variation des conditions d'illumination (Marr et Nishihara, 1978 ; Bierderman, 1987). En effet, d'après ces théories, la reconnaissance de l'objet se fait grâce à la décomposition de sa forme en éléments structuraux (« geons ») qui ne contiennent pas ce type d'information. Des études de psychophysiques ont démontré que la reconnaissance des objets n'est pas influencée par la direction d'illumination ; les sujets humains reconnaissent les objets indépendamment de la direction d'illumination (Biederman et Bar, 1999; Nederhouser *et al.*, 2001). Ce mécanisme d'invariance à la direction d'illumination est présent au sein du cortex IT d'après l'étude électrophysiologique de Vogels et Biederman (2002). Les neurones de IT sont invariants pour différentes directions d'illumination pour leur forme préférée (Vogels et Bierderman, 2002).

Les modèles d'alignement proposent que la direction d'illumination soit encodée dans la représentation de l'objet ou du visage (Ulmann, 1989, Poggio et Edelman, 1990). Des études de psychophysiques soulignent que la reconnaissance d'un visage (le visage peut être considéré comme un objet complexe) ou d'un objet est fonction des conditions d'illumination (Gauthier et Tarr, 1997; Troje et Bulthoff, 1998; Braje, 2003). Des enregistrements électrophysiologiques au sein de l'aire V4 ont démontré que les neurones de V4 sont sensibles à la direction d'illumination lorsqu'ils codent des gradients de textures artificielles (Hanazawa et Komatsu, 2001). Ils ont également en évidence que les neurones présentaient une direction d'illumination préférée, notamment le long de l'axe vertical, pour des sources d'illumination

situées au-dessus et en dessous des patterns texturés. Contrairement aux résultats de Hanazawa et Komatsu, nous n'avons pas retrouvé au sein de notre population 3D une sélectivité à une direction d'illumination donnée même si une majorité des cellules ICS sont influencées par la direction d'illumination et que leur pattern de sélectivité varie selon la forme considérée. En ne s'intéressant qu'aux formes 3D préférées, nous avons également relevé que la population 3D présente une faible modulation de leur taux de réponse selon leur direction d'illumination. Les gradients d'illuminations présents dans cette étude ne génèrent pas un percept tridimensionnel d'une forme complexe alors que dans nos stimuli, il produit des ombrages à la surface d'objet générant une perception tridimensionnelle de sa forme. Nos résultats sont à comparer avec ceux de Vogels et Biederman (2002) qui contrairement à l'étude d'Hanazawa n'étudient pas un gradient de texture mais l'influence de la direction d'illumination sur une forme 3D complexe. La faible modulation dans V4 que nous avons relevée au sein de V4 permettrait d'établir une invariance à la direction d'illumination observée au sein de l'aire IT (Vogels et Biederman, 2002). A l'aide de l'analyse par clusters, nous avons également remarqué que les stimuli illuminés par en dessous sont regroupés au sein d'un même cluster isolé des autres stimuli. Cette direction d'illumination n'est pas naturelle et nos résultats rejoignent les résultats de tests psychophysiques qui indiquent les hommes supposent que la direction d'illumination vient d'au-dessus (Ramachandran, 1988; Kleffner et Ramachandran, 1992).

Nos stimuli sont présentés passivement aux animaux, il n'est absolument pas certain que l'animal perçoit ces stimuli comme 3D, d'où les effets relativement modestes que nous avons relevé au niveau des cellules individuelles. Toutefois, nous avons montré que des mécanismes existent dans V4 qui permettent au niveau de la population de distinguer des stimuli 3D d'autres contrôles 2D. Plusieurs études ont montré que les neurones de V4 ont leur propriétés modulées par un apprentissage perceptif (perceptual learning) (Rainer *et al.*, 2004). On peut raisonnablement penser que si l'animal avait à faire une tâche active (classer les stimuli suivant leur aspect 3D ou non par exemple), on observerait un codage 3D plus évident au niveau des cellules individuelles. On peut supposer que ces 23 cellules 3D que nous avons enregistrées représentent une réelle influence 3D, une signature de l'affordance de ces stimuli; l'objet présente une certaine cavité à sa surface donc je peux le saisir de telle ou telle façon. Peut-être en enregistrerions-nous une plus grande proportion si on disposait d'une phase préliminaire où l'animal aurait à manipuler les objets dont on lui présente les photographies, puisque dans certaines conditions, les singes n'ont pas la même appréhension que les humains des lois physiques de leur environnement (Povinelli, 2000).

Si ces cellules 3D sont la signature du composant dorsal de V4, il reste encore beaucoup à déterminer comment s'établissent les liens fonctionnels entre V4 et les aires pariétales, notamment l'aire CIP, centre hiérarchiquement élevé pour le traitement de la 3D (Tsutsui *et al.*, 2005). Aucune connexion directe n'a été mise en évidence entre V4 et CIP, l'aire V4 présentant des projections minimes avec l'aire LIP (Kennedy et Barone, 1999). On peut supposer une modulation descendante qui véhiculerait l'expérience tactile de l'animal pour arriver à percevoir la 3D dans des images 2D.

Chapitre 5

Short visual latencies in area V4 of the macaque monkey

Ce chapitre décrit une étude sur les latences de réponses d'une population de cellules enregistrée lors du protocole sur les ombrages. Cette étude est partie d'une simple observation effectuée suite aux résultats des études précédentes qui lors d'analyses des réponses des neurones, nous avons remarqué que les réponses des neurones étaient relativement rapides. Nous avons donc entrepris d'étudier plus précisément ces latences de réponses car peu d'études traitent des latences de réponses des neurones chez le singe vigile au sein de l'aire V4. L'originalité de cette étude est qu'elle compare deux méthodes même trois de calcul de latences sur des mêmes neurones. Toutefois, une mauvaise synchronisation verticale des présentations des stimuli est à considérer dans ces résultats (voir analyses complémentaires).

Avertissement au lecteur :

Nous nous sommes rendu compte à posteriori de la présence d'erreurs critiques de timing. (une mauvaise synchronisation verticale lors de la présentation des stimuli). Plus précisément un délai aléatoire apparaît entre le code logiciel qui ordonne l'affichage des stimuli et leur apparition réelle. Ce délai ne peut être mesuré à posteriori pour les fichiers enregistrés. Après vérification avec une photodiode, le délai s'échelonne entre 11 ms et 24 ms. Les valeurs présentées ci-dessus doivent être considérées en prenant en compte cette erreur de synchronisation. Nous avons estimé que la présence d'un délai aléatoire ne satisfait pas les critères d'une publication de qualité et retiré cette publication acceptée sous revue. Toutefois, nous avons effectué des corrections statistiques (voir en fin de ce chapitre) afin de prendre en compte cette erreur.

A. ABSTRACT

The area V4 of the awake macaque monkey is considered as a major component of the parvo-dominated ventral visual pathway. However, latencies of V4 neurons responses to flashed stimuli are very long (Schmolesky et al, 1998). As this princeps but only detailed study has been conducted on anesthetized monkeys, we measured latencies of visual responses in area V4 of two awake macaque monkeys. They were computed from single unit responses to flashed stimuli during passive viewing. Although we observed a 10 ms significant difference between the animals, latencies were remarkably short and ranged between 40 and 80 ms. They were reliably calculated with two different methods and did not depend on receptive field eccentricities. These results contradict the outlier status of area V4 and show that response latencies in this area are in agreement with the general scheme of fast visual processing in the cortical areas of the primates.

B. INTRODUCTION

Visual perception can happen quickly. For example, human and non-human primates can accurately perform sophisticated visual discriminations with very short reaction times: pictures containing an animal can be discriminated within 280 ms by humans (Thorpe et al. 1996) and 180 ms by macaque monkeys (Fabre-Thorpe et al. 1998). Studies on the neuronal counterparts of such abilities show remarkably short latencies of neuronal responses to visual stimuli (Perrett et al. 1982, Vogels 1999) or ERP studies (VanRullen and Thorpe 2001).

Among the numerous brain areas devoted to vision, a large proportion has been studied in the macaque monkey. Single cell recording studies have shown that there is an extensive overlap of latencies between areas. Even more surprisingly, many neurons in area V1 have not yet started to respond when some IT cells have already started responding (reviewed in Nowak and Bullier, 1997). These results invalidate a strictly hierarchical scheme of visual processing.

However, since data often come from different studies that have employed different methods, it is difficult to compare response latencies across visual areas. To overcome this difficulty, some authors (Raiguel et al. 1989; Nowak et al. 1995; Schmolesky et al. 1998; Vogels and Orban 1994) have tried to determine latencies in several areas with the same methodology. In particular, Schmolesky et al. (1998) examined latencies in the parvocellular and magnocellular layers of the dorsal lateral geniculate nucleus and in 7 cortical visual areas (V1, V2, V3, V4, MT, MST and FEF). They found that area V4 was an outlier with remarkably long latencies compared to other areas examined (these latencies were the longest with an average of 104 ms, a large variance ($SD = 23.4$ ms) and a maximum value that extended to 159 ms).

Although area V4 is a well-studied visual area in macaque monkeys, systematic studies of visual response latencies in the awake animal are still lacking. The only existing report to our knowledge is by Maunsell (Maunsell 1987) who showed a population histogram starting at 50 ms and a plateau starting at about 80 ms (see Zhou et al. 2000 for a comparable population histogram in the alert macaque monkey). Other authors working with behaving macaque report long latencies. Tanabe et al (2004) show that latencies of V4 population recorded with binocular stimuli (Random Dot Stereograms) were distributed across a wide range of values (114 +/- 37 ms). In the present paper, our goal was to determine the response latency of V4 neurons to flashed stimuli. We report a detailed analysis of latencies in area V4

of the awake macaque monkey and we provide data to complement the scheme of the timing of visual processing.

C. MATERIAL AND METHODS

1. Animals

Two adult rhesus monkeys (*macaca mulatta*), Monkey 1 (female 10 years old) and Monkey 2 (male 5 years old), weighing between 3 and 5 kg, were used in this study. A first surgery, performed under general anesthesia and sterile conditions, consisted of implanting a head post (CRIST®). A pain reliever (Ketofen, 0.2ml/kg) and systemic antibiotics were administrated just before the surgery started. Induction of anesthesia was performed by an injection of ketamine (16 mg/kg IM). Deeper anesthesia was achieved with a mixture of alphadolone/alphaxolone (Saffan, 15 mg/kg/h IV, rate adjusted if required). A recording chamber (CRIST®) was implanted during a second surgery, which was performed under the same conditions. In addition, corticoids were injected (Solumedrol, 1 mg/kg, IM) to prevent brain edema. We aimed at placing the chamber over the parafoveal representation of the left inferior contralateral quadrant of area V4. The position of the chamber was based on stereotaxic coordinates as described in our previous work (Girard et al. 2002), on skull landmarks and on positions of the sulci that were visible through the intact dura. The bone within the chamber (2-cm diameter) was removed with a trephine. This procedure allows more flexibility for electrode positioning. We left the dura intact within the chamber and we performed sterile cleaning on a daily basis. We used guide tubes to lower the electrodes through the thickened dura. All procedures complied with guidelines of the European Ethics committee on Use and Care of Animals.

2. Task and stimuli

The animals sat in a primate chair, with the head restrained, 57 cm in front of a Iiyama monitor (22", 75 Hz). An ISCAN infrared eye-tracking system (120Hz) monitored horizontal and vertical eye positions by tracking the corneal reflection of a focused infrared LED through a CCTV camera with a 250-mm lens. The precision of our eye position measurement is 0.2-degree. CORTEX software (courtesy of the NIMH Laboratory of Neuropsychology) controlled stimulus presentation, animal behaviour and data acquisition.

The monkeys were first trained to maintain fixation within a 2-degrees square window. The monkey had to keep fixating on a 0.1-degree central grey spot for a variable delay (400 to 600 ms) before one stimulus was flashed for 250 ms. After the stimulus was turned off, the fixation spot remained on the screen for a variable delay (350 to 400 ms). Stimulus presentation was randomized. Each stimulus was presented between 5 to 10 times

with an inter-stimulus interval ranging from 600 to 900 ms. Trials completed without breaking fixation were rewarded with a drop of water. We monitored the weight of the monkey every working day and ad libitum water intake was allowed during the weekend.

The stimuli were designed for studying the coding of 3D shape from shading in V4 and will be described in detail in a subsequent publication. Briefly, they were 2D images of randomly distorted spheres illuminated from 4 directions, leading to a 3D percept. The stimuli were adjusted in size in order to be contained inside the receptive field (RF). The luminance ranged between 9.3 and 15.8 cd/m² (mean 12.6 cd/m²). The mean RMS contrast was 0.7 and background luminance was 11.5 cd/m².

3. Recording

Tungsten-in-glass microelectrodes were used to record extracellular neuronal activity using conventional techniques. Up to three out of five possible independent microelectrodes were lowered in the brain (Trec®, 0.5 to 1.5 MΩms). Action potential waveforms were sorted online with the help of a spike sorting software (AlphaOmega MSD®) and only single units obtained through complete trials were taken into account.

Before running the recordings with the set of stimuli (see below), we used a sparse noise technique to plot online the RFs of the neurons under study. Once the spike waveform was isolated, we ran pseudo-random sequences of black and white squares (1°). Each square was flashed 10 times during 50 ms. These stimuli sampled the left inferior quadrant within a 12° square region centered on 5 degrees of eccentricity, where the RFs were expected to be found. In some cases, the size and the position of the black and white squares were modified in order to refine the RF plotting. Neuronal activity was correlated with each stimulus in order to build receptive field maps online. This method provides us with good estimates of the RF size and boundaries, thus allowing us to center the stimuli in the receptive field. We then used the positions of the RF centers to draw retinotopic maps of the recorded regions for both monkeys. We superimposed these maps on the drawings of sulcal landmarks that could be visible during the surgery at the time of bone opening (Figure 5-1).

4. Measurements of response latencies

Since the purpose of this paper was to measure the latencies of V4 neurons in awake monkey regardless the stimulus, we deliberately chose to pool the responses to all the stimuli. To assess the response latency of each cell, we first computed a peristimulus-time histogram (PSTH) from the whole set of stimuli (average: 230 presentations for MONKEY 1 and 100 for MONKEY 2). Then, two different methods were used to compute the latencies. One is the

method proposed by Maunsell and Gibson (1992). We termed 'onset latencies' the latencies computed with this method. It assumes that spontaneous neuronal activity follows either a Poisson or a Gaussian distribution that allows one to characterize a threshold value. We computed the spontaneous discharge rate (number of spikes) in a 200 ms period starting 400 ms before stimulus presentation. The threshold was defined as the mean value of the spontaneous activity plus 3 standard deviations, which corresponds to a P value of 0.01. The response of the neuron to the stimulus was considered to be significant when: 1) a bin in the histogram crossed the threshold, 2) the next bin stayed above this level, and 3) the second next bin did not fall below the $P=0.05$ level (Maunsell and Gibson 1992). The latency of response then corresponded to the duration between stimulus onset and the first significant bin. We used bin widths of 1 and 2 ms and we finally chose the latency corresponding to the bin with the lowest P value. Because of weak activity, some latencies could only be estimated with a bin width of 5 or 10 ms (26% of the neurons).

We used a second method to assess the latencies by calculating second derivatives (d'') of PSTHs computed from a 5 ms bin width. We term this method 'derivative latency'. Peaks in this function indicate sudden increases or decreases of neuronal discharge rate. We considered the response as significant when the first peak of d'' appeared in a 200 ms period after the onset of the stimulus. As many cells had non-significant bursts of activity during this period, we only considered the peaks in d'' corresponding to bins that were followed by two other bins above the threshold value. The onset latency corresponded to the time to the peak of the d'' .

D. RESULTS

Data were obtained from 144 neurons recorded in two monkeys (115 in MONKEY 1 and 29 in MONKEY 2) which had significant response to stimuli. The response was computed as the mean firing rate over a period of 250 ms period starting 50 ms after stimulus onset and was considered as significant when it was higher (T-test, $p<0.05$) than the spontaneous activity. Examples of visual responses and corresponding onset latencies are displayed in Figure 5-2. The response of a majority of cells (84%, $N=121$) was characterized by a sustained activation. For most of the cells ($N=63$) this activation followed a peak of discharge as shown in Figure 5-2A. Other cells ($N=58$) showed sustained activation without the first peak of activity (Figure 5-2B). More rarely, we observed a phasic activation without sustained activation (Figure 5-2C and D, $N=20$). The 3 remaining cells were inhibited by the stimuli and latencies were not computed for these cells.

For the vast majority of neurons, the latencies computed with both methods gave similar results. For example, the cells in Figure 5-2A and C showed an onset latency of 53.5-50 ms and 50-54.5-50 ms respectively (the first value indicates ‘onset latency’ and the second value ‘derivative latency’). However, for some neurons ($n=21$), the latencies determined with each method differed by at least 10 ms (Figure 5-2B: 82.5-70ms; Fig. 2D: 75.5-55ms).

Table 1 summarizes the descriptive statistics of onset latencies computed with both methods. The shortest latency was 40 ms (Monkey 1) and the longest latency was 84 ms (Monkey 2). At the population level, response latencies of visual responses in V4 appear remarkably short. The medians of the distributions are 50 and 51 ms for Monkey 1 and 61 and 60 ms for monkey 2. The difference between the average latency values obtained with both methods was 3 ms for Monkey 1 and 4 ms for Monkey 2. Since this difference is smaller than the bin size (5 ms) used for the derivative analysis, we considered both methods to be similar and we restricted the rest of the analysis to latencies computed with the onset method.

Figure 5-3 presents the cumulative distributions of latencies for each animal and method. The plots emphasize the presence of short visual latencies in V4 since 50% of the latencies were below 50 and 60 ms for monkeys 1 and 2 respectively. An important result shown in Figure 5-3 is that the latencies of the Monkey 1 population were shorter on average by about 10 ms than those of the Monkey 2 population. The difference between them was significant ($P<0.001$, Mann Whitney U test). One possible explanation for this difference comes from the fact that recordings were obtained at different eccentricities in the visual field for each animal. The range of eccentricities of MONKEY 2 receptive fields was narrower (2,2

to 5°) than that of MONKEY 1 receptive fields (2 to 14°). To examine this possibility, we compared latencies for recordings obtained at comparable eccentricities for both animals: 35 cells recorded in MONKEY 1 possessed receptive fields whose eccentricities were inferior to 5°, equivalent to those obtained for MONKEY 2 neurons (N=27). However, the difference between the two animals was still highly significant (Mann Whitney U test, $P=0.0014$).

Figure 5-4 represents the plot of latencies as a function of the eccentricity of the receptive fields for Monkey 1 only (no Monkey 2 neuron was recorded with an eccentricity larger than 5°). The result of regression analysis clearly shows that latencies did not depend on eccentricity ($R^2=0.014$).

E. DISCUSSION

Our study shows that latencies of V4 single neurons of the awake macaque monkey are short (mean values of 54 and 63 ms for MONKEY 1 and MONKEY 2 respectively) and range between 40 and 80 ms. These results were consistent over two animals and two different methods of measurement. They were also invariant to eccentricity. Of special note is that these latencies are considerably shorter than those recorded on anesthetized monkeys by Schmolesky et al. (1998; mean value=104, SD=23.4 ms). There are many reasons that may explain the large difference between our results and the results of Schmolesky et al (1988). In the subsequent paragraphs we will discuss the possible role of precise location of the recordings, methods of latency measurement, monkeys (age, size) and protocols (stimuli, task) discrepancies.

1. Location of recording sites

As recordings are still ongoing in both monkeys, the histological confirmation of recording sites was not performed. Hence, we provide detailed evidence that we indeed recorded in area V4. The respective positions of vertical meridian (close to the lunate sulcus) and horizontal meridian (anterior most portion of the chamber) are in agreement with previous mapping studies (Gattass et al. 1988; Fize et al. 2003). As in the Gattass et al (1988) study, foveal RFs are latero-anterior whereas peripheral RFs are dorso-posterior. Consistent with their results, the distance between the neurons with a parafoveal RF (2° eccentricity) and those with a 14 degree of eccentricity RF is about 6 mm in monkey 1 (as measured on cortical surface). The possibility that some cells were recorded in area V2 is excluded by several arguments: 1) dorso-posterior RFs reach eccentricities of 14 degrees, which can never be encountered in the portion of dorsal V2 that abuts the lunate sulcus; 2) laterally, V2 and V4 foveal representations in the prelunate gyrus are abutting each other but our central most RFs have two degrees of eccentricity and hence were too peripheral to be in V2; 3) RFs sizes at different eccentricities were compatible with previously published V4 data (Gattass et al. 1988). Although RFs were approximately mapped, we estimate that the surface was about 1 square degree per degree of eccentricity. Finally, one could argue that neurons recorded in monkey 2, which have a less extended range of RFs eccentricities, are in V2. If this were the case, we would expect having similar or even shorter latencies than in Monkey 1, which is the opposite. In addition, sulcals landmarks, neuronal responses and RFs sizes were very similar to those recorded in monkey 1. Then it is very likely that we recorded neurons in area V4.

Schmolesky et al. (1998) recorded 29 neurons in the area V4 of anesthetized monkeys. As they did not show any map of the penetrations, it is difficult to compare the locations of the recording sites between the studies. However, they mention that the RFs eccentricities in their sample of neurons were between 3 and 10 degrees, which is similar to ours. Then, even if we can't exclude this hypothesis, it seems unlikely that the location of the recording sites explains the discrepancy of latencies.

2. Latency measurement

We computed latencies of V4 neurons responses with two different methods. The first method which we called ‘onset latency’, has already been validated in many studies (Maunsell and Gibson 1992, Nowak et al. 1995). The second method, ‘derivative latency’ is new and based on the second derivative of the PSTH. This method detects a significant change in firing rate even if this change is weak or relies on a weak spontaneous activity. The observed differences between the two methods were not functionally relevant. Indeed, the differences computed for the minimum and the maximum latencies for both methods do not exceed 5 ms. In addition, although ‘derivative latencies’ tend to be shorter than ‘onset latencies’, the mean values do not differ by more than 4 ms (2,7 ms for MONKEY 1 and 3,7 ms for MONKEY 2; see Table 1). One should note that these differences are inferior to the bin width used to compute ‘derivative latencies’ (5ms). For the neurons without phasic response and/or with a low discharge rate (21 neurons, examples in Figure 5-2B and D), the difference is significant. Thus, it seems that the methodology may account for small differences only in measurements of the latencies.

3. Discrepancies in procedures (animals, stimuli, anesthesia)

We observed a significant difference of mean latency (10 ms) between animals. This difference does not depend on eccentricity: when restricting the analysis to the cells with receptive fields centered on the 5 central degrees of the visual field, which is the common range for both monkeys, the difference was still significant with the same probability level. This difference may depend on the age of the animal, but unexpectedly the longest latencies were found in the younger animal (Monkey 2). Importantly, both animals were not in the range of senescent monkeys that have particularly delayed visual latencies in the extrastriate cortex (Wang et al. 2005). The mean latency difference may be due to a difference in animal size since the shortest latencies were found in the smaller (3 kg) female and the longest in the heavier (5 kg) male. A similar possibility was already evoked in Maunsell and Gibson (1992) who found about the same range of difference between V1 latencies in monkeys of different

size and age. In any case, we do not think that the size or the age of the monkey may explain the discrepancy observed with the results of Schmolesky et al. (1998).

The range of latencies (40-80 ms) we measured in area V4 is remarkably short. This may come from the fact that we pooled responses obtained with several stimuli that drove efficiently the population of V4 neurons. Schmolesky et al. (1998) used simple flashed stimuli (spots or bars) that were selected according to the preference of the neurons (shape, size, orientation, phase and color). Since the contrast of these stimuli was 80%, we can exclude the possibility that they were less efficient in driving the V4 cells, resulting in the delayed responses they observed. Other authors (Tanabe et al. 2004) have found, in the awake macaque monkey, response latencies in V4 markedly longer than 100 ms. In this study, the neurons were stimulated with random dot stereograms. These results show that long latencies exist in area V4 of awake monkeys. These results may reveal that, in V4, binocular processes require additional processing time compared to monocular ones.

A possible explanation for the long V4 latencies measured by Schmolesky et al. (1988) is that neurons were recorded under general anesthesia. Indeed, V4 neurons are sensitive to drowsiness in awake monkeys (Pigarev et al. 1997) and one can suppose that anesthesia induces a similar effect. This effect might be more pronounced in parvocellular-dominated cortical regions. For example, latencies superior to 100 ms have been encountered in layer 4C β in V1 of the anesthetized monkey (Nowak et al., 1995), and in V2 color selective cells and V2 thin stripe cells in the anaesthetized monkey (Munk et al., 1995). So, the explanation we favor is that area V4, that is closely engaged in active processes like attention and visual search (Connor et al. 1996, McAdams and Maunsell 1999, De Weerd et al. 2003, Bichot et al. 2005) or shape discrimination (Walsh et al. 1992, Schiller 1993, Merigan 1996, Girard et al. 2002), is more susceptible to display longer latencies under anesthesia.

Although a thorough (and difficult) study of visual latencies of parvocellular and magnocellular compartments remains to be done in the awake animal, a possible reason for missing long latencies would be that we only sampled magnocellular dominated neurons. This seems unlikely since V4 neurons usually receive mixed parvo- and magnocellular inputs (Ferrera et al. 1992). Furthermore, the distribution of the latencies that we measured (data not shown) does not show evidence of two separate subpopulations of latencies.

F. CONCLUSION

The short latencies we measured in area V4 of the awake macaque monkey contradict the outlier status of this area. Latencies ranging from 40 to 80 ms fit well the scheme of fast processing of visual information within cortical areas (Nowak and Bullier 1997). These authors have postulated a "10 ms rule" for the visual information to travel from one area to the next. Latencies in successive areas of the visual pathway differ by 5-10 ms, which correspond to the axonal conduction delay (usually < 2 ms, Girard et al. 2001) and to the integration time (6-7 ms in general, Volgushev et al. 1995). Maunsell and Gibson (1992) have measured latencies as short as 20 ms in V1 of the awake macaque monkey with the method that we borrowed from them ('onset latency'). In area V2, which is the next step of visual processing, systematic data are lacking but different authors (Hegde and Van Essen 2004, Zhou et al. 2000) provide evidence of latencies as short as 30 ms in the awake macaque monkey. The following step of visual processing is area V4 where we recorded latencies as short as 40 ms. This is in agreement with the rule of 10 ms steps between cortical areas and compatible with the difference in starting point of population histograms of latencies obtained in areas V1 and V4 by Maunsell (1987).

Theoretical studies (Bar 2003, Bullier 2003) and psychophysical studies coupled with electrophysiology (VanRullen and Thorpe 2001) have established the necessity of a fast feedforward pass of visual information that allows a preprocessing of visual information. According to the authors, this preprocessing may account for the speed up of behavioural performances, thus allowing to perform sophisticated visual analysis, like a categorization of complex natural images, within 150 ms (Thorpe et al. 1996). With a mean latency of about 50 ms, visual processing in area V4 is compatible with such rapid visual abilities. This fast component of visual processing is very likely to be magnocellular since humans have good performances in natural images identification at low contrast (Mace et al. 2005). Area V4 is generally considered to be a major component of the parvo-dominated ventral visual pathway. However, it is firmly established that magnocellular information reach area V4 (DeYoe and Van Essen 1985, Ferrera et al., 1992) and anatomical studies suggest that it is a central node between the dorsal and ventral streams of visual processing (Baizer et al. 1991; Jouve et al. 1998). The presence of short latencies in area V4 emphasizes its relay position between the dorsal and ventral visual streams. Future studies have to determine the respective role of the magnocellular component that travel through the ventral and the dorsal visual streams.

G. ACKNOWLEDGMENTS

We thank R.T. Born and P. Barone for useful comments, L.G. Nowak, J-M. Hupé for help with data processing, F. Lefèvre and S. Aragones for animal care. This work was supported by grants from the Information Society Technologies (INSIGHT2+, #2000 29688, Neuronal basis of coding of 3D shape and material properties for recognition) and the Fondation de France.

H. REFERENCES CITED

- Baizer JS, Ungerleider LG, and Desimone R.** Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *J Neurosci* 11: 168-190., 1991.
- Bar M.** A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15: 600-609., 2003.
- Bichot NP, Rossi AF, and Desimone R.** Parallel and serial neural mechanisms for visual search in macaque area v4. *Science* 308: 529-534, 2005.
- Bullier J.** Hierarchies of cortical areas. In: *The Primate Visual System*, edited by Kaas JH and Collins CE: CRC Press, 2003, p. 181-204.
- Connor CE, Gallant JL, Preddie DC, and Van Essen DC.** Responses in area V4 depend on the spatial relationship between stimulus and attention. *J Neurophysiol* 75: 1306-1308, 1996.
- De Weerd P, Desimone R, and Ungerleider LG.** Generalized deficits in visual selective attention after V4 and TEO lesions in macaques. *Eur J Neurosci* 18: 1671-1691., 2003.
- DeYoe EA and Van Essen DC.** Segregation of efferent connections and receptive field properties in visual area V2 of the macaque. *Nature* 317: 58-61., 1985.
- Fabre-Thorpe M, Richard G, and Thorpe SJ.** Rapid categorization of natural images by rhesus monkeys. *Neuroreport* 9: 303-308, 1998.
- Ferrera VP, Nealey TA, and Maunsell JH.** Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature* 358: 756-761, 1992.
- Fize D, Vanduffel W, Nelissen K, Denys K, Chef d'Hotel C, Faugeras O, and Orban GA.** The retinotopic organization of primate dorsal V4 and surrounding areas: A functional magnetic resonance imaging study in awake monkeys. *J Neurosci* 23: 7395-7406., 2003.
- Gattass R, Sousa AP, and Gross CG.** Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8: 1831-1845, 1988.
- Girard P, Hupe JM, and Bullier J.** Feedforward and feedback connections between areas v1 and v2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85: 1328-1331., 2001.
- Girard P, Lomber SG, and Bullier J.** Shape Discrimination Deficits During Reversible Deactivation of Area V4 in the Macaque Monkey. *Cereb Cortex* 12: 1146-1156., 2002.
- Hegde J and Van Essen DC.** Temporal dynamics of shape analysis in macaque visual area v2. *J Neurophysiol* 92: 3030-3042, 2004.
- Jouve B, Rosenstiehl P, and Imbert M.** A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. *Cereb Cortex* 8: 28-39., 1998.
- Mace MJ, Thorpe SJ, and Fabre-Thorpe M.** Rapid categorization of achromatic natural scenes: how robust at very low contrasts? *Eur J Neurosci* 21: 2007-2018, 2005.
- Maunsell JH and Gibson JR.** Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* 68: 1332-1344., 1992.
- Maunsell JHR.** Physiological evidence for two visual subsystems. In: *Matters of intelligence*, edited by Vaina LM. Dordrecht, Holland.: Reidel Pub. Comp., 1987, p. 59-87.
- McAdams CJ and Maunsell JHR.** Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19: 431-441, 1999.
- Merigan WH.** Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Vis Neurosci* 13: 51-60, 1996.
- Nowak LG and Bullier J.** The timing of information transfer in the visual system. In: *Extrastriate visual cortex in primates*, edited by Rockland KS, Kaas JH and Peters A. New York: Plenum Press, 1997, p. 205-241.
- Nowak LG, Munk MH, Girard P, and Bullier J.** Visual latencies in areas V1 and V2 of the macaque monkey. *Vis Neurosci* 12: 371-384., 1995.

- Perrett DI, Rolls ET, and Caan W.** Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47: 329-342, 1982.
- Pigarev IN, Nothdurft HC, and Kastner S.** Evidence for asynchronous development of sleep in cortical areas. *Neuroreport* 8: 2557-2560, 1997.
- Raiguel SE, Lagae L, Gulyas B, and Orban GA.** Response latencies of visual cells in macaque areas V1, V2 and V5. *Brain Res* 493: 155-159, 1989.
- Schiller PH.** The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis Neurosci* 10: 717-746, 1993.
- Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, and Leventhal AG.** Signal timing across the macaque visual system. *J Neurophysiol* 79: 3272-3278, 1998.
- Tanabe S, Umeda K, and Fujita I.** Rejection of false matches for binocular correspondence in macaque visual cortical area v4. *J Neurosci* 24: 8170-8180, 2004.
- Thorpe S, Fize D, and Marlot C.** Speed of processing in the human visual system. *Nature* 381: 520-522, 1996.
- VanRullen R and Thorpe SJ.** The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci* 13: 454-461., 2001.
- Vogels R.** Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *Eur J Neurosci* 11: 1239-1255, 1999.
- Vogels R and Orban GA.** Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *J Neurophysiol* 71: 1428-1451., 1994.
- Volgushev M, Vidyasagar TR, and Pei X.** Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex. *Vis Neurosci* 12: 621-628., 1995.
- Walsh V, Butler SR, Carden D, and Kulikowski JJ.** The effects of V4 lesions on the visual abilities of macaques: shape discrimination. *Behav Brain Res* 50: 115-126, 1992.
- Wang Y, Zhou Y, Ma Y, and Leventhal AG.** Degradation of signal timing in cortical areas v1 and v2 of senescent monkeys. *Cereb Cortex* 15: 403-408, 2005.
- Zhou H, Friedman HS, and von der Heydt R.** Coding of border ownership in monkey visual cortex. *J Neurosci* 20: 6594-6611., 2000.

FIGURES

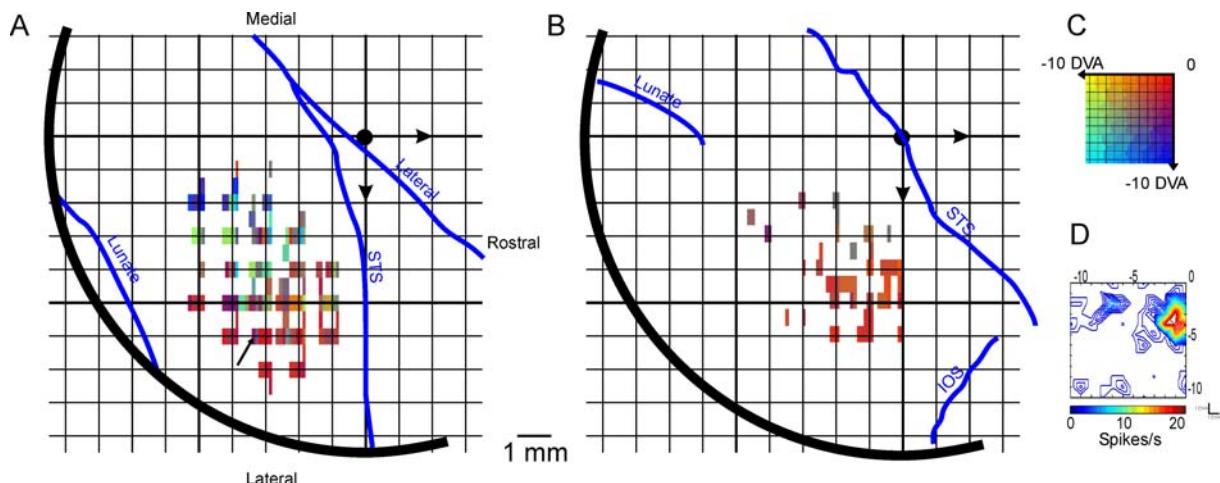


Figure 5-1 : Lateral view of area V4 of both macaque monkeys (A : Monkey 1 and B : Monkey 2) including the positions of the circular chamber (heavy line) and of the sulci (blue lines). Each point in A and B corresponds to the location (x and y coordinates on the vernier mounted on the chamber) of a recorded neuron. The color of a given point indicates the RF position in the visual field according to the color code in C. C: example of a RF mapping. The black arrow indicates on A the penetration on which neuron in C was recorded. One should note that the maps in A and B are in agreement with the maps shown in Gattass et al (1988). DVA stands for degrees of visual angle.

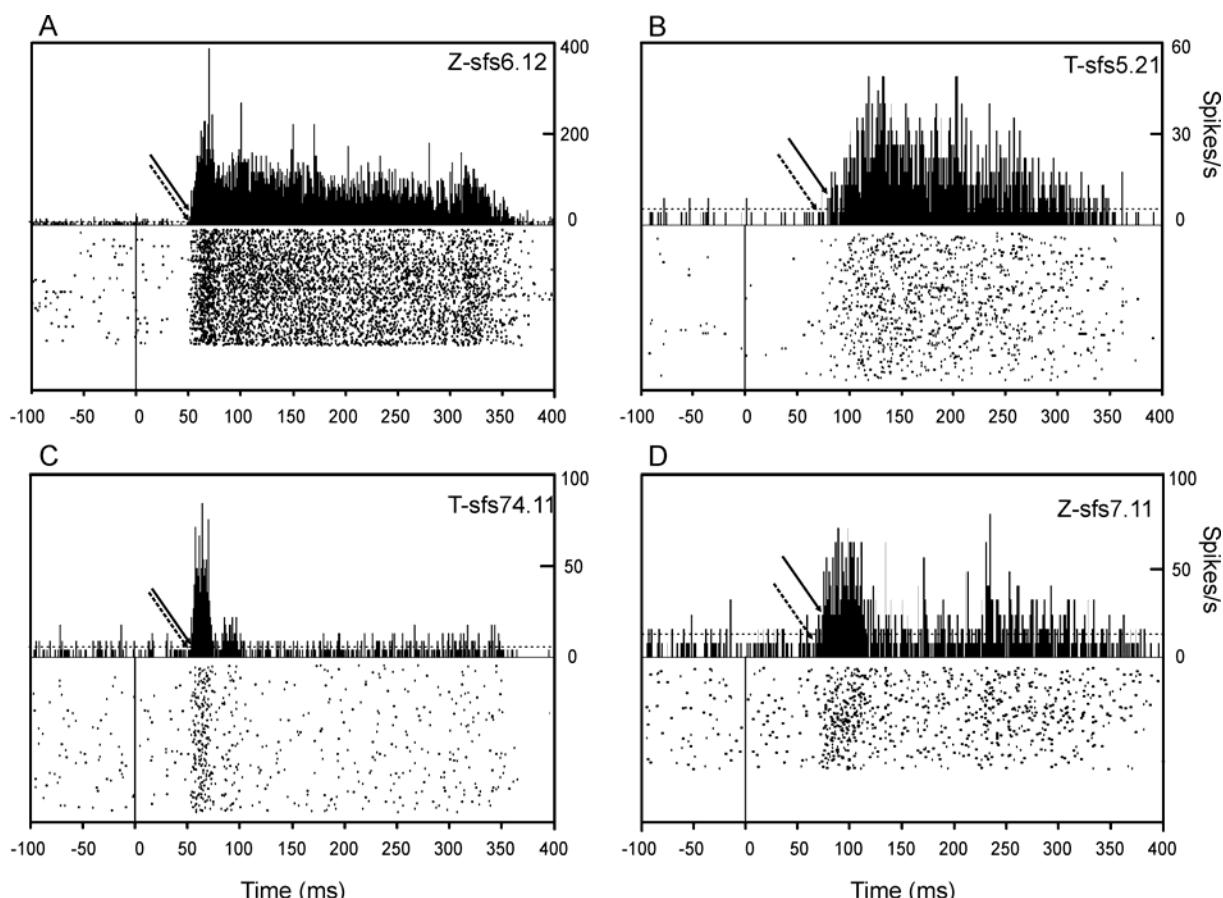


Figure 5-2 : Examples of the various types of visual responses and the corresponding latencies obtained in V4. A - Sustained activity following a transient peak. B - Sustained activity without a transient peak. C - Transient response only. D – Transient response with weak signal to noise ratio. Neurons illustrated in B and C were obtained in monkey MONKEY 1 whereas neurons illustrated in A and D in monkey

MONKEY 2. PSTH and the corresponding raster plots are represented in each panel. The latencies of response are indicated by arrows (solid and dashed arrows for ‘onset’ method and ‘derivative’ method respectively) on the four PSTHs. Corresponding values are mentioned on the top left part of each PSTH. Right upper corner indicates the register name of the neuron. All trials are lined up with the onset of visual stimulation (Time =0 ms). Bin width is 1 ms.

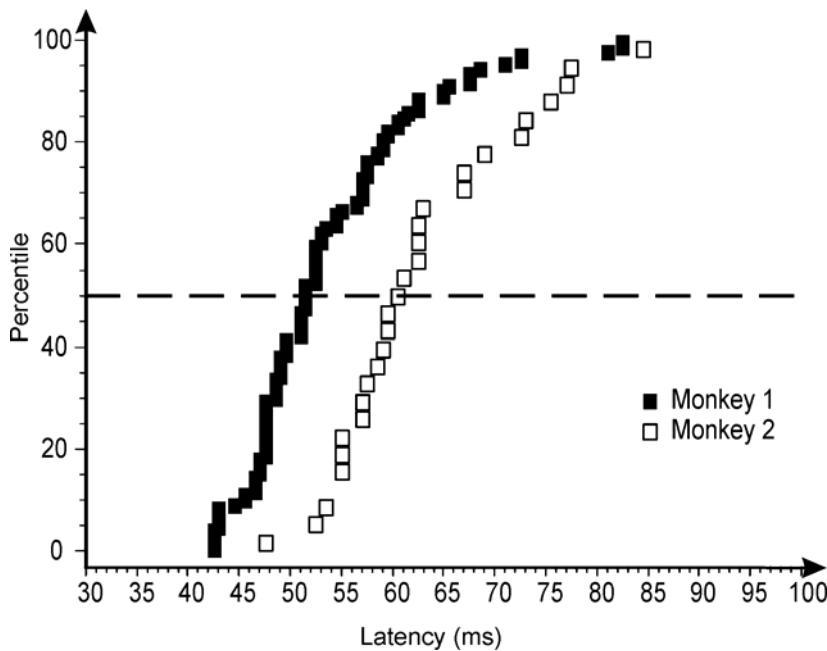


Figure 5-3 : Cumulative distributions of latencies of visual responses. Black squares and open squares represent results of latencies of Monkey 1 neurons and latencies of Monkey 2 neurons computed with ‘onset’ method.

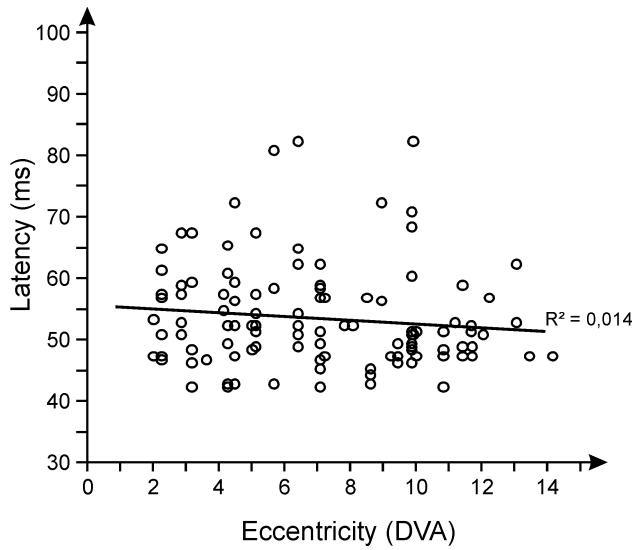


Figure 5- 4 : Absence of correlation between latencies and RFs eccentricities (data from Monkey 1).

TABLE 5-1

Monkey	MONKEY 1		MONKEY 2	
Method	Onset	<u>Derivative</u>	Onset	<u>Derivative</u>
Mean (ms)	54	51	63	59
Std. Dev	8.4	8.6	8.8	9.1
S.E.M	0.8	0.8	1.7	1.7
Maximum	82.5	80	84.5	80
Minimum	42.5	40	47.5	45
Count	113	113	28	28
Median	51	50	61	60

Table 5-1 : Latencies of V4 neurons recorded on monkeys 1 and 2 with both methods.

ANALYSES COMPLEMENTAIRES SUR LES LATENCES DE NEURONES.

1. Une autre méthode : Spike Train Analysis (STA)

J'ai utilisé une autre méthode de calcul de latences pour des analyses complémentaires, la *Spike Train Analysis* (STA). Cette méthode a été utilisée par l'équipe de Schmolesky afin de calculer les latences des neurones de plusieurs aires visuelles chez le macaque. Elle diffère des deux précédentes méthodes (onset et dérivés secondes) car elle ne se base pas sur l'histogramme des réponses (PSTH) mais calcule une valeur de latence pour chaque essai. Les méthodes de calcul sont détaillées dans les publications de Hanes et al (1995), Thompson et al. (1996) et Schmolesky et al. (1998). Les analyses statistiques permettent ensuite d'extraire de ces calculs une latence pour chaque neurone. Le tableau suivant reprend les données des latences calculées avec les 2 précédentes méthodes et y rajoute les valeurs caractéristiques calculées avec le STA.

Animal	MONKEY 1			MONKEY 2		
Méthodes	Onset	Derivative	STA	Onset	Derivative	STA
Moyenne (ms)	54	51	65	63	59	73
Std. Dev	8.4	8.6	13.7	8.8	9.1	10.2
E.S.M	0.8	0.8	1.3	1.7	1.7	1.9
Maximum	82.5	80	116	84.5	80	99
Minimum	42.5	40	49	47.5	45	58
N	113	113	113	28	28	28
Médiane	51	50	61	61	60	70

Tableau 5-2 : même tableau que Table 5-1 mais sont rajoutées pour chaque animal les valeurs caractéristiques de latence calculées avec la méthode STA.

Avec cette nouvelle méthode, nous enregistrons une hausse significative des latences pour les 2 animaux (Test Man-Whitney, $p<0.001$). Toutefois, les premières latences restent relativement précoces (49 ms chez Tima et 58 ms chez Zrem) en comparaison avec les valeurs relevées dans l'étude de Schmolesky

2. Correction de l'erreur aléatoire

Les deux méthodes (Onset et Derivative) calculent les latences à partir du PSTH des neurones où tous les essais sont sommés. Il est donc logique de penser que les essais qui ont une latence plus courte bénéficient d'un décalage de synchronisation minimal de 11 ms. Il faudrait donc retrancher 11 ms aux valeurs caractéristiques des latences pour les deux animaux.

Pour la troisième méthode (STA), une analyse statistique est entreprise afin de prendre en compte l'erreur due au décalage temporel.

Afin de prendre en compte ce décalage temporel dans la méthode de calcul STA des latences, je rajoute cette erreur sur chacun des essais par une méthode de répétition aléatoire (bootstrap).

Le décalage temporel s'échelonne de 11 ms à 24 ms. Je crée une matrice de taille N (N = nombre d'essais pris en compte pour l'analyse STA) à l'intérieur de laquelle sont distribuées aléatoirement des valeurs du décalage allant de 11 ms à 24 ms. La distribution des valeurs de cette matrice doit être normale de sorte que la médiane des valeurs soit de 18 ms. Puis pour chaque essai, je raccourcis l'essai de la valeur correspondante de la matrice. Cette nouvelle matrice est ensuite analysée par la méthode de STA afin d'en sortir la latence de réponse du neurone. Cette opération est réalisée 1000 fois afin d'exclure toute variabilité. La valeur (STA_bstrp) est la médiane des 1000 valeurs (la distribution de ces 1000 valeurs est normale). Le Tableau 5-3 reprend les valeurs caractéristiques des latences calculées avec la méthode STA et en prenant en compte l'erreur de synchronisation.

Animal	Timax	Zrem
Moyenne	48	57
Ecart-type	14.1	12.6
E.S.M	1.3	2.4
Maximum	100.5	97.5
Minimum	33	41.5
N	113	28
Médiane	43	52

Tableau 5-3 : Valeurs caractéristiques des latences des neurones après rectification de la synchronisation d'après la méthode STA.

Les latences moyennes calculées en prenant compte l'erreur statistique sont significativement inférieures à celles relevées précédemment avec la méthode STA (test Man-Whitney, $p<0,001$). Elles sont raccourcies de 16 à 17ms en moyenne. Les corrections statistiques ne remettent pas en cause les conclusions de l'article. Elles laissent entrevoir la possibilité de latences très courtes dans l'aire V4 qui demanderaient à être vérifiées sans erreur.

Chapitre 6

Synthèse des principaux résultats et discussion générale

Ce dernier chapitre présente une synthèse des principaux résultats des 3 études de cette thèse. Puisque les résultats de chacune de ces études ont été discutés en détail à la fin de chaque chapitre, je ne reprendrai ici que les résultats majeurs afin de proposer un rôle fonctionnel de l'aire V4 dans la perception tridimensionnelle des objets.

A. PERTINENCE DE L'ETUDE

L'aire V4 dont nous avons étudié les propriétés neuronales fait partie du système ventral impliqué dans la reconnaissance des objets. Cette aire participe activement au traitement des indices bidimensionnels tels que la forme et la couleur de l'objet. En effet, au niveau de la couleur de l'objet, cette aire joue un rôle primordial dans la constance de la couleur en nous permettant de maintenir constante notre perception de la couleur de l'objet indépendamment des conditions d'illumination (Kusunoki *et al.*, 2006). Quant à la forme 2D de l'objet, les neurones de V4 seraient sélectifs à une combinaison d'éléments (Pasupathy et Connor, 1999; Pasupathy et Connor, 2001). Ces résultats sont en accord avec les théories de Biederman (1987) et Riensuhuber et Poggio (1999) qui énoncent qu'une forme complexe est représentée par une combinaison d'éléments simples : les primitives (Biederman, 1987; Riesenhuber et Poggio, 1999). Ainsi, la population de neurones de V4 coderait tous les éléments d'un alphabet de primitives permettant l'analyse de n'importe quelle forme complexe. L'élément de texture a souvent été testé dans des études de segmentation qui ont démontré que l'aire V4 participe à une segmentation d'objets sur la base de ces textures (Kastner *et al.*, 2000; Merigan, 2000). Peu d'études se sont intéressées au rôle fonctionnel de la texture en tant que propriété matérielle de l'objet nécessaire à sa reconnaissance même si l'apport de la texture améliore la reconnaissance de l'objet (Price et Humphreys, 1989; Rossion et Pourtois, 2004).

Outre les indices bidimensionnels, l'aire V4 traite également des indices tridimensionnels de l'objet nous permettant d'appréhender sa structure tridimensionnelle ou l'orientation de sa surface 3D. Parmi ces indices, des neurones de V4 présentent une sélectivité au gradient de disparité et au gradient d'illumination (Hanazawa et Komatsu, 2001; Hinkle et Connor, 2001; Hinkle et Connor, 2002). Il a été également démontré, chez les humains et chez le singe, que le cortex inférot temporal vers lequel l'aire V4 projette massivement, participe au traitement d'indices monoculaires statiques tels que les ombrages (Moore et Engel, 2001; Sereno *et al.*, 2002; Vogels et Biederman, 2002). .

Mon travail de thèse s'est intéressé à déterminer des bases neuronales du traitement de la texture naturelle d'un objet et de sa structure tridimensionnelle sur la base d'ombrages. Les principaux résultats de ces études sont synthétisés dans cette partie afin de proposer un rôle fonctionnel de l'aire V4 dans le codage de ces deux indices.

B. CODAGE D'INDICES TRIDIMENSIONNELS AU SEIN DE V4

Synthèse générale

Au cours de mon doctorat, nous avons réalisé deux principales études au sein de l'aire V4 chez le macaque vigile ; une sur le traitement de la texture et l'autre sur l'encodage des ombrages. Les résultats de ces études semblent indiquer que l'aire V4 joue un rôle primordial dans le codage de ces indices monoculaires pour une perception tridimensionnelle de l'objet. Ce nouveau rôle fonctionnel de l'aire V4 est à rajouter aux rôles de cette aire dans le traitement d'indices bidimensionnels tels que la couleur et la forme des objets ou d'indices tridimensionnels tels que la disparité ou les gradients de textures.

La première étude a démontré que les neurones de V4 présentent une sélectivité à la texture naturelle et codent pour un ensemble de textures selon leurs paramètres statistiques notamment leur luminance moyenne et leur hétérogénéité. De plus, ces neurones sélectifs présentent une relative invariance de leur taux de décharge pour leur texture préférée selon les conditions d'illumination. Nos résultats ont également démontré que les cellules de V4 se comportent comme un ensemble de filtres GABOR pour une tâche de classification de textures. La population de neurones de V4 contient donc les mécanismes de base permettant de classifier les textures naturelles entre elles. Une propriété matérielle de l'objet, sa texture naturelle, semble donc être traitée au sein de la voie visuelle ventrale, dans l'aire V4 et pourrait ainsi aider à sa reconnaissance.

De nouvelles propriétés fonctionnelles de V4

Cette conclusion s'accorde avec les résultats d'études neuropsychologiques qui ont démontré le rôle fonctionnel de V4 dans le traitement de plusieurs propriétés de l'objet. Concernant la couleur, les études de lésions corticales ont démontré que l'aire V4 joue un rôle primordial dans la constance de la couleur indépendamment des conditions d'illumination. Suite à nos résultats, nous proposons donc que l'aire V4 participe également à la constance d'une autre propriété matérielle de l'objet, la texture. La texture d'un objet est un attribut intrinsèque de l'objet qui nous aide à sa reconnaissance au même titre que la couleur même si les effets de la couleur sur la reconnaissance d'objets sont plus controversés (Price et Humphreys, 1989; Rossion et Pourtois, 2004). En effet, nos résultats montrent qu'un grand nombre de neurones sélectifs aux textures naturelles présente une relative invariance à l'angle d'incidence de la source d'illumination pour leur texture préférée. Toutefois, l'étude de

Hanazawa et Komatsu (2001) semble contredire notre conclusion. Elle souligne en effet que les neurones de V4 codant des gradients de texture montrent une préférence pour une source d'illumination située le long d'un axe vertical. Nos résultats sur les ombrages soulignent que la population de neurones est sélective à la direction d'illumination même si nous ne retrouvons pas une préférence unidirectionnelle. Il est donc possible que les neurones de V4 soient sélectifs à la direction d'illumination tout en étant invariants aux angles d'incidence de la source d'illumination pour une direction donnée. En effet, les textures naturelles que nous avons utilisées sont éclairées selon différents angles d'incidence dans une même direction (vers la droite, cf. Figure 2-). En conséquence, les ombrages dus aux aspérités de la texture ne changent pas de position à l'intérieur de la structure, seule leur taille varie. L'aspect de la texture évolue donc peu avec ces modifications des angles d'incidence. A l'inverse, les formes tridimensionnelles que nous avons utilisées sont illuminées selon quatre directions différentes (cf. Figure 2-). L'aspect de ces formes change donc beaucoup plus avec ces modifications d'éclairage. On peut donc penser que si nous avions plusieurs angles d'incidence pour une direction donnée, notre population de neurones serait invariante aux angles d'incidence car l'aspect de la forme tridimensionnel serait peu modifié, comme c'est le cas avec les textures. Pour le cas inverse, si nos textures naturelles étaient illuminées selon différentes directions, les patterns d'ombrages au sein-même de la texture seraient alors fortement modifiés. Cependant les humains peuvent reconnaître une texture naturelle indépendamment des conditions d'illumination. Les données bibliographiques manquent à ce sujet mais dans notre entourage visuel, nous reconnaissions la texture naturelle de l'objet (par exemple, l'écorce d'un arbre) à n'importe quelle heure de la journée alors que les conditions d'illumination varient. Les réponses des neurones seraient alors invariantes à la direction d'illumination permettant d'extraire de la constance de la texture.

Les concavités et les convexités qui définissent les formes tridimensionnelles que nous avons utilisées entraînent des gradients d'ombrages fonction de la direction d'illumination. Les particularités de la forme sont toujours présentes mais définies différemment selon les gradients d'ombrages. Notre population de neurones ne présente pas de sélectivité importante aux formes tridimensionnelles qui leur sont présentées. La plupart des neurones ont un même taux de décharge quelle que soit la forme considérée. Cette population ne coderait donc pas la présence d'indices pertinents de la forme tridimensionnelle sur la base des ombrages, comme par exemple la présence d'une concavité sur la partie supérieure de la forme. Alors qu'une sélectivité à un élément particulier de la forme 2D de l'objet est présente au sein de l'aire V4 (Pasupathy et Connor, 2001), nos résultats semblent indiquer qu'une telle sélectivité n'est pas

retrouvée pour des éléments tridimensionnels définis par des ombrages telles que (concavité ou convexités). Cependant, ces résultats sont en accord avec de l'étude de Hegde et Van Essen (2005) qui démontrent que les neurones de V4 ne sont pas sélectifs à des éléments de courbures tridimensionnels sur la base de gradient de disparité (Hegde et Van Essen, 2005). Cette sélectivité s'installerait au niveau d'une aire antérieure du cortex inférotemporal, l'aire TEs (Janssen *et al.*, 2001). La sélectivité aux courbures tridimensionnelles de l'objet sur la base des ombrages pourrait donc également se mettre en place au niveau du cortex inférotemporal. L'étude de Vogels et Biederman (2002) rapporte que leur population de neurones de l'aire TE est très sélective à la structure tridimensionnelle de l'objet basée sur les ombrages (Vogels et Biederman, 2002) même s'ils utilisent comme stimuli des formes très différentes des nôtres créées par des assemblages de 2 ou 3 géons. Une analyse complémentaire portant spécifiquement sur les réponses des neurones de V4 aux éléments 3D de la forme en utilisant un set élargi de formes tridimensionnelles bien définies (par exemple, un set de stimuli comme celui présenté dans l'étude de Pasupathy *et al.* –1999- mais avec des formes 3D) permettrait de mieux caractériser la sélectivité des neurones à des indices caractéristiques de la forme 3D.

Nous avons vu en introduction que l'aire V4 participe également au traitement d'indices tridimensionnels tels que la disparité, l'interposition ou les gradients de texture (Zhou *et al.*, 2000; Hanazawa et Komatsu, 2001; Hinkle et Connor, 2001; Hinkle et Connor, 2002). Ces indices une fois combinés permettent de percevoir la forme tridimensionnelle de l'objet. Les résultats de notre seconde étude démontrent qu'un indice monoclulaire statique tel que la distribution des ombrages sur une surface est également traité au sein de V4. En effet, nous avons mis en évidence qu'une population de neurones de V4 permet de caractériser le pattern interne tridimensionnel de nos stimuli en les dissociant des stimuli contrôles bidimensionnels. Ainsi cette aire enverrait ce type d'information tridimensionnelle vers le cortex inférotemporal vers lequel elle projette massivement. De même, de nombreuses études ont démontré chez l'Homme et chez l'animal que cette zone corticale (IT ou LOC son équivalent chez l'Homme) reconnaît l'aspect tridimensionnel des objets en se basant sur les ombrages (Moore et Engel, 2001; Sereno *et al.*, 2002; Vogels et Biederman, 2002).

Codage réel de la 3D

Les conclusions de nos deux études rapportent qu'une population de neurones est intéressée par l'aspect tridimensionnel de nos stimuli. Mais s'agit-il réellement de la 3D ? Dans le cas des textures, l'aspect 3D apparaît dans leurs versions originales lorsque

l’agencement des phases de Fourier est respecté. L’image originale contient alors plus de primitives ou « textons » que son image contrôle. Comme nous les présentons en situation passive, deux aspects importants restent à démontrer ; est-ce au niveau de V4 elle-même que la distribution régulière des textons dans une image est identifiée comme une texture ? En d’autres termes existe-t-il un liage (ou binding) de tous les textons ou bien seulement de quelques éléments discriminants qui définissent une réponse du neurone à un indice local de la texture ? Nos résultats sur une faible population de neurones ($n=10$) pour lesquels nous avons déplacé les stimuli à l’intérieur du CR du neurone nous indiquent que cette population de neurone est invariante au déplacement des stimuli. La sélectivité des neurones persiste en dépit du déplacement des stimuli. Les neurones de V4 ne coderaient donc pas les textures sous forme d’indices locaux présents à l’intérieur de la texture mais plutôt comme une entité globale. En traitant d’une autre problématique concernant la texture, le modèle de Thielscher et Neumann (2005) propose que les cellules de V4 représentent le plus haut niveau hiérarchique pour une segmentation de textures. Les neurones détecteraient les bordures des textures et non les textons individuels (Thielscher et Neumann, 2005). Cette hypothèse est en accord avec les résultats d’études en neuroimagerie ou d’études de lésions chez le singe qui démontrent que V4 participe au traitement de bordures sur la texture (De Weerd *et al.*, 1996; Kastner *et al.*, 2000; Merigan, 2000). Le second aspect qui paraît très difficile à démontrer concerne la reconnaissance même de la 3D. Le *quale* de la 3D est-il bien représenté dans l’aire V4 ? Les textures naturelles et les formes tridimensionnelles nous donnent une sensation forte de la profondeur mais est-ce que les différences de réponses de notre population de neurones suffisent à affirmer que c’est au sein de l’aire V4 que se crée cette sensation ? Nous pouvons remarquer que ce même problème se pose avec la couleur. Alors qu’il est clairement établi que l’aire V4 joue un rôle primordial dans la constance de la couleur, aucune étude ne conclut que cette aire est également le siège des *qualia* de la couleur.

Au cours de l’introduction, nous avons noté que des indices tridimensionnels se retrouvent également traités au niveau d’aires corticales dorsales. En effet, le groupe de Tsutsui a enregistré à l’intérieur du sillon intra-pariétal (dans l’aire CIP) des neurones sélectifs au gradient de disparité mais aussi au gradient de textures. Tsutsui définit cette aire comme un centre hiérarchiquement élevé pour le traitement de la profondeur puisque des informations concernant la 3D convergent ici (Taira *et al.*, 2001; Tsutsui *et al.*, 2002; Tsutsui *et al.*, 2003; Tsutsui *et al.*, 2005). A ma connaissance, aucune connexion directe n’a été démontrée entre l’aire CIP et l’aire V4. L’aire V4 est connectée réciproquement au niveau de la voie dorsale aux aires MT et LIP. Toutefois, ces projections sont minimes même si la proportion des

projections vers MT est plus importante que celle vers LIP (Kennedy et Barone, 1999). L'aire CIP reçoit principalement des projections des aires V3 et V3A, où les neurones sont particulièrement intéressés par la disparité. Les aires V3 et V3A projettent également vers V4 et pourraient ainsi véhiculer les informations relatives à la disparité. En adéquation avec les fonctions attribuées aux voies ventrales et dorsales, on peut supposer que les informations tridimensionnelles au niveau des aires dorsales permettraient d'appréhender la structure tridimensionnelle de l'objet afin qu'on puisse s'en saisir. Le codage d'indices tridimensionnels au sein des aires de la voie visuelle ventrale permettrait de définir sa structure tridimensionnelle afin de reconnaître l'objet. Plusieurs études rapportent que les neurones du cortex inférot temporal sont sélectifs aux indices tridimensionnels que sont les gradients de disparité, les gradients de textures et les ombrages (Janssen *et al.*, 2000; Janssen *et al.*, 2001; Moore et Engel, 2001; Sereno *et al.*, 2002; Liu *et al.*, 2004). L'aire V4 serait alors une première étape de traitement de ces indices tridimensionnels tout en étant différente de celle effectuée au niveau du cortex inférot temporal. Par exemple, au sujet du gradient de disparité, l'aire V4 présente une majorité de cellules sélectives à la disparité, ces cellules permettent de coder l'orientation d'une barre tridimensionnelle mais pas la courbure de contours comme cela a été démontré au niveau de l'aire TEs (Janssen *et al.*, 2000; Hegde et Van Essen, 2005). De même, au niveau du codage des ombrages, nos résultats soulignent que les neurones de V4 sont sensibles à ces ombrages afin de coder la forme tridimensionnelle de l'objet tout en montrant une variation de leur réponse en fonction de la direction d'illumination. L'invariance à la direction d'illumination n'a donc pas été observée au sein de l'aire V4 et s'installerait au niveau du cortex inférot temporal, (Vogels et Biederman, 2002). Cette invariance à la direction d'illumination s'ajoutera à l'invariance à la taille et la position de l'objet démontrée au sein du cortex IT par l'étude de du groupe de Ito (Ito *et al.*, 1995). Ceci permettrait une meilleure reconnaissance de l'objet et cela sous différentes conditions de présentation (illumination, position ou taille).

Rôle de la texture pour une reconnaissance d'objet

Une interrogation intéressante concerne l'apport de la texture dans la reconnaissance de l'objet. La texture peut-elle être définie comme une information de détail de l'objet et est-elle traitée comme la forme globale de l'objet ? D'après les études du groupe de Maunsell (Ferrera *et al.*, 1992; Ferrera *et al.*, 1994), nous savons que l'aire V4 reçoit à la fois des projections magno- et parvocellulaires et qu'un même neurone de V4 peut recevoir à la fois des projections des deux voies. La voie magnocellulaire est une voie à conduction rapide et

véhicule principalement des informations telles que les fréquences spatiales basses ou des contrastes faibles présents dans les images. Cette voie magnocellulaire permettrait une description assez grossière de l'image donc plutôt de la forme globale de l'objet. La voie parvocellulaire est caractérisée par une vitesse de conduction plus lente que celle de la voie magnocellulaire. Cette voie parvocellulaire véhiculerait les informations de détails à fréquence spatiale élevée et contraste assez fort. Ces deux voies de conduction en convergeant vers une même aire entraînent donc un décalage temporel dans les activations des neurones. Ce décalage se retrouve dans l'arrivée de deux sortes d'information: une information globale et une information fine. Au moyen de tests comportementaux chez l'homme, cette distinction vision globale - vision fine a été testée par Navon (1977) qui démontre que la forme globale de l'objet est discriminée plus rapidement que ses détails internes (Navon, 1977). Il nomme ce phénomène la précédence globale. Toutefois, du fait de la construction hiérarchique des stimuli (une grande lettre composée de petites lettres), les aspects local et global d'un objet sont étroitement liés générant une interférence du global sur le local et il paraît difficile de pouvoir séparer ces deux aspects. Chez le macaque, l'étude de Tanaka et Fujita (2000) montre qu'une discrimination sur la forme globale des stimuli est plus rapide qu'une discrimination sur les détails (Tanaka et Fujita, 2000). Les stimuli utilisés pour cette étude sont des lettres orientées qui assemblées d'une certaine manière forment une lettre, un résultat contraire étant observé par le groupe de J. Fagot sur des babouins (Fagot et Deruelle, 1997). Au point de vue neurophysiologique, une inactivation réversible de l'aire IT entraîne, chez le macaque, une mauvaise discrimination de détails alors que la discrimination de formes globales est toujours possible pour les animaux (Horel, 1994). Ainsi, la sélectivité aux textures naturelles observées au sein de notre étude pourrait être issue du traitement des détails effectués dans l'aire IT car l'aire V4 reçoit une proportion importante de connexions feedback qui proviennent de IT. En effet, nous avons observé que la distinction entre les stimuli originaux et les stimuli *scrambled* apparaissait à une période tardive de la réponse (cf. Figure 3-22). Une inactivation réversible par refroidissement dans IT tout en enregistrant les réponses des neurones de V4 permettrait de tester si la sélectivité aux textures naturelles est due aux connexions feedback de IT vers V4.

Rapidité des informations au sein de V4

Dans la troisième étude, nous avons enregistré des latences de réponses relativement courtes (40 ms). Ces latences s'accordent avec la théorie du traitement rapide des informations visuelles de Nowak et Bullier (1997) à travers les aires corticales (Nowak et

Bullier, 1997). Selon ces auteurs, les informations visuelles mettent un délai minimum 10 ms pour être transmises d'une aire à la suivante dans une organisation hiérarchique. Or, l'étude de Maunsell et Gibson (1992) révèle des latences de 20 ms dans l'aire V1 du macaque vigile (Maunsell et Gibson, 1992). Au niveau de l'aire V2, étape suivante de la voie visuelle ventrale, certaines études rapportent des latences de 30 ms (Zhou *et al.*, 2000; Hegde et Van Essen, 2004). Puisque V2 projette massivement sur V4, il semble cohérent, d'après les observations de Nowak et Bullier, de retrouver des latences de 40ms dans V4.

Les modèles théoriques (Bar, 2003; Bullier, 2003) et des études psychophysiques couplées avec de l'électrophysiologie (VanRullen et Thorpe, 2001) ont démontré la nécessité d'un traitement rapide des informations visuelles qui permettrait une catégorisation d'images naturelles (Thorpe *et al.*, 1996). Nos latences courtes dans V4 s'accordent avec ces théories et seraient ainsi dues aux connexions magnocellulaires qui convergent vers V4. L'aire V4 est traditionnellement incluse dans la voie visuelle ventrale mais les strictes dichotomies parvocellulaire / magnocellulaire et voies ventrale / dorsale ne sont plus fondées puisqu'un mélange des entrées parvo- et magnocellulaires est montré au moins dès V2. Certaines études soulignent même que les voies parvo- et magnocellulaires convergent vers V4 (Ferrera *et al.*, 1992; Ferrera *et al.*, 1994). L'aire V4 serait alors un centre d'intégration des informations issues des voies parvo/magnocellulaires, des informations issues des aires dorsales (MT, LIP) y convergent également, leurs rôles fonctionnels restent à déterminer.

V4 centre d'intégration ?

Une autre propriété des neurones de V4 peut nous aider à mieux comprendre sa situation, si nous considérons l'aire V4 en tant que carrefour des voies dorsales et ventrales. Ses neurones sont fortement sensibles à l'apprentissage (perceptual learning, (Rainer *et al.*, 2004). Imaginons la situation, volontairement simpliste, dans laquelle un primate, animal qui interagit fortement avec son environnement du point de vue manuel, rencontre un nouvel objet, par exemple un fruit tombé d'un arbre. Il peut reconnaître la forme ronde du fruit grâce à sa voie visuelle ventrale dont fait partie V4. Il devine qu'il peut se saisir du fruit, le faire rouler, etc., car V4 interagit avec la voie dorsale pour spécifier l'affordance de l'objet. De même, en parallèle de sa forme, il reconnaît sa texture, indépendamment des variations d'éclairage qui se produisent lorsqu'il se déplace autour grâce à sa voie ventrale. Puis, lorsqu'il saisit l'objet, il peut comprendre s'il est comestible ou non, s'il doit être épluché ou non grâce aux interactions visuo-tactiles de V4 avec la voie dorsale. Le temps passé à observer cet objet

et le manipuler permet une reconnaissance plus efficace, plus adaptée lors de nouvelles présentations avec ce même type d'objet. Bien sûr, V4 ne fait pas tout mais la notion importante ici à retenir est qu'il faut apprendre à voir.

Ce type d'hypothèse met le doigt sur la nécessité d'étapes complémentaires à notre étude. En effet, nos travaux montrent que l'aire V4 participe au traitement d'indices monoculaires de la perception tridimensionnelle mais la tâche proposée aux animaux est une tâche de fixation passive. Nous n'avons donc aucune preuve de la perception tridimensionnelle de nos stimuli. Une tâche d'apprentissage nous permettrait d'avoir un feedback sur la perception tridimensionnelle de l'animal. De plus, il a été démontré que les neurones de V4 sont influencés par la congruence visuo-somesthésique des textures (Haenny et Schiller, 1988), il serait donc intéressant de tester les réponses des cellules à des informations visuo-tactiles congruentes ou pas d'une texture naturelle.

Au niveau des ombrages, il n'est pas clairement établi que les singes perçoivent nos formes patatoïdes définies par des ombrages comme des objets 3D. Une première étape d'apprentissage dans une étude comportementale permettrait de vérifier la perception tridimensionnelle de nos formes. De plus, il a été démontré que les neurones de V4 ont leur propriétés modulées par un apprentissage perceptif (perceptual learning) (Rainer *et al.*, 2004). Après un tel apprentissage, la dichotomie 3D/2D pourrait être plus prononcée au niveau de nos décharges neuronales. Peut-être en enregistrerions-nous une plus grande dichotomie si l'animal avait à manipuler les objets dont on lui présente les photographies. Si les neurones de V4 peuvent être influencés par la voie dorsale, il reste encore beaucoup à faire pour déterminer comment s'établissent les liens fonctionnels entre V4 et les aires pariétales, notamment l'aire CIP, centre hiérarchiquement élevé pour le traitement de la 3D (Tsutsui *et al.*, 2005). On peut supposer une modulation descendante qui véhiculerait l'expérience tactile de l'animal pour arriver à percevoir la 3D dans des images 2D.

C. PERSPECTIVES

Les images de l'environnement visuel qui nous entourent arrivent à nos rétines sous la forme d'images rétiniennes planes. Afin de traiter la profondeur présente, notre système visuel utilise une variété d'indices monoculars (ombrages, perspectives, interposition, gradient de texture) et binoculaires (disparités rétiniennes). Cette perception tridimensionnelle permettra au sujet de se mouvoir dans son environnement et de reconnaître les objets et de planifier des actes moteurs en direction de ces objets.

Nos études ont démontré que l'aire V4 chez le singe semble participer au traitement d'indices monoculars tels que la texture et les ombrages. Des études électrophysiologiques comparatives dans d'autres aires corticales, notamment V2 et IT, permettraient de situer la contribution de V4 dans le traitement de ces paramètres. De même, l'enregistrement d'autres aires au niveau dorsal permettrait d'étudier les interactions visuo-tactiles qui permettent une meilleure reconnaissance des textures ou de déterminer la forme tridimensionnelle de l'objet.

Durant les tâches expérimentales de cette thèse, l'animal est confronté passivement aux stimuli. Ceci ne permet pas de conclure sur sa perception réelle de la 3D. Une période d'apprentissage des textures ou de formes tridimensionnelles (classification ou maniement d'objets) permettrait de mieux évaluer sa perception tridimensionnelle et aussi de relever l'activité des neurones pendant cet apprentissage. Dans la première étude, on pourrait par exemple apprendre à l'animal à classer les textures selon leur percept tridimensionnel ou lui apprendre à discriminer le stimulus original de leur contrôle *scrambled*. Pour la seconde étude, on pourrait demander au singe de classer les stimuli selon leur aspect 3D comme nous l'avons testé chez les humains (cf.annexe). Afin de tester si la texture (information fine) est codée de la même manière que la forme globale de l'objet, l'animal pourrait apprendre à catégoriser les stimuli en fonction de leur texture naturelle ou de leur forme 2D. Ces tâches ont déjà été apprises par nos animaux avec un taux de réussite assez important. Ce travail a été réalisé par Valérie Marchat dans le cadre d'un stage de DEA. Ses résultats comportementaux sont concluants et montrent un très bon apprentissage de la tâche. Toutefois, il reste à déterminer les bases neurophysiologiques de cet apprentissage.

Enfin, les enregistrements cellulaires effectués au cours de nos études permettent de corrélérer l'activité des neurones avec les images qui sont présentées à l'animal. Ce protocole expérimental pourrait être complété à l'avenir par une étude fonctionnelle en neuroimagerie. Ces travaux en imagerie fonctionnelle permettraient dans un premier temps de mieux cibler

les enregistrements cellulaires vers les régions cérébrales souhaitées mais également de comparer ces procédés de traitement entre l'homme et le singe dans une tâche de perception simple de stimuli ou des tâches plus « cognitives » comme une catégorisation de textures ou une reconnaissance tactile ou visuelle des textures ou formes tridimensionnelles.

BIBLIOGRAPHIE

- Abramov, I. et Gordon, J. (1994). Color appearance: on seeing red--or yellow, or green, or blue. *Annu Rev Psychol* 45: 451-85.
- Adams, D. L. et Zeki, S. (2001). Functional organization of macaque v3 for stereoscopic depth. *J Neurophysiol* 86(5): 2195-203.
- Adams, N. C., Lozsadi, D. A. et Guillory, R. W. (1997). Complexities in the thalamocortical and corticothalamic pathways. *Eur J Neurosci* 9(2): 204-9.
- Allman, J. M. et Kaas, J. H. (1971). Representation of the visual field in striate and adjoining cortex of the owl monkey (*Aotus trivirgatus*). *Brain Res* 35(1): 89-106.
- Allman, J. M. et Kaas, J. H. (1974). The organization of the second visual area (V II) in the owl monkey: a second order transformation of the visual hemifield. *Brain Res* 76(2): 247-65.
- Andrey, P. et Tarroux, P. (1998). Unsupervised segmentation of Markov random field modeled textured images using selectionist relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20: 252-263.
- Baizer, J. S., Robinson, D. L. et Dow, B. M. (1977). Visual responses of area 18 neurons in awake, behaving monkey. *J Neurophysiol* 40(5): 1024-37.
- Baizer, J. S., Ungerleider, L. G. et Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *J Neurosci* 11(1): 168-90.
- Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J Comp Neurol* 276(3): 313-42.
- Barbas, H. et Mesulam, M. M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *J Comp Neurol* 200(3): 407-31.
- Barone, P., Batardiere, A., Knoblauch, K. et Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J Neurosci* 20(9): 3263-81.
- Baumann, R., van der Zwan, R. et Peterhans, E. (1997). Figure-ground segregation at contours: a neural mechanism in the visual cortex of the alert monkey. *Eur J Neurosci* 9(6): 1290-303.
- Baylis, G. C., Rolls, E. T. et Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. *J Neurosci* 7(2): 330-42.
- Beason-Held, L. L., Purpura, K. P., Krasuski, J. S., Maisog, J. M., Daly, E. M., Mangot, D. J., Desmond, R. E., Optican, L. M., Schapiro, M. B. et VanMeter, J. W. (1998). Cortical regions involved in visual texture perception: a fMRI study. *Brain Res Cogn Brain Res* 7(2): 111-118.
- Beck, J. (1972). Similarity grouping and peripheral discrimination under uncertainty. *Am J psychol* 85: 1-19.
- Beck, J. (1982). Organization and Representation in Perception, Erlbaum.
- Bhushan, N., Rao, A. R. et Lohse, G. L. (1997). The Texture Lexicon: Understanding the categorization of visual texture terms and their relationship to texture images. *Cogn Sci* 21(2): 219-246.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94(2): 115-47.
- Biederman, I. et Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Res* 39(17): 2885-99.

Bibliographie

- Biederman, I. et Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance [published erratum appears in *J Exp Psychol Hum Percept Perform* 1994 Feb;20(1):80] [see comments]. *J Exp Psychol Hum Percept Perform* 19(6): 1162-82."
- Biederman, I. et Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognit Psychol* 20(1): 38-64.
- Boucart, M. et Humphreys, G. W. (1997). Integration of physical and semantic information in object processing. *Perception* 26(9): 1197-209.
- Boussaoud, D., Desimone, R. et Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *J Comp Neurol* 306(4): 554-75.
- Boussaoud, D., Ungerleider, L. G. et Desimone, R. (1990). Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J Comp Neurol* 296(3): 462-95.
- Braje, W. L. (2003). Illumination encoding in face recognition: effect of position shift. *J Vis* 3(2): 161-70.
- Braje, W. L., Kersten, D., Tarr, M. J. et Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology* 26(4): 371-380.
- Braje, W. L., Legge, G. E. et Kersten, D. (2000). Invariant recognition of natural objects in the presence of shadows. *Perception* 29(4): 383-98.
- Bremmer, F. (2000). Eye position effects in macaque area V4. *Neuroreport* 11(6): 1277-83.
- Brewer, A. A., Press, W. A., Logothetis, N. K. et Wandell, B. A. (2002). Visual areas in macaque cortex measured using functional magnetic resonance imaging. *J Neurosci* 22(23): 10416-26.
- Brodatz, P. (1966). *Textures--A photographic Album for Artists and Designers*. New York, Dover publications.
- Bullier, J. et Barone, P. (1997). Voies optiques intracrâniennes et lobe occipital : anatomie, fonction, développement. *Encyclopédie Médico-Chirurgicale - Ophtalmologie*. Paris, Elsevier: 1-11.
- Bullier, J. et Kennedy, H. (1983). Projection of the lateral geniculate nucleus onto cortical area V2 in the macaque monkey. *Exp Brain Res* 53(1): 168-72.
- Bullier, J. et Nowak, L. G. (1995). Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Curr Opin Neurobiol* 5(4): 497-503.
- Bullier, J., Girard, P. et Salin, P. A. (1994). The Role of area 17 in the transfer of information to extrastriate visual cortex. *Primary visual cortex in primates*. A. Peters et K. S. Rockland. New-York, Plenum press. 10: 301-330.
- Chubb, C. et Landy, M. S. (1991). *Orthogonal distribution analysis : A new approach to the study of texture perception*. Cambridge.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L. et Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J Neurosci* 11(8): 2383-402.
- Cowey, A. (1964). Projection of the retina on to striate and prestriate cortex in the squirrel monkey, *Saimiri Sciureus*. *J Neurophysiol* 27: 366-393.
- Damasio, A., Yamada, T., Damasio, H., Corbett, J. et McKee, J. (1980). Central achromatopsia: behavioral, anatomic, and physiologic aspects. *Neurology* 30(10): 1064-71.
- Dana, K. J., Van Ginneken, B., Nayar, S. K. et Koenderink, J. J. (1999). Refelctance and Texture of Real World Surfaces. *ACM Transactions on Graphics (TOG)* 18(1): 1-34.
- Daubechies, I. (1992). *Ten lectures on Wavelets*. CBMS-NSF Lecture Notes 61.

Bibliographie

- De Weerd, P., Desimone, R. et Ungerleider, L. G. (1996). Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis Neurosci* 13(3): 529-38.
- De Weerd, P., Peralta, M. R., 3rd, Desimone, R. et Ungerleider, L. G. (1999). Loss of attentional stimulus selection after extrastriate cortical lesions in macaques [published erratum appears in *Nat Neurosci* 2000 Apr;3(4):409]. *Nat Neurosci* 2(8): 753-8.
- Dean, P. (1976). Effects of inferotemporal lesions on the behavior of monkeys. *Psychol Bull* 83(1): 41-71.
- Dean, P. (1979). Visual cortex ablation and thresholds for successively presented stimuli in rhesus monkeys: II. Hue. *Exp Brain Res* 35(1): 69-83.
- DeAngelis, G. C. et Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *J Neurosci* 19(4): 1398-415.
- DeAngelis, G. C., Cumming, B. G. et Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature* 394(6694): 677-80.
- Delorme, A., Richard, G. et Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Res* 40(16): 2187-200.
- Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Van Essen, D. et Orban, G. A. (2004). The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *J Neurosci* 24(10): 2551-65.
- Desimone, R. et Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Res* 178(2-3): 363-80.
- Desimone, R. et Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57(3): 835-68.
- Desimone, R. et Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol* 248(2): 164-89.
- Desimone, R., Fleming, J. et Gross, C. G. (1980). Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Res* 184(1): 41-55.
- Desimone, R., Schein, S. J., Moran, J. et Ungerleider, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Res* 25(3): 441-52.
- DeYoe, E. A. et Van Essen, D. C. (1985). Segregation of efferent connections and receptive field properties in visual area V2 of the macaque. *Nature* 317(6032): 58-61.
- DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., Miller, D. et Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci U S A* 93(6): 2382-6.
- DeYoe, E. A., Felleman, D. J., Van Essen, D. C. et McClendon, E. (1994). Multiple processing streams in occipitotemporal visual cortex. *Nature* 371(6493): 151-4.
- Distler, C., Boussaoud, D., Desimone, R. et Ungerleider, L. G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. *J Comp Neurol* 334(1): 125-50.
- Dobkins, K. R., Thiele, A. et Albright, T. D. (2000). Comparison of red-green equiluminance points in humans and macaques: evidence for different L:M cone ratios between species. *J Opt Soc Am A Opt Image Sci Vis* 17(3): 545-56.
- Edelman, S., Bulthoff, H. H. et Bulthoff, I. (1999). Effects of parametric manipulation of inter-stimulus similarity on 3D object categorization. *Spat Vis* 12(1): 107-23.
- Fabre-Thorpe, M., Fize, D., Richard, G. et Thorpe, S. J. (1998a). Rapid categorization of extrafoveal natural images : implications for biological models. *Comput Neurosci*. Bower, Plenum Press: 7-12.

Bibliographie

- Fabre-Thorpe, M., Richard, G. et Thorpe, S. (1998b). On the speed of natural scene categorisation in human and non-human primates. *Curr Psychol Cogn* 17(4-5): 791-805.
- Fabre-Thorpe, M., Richard, G. et Thorpe, S. J. (1998c). Rapid categorization of natural images by rhesus monkeys. *Neuroreport* 9(2): 303-8.
- Farah, M. J. (1990). *Visual Agnosia*. Cambridge, MIT Press.
- Felleman, D. J. et Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1(1): 1-47.
- Felleman, D. J., Burkhalter, A. et Van Essen, D. C. (1997a). Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex. *J Comp Neurol* 379(1): 21-47.
- Felleman, D. J., Xiao, Y. et McClendon, E. (1997b). Modular organization of occipito-temporal pathways: cortical connections between visual area 4 and visual area 2 and posterior inferotemporal ventral area in macaque monkeys. *J Neurosci* 17(9): 3185-200.
- Ferrera, V. P., Nealey, T. A. et Maunsell, J. H. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature* 358(6389): 756-61.
- Ferrera, V. P., Nealey, T. A. et Maunsell, J. H. (1994). Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LGN pathways. *J Neurosci* 14(4): 2080-8.
- Fize, D., Vanduffel, W., Nelissen, K., Denys, K., Chef d'Hotel, C., Faugeras, O. et Orban, G. A. (2003). The retinotopic organization of primate dorsal V4 and surrounding areas: A functional magnetic resonance imaging study in awake monkeys. *J Neurosci* 23(19): 7395-406.
- Fogel, I. et Sagi, D. (1989). Gabor filters as texture discriminator. *Biological cybernetics* 61: 103-113.
- Foster, D. H. et Gilson, S. J. (2002). Recognizing novel three-dimensional objects by summing signals from parts and views. *Proc R Soc Lond B Biol Sci* 269(1503): 1939-47.
- Fries, W. (1981). The projection from the lateral geniculate nucleus to the prestriate cortex of the macaque monkey. *Proc R Soc Lond B Biol Sci* 213(1190): 73-86.
- Gallant, J. L., Braun, J. et Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259(5091): 100-3.
- Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W. et Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76(4): 2718-39.
- Gallant, J. L., Shoup, R. E. et Mazer, J. A. (2000). A human extrastriate area functionally homologous to macaque V4. *Neuron* 27(2): 227-35.
- Gattass, R., Sousa, A. P. et Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8(6): 1831-45.
- Gauthier, I. et Tarr, M. J. (1997). Orientation priming of novel shapes in the context of viewpoint- dependent recognition. *Perception* 26(1): 51-73.
- Gegenfurtner, K. R., Kiper, D. C. et Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Vis Neurosci* 13: 161-172.
- Gegenfurtner, K. R., Kiper, D. C. et Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *J Neurophysiol* 77(4): 1906-23.
- Gibson, J. J. (1950a). *The perception of the visual world*. Boston.
- Gibson, J. J. (1950b). The perception of visual surfaces. *Am J Psychol.* 63: 367-84.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, Houghton Mifflin.

Bibliographie

- Girard, P. et Morrone, M. C. (1995). Spatial Structure of Chromatically Opponent Receptive Fields in the Human Visual System. *Vis Neurosci* 12: 103-116.
- Girard, P., Hupe, J. M. et Bullier, J. (2001). Feedforward and feedback connections between areas v1 and v2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85(3): 1328-31.
- Girard, P., Lomber, S. G. et Bullier, J. (2002). Shape Discrimination Deficits During Reversible Deactivation of Area V4 in the Macaque Monkey. *Cereb Cortex* 12(11): 1146-1156.
- Goodale, M. A. et Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci* 15(1): 20-25.
- Grill-Spector, K., Kourtzi, Z. et Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res* 41(10-11): 1409-22.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y. et Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24(1): 187-203.
- Grill-Spector, K., Kushnir, T., Hendler, T. et Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 3(8): 837-43.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzhak, Y. et Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum Brain Mapp* 6(4): 316-28.
- Gross, C. G., Rocha-Miranda, C. E. et Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol* 35(1): 96-111.
- Gruber, H. E. et Clark, W. C. (1956). Perception of slant surfaces. *Percept Mot Skills* 6: 97-107.
- Gulyas, B. et Roland, P. E. (1994). Processing and analysis of form, colour and binocular disparity in the human brain: functional anatomy by positron emission tomography. *Eur J Neurosci* 6: 1811-1828.
- Gulyas, B., Cowey, A., Heywood, C. A., Popplewell, D. et Roland, P. E. (1998). Visual form discrimination from texture cues: a PET study. *Hum Brain Mapp* 6(3): 115-27.
- Guzman, A. (1968). Decomposition of a visual scene into three-dimensional bodies. New York.
- Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P. et Tootell, R. B. (1998). Retinotopy and color sensitivity in human visual cortical area V8 [see comments]. *Nat Neurosci* 1(3): 235-41.
- Hanazawa, A. et Komatsu, H. (2001). Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. *J Neurosci* 21(12): 4490-7.
- Hanes, D. P., Thompson, K. G. et Schall, J. D. (1995). Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. *Exp Brain Res* 103(1): 85-96.
- Haralick, R. M. (1979). Statistical and Structural Approaches to Texture. *Proc IEEE* 67(5): 786-804.
- Hayman, E., Caputo, B., Fritz, M. et Eklundh, J. (2004). On the significance of real-world conditions for material classification. *Proc ECCV*.
- Hayward, W. G. et Williams, P. (2000). Viewpoint dependence and object discriminability. *Psychol Sci* 11(1): 7-12.
- Hegde, J. et Van Essen, D. (2006). A comparative study of shape representation in macaque visual areas V2 and V4. *Cerebral Cortex* Online publication.
- Hegde, J. et Van Essen, D. C. (2003). Strategies of shape representation in macaque visual area V2. *Vis Neurosci* 20(3): 313-28.

Bibliographie

- Hegde, J. et Van Essen, D. C. (2005). Stimulus dependence of disparity coding in primate visual area v4. *J Neurophysiol* 93(1): 620-6.
- Heider, B., Spillmann, L. et Peterhans, E. (2002). Stereoscopic illusory contours--cortical neuron responses and human perception. *J Cogn Neurosci* 14(7): 1018-29.
- Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E. et Kubler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Res* 32(5): 963-81.
- Heller, M. A. (1989). Texture perception in sighted and blind observers. *Perception & Psychophysics* 45: 49-54.
- Helmholtz, H. v. (1910). *Physiological optics*. New York.
- Hershberger, W. (1970). Attached-shadow orientation perceived as depth by chickens reared in an environment illuminated from below. *J Comp Physiol Psychol* 73(3): 407-411.
- Heywood, C. A. et Cowey, A. (1987). Effects on visual search of lesions of the superior colliculus in infant or adult rats. *Exp Brain Res* 65(2): 465-70.
- Heywood, C. A., Gadotti, A. et Cowey, A. (1992). Cortical area V4 and its role in the perception of color. *J Neurosci* 12(10): 4056-65.
- Hietanen, J. K., Perrett, D. I., Oram, M. W., Benson, P. J. et Dittrich, W. H. (1992). The effects of lighting conditions on responses of cells selective for face views in the macaque temporal cortex. *Exp Brain Res* 89(1): 157-71.
- Hilgetag, C. C., O'Neill, M. A. et Young, M. P. (1996). Indeterminate organization of the visual system. *Science* 271(5250): 776-7.
- Hill, H. et Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*. 22(4): 986-1004.
- Hinkle, D. A. et Connor, C. E. (2001). Disparity tuning in macaque area V4. *Neuroreport* 12(2): 365-9.
- Hinkle, D. A. et Connor, C. E. (2002). Three-dimensional orientation tuning in macaque area V4. *Nat Neurosci* 5(7): 665-70.
- Hinkle, D. A. et Connor, C. E. (2005). Quantitative characterization of disparity tuning in ventral pathway area v4. *J Neurophysiol* 94(4): 2726-37.
- Hollins, M., Faldowski, R., Rao, S. et Young, F. (1993). Perceptual dimensions of tactile surface texture: A multidimensional scaling analysis. *Perception & Psychophysics* 54: 697-705.
- Horton, J. C. et Hubel, D. H. (1981). Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey. *Nature* 292(5825): 762-4.
- Howard, I. P. (2002). *Seeing in depth -Depth perception*. Toronto, I Porteous.
- Howard, I. P., Bergström, S. S. et Ohmi, M. (1990). Shape-from-shading in different frames of references. *Perception* 19: 523-530.
- Hubel, D. H. et Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *J Neurosci* 7(11): 3378-415.
- Hubel, D. H. et Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195(1): 215-43.
- Hubel, D. H. et Wiesel, T. N. (1969). Anatomical demonstration of columns in the monkey striate cortex. *Nature* 221(182): 747-50.
- Hubel, D. H. et Wiesel, T. N. (1970). Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225(227): 41-2.

Bibliographie

- Hubel, D. H. et Wiesel, T. N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *J Comp Neurol* 146(4): 421-50.
- Hubel, D. H. et Wiesel, T. N. (1974). Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J Comp Neurol* 158(3): 267-93.
- Huffman, D. A. (1971). Impossible objects in nonsense sentences. Edinburgh.
- Hummel, J. E. et Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychol Rev* 99(3): 480-517.
- Humphrey, G. K., Goodale, M. A., Bowen, C. V., Gati, J. S., Vilis, T., Rutt, B. K. et Menon, R. S. (1997). Differences in perceived shape from shading correlate with activity in early visual areas. *Curr Biol* 7(2): 144-7.
- Humphrey, G. K., Goodale, M. A., Jakobson, L. S. et Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. *Perception* 23(12): 1457-81.
- Humphrey, G. K., Symons, L. A., Herbert, A. M. et Goodale, M. A. (1996). A neurological dissociation between shape from shading and shape from edges. *Behav Brain Res* 76(1-2): 117-25.
- Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. et Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394(6695): 784-7.
- Jackson, D. A. (1995). PROTEST: a PROcrustean Randomization TEST of community environment concordance. *Ecoscience* 2: 297-303.
- James, T. W., Culham, J., Humphrey, G. K., Milner, A. D. et Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain* 126(Pt 11): 2463-75.
- Janssen, P., Vogels, R. et Orban, G. A. (1999). Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc Natl Acad Sci U S A* 96(14): 8217-22.
- Janssen, P., Vogels, R. et Orban, G. A. (2000). Three-dimensional shape coding in inferior temporal cortex. *Neuron* 27(2): 385-97.
- Janssen, P., Vogels, R., Liu, Y. et Orban, G. A. (2001). Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *J Neurosci* 21(23): 9419-29.
- Jenkin, H. L., Jenkin, M. R., Dyde, R. T. et Harris, L. R. (2004). Shape-from-shading depends on visual, gravitational, and body-orientation cues. *Perception* 33: 1453-1461.
- Jones, B. et O'Neil, S. (1985). Combining vision and touch in texture perception. *Perception & Psychophysics* 37: 66-72.
- Jones, J. P. et Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58(6): 1187-211.
- Julesz, B. (1975). Experiments in the visual perception of texture. *Sci Am* 232(4): 34-43.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature* 290(5802): 91-7.
- Julesz, B. (1986). Texton gradients: the texton theory revisited. *Biol Cybern* 54: 245-251.
- Julesz, B. et Kroese, B. (1988). Visual texture perception. Features and spatial filters. *Nature* 333(6171): 302-3.
- Kaas, J. H., Huerta, M. F., Weber, J. T. et Harting, J. K. (1978). Patterns of retinal terminations and laminar organization of the lateral geniculate nucleus of primates. *J Comp Neurol* 182(3): 517-53.
- Kanizsa, G. (1976). Subjective contours. *Sci Am* 234: 48-52.
- Kanwisher, N., Chun, M. M., McDermott, J. et Ledden, P. J. (1996). Functional imagining of human visual recognition. *Brain Res Cogn Brain Res* 5(1-2): 55-67.

Bibliographie

- Kanwisher, N., McDermott, J. et Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17(11): 4302-11.
- Kastner, S., De Weerd, P. et Ungerleider, L. G. (2000). Texture segregation in the human visual cortex: A functional MRI study. *J Neurophysiol* 83(4): 2453-7.
- Kennedy, H. et Barone, P. (1999). Relative contribution of cortical projections to area V1 and V4. *Soc. Neuro. Abst* 25 :1424.
- Kiper, D. C., Fenstemaker, S. B. et Gegenfurtner, K. R. (1997). Chromatic properties of neurons in macaque area V2. *Vis Neurosci* 14(6): 1061-72.
- Kleffner, D. A. et Ramachandran, V. S. (1992). On the perception of shape from shading. *Percept Psychophys* 52(1): 18-36.
- Knierim, J. J. et Van Essen, D. C. (1992). Visual cortex: cartography, connectivity, and concurrent processing. *Curr Opin Neurobiol* 2(2): 150-5.
- Kobatake, E. et Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71(3): 856-67.
- Koenderink, J. J. et van Doorn, A. J. (1976). Geometry of binocular vision and a model for stereopsis. *Biol Cybern* 21(1): 29-35.
- Koenderink, J. J., Kappers, A. M., Todd, J. T., Norman, J. F. et Phillips, F. (1996a). Surface range and attitude probing in stereoscopically presented dynamic scenes. *J Exp Psychol Hum Percept Perform* 22(4): 869-78.
- Koenderink, J. J., van Doorn, A. J. et Kappers, A. M. (1996b). Pictorial surface attitude and local depth comparisons. *Percept Psychophys* 58: 163-173.
- Komatsu, H. et Ideura, Y. (1993). Relationships between color, shape, and pattern selectivities of neurons in the inferior temporal cortex of the monkey. *J Neurophysiol* 70(2): 677-94.
- Kourtzi, Z. et Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *J Neurosci* 20(9): 3310-8.
- Kruskal, J. B. et Wish, M. (1978). *Multidimensional Scaling*. Beverly Hills, CA, Sage Publications.
- Kulikowski, J. J., Walsh, V., McKeefry, D., Butler, S. R. et Carden, D. (1994). The electrophysiological basis of colour processing in macaques with V4 lesions. *Behav Brain Res* 60(1): 73-8.
- Kusunoki, M., Moutoussis, K. et Zeki, S. (2006). Efect of Background Colors on the Tuning of color-Selective Cells In Monkey Area V4. *J Neurophysiol* 95: 3047-3059.
- Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15(2): 1605-15.
- Lamme, V. A., Van Dijk, B. W. et Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Res* 32(5): 797-807.
- Land, E. H. (1977). The retinex theory of color vision. *Sci Am* 237(6): 108-28.
- Landy, M. S. et Graham, N. (2002). Visual Perception of Texture. *The Visual Neurosciences*. L. M. Chalupa et J. S. Werner. Cambridge MA. 2.
- Langer, M. S. et Bulthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception* 29(6): 649-60.
- Lee, T. S., Yang, C. F., Romero, R. D. et Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nat Neurosci* 5(6): 589-97.

Bibliographie

- Lennie, P. (1998). Single units and visual cortical organization. *Perception* 27(8): 889-935.
- Leung, T. et Malik, J. (2001). Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons. *Int J Comput Vision* 43(1): 29-44.
- Leventhal, A. G., Thompson, K. G., Liu, D., Zhou, Y. et Ault, S. J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *J Neurosci* 15(3 Pt 1): 1808-18.
- Lewis, J. W. et Van Essen, D. C. (2000). Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. *J Comp Neurol* 428(1): 79-111.
- Li, S. Z. (1994). Markov random field model for object matching under contextual constraints. *Proceedings of the IEEE Computer Society Conference on computer Vision and Pattern recognition.*: 866-869.
- Liu, L., Stevenson, S. B. et Schor, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence. *Nature* 367(6458): 66-9.
- Liu, Y., Vogels, R. et Orban, G. A. (2004). Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. *J Neurosci* 24(15): 3795-800.
- Livingstone, M. S. et Hubel, D. H. (1984a). Anatomy and physiology of a color system in the primate visual cortex. *J Neurosci* 4(1): 309-56.
- Livingstone, M. S. et Hubel, D. H. (1984b). Specificity of intrinsic connections in primate primary visual cortex. *J Neurosci* 4(11): 2830-5.
- Livingstone, M. S. et Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J Neurosci* 7(11): 3416-68.
- Logothetis, N. K. et Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex* 5(3): 270-88.
- Maguire, W. M. et Baizer, J. S. (1984). Visuotopic organization of the prelunate gyrus in rhesus monkey. *J Neurosci* 4(7): 1690-704.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R. et Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92(18): 8135-9.
- Malik, J. et Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *J Opt Soc Am A* 7(5): 923-32.
- Mallat, S. (1989). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 674-693.
- Mamassian, P., Knill, D. C. et Kersten, D. (1998). The perception of cast shadows. *Trends Cogn Sci* 2(8): 288-295.
- Marr, D. (1976). Early processing of visual information. *Philos Trans R Soc Lond B* 275: 483-519.
- Marr, D. (1982). Vision. San Francisco.
- Marr, D. et Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Philos Trans R Soc Lond B* (200): 269-294.
- Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P. et Martin, K. E. (1996). Occlusion edge blur: a cue to relative visual depth. *J Opt Soc AM* 13(681-688).
- Maunsell, J. H. et Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* 68(4): 1332-44.

Bibliographie

- Maunsell, J. H. et Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci* 10: 363-401.
- Maunsell, J. H. et van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3(12): 2563-86.
- Mayhew, J. (1982). The interpretation of stereo-disparity information: the computation of surface orientation and depth. *Perception* 11(4): 387-403.
- Mayhew, J. E. et Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature* 297(5865): 376-8.
- Mendola, J. D. et Corkin, S. (1999). Visual discrimination and attention after bilateral temporal-lobe lesions: a case study. *Neuropsychologia* 37(1): 91-102.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K. et Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J Neurosci* 19(19): 8560-72.
- Merigan, W. H. (1996). Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Vis Neurosci* 13(1): 51-60.
- Merigan, W. H. (2000). Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis Neurosci* 17(6): 949-58.
- Merigan, W. H. et Pham, H. A. (1998). V4 lesions in macaques affect both single- and multiple-viewpoint shape discriminations. *Vis Neurosci* 15(2): 359-67.
- Merigan, W. H., Nealey, T. A. et Maunsell, J. H. (1993). Visual effects of lesions of cortical area V2 in macaques. *J Neurosci* 13(7): 3180-91.
- Merigan, W., Freeman, A. et Meyers, S. P. (1997). Parallel processing streams in human visual cortex. *Neuroreport* 8(18): 3985-91.
- Milner, A. D. et Goodale, M. A. (1993). Visual pathways to perception and action. *Progress in Brain Research*. S. M. T. O. T.P. Hicks, Elsevier. 95: 317-337.
- Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E. et al. (1991). Perception and action in 'visual form agnosia'. *Brain* 114(Pt 1B): 405-28.
- Mishkin, M., Ungerleider, L. G. et Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci* oct.: 414-417.
- Moore, C. et Engel, S. A. (2001). Neural response to perception of volume in the lateral occipital complex. *Neuron* 29(1): 277-86.
- Moore, T. et Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421(6921): 370-3.
- Morel, A. et Bullier, J. (1990). Anatomical segregation of two cortical visual pathways in the macaque monkey. *Vis Neurosci* 4(6): 555-78.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M. et Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol* 83(5): 2580-601.
- Murray, S. O., Olshausen, B. A. et Woods, D. L. (2003). Processing Shape, Motion and Three-dimensional Shape-from-motion in the Human Cortex. *Cereb Cortex* 13(5): 508-16.
- Nakamura, H., Gattass, R., Desimone, R. et Ungerleider, L. G. (1993). The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J Neurosci* 13(9): 3681-91.

Bibliographie

- Norman, J. F. et Todd, J. T. (1995). The perception of 3-D structure from contradictory optical patterns. *Percept Psychophys* 57(6): 826-34.
- Norman, J. F., Todd, J. T. et Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychol Sci* 15(8): 565-70.
- Nothdurft, H. C. et Li, C. Y. (1985). Texture discrimination: representation of orientation and luminance differences in cells of the cat striate cortex. *Vision Res* 25(1): 99-113.
- Nowak, L. G., Sanchez-Vives, M. V. et McCormick, D. A. (1997). Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cereb Cortex* 7(6): 487-501.
- Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., Todd, J. et Vanduffel, W. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia* 41(13): 1757-1768.
- Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P. et Marchal, G. (1999). Human cortical regions involved in extracting depth from motion. *Neuron* 24(4): 929-40.
- Orban, G. A., Van Essen, D. et Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn Sci* 8(7): 315-324.
- Ostergaard, A. L. et Davidoff, J. B. (1985). Some effects of color on naming and recognition of objects. *J Exp Psychol Learn Mem Cogn* 11(3): 579-87.
- Pasupathy, A. et Connor, C. E. (1999). Responses to contour features in macaque area V4. *J Neurophysiol* 82(5): 2490-502.
- Pasupathy, A. et Connor, C. E. (2001). Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol* 86(5): 2505-19.
- Pasupathy, A. et Connor, C. E. (2002). Population coding of shape in area V4. *Nat Neurosci* 5(12): 1332-1338.
- Perry, V. H. et Cowey, A. (1984). Retinal ganglion cells that project to the superior colliculus and pretectum in the macaque monkey. *Neuroscience* 12: 1125-1137.
- Perry, V. H., Oehler, R. et Cowey, A. (1984). Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience* 12(4): 1101-23.
- Peterhans, E. et von der Heydt, R. (1993). Functional organization of area V2 in the alert macaque. *Eur J Neurosci* 5(5): 509-24.
- Phillips, F. et Todd, J. T. (1996). Perception of local three-dimensional shape. *J Exp Psychol Hum Percept Perform* 22(4): 930-44.
- Picard, D., Dacremont, C., Valentin, D. et Giboreau, A. (2003). Perceptual dimensions of tactiles textures. *Acta Psychologica* 114: 164-185.
- Pinon, M. C., Gattass, R. et Sousa, A. P. (1998). Area V4 in Cebus monkey: extent and visuotopic organization. *Cereb Cortex* 8(8): 685-701.
- Poggio, G. F. et Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J Neurophysiol* 40(6): 1392-405.
- Poggio, T. et Edelman, S. (1990). A network that learns to recognize three-dimensional objects [see comments]. *Nature* 343(6255): 263-6.
- Portilla, J. et Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal Computation vision* 40: 49-71.
- Povinelli, D. J. (2000). Folk physics for apes: the chimpanzee's theory of how the world works. Oxford; New York, Oxford University Press."

- Price, C. J. et Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Q J Exp Psychol A* 41(4): 797-827.
- Puce, A., Allison, T., Asgari, M., Gore, J. C. et McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* 16(16): 5205-15.
- Purpura, K. P., Victor, J. D. et Katz, E. (1994). Striate cortex extracts higher-order spatial correlations from visual textures. *Proc Natl Acad Sci U S A* 91(18): 8482-6.
- Rainer, G., Lee, H. K. et Logothetis, N. K. (2004). The Effect of Learning on the Function of Monkey Extrastriate Visual Cortex. *PLoS Biology* 2(2): 275-283.
- Ramachandran, V. S. (1988). Perceiving shape from shading. *Sci Am* 259(2): 76-83.
- Rao, A. R. et Lohse, G. L. (1996). Toward a texture naming system: Identifying relevant dimensions of texture. *Vision res* 36(11): 1649-1669.
- Regan, D. M. (2000). Human perception of Objects: Early Visual Processing of Spatial Forme Defined by Luminance, Color, Texture, Motion, and Sinauer Assoc.
- Riesenhuber, M. et Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci* 2(11): 1019-25.
- Rivest, J. et Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Res* 36(1): 53-66.
- Rockland, K. S. (1994). The organization of feedback connections from area V2 (18) to V1 (17). *Cerebral Cortex*. A. Peters et K. S. Rockland. 10: 261-299.
- Rockland, K. S., Saleem, K. S. et Tanaka, K. (1994). Divergent feedback connections from areas V4 and TEO in the macaque. *Vis Neurosci* 11(3): 579-600.
- Roe, A. W. et Ts'o, D. Y. (1995). Visual topography in primate V2: multiple representation across functional stripes. *J Neurosci* 15(5 Pt 2): 3689-715.
- Rossion, B. et Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception* 33(2): 217-36.
- Rubner, Y. et Tomasi, C. (1998). Texture metrics. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.*: 4601-4607.
- Sakata, H., Tsutsui, K. et Taira, M. (2005). Toward an understanding of the neural processing for 3D shape perception. *Neuropsychologia* 43(2): 151-61.
- Schein, S. J. et Desimone, R. (1990). Spectral properties of V4 neurons in the macaque. *J Neurosci* 10(10): 3369-89.
- Schein, S. J., Marrocco, R. T. et de Monasterio, F. M. (1982). Is there a high concentration of color-selective cells in area V4 of monkey visual cortex? *J Neurophysiol* 47(2): 193-213.
- Schiller, P. H. (1993). The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis Neurosci* 10(4): 717-46.
- Schiller, P. H. (1995). Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature* 376(6538): 342-4.
- Schiller, P. H. et Lee, K. (1991). The role of the primate extrastriate area V4 in vision. *Science* 251(4998): 1251-3.

Bibliographie

- Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D. et Leventhal, A. G. (1998). Signal timing across the macaque visual system. *J Neurophysiol* 79(6): 3272-8.
- Schriever, W. (1924). Experimentelle Studien über stereoskopisches Sehen. *Z Psychol* 96: 113-170.
- Seltzer, B. et Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res* 149(1): 1-24.
- Sereno, M. E., Trinath, T., Augath, M. et Logothetis, N. K. (2002). Three-dimensional shape representation in monkey cortex. *Neuron* 33(4): 635-52.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R. et Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268(5212): 889-93.
- Severac Cauquil, A., Trotter, Y. et Taylorn, M. J. (2005). At what stage neural processing do perspective depth cues make a difference? *Exp Brain Res*.
- Shikata, E., Hamzei, F., Glauche, V., Knab, R., Dettmers, C., Weiller, C. et Buchel, C. (2001). Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an event-related fMRI study. *J Neurophysiol* 85(3): 1309-14.
- Shipp, S. (2001). Corticopulvinar connections of areas V5, V4, and V3 in the macaque monkey: a dual model of retinal and cortical topographies. *J Comp Neurol* 439(4): 469-90.
- Shipp, S. et Zeki, S. (1985). Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. *Nature* 315(6017): 322-5.
- Shipp, S. et Zeki, S. (1989). The organization of connections between areas V5 and V1 in macaque monkey visual cortex. *Eur J Neurosci* 1(4): 309-332.
- Shipp, S. et Zeki, S. (1995). Segregation and convergence of specialised pathways in macaque monkey visual cortex. *J Anat* 187(Pt 3): 547-62.
- Shiwa, T. (1987). Corticocortical projections to the monkey temporal lobe with particular reference to the visual processing pathways. *Arch Ital Biol* 125(2): 139-54.
- Sincich, L. C. et Horton, J. C. (2005a). The circuitry of V1 and V2: Integration of Color, Form, and Motion. *Annu Rev Neurosci* 28: 303-26.
- Sincich, L. C. et Horton, J. C. (2005b). Input to V2 thin stripes arises from V1 cytochrome oxidase patches. *J Neurosci* 25(44): 10087-93.
- Sun, J. Y. et Perona, P. (1998). Where is the sun? *Nat Neurosci* 1: 183-184.
- Tadmor, Y. et Tolhurst, D. J. (1993). Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision res* 33: 141-145.
- Taira, M., Nose, I., Inoue, K. et Tsutsui, K. (2001). Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. *Neuroimage* 14(5): 959-66.
- Taira, M., Tsutsui, K. I., Jiang, M., Yara, K. et Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 83(5): 3140-6.
- Tanabe, S., Doi, T., Umeda, K. et Fujita, I. (2005). Disparity-tuning characteristics of neuronal responses to dynamic random-dot stereograms in macaque visual area v4. *J Neurophysiol* 94(4): 2683-99.
- Tanaka, H., Uka, T., Yoshiyama, K., Kato, M. et Fujita, I. (2001a). Processing of shape defined by disparity in monkey inferior temporal cortex. *J Neurophysiol* 85(2): 735-44.
- Tanaka, J., Weiskopf, D. et Williams, P. (2001b). The role of color in high-level vision. *Trends Cogn Sci* 5(5): 211-215.

Bibliographie

- Tanaka, K. (2002). Inferotemporal Response Properties The visual Neurosciences. L. M. Chalupa et J. S. Werner. Cambridge, MA, MIT press. 2.
- Tanaka, K., Saito, H., Fukada, Y. et Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophysiol* 66(1): 170-89.
- Tanaka, M., Weber, H. et Creutzfeldt, O. D. (1986). Visual properties and spatial distribution of neurones in the visual association area on the prelunate gyrus of the awake monkey. *Exp Brain Res* 65(1): 11-37.
- Tarr, M. J. et Bulthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *J Exp Psychol Hum Percept Perform* 21(6): 1494-505.
- Tarr, M. J., Kersten, D. et Bulthoff, H. H. (1998a). Why the visual recognition system might encode the effects of illumination. *Vision Res* 38(15-16): 2259-75.
- Tarr, M. J., Williams, P., Hayward, W. G. et Gauthier, I. (1998b). Three-dimensional object recognition is viewpoint dependent. *Nat Neurosci* 1(4): 275-7.
- Thielscher, A. et Neumann, H. (2005). Neural mechanisms of human texture processing: texture boundary detection and visual search. *Spat Vis* 18(2): 227-57.
- Todd, J. T. (2004). The visual perception of 3D shape. *Trends Cogn Sci* 8(3): 115-121.
- Todd, J. T. et Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *J Exp Psychol* 9: 583-595.
- Todd, J. T., Oomes, A. H., Koenderink, J. J. et Kappers, A. M. (2001). On the affine structure of perceptual space. *Psychological Science* 12(3): 191-196.
- Tootell, R. B. et Hadjikhani, N. (2001). Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb Cortex* 11(4): 298-311.
- Tootell, R. B., Dale, A. M., Sereno, M. I. et Malach, R. (1996). New images from human visual cortex. *Trends Neurosci* 19(11): 481-9.
- Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., Sereno, M. I. et Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *J Neurosci* 17(18): 7060-78.
- Tootell, R. B., Silverman, M. S., De Valois, R. L. et Jacobs, G. H. (1983). Functional organization of the second cortical visual area in primates. *Science* 220(4598): 737-9.
- Troje, N. F. et Bulthoff, H. H. (1998). How is bilateral symmetry of human faces used for recognition of novel views? *Vision Res* 38(1): 79-89.
- Trotter, Y. (1995). Cortical representation of visual three-dimensional space. *Perception* 24(3): 287-98.
- Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B. et Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nat Neurosci* 6(9): 989-95.
- Ts'o, D. Y. et Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *J Neurosci* 8(5): 1712-27.
- Tsutsui, K., Jiang, M., Yara, K., Sakata, H. et Taira, M. (2001). Integration of perspective and disparity cues in surface-orientation- selective neurons of area CIP. *J Neurophysiol* 86(6): 2856-67.
- Tsutsui, K., Sakata, H., Naganuma, T. et Taira, M. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298(5592): 409-12.
- Tsutsui, K., Taira, M. et Sakata, H. (2005). Neural mechanisms of three-dimensional vision. *Neurosci Res* 51(3): 221-9.

Bibliographie

- Turner, M. R. (1986). Texture discrimination by Gabor functions. *Biol Cybern* 55(2-3): 71-82.
- Uka, T., Tanaka, H., Yoshiyama, K., Kato, M. et Fujita, I. (2000). Disparity selectivity of neurons in monkey inferior temporal cortex. *J Neurophysiol* 84(1): 120-32.
- Ullman, S. (1989). Aligning pictorial description ; an approach to object recognition. *Cognition* 32: 193-254."
- Ungerleider, L. G. et Desimone, R. (1986). Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. *J Comp Neurol* 248(2): 147-63.
- van der Schaaf, A. et van Hateren, J. H. (1996). Modelling the power spectra of natural images: statistics and information. *Vision Res* 36(17): 2759-70.
- Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. *The Visual Neurosciences*. L. M. Chalupa et J. S. Werner: 507-521.
- Van Essen, D. C. et Zeki, S. M. (1978). The topographic organization of rhesus monkey prestriate cortex. *J Physiol* 277: 193-226.
- Van Essen, D. C., Felleman, D. J., DeYoe, E. A., Olavarria, J. et Knierim, J. (1990). Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. *Cold Spring Harb Symp Quant Biol* 55: 679-96.
- Van Essen, D. C., Newsome, W. T. et Maunsell, J. H. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Res* 24(5): 429-48.
- Van Essen, D. C., Newsome, W. T., Maunsell, J. H. et Bixby, J. L. (1986). The projections from striate cortex (V1) to areas V2 and V3 in the macaque monkey: asymmetries, areal boundaries, and patchy connections. *J Comp Neurol* 244(4): 451-80.
- Varma, M. et Zisserman, A. (2002). Statistical approaches to material classification. *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing*.
- Vogels, R. et Biederman, I. (2002). Effects of illumination intensity and direction on object coding in macaque inferior temporal cortex. *Cereb Cortex* 12(7): 756-66.
- Von Bonin, G. (1947). The frontal lobe of primates: cytoarchitectural studies: 67-83 (Ch. 3).
- von der Heydt, R. et Peterhans, E. (1989). Cortical contour mechanisms and geometrical illusions. *Neural mechanisms of visual perception*. D. M. K. L. C. D. Gilbert, Portfolio Publ. Company: 157-170.
- von der Heydt, R., Peterhans, E. et Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science* 224(4654): 1260-2.
- Wallis, G. et Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Prog Neurobiol* 51(2): 167-94.
- Walsh, V., Carden, D., Butler, S. R. et Kulikowski, J. J. (1993). The effects of V4 lesions on the visual abilities of macaques: hue discrimination and colour constancy. *Behav Brain Res* 53(1-2): 51-62.
- Walsh, V., Kulikowski, J. J., Butler, S. R. et Carden, D. (1992). The effects of lesions of area V4 on the visual abilities of macaques: colour categorization. *Behav Brain Res* 52(1): 81-9.
- Wandell, B. A., Poirson, A. B., Newsome, W. T., Baseler, H. A., Boynton, G. M., Huk, A., Gandhi, S. et Sharpe, L. T. (1999). Color signals in human motion-selective cortex. *Neuron* 24(4): 901-9.
- Warrington, E. K. (1985). Agnosia: the impairment of object recognition. *Handbook of Clinical Neurology* - Vol. 1 *Clinical Neuropsychology*. G. W. B. H. K. P.J. Vinken, Elsevier. 1: 333-349.
- Watanabe, M., Tanaka, H., Uka, T. et Fujita, I. (2002). Disparity-selective neurons in area V4 of macaque monkeys. *J Neurophysiol* 87(4): 1960-73.

Bibliographie

- Webster, M. J., Bachevalier, J. et Ungerleider, L. G. (1993). Subcortical Connections of Inferior Temporal Areas TE and TEO in Macaque Monkeys. *J Comp Neurol* 335(1): 73-91.
- Weiskrantz, L. et Saunders, R. C. (1984). Impairments of visual object transforms in monkeys. *Brain* 107(Pt 4): 1033-72.
- Wheatstone, C. (1838). Contribution of the physiology of vision - Part the first On some remarkable and hitherto unobserved phenomena of binocular vision. *Philos Trans R Soc Lond B Biol Sci* 128: 371-394.
- Wilson, M. E. et Cragg, B. G. (1967). Projections from the lateral geniculate nucleus in the cat and monkey. *J Anat* 101(4): 677-92.
- Won, C. S. et Derin, H. (1992). Unsupervised segmentation of noisy and textured images y using Markov random fields. *CVGIP:Graphics Model and Image processing* 54: 308-328.
- Wurm, L. H., Legge, G. E., Isenberg, L. M. et Luebker, A. (1993). Color improves object recognition in normal and low vision. *J Exp Psychol Hum Percept Perform* 19(4): 899-911.
- Xiao, Y., Zych, A. et Felleman, D. J. (1999). Segregation and convergence of functionally defined V2 thin stripe and interstripe compartment projections to area V4 of macaques. *Cereb Cortex* 9(8): 792-804.
- Yonas, A., Kuskowski, M. et S., S. (1979). The role of reference in the development of responsiveness to shading. *Child development* 50: 495-500.
- Yoshida, M. (1968). Dimensions of tactual impressions. *Japanese Psychological Research* 10: 157-173.
- Yukie, M. et Iwai, E. (1985). Laminar origin of direct projection from cortex area V1 to V4 in the rhesus monkey. *Brain Res* 346(2): 383-6.
- Yukie, M. et Iwai, E. (1988). Direct projections from the ventral TE area of the inferotemporal cortex to hippocampal field CA1 in the monkey. *Neurosci Lett* 88(1): 6-10.
- Zeki, S. (1974a). The mosaic organization of the visual cortex in the monkey. *Essays on the nervous system*. R. B. E. G. Gray, Clarendon Press: 327-343.
- Zeki, S. (1983a). Colour coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelenghts and colors. *Neuroscience* 9: 741-765.
- Zeki, S. (1983b). The relationship between wavelength and color studied in single cells of monkey striate cortex. *Prog Brain Res* 58: 219-27.
- Zeki, S. et Shipp, S. (1989). Modular connections between areas V2 and V4 of macaque monkey visual cortex. *Eur J Neurosci* 1: 494-506.
- Zeki, S. M. (1969). Representation of central visual fields in prestriate cortex of monkey. *Brain Res* 14(2): 271-91.
- Zeki, S. M. (1973a). Colour coding in rhesus monkey prestriate cortex. *Brain Res* 53(2): 422-7.
- Zeki, S. M. (1973b). Comparison of the cortical degeneration in the visual regions of the temporal lobe of the monkey following section of the anterior commissure and the splenium. *J Comp Neurol* 148(2): 167-75.
- Zeki, S. M. (1974b). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol* 236(3): 549-73.
- Zeki, S. M. (1977). Colour coding in the superior temporal sulcus of rhesus monkey visual cortex. *Proc R Soc Lond B Biol Sci* 197(1127): 195-223.
- Zeki, S. M. (1978a). The cortical projections of foveal striate cortex in the rhesus monkey. *J Physiol* 277: 227-244.

Bibliographie

Zeki, S. M. (1978b). Functional specialisation in the visual cortex of the rhesus monkey. *Nature* 274(5670): 423-428.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C. et Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11(3): 641-9.

Zhou, H., Friedman, H. S. et von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *J Neurosci* 20(17): 6594-611.

Zhu, S. C., Wu, Y. N. et Mumford, D. B. (1998). Filters, random fields and maximum entropy (FRAME)-towards a unified theory for texture modeling. *International Journal Computation vision* 27: 107-126.

Zimmerman, G. L., Legge, G. E. et Cavanagh, P. (1995). Pictorial depth cues: a new slant. *J Opt Soc AM* 12: 17-29.

Glossaire

Cellule 3D : Dans le chapitre sur les ombrages, une cellule 3D dissocie les stimuli 3D des stimuli 2D blobs en ayant une réponse moyenne supérieure pour les stimuli 3D (Test TUKEY, $p<0.05$).

Cellule 2D blobs: Dans le chapitre sur les ombrages, une cellule 2D blobs dissocie les stimuli 2D blobs des stimuli 3D en ayant une réponse moyenne supérieure pour les stimuli 2D blobs (Test TUKEY, $p<0.05$).

Cellule 2D : Dans le chapitre sur les ombrages, une cellule 2D ne dissocie pas les stimuli 3D des stimuli 2D blobs (Test TUKEY, $p>0.05$).

Cellule TS : Cellule « Texture selective ». Une cellule TS présente un effet principal de la texture naturelle sur son taux de décharge (ANOVA, $p<0.05$). Pour les sets A et C, l'ANOVA comprend 3 facteurs (Texture, Phase et angle d'incidence). Pour le set B, l'ANOVA ne comprend qu'un seul facteur (Texture).

Cellule ICS : Cellule « Inner content selective ». Les stimuli du protocole sur les ombrages sont définis par plusieurs patterns internes ou type (3D, 2D blobs, random ou postérisé). Une cellule ICS présente un effet principal du type des stimuli sur son taux de décharge (ANOVA, $p<0.05$).

DAV : Degré d'angle visuel

DMI : Direction modulation index.

Forme 3D optimale : Forme 3D qui présente un I_{3D} maximal. C'est la forme la plus dissociée de son contrôle 2D blobs correspondant.

Forme 3D préférée : Forme 3D qui engendre une réponse moyenne maximale du neurone.

I_{max} : dans le chapitre « Textures », cet indice permet de déterminer les textures naturelles auxquelles le neurone répond. Ce sont les textures naturelles préférées du neurone. Pour calcul, voir p.143.

I_{3D} : Index de dissociation entre les stimuli 3D et 2D blobs. Pour calcul, voir p.143.

I_{rand} : Index de dissociation entre les stimuli 3D et random. Pour calcul, voir p.143.

I_{post} : Index de dissociation entre les stimuli 3D et postérisés. Pour calcul, voir p.143.

PSTHs : Peri stimuli Temporal Histograms

RMS : Root mean square. Contraste moyen.

SMI : Shape modulation index.

SNR : Signal to noise ratio.

Texture maximale : il s'agit de la texture qui engendre une réponse moyenne maximale du neurone.

Textures préférées : ce sont les textures pour lesquelles le neurone répond préférentiellement d'après le critère I_{max}

Annexes de psychophysiques

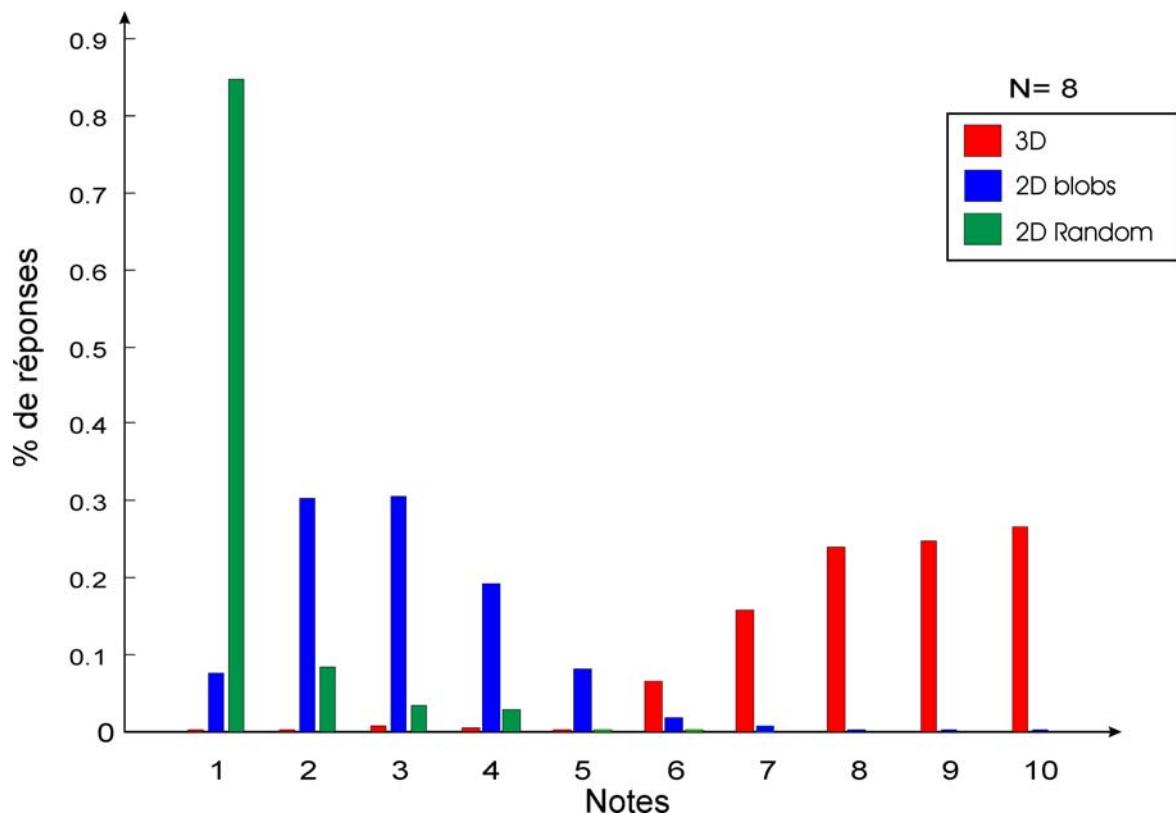
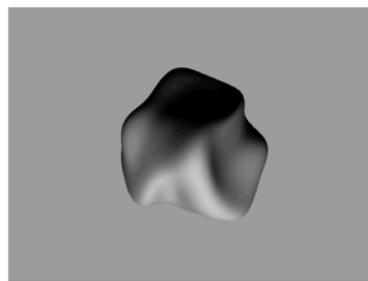


Figure 1 : Test psychophysique sur la perception tridimensionnelle des stimuli 3D, 2D blobs et 2D random. Les sujets (N=8) doivent estimer le percept tridimensionnelle du stimulus qui leur est présenté fovealement. Ils attribuent une note au stimulus en fonction du percept tridimensionnel, la note 10 correspond à un très bon percept 3D. Ici sont représentés les pourcentages moyens des réponses pour chacune des notes. Ainsi, dans 85% des réponses, les sujets attribuent une note de 1 pour les stimuli 2D random. Les stimuli 3D ont des notes au-dessus de 6. Ils sont perçus comme tridimensionnels.

Forme: 3
Dir. illu.: En dessous



Niveaux de gris (n) :

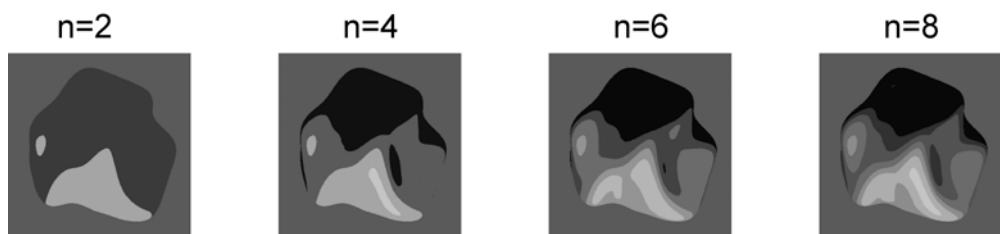


Figure 2 : Exemples de stimuli postérisés pour le stimulus représentant la forme n° 3 illuminée par en dessous. La postérisation a été testée avec 2, 4, 6 et 8 niveaux de gris. Ces stimuli sont testés en psychophysiques afin de savoir à partir de combien de niveaux de gris le percept 3D disparaît.

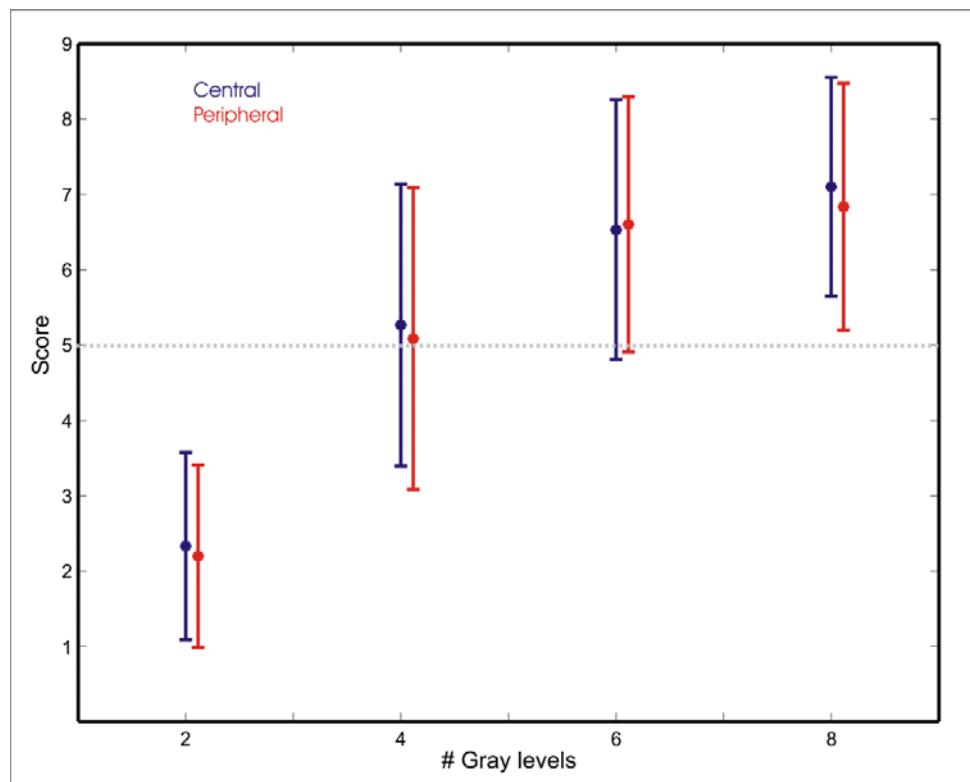
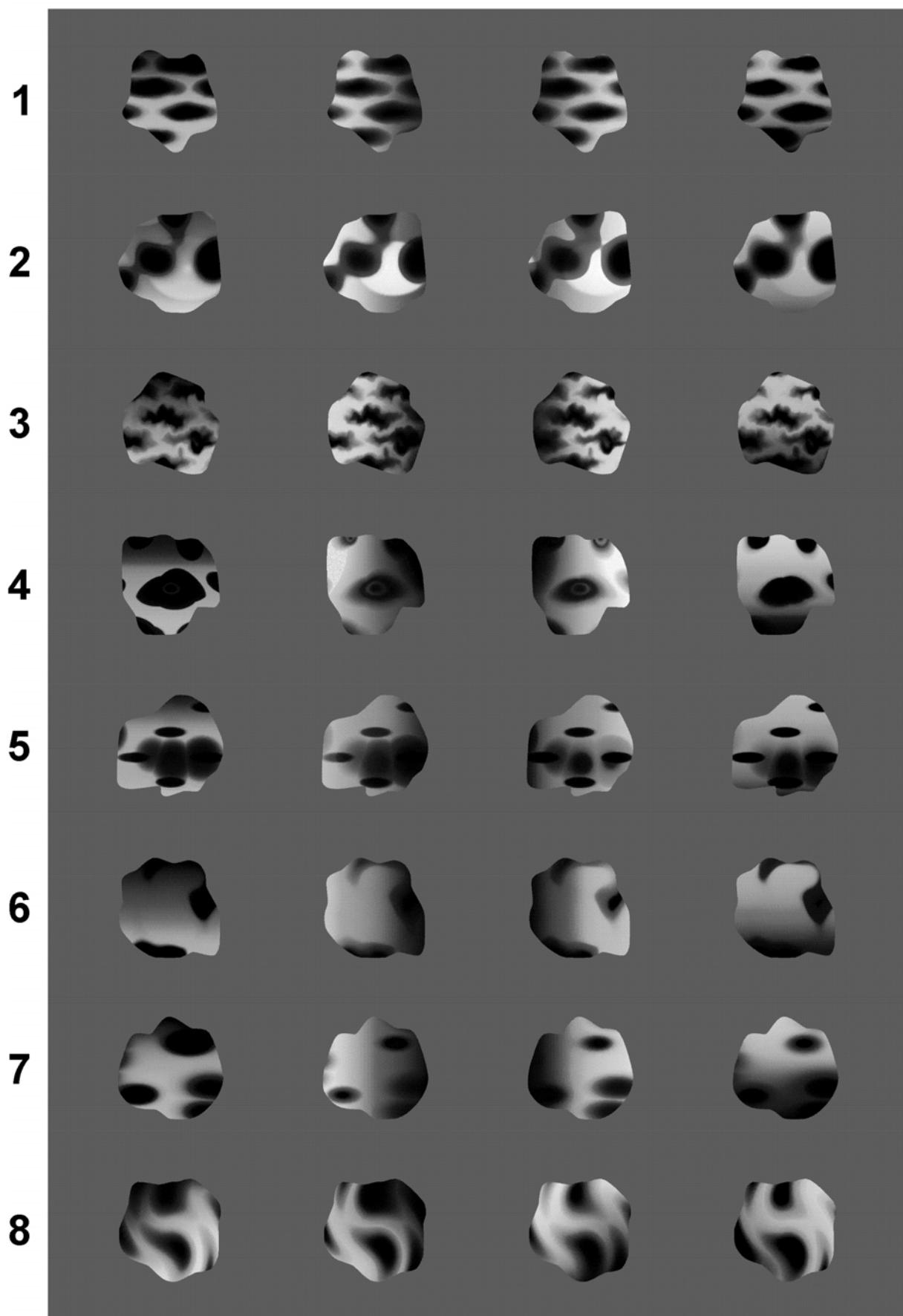


Figure 3 : Résultat du test de psychophysique sur les stimuli postérisés. Les stimuli sont présentés soit en position centrale soit en position périphérique (excentricité : 10°). Les sujets (n=6) doivent évaluer leur percept 3D en le notant sur 10 pour chacun des stimuli présentés (la note 10 correspondant à un très bon percept 3D). Les notes sont moyennées en fonction du nombre des niveaux de gris et de la position de présentation des stimuli. Seuls les stimuli réalisés avec 2 niveaux de gris présentent un mauvais percept 3D en position périphérique et centrale.

En dessous **Gauche** **Droite** **Au dessus**

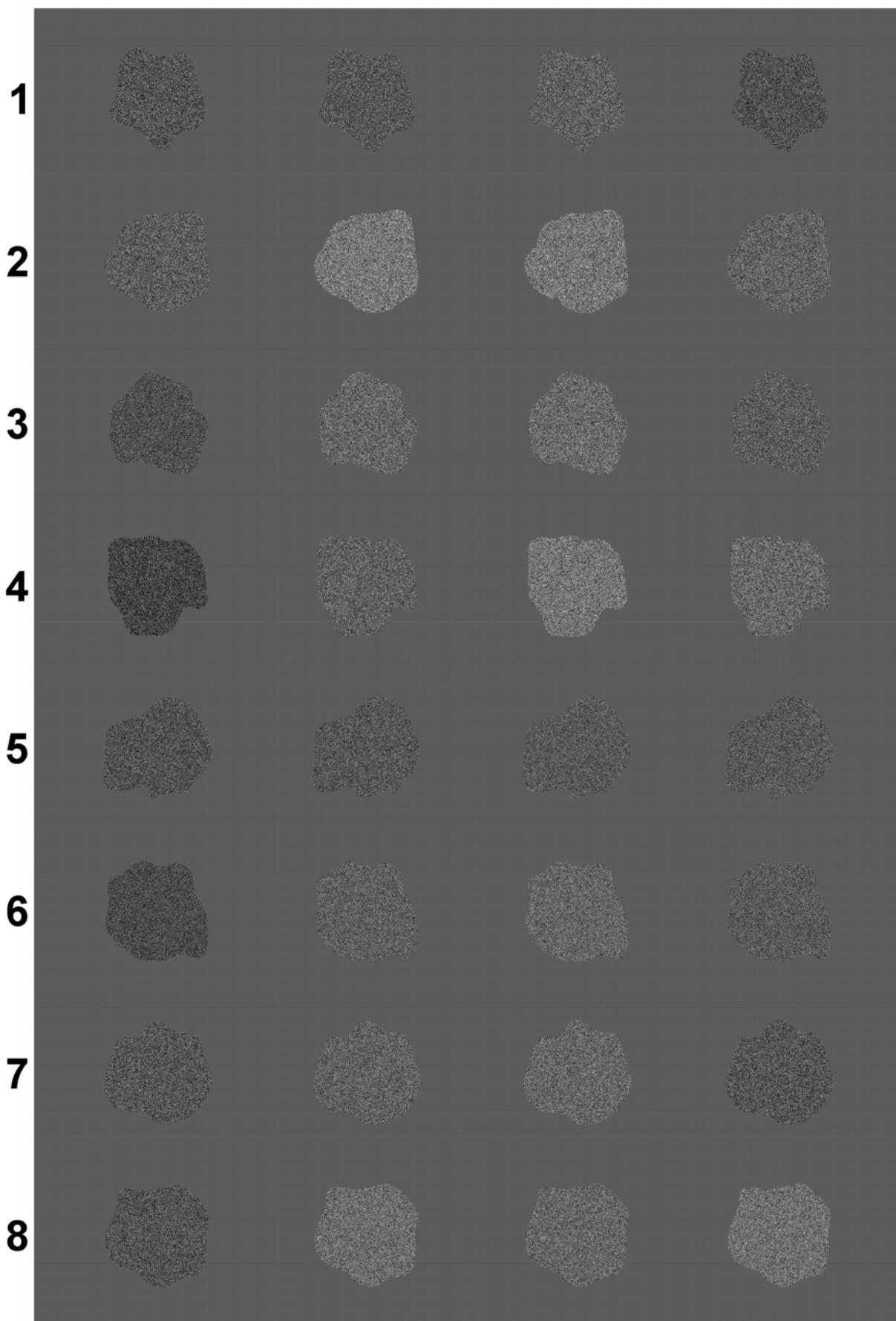


En dessous

Gauche

Droite

Au dessus

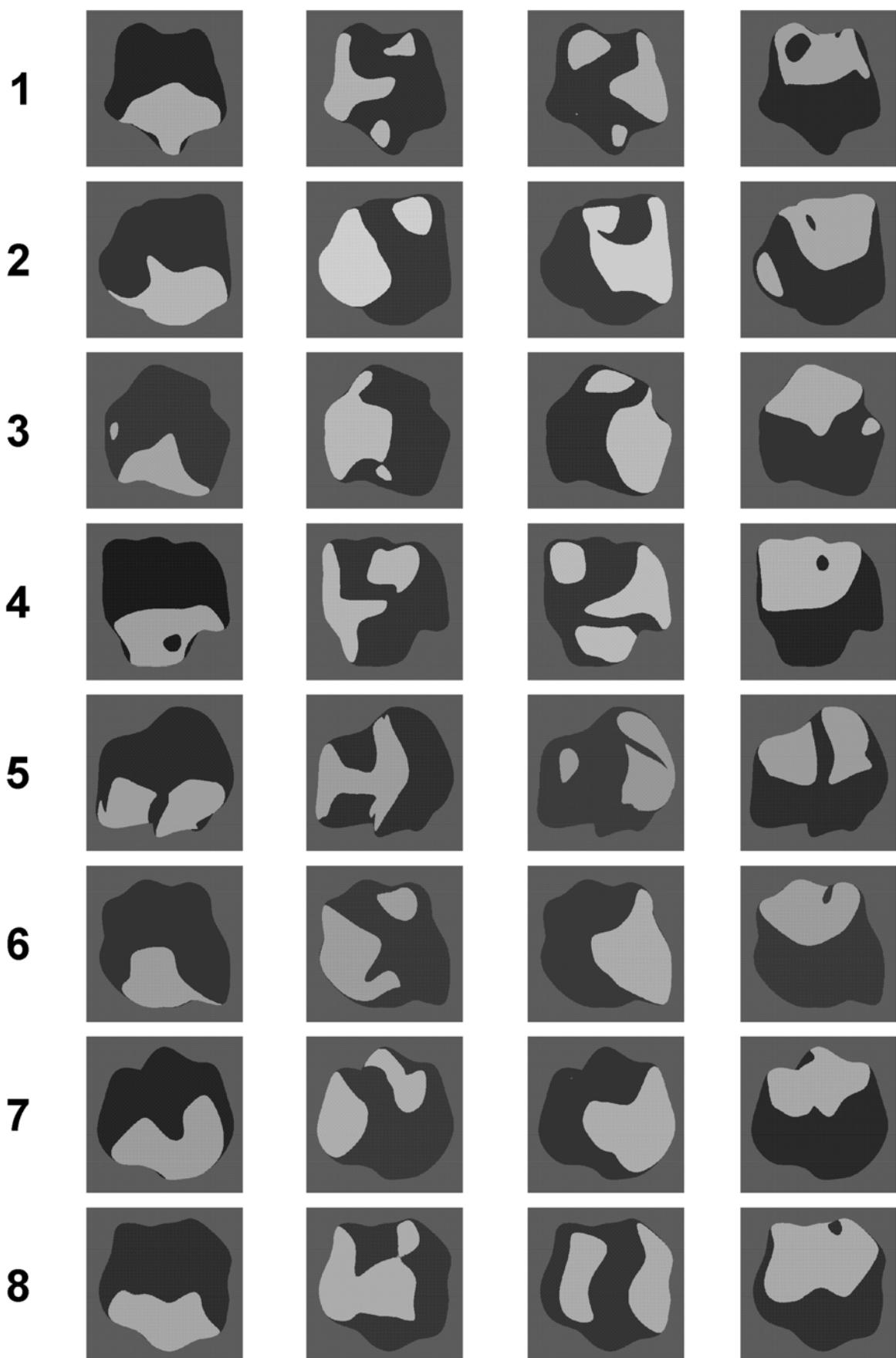


En dessous

Gauche

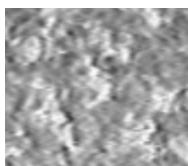
Droite

Au dessus



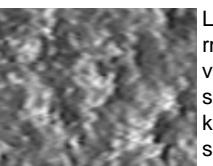
SET A (Originaux)

Velours (#03)



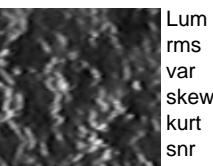
Lum 0.63
rms 0.19
var 0.01
skew 0.38
kurt 2.89
snr 5.15

L2

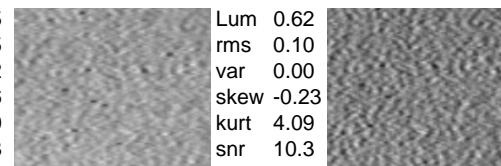


Lum 0.47
rms 0.30
var 0.02
skew 0.54
kurt 2.98
snr 3.34

L3

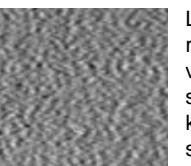


Papier de verre (#06)



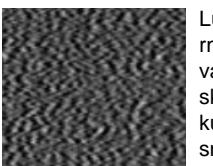
Lum 0.25
rms 0.55
var 0.02
skew 1.36
kurt 4.80
snr 1.83

L2



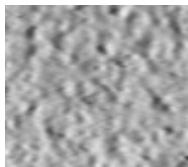
Lum 0.62
rms 0.10
var 0.00
skew -0.23
kurt 4.09
snr 10.3

L3

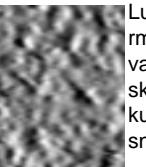


Lum 0.50
rms 0.18
var 0.01
skew -0.02
kurt 3.09
snr 5.68

Plâtre (#10)

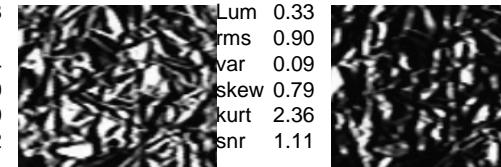


Lum 0.64
rms 0.14
var 0.01
skew -0.50
kurt 3.44
snr 6.93

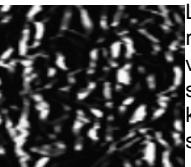


Lum 0.50
rms 0.34
var 0.03
skew -0.05
kurt 2.81
snr 2.93

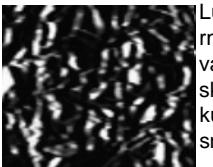
Aluminium (#15)



Lum 0.28
rms 0.71
var 0.04
skew 0.70
kurt 2.90
snr 1.42

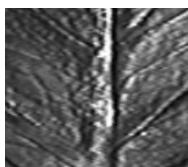


Lum 0.33
rms 0.90
var 0.09
skew 0.79
kurt 2.36
snr 1.11

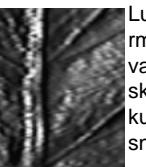


Lum 0.20
rms 1.25
var 0.06
skew 1.62
kurt 4.60
snr 0.80

Laitue (#23)

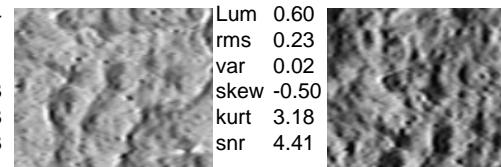


Lum 0.34
rms 0.46
var 0.03
skew 1.26
kurt 5.30
snr 2.17

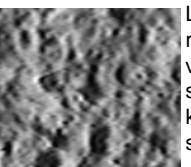


Lum 0.26
rms 0.59
var 0.02
skew 1.49
kurt 6.14
snr 1.68

Plâtre_z (#30)



Lum 0.14
rms 0.81
var 0.01
skew 2.23
kurt 11.3
snr 1.23

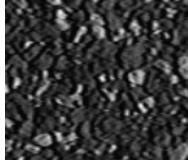


Lum 0.60
rms 0.23
var 0.02
skew -0.50
kurt 3.18
snr 4.41

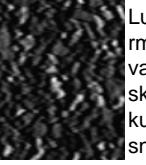


Lum 0.45
rms 0.41
var 0.03
skew -0.18
kurt 2.33
snr 2.43

Galets de toit (#32)

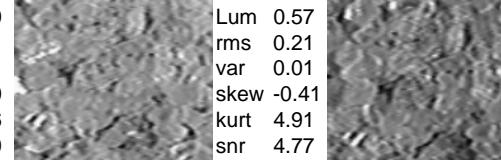


Lum 0.21
rms 0.72
var 0.02
skew 1.33
kurt 5.06
snr 1.39

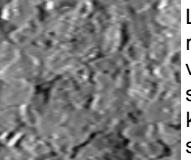


Lum 0.17
rms 0.87
var 0.02
skew 1.60
kurt 6.09
snr 1.16

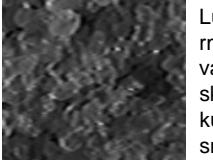
Sel (cristaux) (#43)



Lum 0.10
rms 1.01
var 0.01
skew 2.00
kurt 7.66
snr 0.99

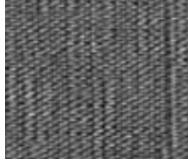


Lum 0.57
rms 0.21
var 0.01
skew -0.41
kurt 4.91
snr 4.77

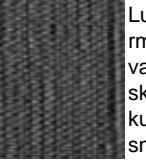


Lum 0.47
rms 0.26
var 0.02
skew -0.03
kurt 4.18
snr 3.83

Lin (#44)

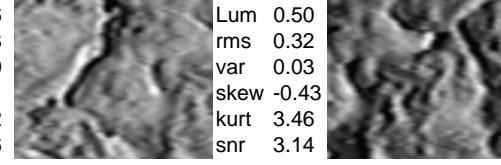


Lum 0.37
rms 0.24
var 0.01
skew 0.16
kurt 2.91
snr 4.17



Lum 0.27
rms 0.27
var 0.01
skew 0.13
kurt 2.99
snr 3.77

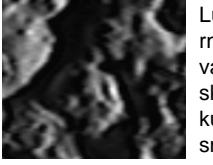
Béton (#45)



Lum 0.16
rms 0.38
var 0.00
skew 0.11
kurt 2.92
snr 2.66

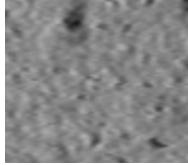


Lum 0.50
rms 0.32
var 0.03
skew -0.43
kurt 3.46
snr 3.14

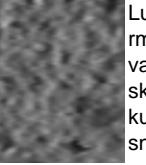


Lum 0.39
rms 0.49
var 0.04
skew 0.15
kurt 2.35
snr 2.04

Mie de pain (#52)

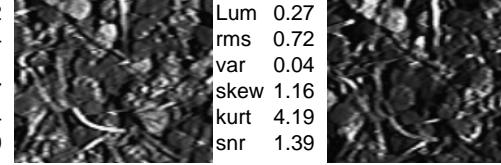


Lum 0.52
rms 0.12
var 0.00
skew -1.38
kurt 7.78
snr 8.46



Lum 0.40
rms 0.21
var 0.01
skew -0.50
kurt 3.55
snr 4.66

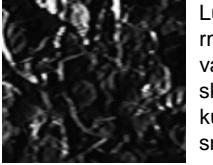
Plante Soleirolia (#53)



Lum 0.22
rms 0.44
var 0.01
skew 0.47
kurt 2.74
snr 2.29



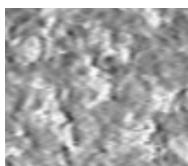
Lum 0.27
rms 0.72
var 0.04
skew 1.16
kurt 4.19
snr 1.39



Lum 0.19
rms 0.92
var 0.03
skew 1.82
kurt 6.59
snr 1.08

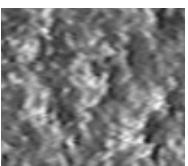
SET A (Originaux) N=Energie, In=Inertie, En=Entropie, Ho=Homogénéité

Velours (#03)



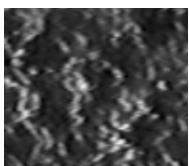
N = 3.3E-04
In = 132.4
En = 8.34
Ho = 0.12

L2



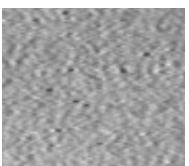
N = 2.8E-04
In = 161.8
En = 8.51
Ho = 0.12

L3



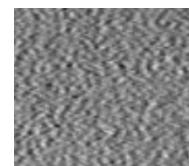
N = 5.9E-04
In = 181.2
En = 8.11
Ho = 0.16

Papier de verre (#06)



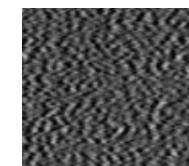
N = 7.1E-04
In = 91.1
En = 7.57
Ho = 0.13

L2



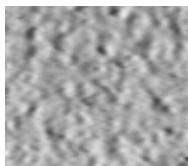
N = 3.1E-04
In = 233.7
En = 8.35
Ho = 0.08

L3

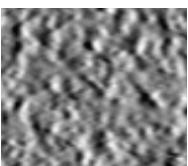


N = 1.8E-04
In = 427.6
En = 8.81
Ho = 0.06

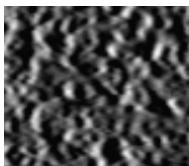
Plâtre (#10)



N = 5.0E-04
In = 77.8
En = 7.91
Ho = 0.14

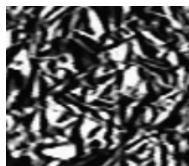


N = 1.7E-04
In = 262.7
En = 8.90
Ho = 0.08

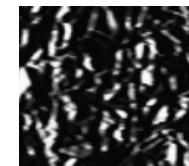


N = 4.8E-04
In = 343.2
En = 8.72
Ho = 0.11

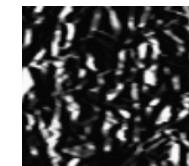
Aluminium (#15)



N = 6.3E-04
In = 1398.2
En = 8.82
Ho = 0.11

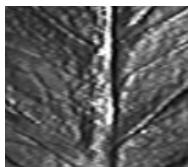


N = 3.6E-03
In = 823.4
En = 7.64
Ho = 0.20

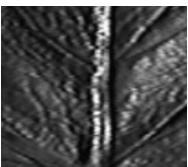


N = 1.7E-02
In = 382.8
En = 5.67
Ho = 0.37

Laitue (#23)



N = 2.5E-04
In = 370.0
En = 8.66
Ho = 0.10

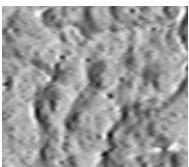


N = 3.2E-04
In = 298.1
En = 8.46
Ho = 0.12

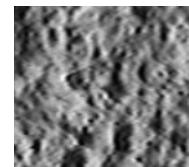


N = 1.2E-03
In = 164.2
En = 7.61
Ho = 0.20

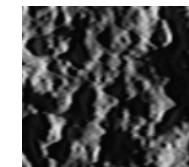
Plâtre_z (#30)



N = 2.6E-04
In = 220.1
En = 8.57
Ho = 0.10

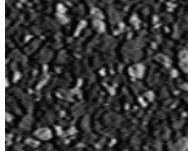


N = 1.5E-04
In = 341.7
En = 9.03
Ho = 0.08

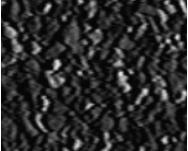


N = 9.5E-04
In = 383.6
En = 8.28
Ho = 0.16

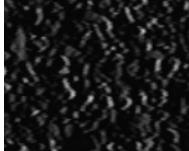
Galets de toit (#32)



N = 2.8E-04
In = 494.7
En = 8.63
Ho = 0.08

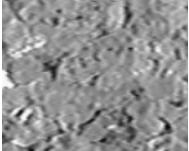


N = 8.6E-04
In = 468.7
En = 8.22
Ho = 0.12

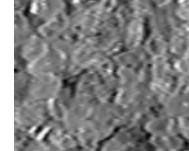


N = 4.5E-03
In = 233.8
En = 7.06
Ho = 0.21

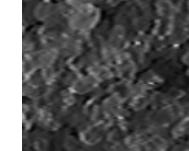
Sel (cristaux) (#43)



N = 3.4E-04
In = 273.5
En = 8.41
Ho = 0.11



N = 2.8E-04
In = 268.2
En = 8.55
Ho = 0.10



N = 4.8E-04
In = 141.8
En = 8.05
Ho = 0.15

Lin (#44)



N = 2.5E-04
In = 397.4
En = 8.53
Ho = 0.06

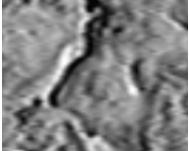


N = 3.6E-04
In = 235.9
En = 8.17
Ho = 0.08

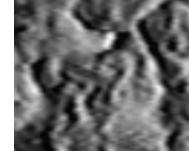


N = 5.7E-04
In = 131.7
En = 7.74
Ho = 0.11

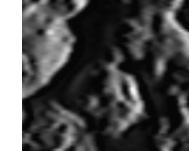
Béton (#45)



N = 2.8E-04
In = 153.0
En = 8.52
Ho = 0.13

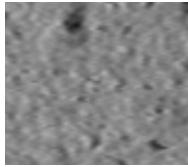


N = 2.0E-04
In = 178.9
En = 8.78
Ho = 0.12

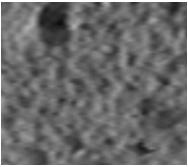


N = 1.2E-03
In = 152.4
En = 8.00
Ho = 0.21

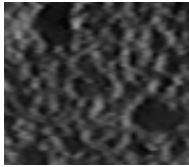
Mie de pain (#52)



N = 1.4E-03
In = 32.6
En = 7.01
Ho = 0.22

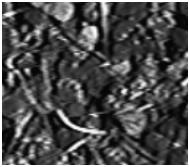


N = 6.3E-04
In = 53.1
En = 7.68
Ho = 0.17

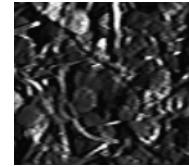


N = 5.1E-04
In = 78.1
En = 7.91
Ho = 0.17

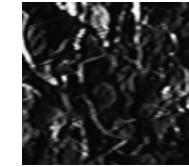
Plante Soleirolia (#53)



N = 2.3E-04
In = 742.3
En = 8.88
Ho = 0.09



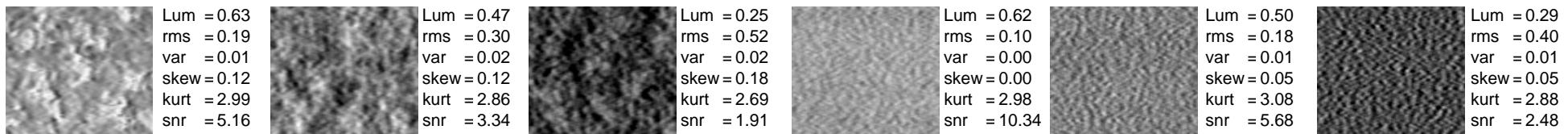
N = 6.1E-04
In = 567.9
En = 8.32
Ho = 0.13



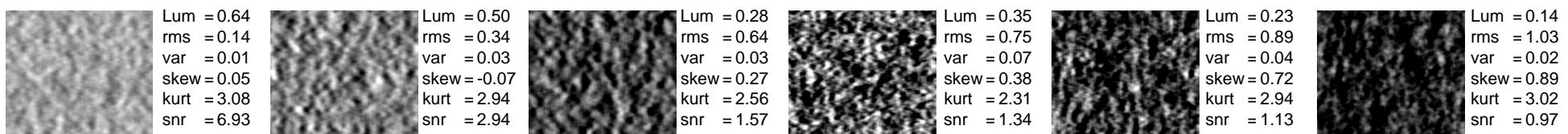
N = 2.0E-03
In = 396.4
En = 7.48
Ho = 0.20

SET A (Scrambled)

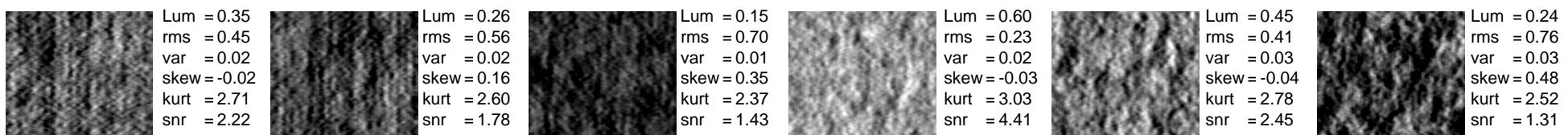
Velours (#03)



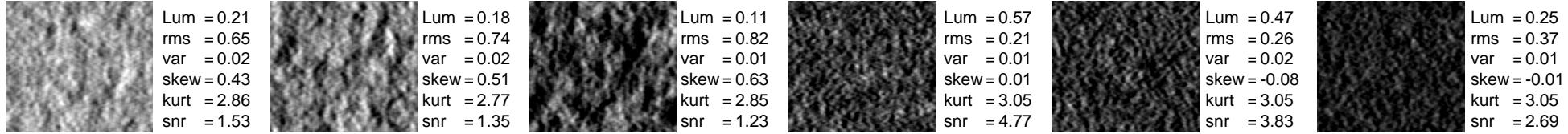
Plâtre (#10)



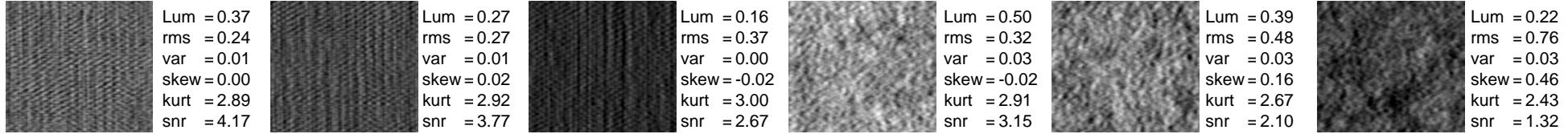
Laitue (#23)



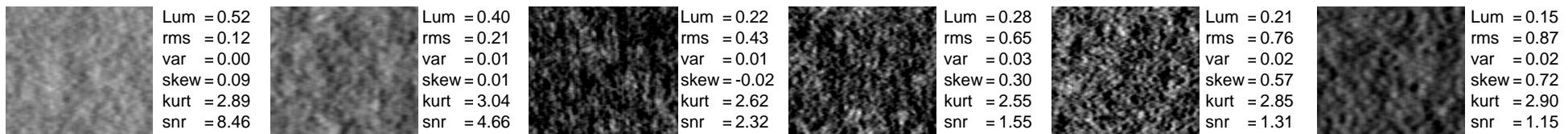
Galets de toit (#32)



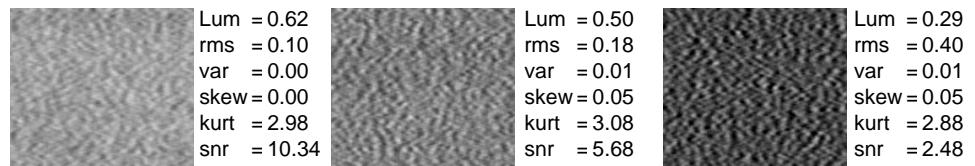
Lin (#44)



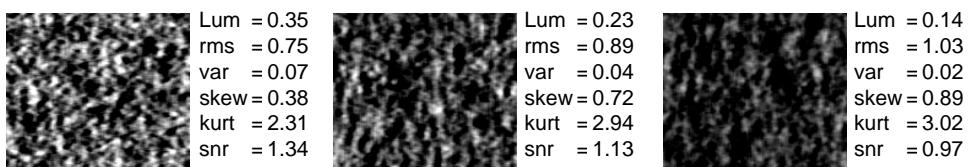
Mie de pain (#52)



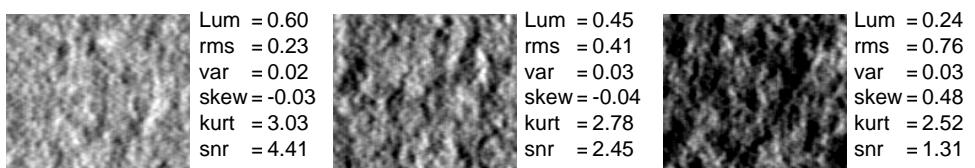
Papier de verre (#06)



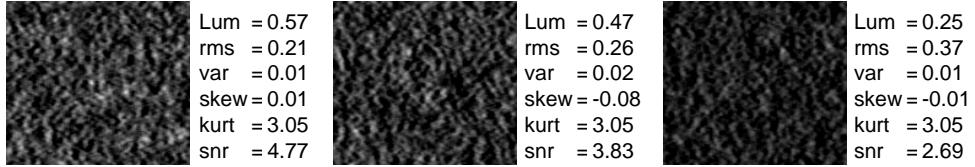
Aluminium (#15)



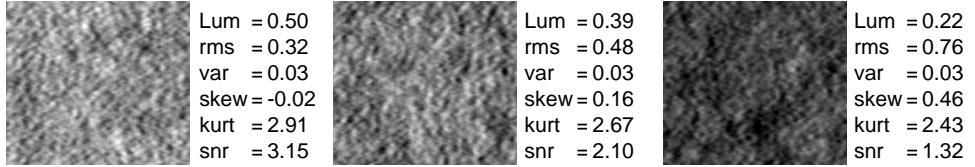
Plâtre_z (#30)



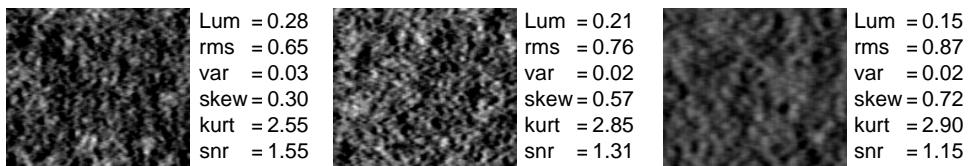
Sel (cristaux) (#43)



Béton (#45)



Plante Soleirolia (#53)

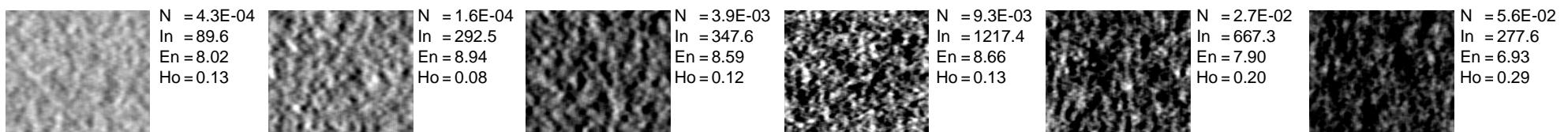


SET A (Scrambled)

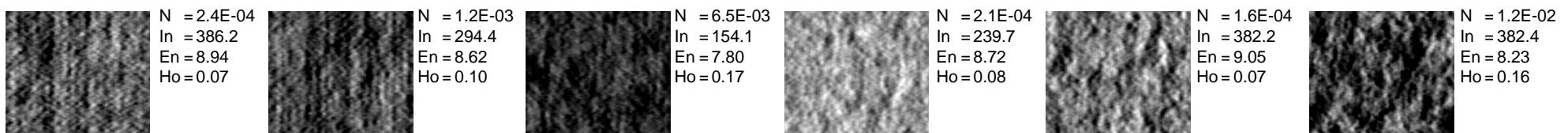
Velours (#03)



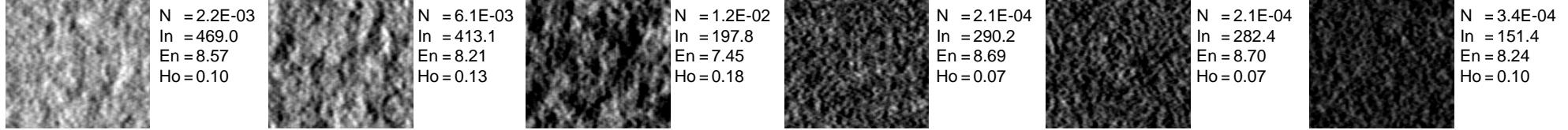
Plâtre (#10)



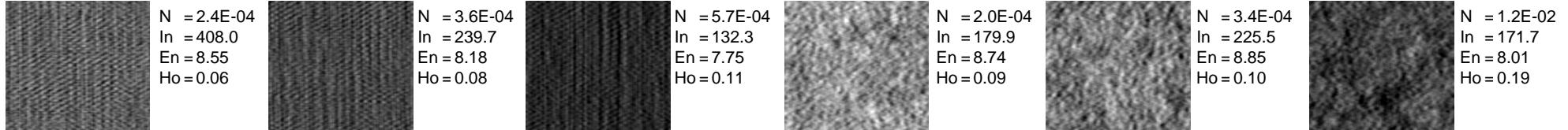
Laitue (#23)



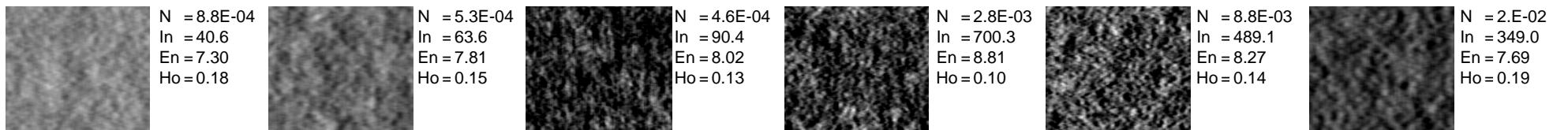
Galets de toit (#32)



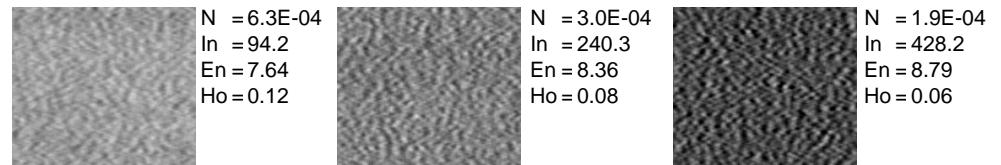
Lin (#44)



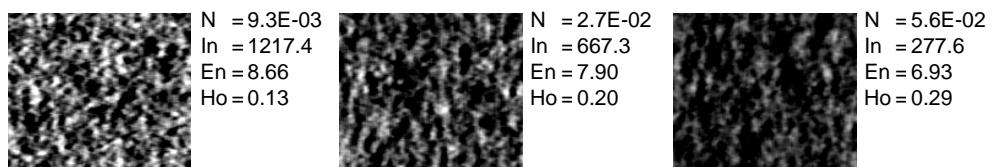
Mie de pain (#52)



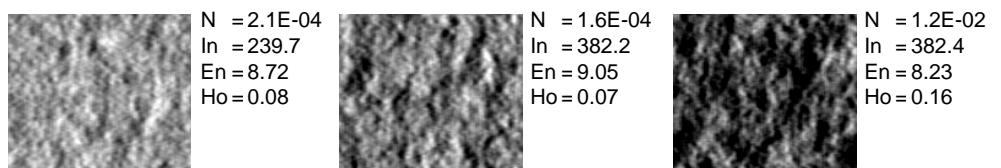
Papier de verre (#06)



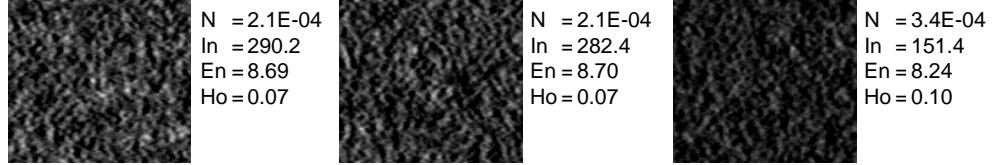
Aluminium (#15)



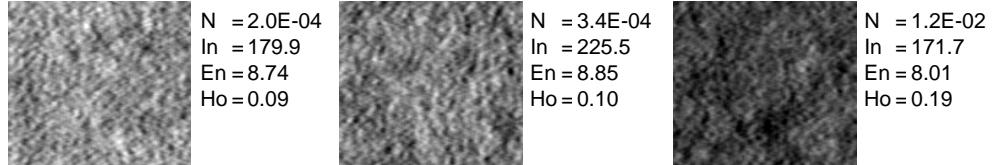
Plâtre_z (#30)



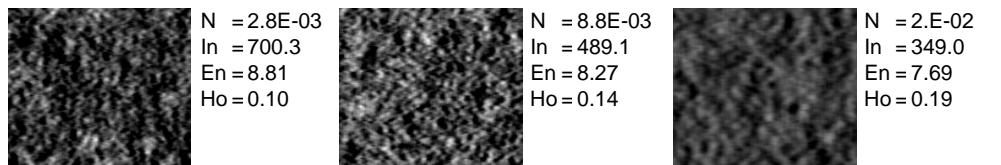
Sel (cristaux) (#43)



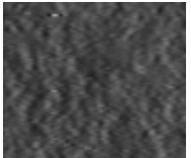
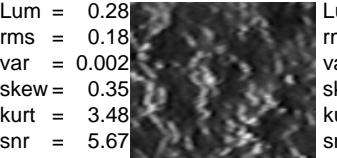
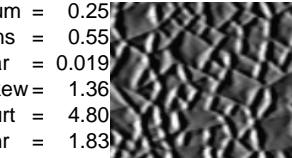
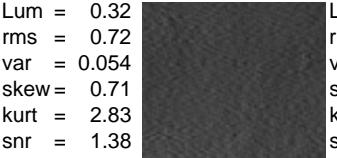
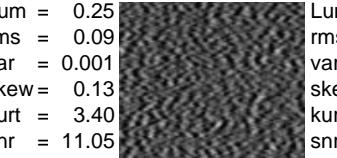
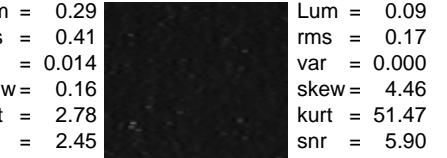
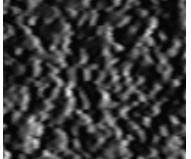
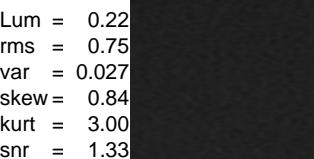
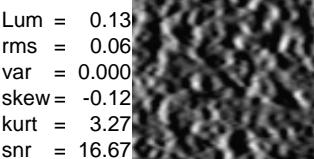
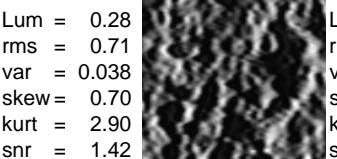
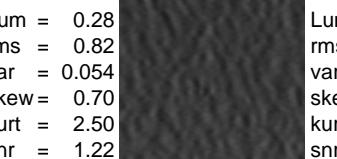
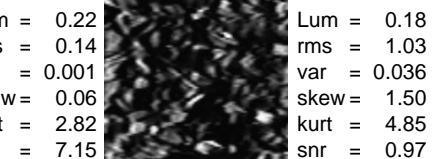
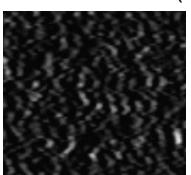
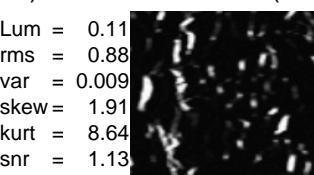
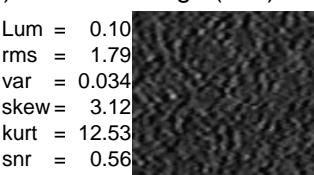
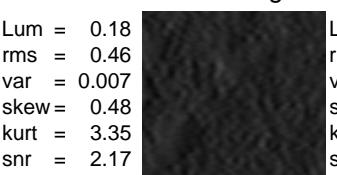
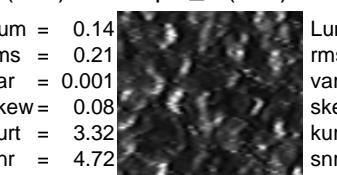
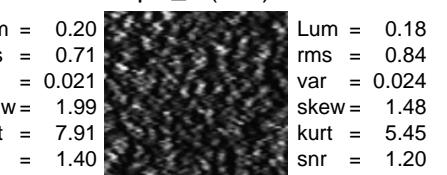
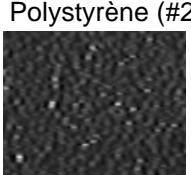
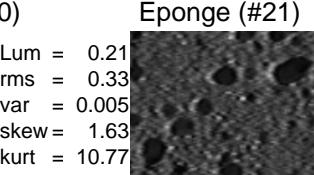
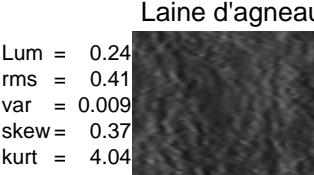
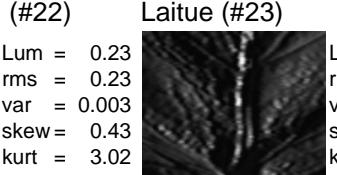
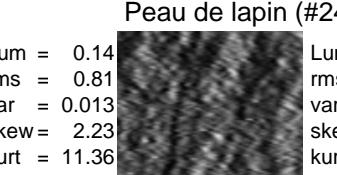
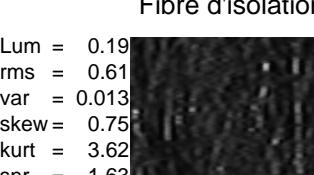
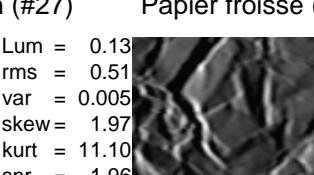
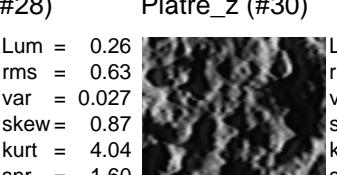
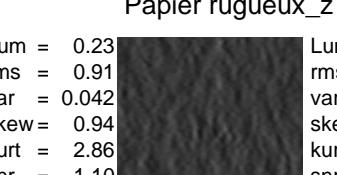
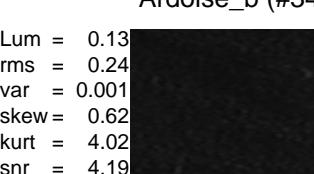
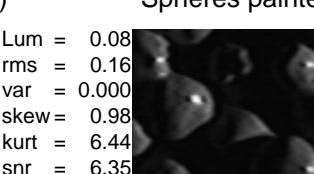
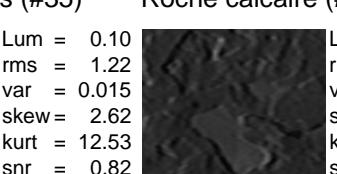
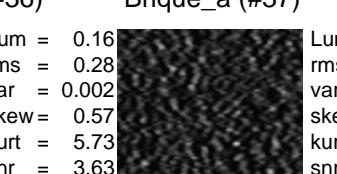
Béton (#45)



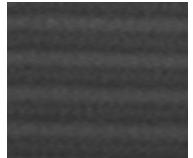
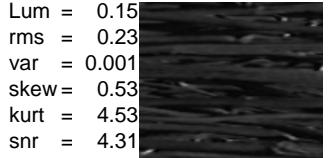
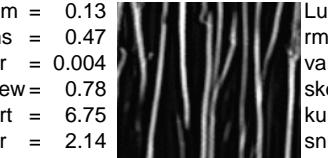
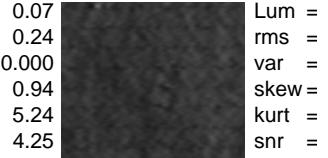
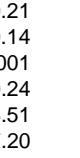
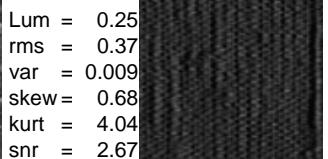
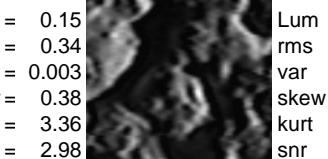
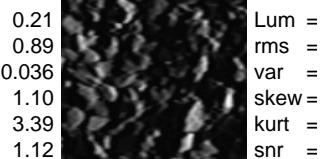
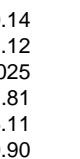
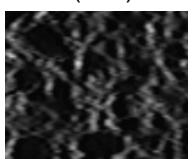
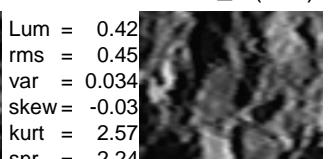
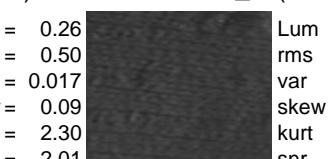
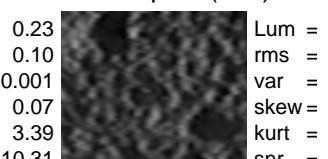
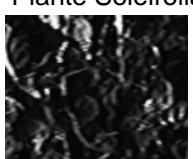
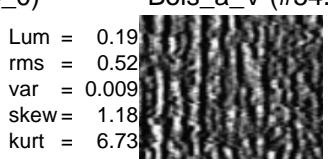
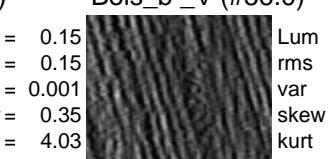
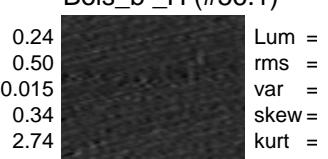
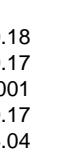
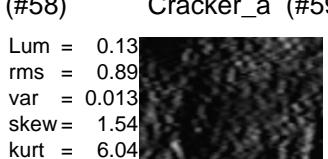
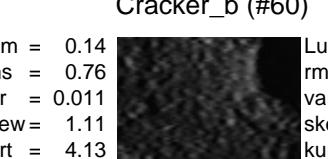
Plante Soleirolia (#53)



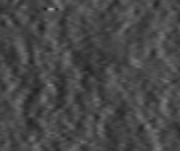
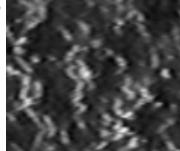
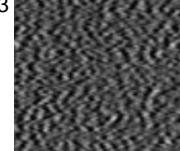
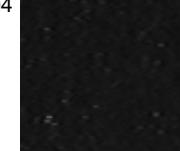
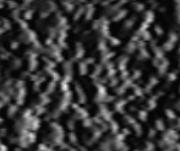
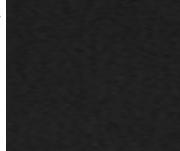
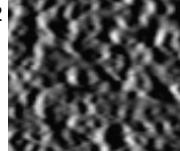
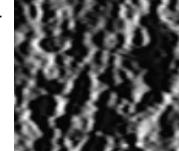
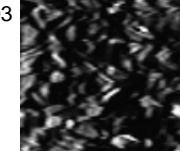
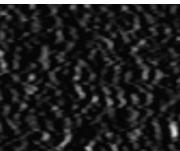
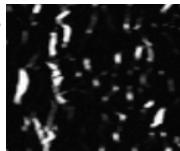
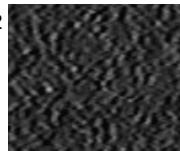
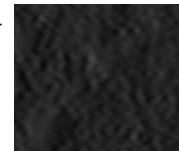
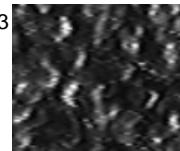
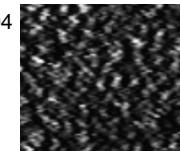
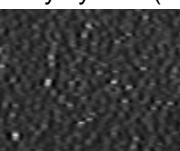
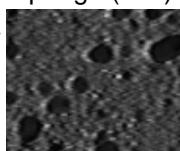
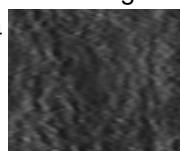
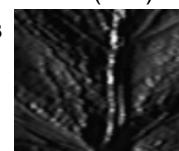
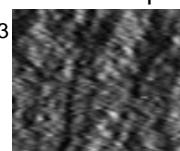
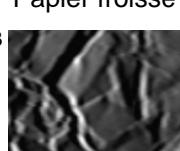
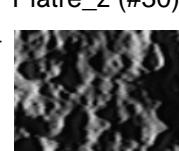
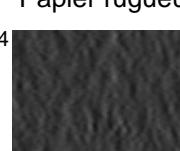
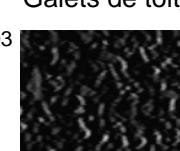
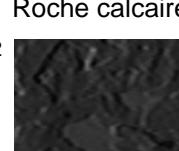
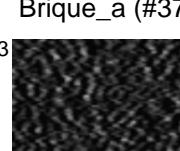
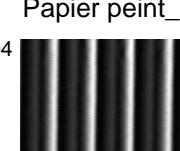
SET B (1/2)

Bitume (#01)	Velours (#03)	Plastique rugueux (#04)	Cuir (#05)	Papier de verre (#06)	Velours_b (#07)
					
Lum = 0.28 rms = 0.18 var = 0.002 skew = 0.35 kurt = 3.48 snr = 5.67	Lum = 0.25 rms = 0.55 var = 0.019 skew = 1.36 kurt = 4.80 snr = 1.83	Lum = 0.32 rms = 0.72 var = 0.054 skew = 0.71 kurt = 2.83 snr = 1.38	Lum = 0.25 rms = 0.09 var = 0.001 skew = 0.13 kurt = 3.40 snr = 11.05	Lum = 0.29 rms = 0.41 var = 0.014 skew = 0.16 kurt = 2.78 snr = 2.45	Lum = 0.09 rms = 0.17 var = 0.000 skew = 4.46 kurt = 51.47 snr = 5.90
Cailloux (#08)	Verre dépoli (#9)	Plâtre_a (#10)	Plâtre_b (#11)	Papier crépi (#12)	Herbe artificielle (#13)
					
Lum = 0.22 rms = 0.75 var = 0.027 skew = 0.84 kurt = 3.00 snr = 1.33	Lum = 0.13 rms = 0.06 var = 0.000 skew = -0.12 kurt = 3.27 snr = 16.67	Lum = 0.28 rms = 0.71 var = 0.038 skew = 0.70 kurt = 2.90 snr = 1.42	Lum = 0.28 rms = 0.82 var = 0.054 skew = 0.70 kurt = 2.50 snr = 1.22	Lum = 0.22 rms = 0.14 var = 0.001 skew = 0.06 kurt = 2.82 snr = 7.15	Lum = 0.18 rms = 1.03 var = 0.036 skew = 1.50 kurt = 4.85 snr = 0.97
Galets de toit (#14)	Aluminium (#15)	Liège (#16)	Carreau rugueux (#17)	Tapis_a (#18)	Tapis_b (#19)
					
Lum = 0.11 rms = 0.88 var = 0.009 skew = 1.91 kurt = 8.64 snr = 1.13	Lum = 0.10 rms = 1.79 var = 0.034 skew = 3.12 kurt = 12.53 snr = 0.56	Lum = 0.18 rms = 0.46 var = 0.007 skew = 0.48 kurt = 3.35 snr = 2.17	Lum = 0.14 rms = 0.21 var = 0.001 skew = 0.08 kurt = 3.32 snr = 4.72	Lum = 0.20 rms = 0.71 var = 0.021 skew = 1.99 kurt = 7.91 snr = 1.40	Lum = 0.18 rms = 0.84 var = 0.024 skew = 1.48 kurt = 5.45 snr = 1.20
Polystyrène (#20)	Eponge (#21)	Laine d'agneau (#22)	Laitue (#23)	Peau de lapin (#24)	Tuile (#25)
					
Lum = 0.21 rms = 0.33 var = 0.005 skew = 1.63 kurt = 10.77 snr = 3.07	Lum = 0.24 rms = 0.41 var = 0.009 skew = 0.37 kurt = 4.04 snr = 2.45	Lum = 0.23 rms = 0.23 var = 0.003 skew = 0.43 kurt = 3.02 snr = 4.32	Lum = 0.14 rms = 0.81 var = 0.013 skew = 2.23 kurt = 11.36 snr = 1.23	Lum = 0.29 rms = 0.46 var = 0.018 skew = 0.56 kurt = 3.05 snr = 2.17	Lum = 0.10 rms = 0.15 var = 0.000 skew = 0.49 kurt = 4.94 snr = 6.53
Loofa (#26)	Fibre d'isolation (#27)	Papier froissé (#28)	Plâtre_z (#30)	Papier rugueux_z (#31)	Galets de toit_z (#32)
					
Lum = 0.19 rms = 0.61 var = 0.013 skew = 0.75 kurt = 3.62 snr = 1.63	Lum = 0.13 rms = 0.51 var = 0.005 skew = 1.97 kurt = 11.10 snr = 1.96	Lum = 0.26 rms = 0.63 var = 0.027 skew = 0.87 kurt = 4.04 snr = 1.60	Lum = 0.23 rms = 0.91 var = 0.042 skew = 0.94 kurt = 2.86 snr = 1.10	Lum = 0.22 rms = 0.16 var = 0.001 skew = 0.14 kurt = 3.13 snr = 6.21	Lum = 0.10 rms = 1.01 var = 0.010 skew = 2.00 kurt = 7.66 snr = 0.99
Ardoise_a (#33)	Ardoise_b (#34)	Sphères peintes (#35)	Roche calcaire (#36)	Brique_a (#37)	Papier peint_V (#38.0)
					
Lum = 0.13 rms = 0.24 var = 0.001 skew = 0.62 kurt = 4.02 snr = 4.19	Lum = 0.08 rms = 0.16 var = 0.000 skew = 0.98 kurt = 6.44 snr = 6.35	Lum = 0.10 rms = 1.22 var = 0.015 skew = 2.62 kurt = 12.53 snr = 0.82	Lum = 0.16 rms = 0.28 var = 0.002 skew = 0.57 kurt = 5.73 snr = 3.63	Lum = 0.15 rms = 0.71 var = 0.011 skew = 1.15 kurt = 4.62 snr = 1.42	Lum = 0.25 rms = 1.02 var = 0.067 skew = 1.43 kurt = 3.84 snr = 0.98

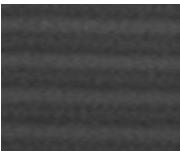
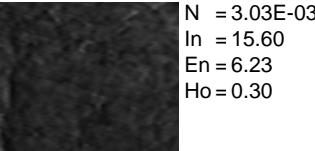
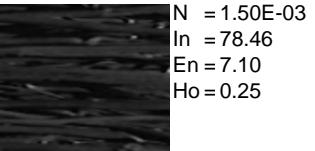
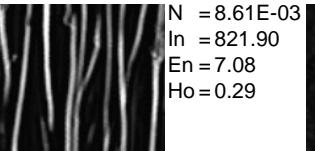
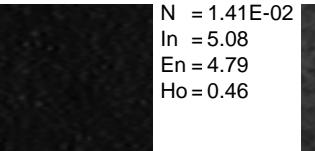
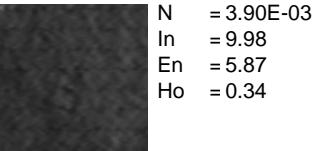
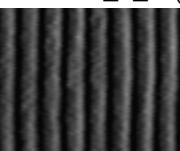
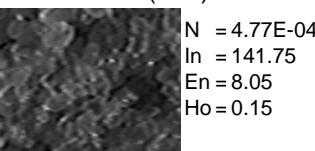
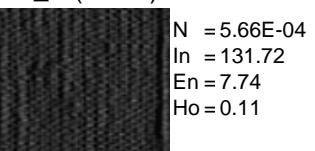
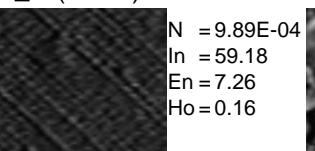
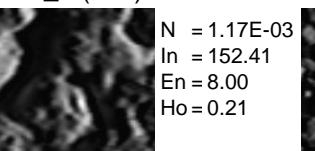
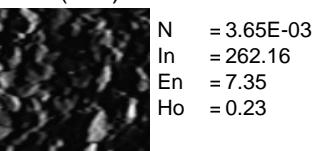
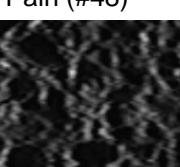
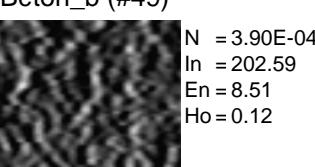
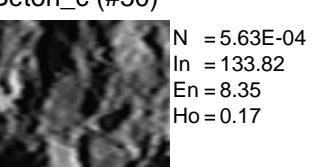
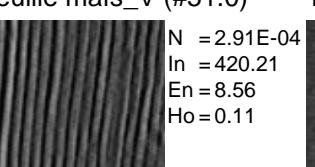
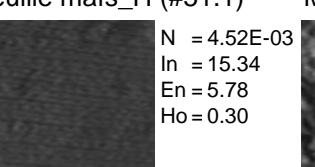
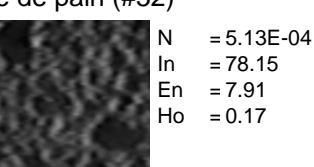
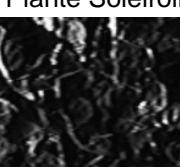
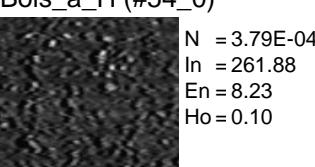
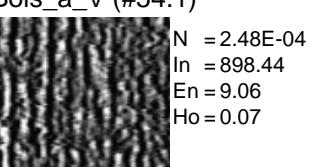
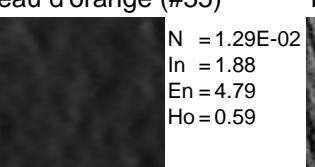
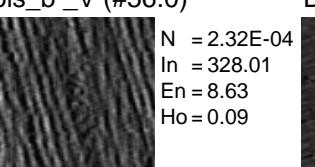
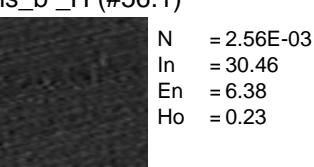
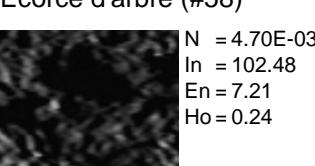
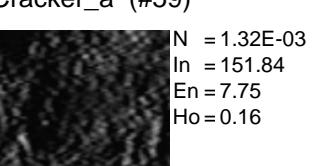
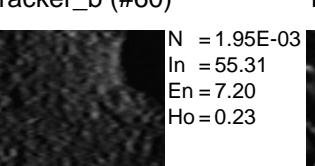
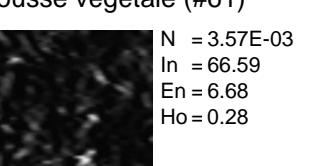
SET B (2/2)

Papier peint_H (#38.1)	Peau humaine (#39)	Paille_H (#40.0)	Paille_V (#40.1)	Brique_b (#41.0)	Velours_b_H (#42.0)
					
Lum = 0.25 rms = 0.10 var = 0.00 skew = 0.42 kurt = 2.42 snr = 10.06	Lum = 0.15 rms = 0.23 var = 0.001 skew = 0.53 kurt = 4.53 snr = 4.31	Lum = 0.13 rms = 0.47 var = 0.004 skew = 0.78 kurt = 6.75 snr = 2.14	Lum = 0.19 rms = 1.25 var = 0.057 skew = 1.45 kurt = 4.02 snr = 0.80	Lum = 0.07 rms = 0.24 var = 0.000 skew = 0.94 kurt = 5.24 snr = 4.25	Lum = 0.21 rms = 0.14 var = 0.001 skew = 0.24 kurt = 3.51 snr = 7.20
Velours_b_V (#42.1)	Cristaux sel (#43)	Lin_V (#44.0)	Lin_H (#44.1)	Béton_a (#45)	Pierres (#47)
					
Lum = 0.21 rms = 0.61 var = 0.016 skew = 0.15 kurt = 1.91 snr = 1.64	Lum = 0.25 rms = 0.37 var = 0.009 skew = 0.68 kurt = 4.04 snr = 2.67	Lum = 0.16 rms = 0.38 var = 0.004 skew = 0.11 kurt = 2.92 snr = 2.66	Lum = 0.15 rms = 0.34 var = 0.003 skew = 0.38 kurt = 3.36 snr = 2.98	Lum = 0.21 rms = 0.89 var = 0.036 skew = 1.10 kurt = 3.39 snr = 1.12	Lum = 0.14 rms = 1.12 var = 0.025 skew = 1.81 kurt = 6.11 snr = 0.90
Pain (#48)	Béton_b (#49)	Béton_c (#50)	Feuille maïs_V (#51.0)	Feuille maïs_H (#51.)	Mie de pain (#52)
					
Lum = 0.15 rms = 0.79 var = 0.014 skew = 1.16 kurt = 4.15 snr = 1.26	Lum = 0.42 rms = 0.45 var = 0.034 skew = -0.03 kurt = 2.57 snr = 2.24	Lum = 0.27 rms = 0.68 var = 0.034 skew = 0.46 kurt = 2.45 snr = 1.46	Lum = 0.26 rms = 0.50 var = 0.017 skew = 0.09 kurt = 2.30 snr = 2.01	Lum = 0.23 rms = 0.10 var = 0.001 skew = 0.07 kurt = 3.39 snr = 10.31	Lum = 0.22 rms = 0.44 var = 0.009 skew = 0.47 kurt = 2.74 snr = 2.29
Plante Soleirolia (#53)	Bois_a_H (#54.0)	Bois_a_V (#54.1)	Peau d'orange (#55)	Bois_b_V (#56.0)	Bois_b_H (#56.1)
					
Lum = 0.14 rms = 1.18 var = 0.025 skew = 2.64 kurt = 10.94 snr = 0.85	Lum = 0.19 rms = 0.52 var = 0.009 skew = 1.18 kurt = 6.73 snr = 1.94	Lum = 0.31 rms = 0.75 var = 0.052 skew = 0.76 kurt = 2.84 snr = 1.34	Lum = 0.15 rms = 0.15 var = 0.001 skew = 0.35 kurt = 4.03 snr = 6.47	Lum = 0.24 rms = 0.50 var = 0.015 skew = 0.34 kurt = 2.74 snr = 2.01	Lum = 0.18 rms = 0.17 var = 0.001 skew = 0.17 kurt = 4.04 snr = 5.81
Queue de paon (#57)	Ecorce d'arbre (#58)	Cracker_a (#59)	Cracker_b (#60)	Mousse végétale (#61)	
					
Lum = 0.07 rms = 0.99 var = 0.004 skew = 2.51 kurt = 10.06 snr = 1.01	Lum = 0.13 rms = 0.89 var = 0.013 skew = 1.54 kurt = 6.04 snr = 1.12	Lum = 0.14 rms = 0.76 var = 0.011 skew = 1.11 kurt = 4.13 snr = 1.32	Lum = 0.12 rms = 0.60 var = 0.005 skew = 0.95 kurt = 4.13 snr = 1.66	Lum = 0.09 rms = 0.95 var = 0.008 skew = 2.93 kurt = 16.52 snr = 1.06	

SET B (1/2)

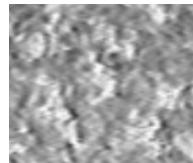
Bitume (#01)	Velours (#03)	Plastique rugueux (#04)	Cuir (#05)	Papier de verre (#06)	Velours_b (#07)
					
N = 1.28E-03 In = 35.69 En = 7.00 Ho = 0.21	N = 5.87E-04 In = 181.22 En = 8.11 Ho = 0.16	N = 4.19E-04 In = 775.49 En = 8.94 Ho = 0.10	N = 3.69E-03 In = 20.67 En = 5.98 Ho = 0.25	N = 1.83E-04 In = 427.58 En = 8.81 Ho = 0.06	N = 2.43E-02 In = 6.59 En = 4.36 Ho = 0.51
Cailloux (#08)	Verre dépoli (#9)	Plâtre_a (#10)	Plâtre_b (#11)	Papier crépi (#12)	Herbe artificielle (#13)
					
N = 4.81E-04 In = 334.32 En = 8.52 Ho = 0.11	N = 3.57E-02 In = 2.07 En = 3.72 Ho = 0.57	N = 4.79E-04 In = 343.17 En = 8.72 Ho = 0.11	N = 7.58E-04 In = 494.21 En = 8.65 Ho = 0.12	N = 3.14E-03 In = 13.46 En = 6.09 Ho = 0.30	N = 1.79E-03 In = 455.00 En = 7.97 Ho = 0.17
Galets de toit (#14)	Aluminium (#15)	Liège (#16)	Carreau rugueux (#17)	Tapis_a (#18)	Tapis_b (#19)
					
N = 2.26E-03 In = 173.87 En = 7.39 Ho = 0.17	N = 1.68E-02 In = 382.83 En = 5.67 Ho = 0.37	N = 3.74E-04 In = 179.67 En = 8.16 Ho = 0.10	N = 3.76E-03 In = 10.86 En = 5.94 Ho = 0.33	N = 5.54E-04 In = 244.89 En = 8.10 Ho = 0.15	N = 4.77E-04 In = 496.56 En = 8.42 Ho = 0.10
Polystyrène (#20)	Eponge (#21)	Laine d'agneau (#22)	Laitue (#23)	Peau de lapin (#24)	Tuile (#25)
					
N = 7.32E-04 In = 125.61 En = 7.59 Ho = 0.14	N = 3.70E-04 In = 198.11 En = 8.22 Ho = 0.11	N = 1.16E-03 In = 39.12 En = 7.10 Ho = 0.20	N = 1.19E-03 In = 164.20 En = 7.61 Ho = 0.20	N = 2.37E-04 In = 215.33 En = 8.59 Ho = 0.09	N = 9.68E-03 In = 8.06 En = 5.10 Ho = 0.38
Loofa (#26)	Fibre d'isolation (#27)	Papier froissé (#28)	Plâtre_z (#30)	Papier rugueux_z (#31)	Galets de toit_z (#32)
					
N = 3.80E-04 In = 228.15 En = 8.33 Ho = 0.11	N = 1.17E-03 In = 74.60 En = 7.21 Ho = 0.21	N = 5.30E-04 In = 232.91 En = 8.34 Ho = 0.18	N = 9.49E-04 In = 383.61 En = 8.28 Ho = 0.16	N = 2.31E-03 In = 22.16 En = 6.42 Ho = 0.25	N = 4.47E-03 In = 233.81 En = 7.06 Ho = 0.21
Ardoise_a (#33)	Ardoise_b (#34)	Sphères peintes (#35)	Roche calcaire (#36)	Brique_a (#37)	Papier peint_V (#38.0)
					
N = 2.60E-03 In = 29.43 En = 6.35 Ho = 0.23	N = 1.76E-02 In = 4.14 En = 4.47 Ho = 0.48	N = 1.69E-02 In = 53.54 En = 5.86 Ho = 0.46	N = 2.60E-03 In = 28.15 En = 6.60 Ho = 0.29	N = 6.54E-04 In = 239.92 En = 8.06 Ho = 0.11	N = 3.83E-03 In = 433.10 En = 7.41 Ho = 0.21

SET B (2/2)

Papier peint_H (#38.1)	Peau humaine (#39)	Paille_H (#40.0)	Paille_V (#40.1)	Brique_b (#41.0)	Velours_b_H (#42.0)
 N = 6.61E-03 In = 4.93 En = 5.33 Ho = 0.43	 N = 3.03E-03 In = 15.60 En = 6.23 Ho = 0.30	 N = 1.50E-03 In = 78.46 En = 7.10 Ho = 0.25	 N = 8.61E-03 In = 821.90 En = 7.08 Ho = 0.29	 N = 1.41E-02 In = 5.08 En = 4.79 Ho = 0.46	 N = 3.90E-03 In = 9.98 En = 5.87 Ho = 0.34
Velours_b_V (#42.1)	Cristaux sel (#43)	Lin_V (#44.0)	Lin_H (#44.1)	Béton_a (#45)	Pierres (#47)
 N = 7.25E-04 In = 120.56 En = 7.78 Ho = 0.15	 N = 4.77E-04 In = 141.75 En = 8.05 Ho = 0.15	 N = 5.66E-04 In = 131.72 En = 7.74 Ho = 0.11	 N = 9.89E-04 In = 59.18 En = 7.26 Ho = 0.16	 N = 1.17E-03 In = 152.41 En = 8.00 Ho = 0.21	 N = 3.65E-03 In = 262.16 En = 7.35 Ho = 0.23
Pain (#48)	Béton_b (#49)	Béton_c (#50)	Feuille maïs_V (#51.0)	Feuille maïs_H (#51.1)	Mie de pain (#52)
 N = 1.12E-03 In = 156.14 En = 7.83 Ho = 0.18	 N = 3.90E-04 In = 202.59 En = 8.51 Ho = 0.12	 N = 5.63E-04 In = 133.82 En = 8.35 Ho = 0.17	 N = 2.91E-04 In = 420.21 En = 8.56 Ho = 0.11	 N = 4.52E-03 In = 15.34 En = 5.78 Ho = 0.30	 N = 5.13E-04 In = 78.15 En = 7.91 Ho = 0.17
Plante Soleirolia (#53)	Bois_a_H (#54.0)	Bois_a_V (#54.1)	Peau d'orange (#55)	Bois_b_V (#56.0)	Bois_b_H (#56.1)
 N = 2.05E-03 In = 396.37 En = 7.48 Ho = 0.20	 N = 3.79E-04 In = 261.88 En = 8.23 Ho = 0.10	 N = 2.48E-04 In = 898.44 En = 9.06 Ho = 0.07	 N = 1.29E-02 In = 1.88 En = 4.79 Ho = 0.59	 N = 2.32E-04 In = 328.01 En = 8.63 Ho = 0.09	 N = 2.56E-03 In = 30.46 En = 6.38 Ho = 0.23
Queue de paon (#57)	Ecorce d'arbre (#58)	Cracker_a (#59)	Cracker_b (#60)	Mousse végétale (#61)	
 N = 2.37E-02 In = 58.54 En = 5.38 Ho = 0.45	 N = 4.70E-03 In = 102.48 En = 7.21 Ho = 0.24	 N = 1.32E-03 In = 151.84 En = 7.75 Ho = 0.16	 N = 1.95E-03 In = 55.31 En = 7.20 Ho = 0.23	 N = 3.57E-03 In = 66.59 En = 6.68 Ho = 0.28	

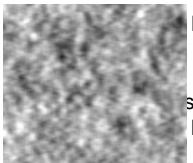
SET C (1/4)

Velours (#03)



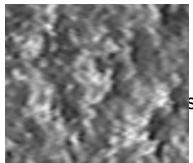
Lum = 0.63
rms = 0.19
var = 0.01
skew = 0.38
kurt = 2.89
snr = 5.15

L2



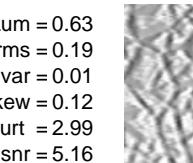
Lum = 0.47
rms = 0.30
var = 0.02
skew = 0.54
kurt = 2.98
snr = 3.34

Scrambled L1



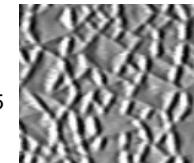
Lum = 0.63
rms = 0.19
var = 0.01
skew = 0.12
kurt = 2.99
snr = 5.16

Plastique rugueux (#04)



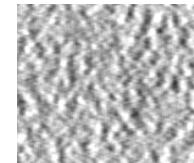
Lum = 0.72
rms = 0.21
var = 0.02
skew = -0.25
kurt = 2.53
snr = 4.86

L2



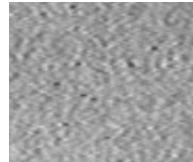
Lum = 0.53
rms = 0.43
var = 0.05
skew = 0.07
kurt = 2.54
snr = 2.32

Scrambled L1

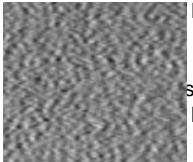


Lum = 0.72
rms = 0.20
var = 0.02
skew = -0.10
kurt = 2.63
snr = 4.96

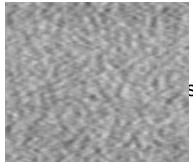
Papier de verre (#06)



Lum = 0.62
rms = 0.10
var = 0.00
skew = -0.23
kurt = 4.09
snr = 10.34

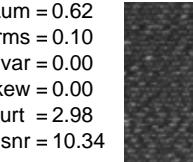


Lum = 0.50
rms = 0.18
var = 0.01
skew = -0.02
kurt = 3.09
snr = 5.68

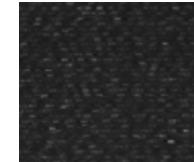


Lum = 0.62
rms = 0.10
var = 0.00
skew = 0.00
kurt = 2.98
snr = 10.34

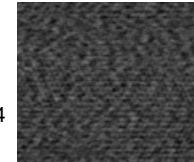
Velours_b (#07)



Lum = 0.25
rms = 0.23
var = 0.00
skew = 1.22
kurt = 5.08
snr = 4.34

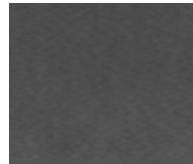


Lum = 0.15
rms = 0.20
var = 0.00
skew = 1.95
kurt = 10.34
snr = 5.13

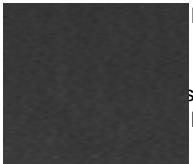


Lum = 0.25
rms = 0.23
var = 0.00
skew = -0.01
kurt = 3.07
snr = 4.33

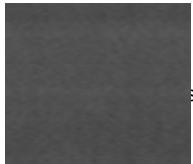
Verre dépoli (#9)



Lum = 0.31
rms = 0.04
var = 0.00
skew = 0.26
kurt = 2.67
snr = 25.59

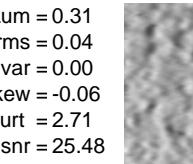


Lum = 0.22
rms = 0.04
var = 0.00
skew = 0.29
kurt = 3.49
snr = 23.10

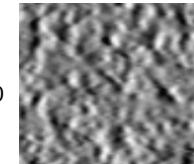


Lum = 0.31
rms = 0.04
var = 0.00
skew = -0.06
kurt = 2.71
snr = 25.48

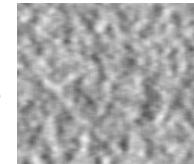
Plâtre_a (#10)



Lum = 0.64
rms = 0.14
var = 0.01
skew = -0.50
kurt = 3.44
snr = 6.93

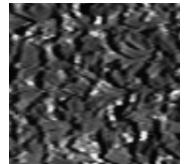


Lum = 0.50
rms = 0.34
var = 0.03
skew = -0.05
kurt = 2.81
snr = 2.93

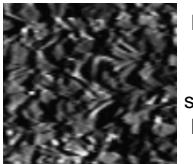


Lum = 0.64
rms = 0.14
var = 0.01
skew = 0.05
kurt = 3.08
snr = 6.93

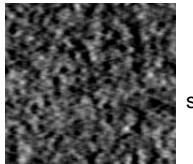
Herbe artificielle (#13)



Lum = 0.26
rms = 0.59
var = 0.02
skew = 1.00
kurt = 4.64
snr = 1.69

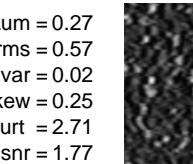


Lum = 0.23
rms = 0.74
var = 0.03
skew = 0.89
kurt = 3.49
snr = 1.35

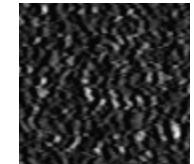


Lum = 0.27
rms = 0.57
var = 0.02
skew = 0.25
kurt = 2.71
snr = 1.77

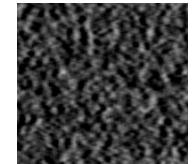
Galets de toit (#14)



Lum = 0.21
rms = 0.66
var = 0.02
skew = 1.34
kurt = 5.28
snr = 1.51

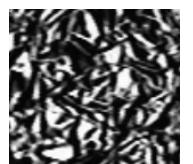


Lum = 0.16
rms = 0.76
var = 0.02
skew = 1.52
kurt = 6.28
snr = 1.32

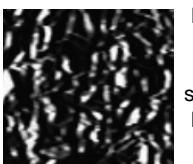


Lum = 0.21
rms = 0.61
var = 0.02
skew = 0.29
kurt = 2.70
snr = 1.64

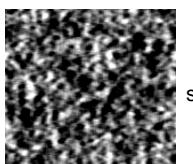
Aluminium (#15)



Lum = 0.33
rms = 0.90
var = 0.09
skew = 0.79
kurt = 2.36
snr = 1.11

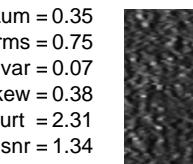


Lum = 0.20
rms = 1.25
var = 0.06
skew = 1.62
kurt = 4.60
snr = 0.80

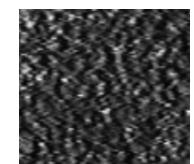


Lum = 0.35
rms = 0.75
var = 0.07
skew = 0.38
kurt = 2.31
snr = 1.34

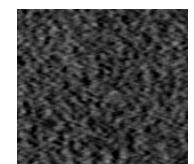
Tapis_b (#19)



Lum = 0.22
rms = 0.45
var = 0.01
skew = 1.14
kurt = 5.42
snr = 2.24

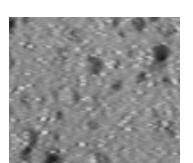


Lum = 0.22
rms = 0.59
var = 0.02
skew = 1.17
kurt = 4.87
snr = 1.70

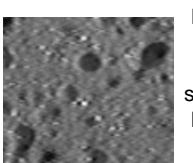


Lum = 0.22
rms = 0.44
var = 0.01
skew = 0.08
kurt = 2.85
snr = 2.28

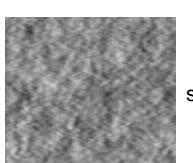
Eponge (#21)



Lum = 0.53
rms = 0.15
var = 0.01
skew = -0.92
kurt = 6.31
snr = 6.57

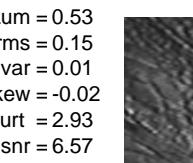


Lum = 0.41
rms = 0.23
var = 0.01
skew = -0.42
kurt = 4.59
snr = 4.33

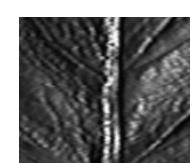


Lum = 0.53
rms = 0.15
var = 0.01
skew = -0.02
kurt = 2.93
snr = 6.57

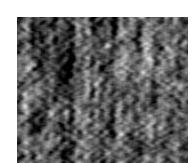
Laitue (#23)



Lum = 0.34
rms = 0.46
var = 0.03
skew = 1.26
kurt = 5.30
snr = 2.17



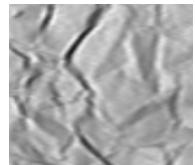
Lum = 0.26
rms = 0.59
var = 0.02
skew = 1.49
kurt = 6.14
snr = 1.68



Lum = 0.35
rms = 0.45
var = 0.02
skew = -0.02
kurt = 2.71
snr = 2.22

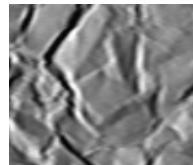
SET C (2/4)

Papier froissé (#28)



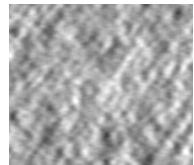
Lum = 0.65
rms = 0.17
var = 0.01
skew = -0.82
kurt = 4.87
snr = 6.06

L2



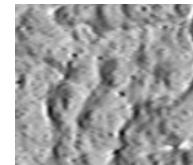
Lum = 0.50
rms = 0.31
var = 0.02
skew = -0.19
kurt = 3.76
snr = 3.18

Scrambled L1



Lum = 0.65
rms = 0.17
var = 0.01
skew = -0.01
kurt = 3.08
snr = 6.06

Plâtre_z (#30)



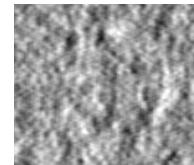
Lum = 0.60
rms = 0.23
var = 0.02
skew = -0.50
kurt = 3.18
snr = 4.41

L2



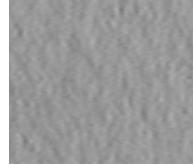
Lum = 0.45
rms = 0.41
var = 0.03
skew = -0.18
kurt = 2.33
snr = 2.43

Scrambled L1

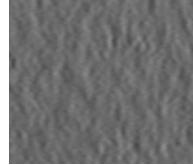


Lum = 0.60
rms = 0.23
var = 0.02
skew = -0.03
kurt = 3.03
snr = 4.41

Papier rugueux_z (#31)



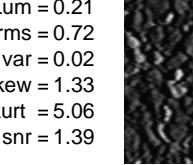
Lum = 0.59
rms = 0.04
var = 0.00
skew = -0.13
kurt = 3.16
snr = 25.09



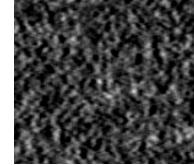
Lum = 0.42
rms = 0.08
var = 0.00
skew = -0.01
kurt = 3.18
snr = 12.94

Lum = 0.59
rms = 0.04
var = 0.00
skew = -0.01
kurt = 3.01
snr = 25.08

Galets de toit_z (#32)

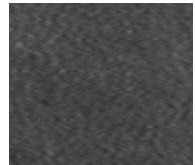


Lum = 0.21
rms = 0.72
var = 0.02
skew = 1.33
kurt = 5.06
snr = 1.39

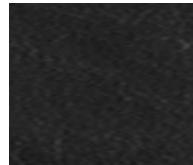


Lum = 0.21
rms = 0.65
var = 0.02
skew = 0.43
kurt = 2.86
snr = 1.53

Ardoise_b (#34)



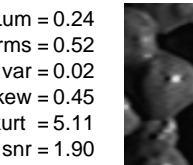
Lum = 0.29
rms = 0.09
var = 0.00
skew = 0.35
kurt = 3.89
snr = 10.98



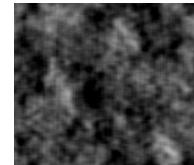
Lum = 0.15
rms = 0.11
var = 0.00
skew = 0.75
kurt = 5.41
snr = 9.34

Lum = 0.29
rms = 0.09
var = 0.00
skew = -0.04
kurt = 2.98
snr = 10.96

Sphères peintes (#35)



Lum = 0.24
rms = 0.52
var = 0.02
skew = 0.45
kurt = 5.11
snr = 1.90



Lum = 0.24
rms = 0.51
var = 0.02
skew = 0.22
kurt = 2.81
snr = 1.95

Papier peint_V (#38.0)



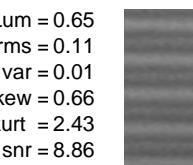
Lum = 0.62
rms = 0.26
var = 0.03
skew = 0.23
kurt = 2.32
snr = 3.89



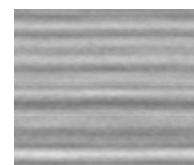
Lum = 0.44
rms = 0.50
var = 0.05
skew = 0.01
kurt = 1.79
snr = 1.99

Lum = 0.62
rms = 0.25
var = 0.02
skew = 0.52
kurt = 2.40
snr = 3.96

Papier peint_H (#38.1)

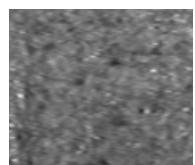


Lum = 0.65
rms = 0.11
var = 0.01
skew = 0.66
kurt = 2.43
snr = 8.86

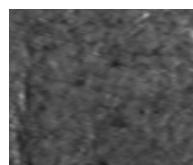


Lum = 0.65
rms = 0.11
var = 0.01
skew = 0.19
kurt = 2.55
snr = 8.86

Peau humaine (#39)



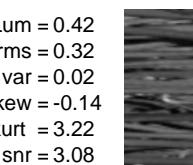
Lum = 0.43
rms = 0.15
var = 0.00
skew = 0.32
kurt = 4.80
snr = 6.67



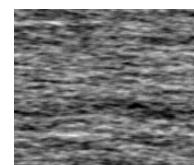
Lum = 0.31
rms = 0.16
var = 0.00
skew = 0.23
kurt = 4.86
snr = 6.39

Lum = 0.43
rms = 0.15
var = 0.00
skew = 0.01
kurt = 2.91
snr = 6.67

Paille_H (#40.0)

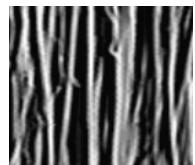


Lum = 0.42
rms = 0.32
var = 0.02
skew = -0.14
kurt = 3.22
snr = 3.08

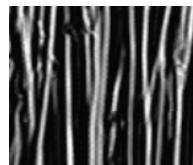


Lum = 0.42
rms = 0.32
var = 0.02
skew = -0.06
kurt = 2.92
snr = 3.09

Paille_V (#40.1)



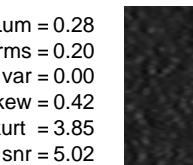
Lum = 0.35
rms = 0.83
var = 0.09
skew = 0.43
kurt = 1.84
snr = 1.21



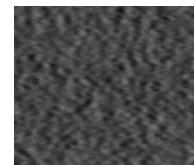
Lum = 0.28
rms = 1.00
var = 0.08
skew = 0.82
kurt = 2.34
snr = 1.00

Lum = 0.37
rms = 0.72
var = 0.07
skew = 0.32
kurt = 2.25
snr = 1.39

Brique_b (#41.0)

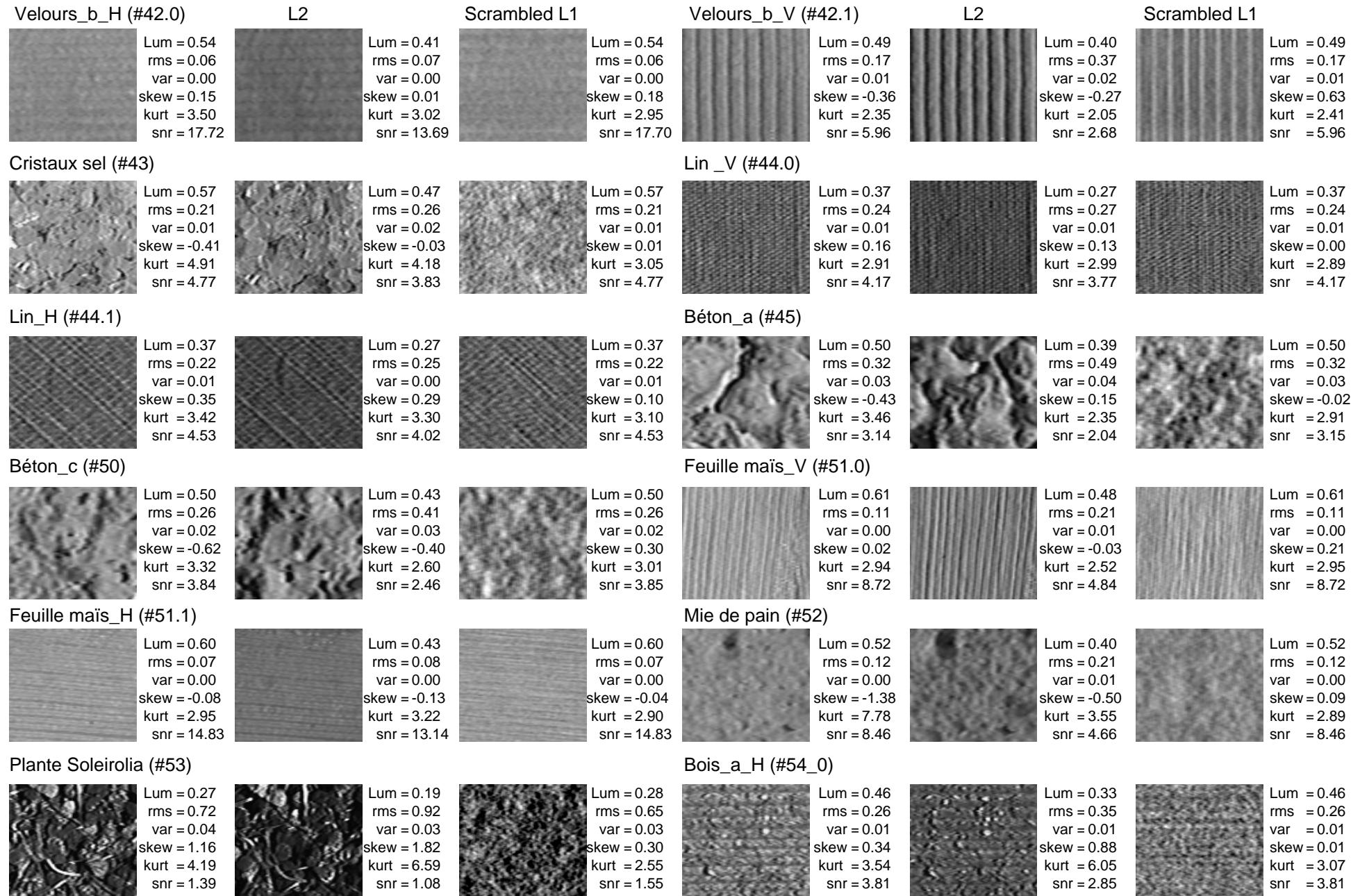


Lum = 0.28
rms = 0.20
var = 0.00
skew = 0.42
kurt = 3.85
snr = 5.02



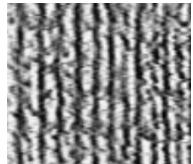
Lum = 0.14
rms = 0.23
var = 0.00
skew = 0.88
kurt = 5.11
snr = 4.37

SET C (3/4)



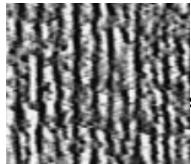
SET C (4/4)

Bois_a_V (#54.1)



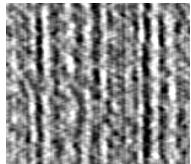
Lum = 0.53
rms = 0.46
var = 0.06
skew = -0.06
kurt = 1.91
snr = 2.18

L2



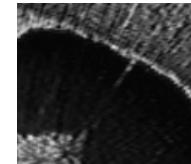
Lum = 0.45
rms = 0.58
var = 0.07
skew = 0.19
kurt = 1.94
snr = 1.73

Scrambled L1



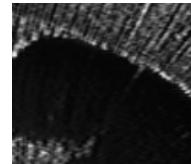
Lum = 0.53
rms = 0.44
var = 0.06
skew = -0.06
kurt = 2.46
snr = 2.25

Queue de paon (#57)



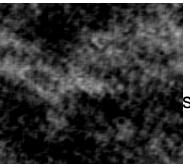
Lum = 0.16
rms = 1.00
var = 0.03
skew = 1.35
kurt = 4.05
snr = 1.00

L2



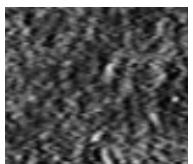
Lum = 0.17
rms = 0.82
var = 0.02
skew = 0.58
kurt = 2.50
snr = 1.21

Scrambled L1

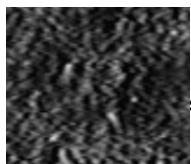


Lum = 0.13
rms = 1.01
var = 0.02
skew = 1.59
kurt = 5.42
snr = 0.99

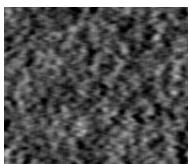
Cracker_a (#59)



Lum = 0.28
rms = 0.44
var = 0.02
skew = 0.62
kurt = 3.45
snr = 2.27

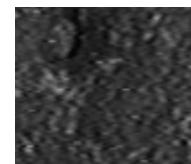


Lum = 0.21
rms = 0.56
var = 0.01
skew = 0.79
kurt = 3.70
snr = 1.78

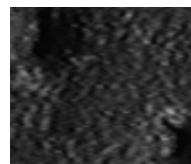


Lum = 0.28
rms = 0.43
var = 0.02
skew = 0.07
kurt = 2.81
snr = 2.30

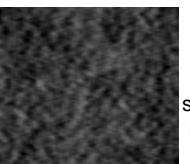
Cracker_b (#60)



Lum = 0.21
rms = 0.32
var = 0.00
skew = 0.87
kurt = 5.33
snr = 3.08

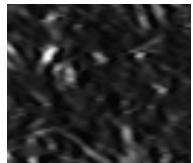


Lum = 0.21
rms = 0.32
var = 0.00
skew = -0.02
kurt = 2.94
snr = 3.09



Lum = 0.17
rms = 0.46
var = 0.01
skew = 0.65
kurt = 4.21
snr = 2.20

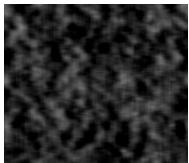
Mousse végétale (#61)



Lum = 0.14
rms = 0.72
var = 0.01
skew = 2.61
kurt = 13.64
snr = 1.39



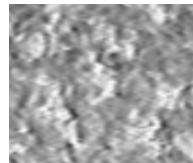
Lum = 0.15
rms = 0.94
var = 0.02
skew = 2.65
kurt = 11.78
snr = 1.06



Lum = 0.15
rms = 0.66
var = 0.01
skew = 0.35
kurt = 2.46
snr = 1.52

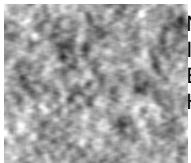
SET C (1/4)

Velours (#03)



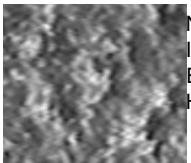
N = 3.3E-04
In = 132.39
En = 8.34
Ho = 0.12

L2



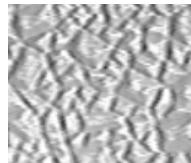
N = 2.8E-04
In = 161.81
En = 8.51
Ho = 0.12

Scrambled L1



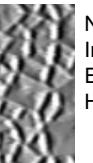
N = 2.77E-04
In = 143.17
En = 8.44
Ho = 0.10

Plastique rugueux (#04)



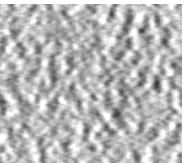
N = 1.79E-04
In = 353.25
En = 8.87
Ho = 0.09

L2



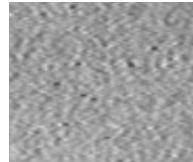
N = 1.40E-04
In = 769.36
En = 9.25
Ho = 0.08

Scrambled L1

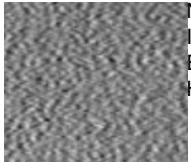


N = 4.05E-04
In = 364.54
En = 8.81
Ho = 0.08

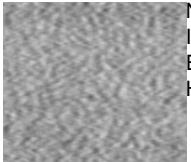
Papier de verre (#06)



N = 7.12E-04
In = 91.06
En = 7.57
Ho = 0.13

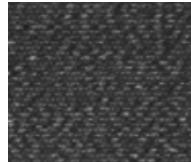


N = 3.05E-04
In = 233.65
En = 8.35
Ho = 0.08

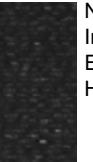


N = 6.28E-04
In = 94.20
En = 7.64
Ho = 0.12

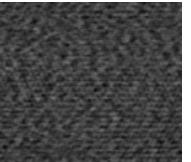
Velours_b (#07)



N = 9.09E-04
In = 134.56
En = 7.43
Ho = 0.13

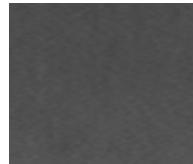


N = 4.44E-03
In = 33.30
En = 6.01
Ho = 0.28

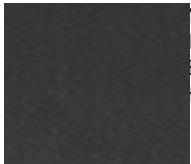


N = 5.86E-04
In = 135.71
En = 7.72
Ho = 0.11

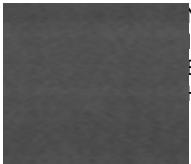
Verre dépoli (#9)



N = 2.14E-02
In = 2.15
En = 4.23
Ho = 0.57

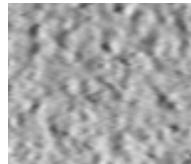


N = 2.73E-02
In = 2.36
En = 4.03
Ho = 0.55



N = 1.66E-02
In = 2.69
En = 4.39
Ho = 0.51

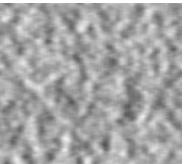
Plâtre_a (#10)



N = 5.01E-04
In = 77.83
En = 7.91
Ho = 0.14

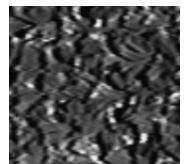


N = 1.73E-04
In = 262.66
En = 8.90
Ho = 0.08

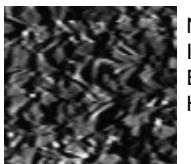


N = 4.32E-04
In = 89.56
En = 8.02
Ho = 0.13

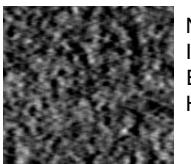
Herbe artificielle (#13)



N = 2.26E-04
In = 401.49
En = 8.71
Ho = 0.10

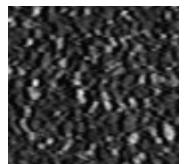


N = 4.28E-04
In = 437.36
En = 8.60
Ho = 0.11

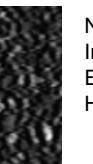


N = 8.44E-04
In = 404.16
En = 8.78
Ho = 0.09

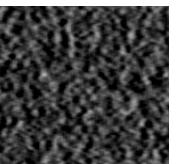
Galets de toit (#14)



N = 2.79E-04
In = 333.89
En = 8.51
Ho = 0.09

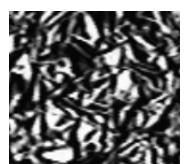


N = 5.36E-04
In = 302.98
En = 8.22
Ho = 0.11

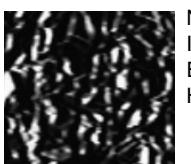


N = 1.90E-03
In = 320.14
En = 8.46
Ho = 0.11

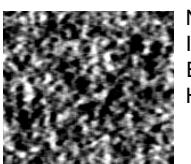
Aluminium (#15)



N = 6.35E-04
In = 1398.15
En = 8.82
Ho = 0.11

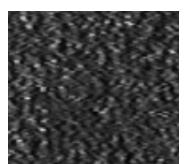


N = 3.57E-03
In = 823.37
En = 7.64
Ho = 0.20



N = 9.34E-03
In = 1217.43
En = 8.66
Ho = 0.13

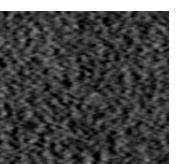
Tapis_b (#19)



N = 3.49E-04
In = 267.69
En = 8.27
Ho = 0.10

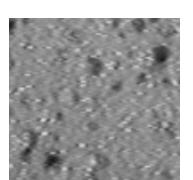


N = 2.47E-04
In = 414.70
En = 8.61
Ho = 0.08

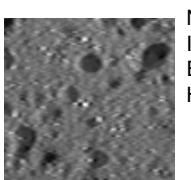


N = 2.88E-04
In = 268.74
En = 8.44
Ho = 0.08

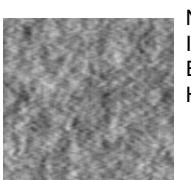
Eponge (#21)



N = 7.19E-04
In = 123.97
En = 7.76
Ho = 0.14

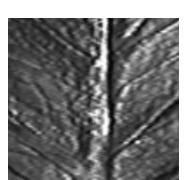


N = 4.78E-04
In = 150.74
En = 8.06
Ho = 0.12

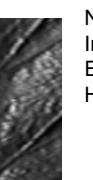


N = 4.02E-04
In = 134.91
En = 8.07
Ho = 0.11

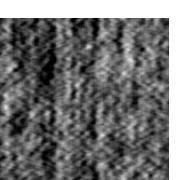
Laitue (#23)



N = 2.52E-04
In = 369.96
En = 8.66
Ho = 0.10

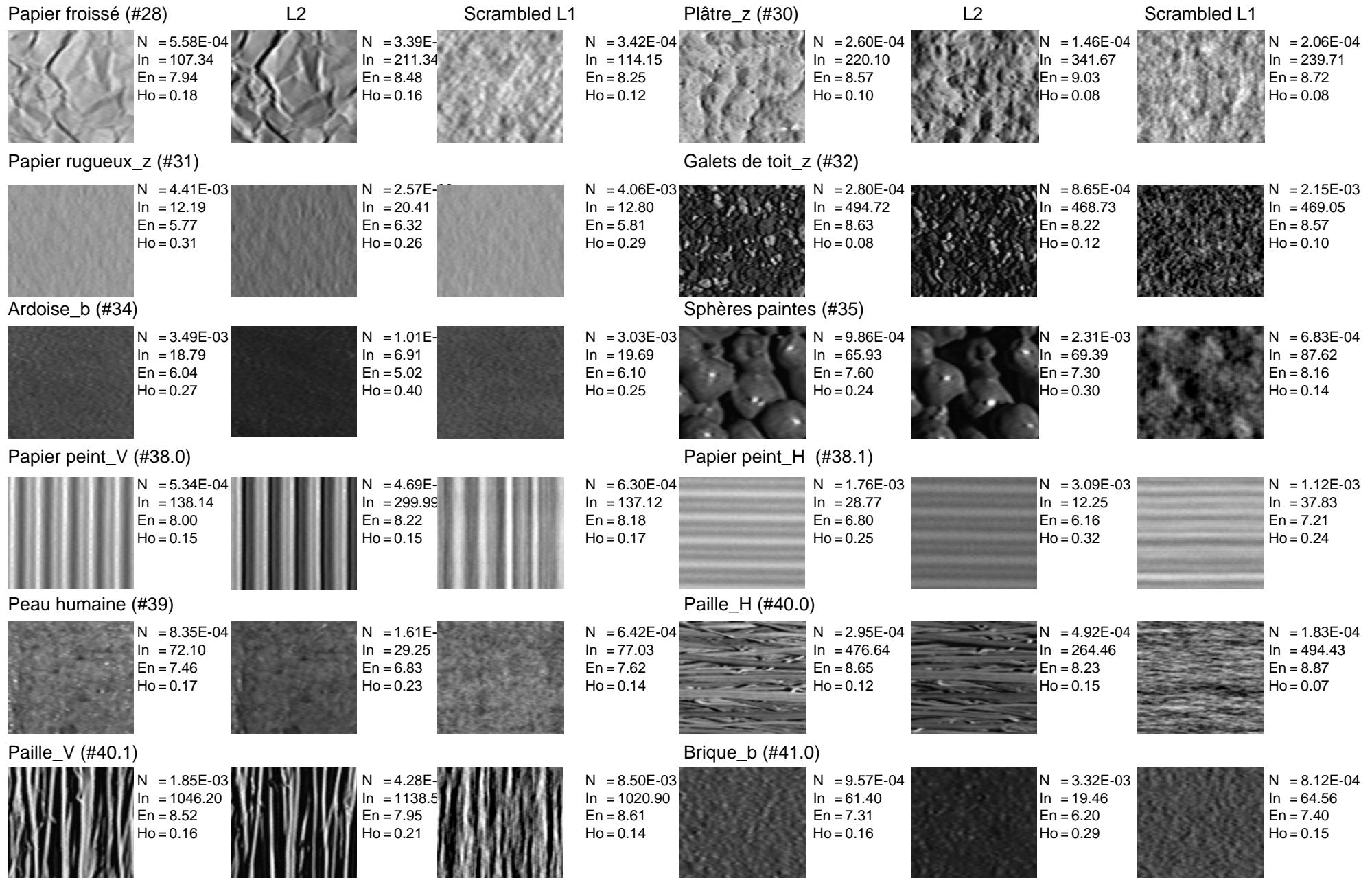


N = 3.20E-04
In = 298.15
En = 8.46
Ho = 0.12

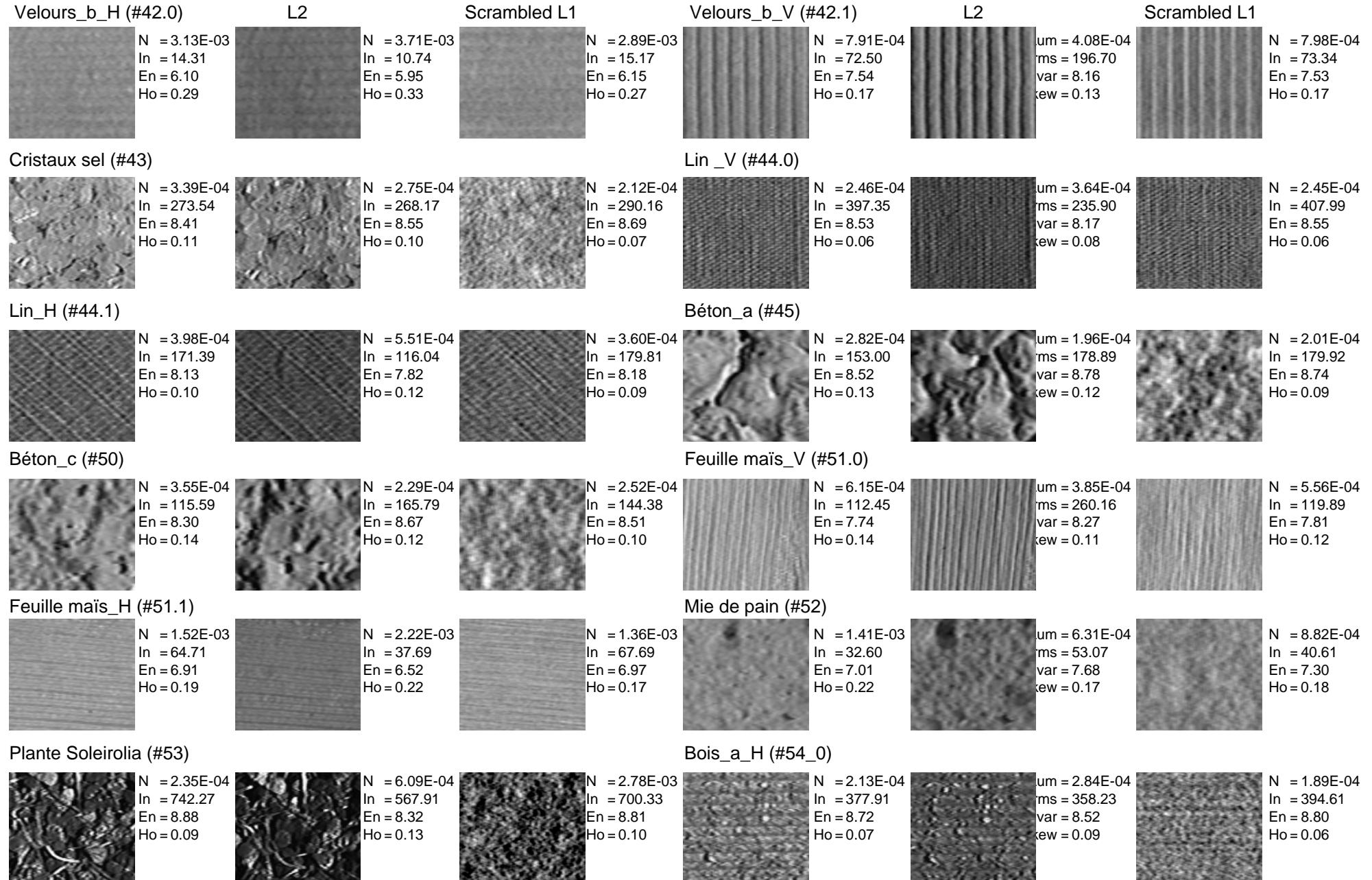


N = 2.40E-04
In = 386.23
En = 8.94
Ho = 0.07

SET C (2/4)

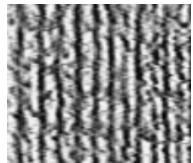


SET C (3/4)



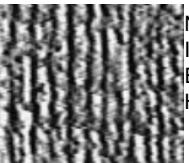
SET C (4/4)

Bois_a_V (#54.1)



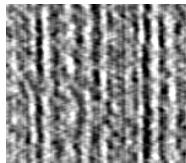
N = 9.57E-05
In = 988.07
En = 9.42
Ho = 0.05

L2



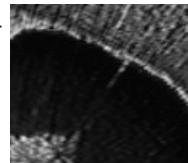
N = 9.99E-05
In = 1119.64
En = 9.42
Ho = 0.05

Scrambled L1



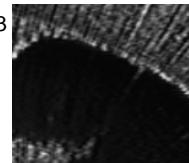
N = 3.97E-04
In = 1036.72
En = 9.30
Ho = 0.06

Queue de paon (#57)



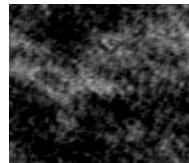
N = 3.02E-03
In = 179.29
En = 7.23
Ho = 0.29

L2



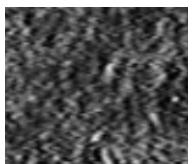
N = 5.58E-03
In = 196.30
En = 6.86
Ho = 0.33

Scrambled L1

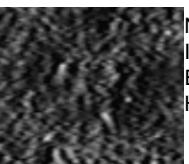


N = 1.53E-02
In = 187.16
En = 7.76
Ho = 0.20

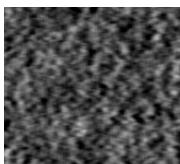
Cracker_a (#59)



N = 2.37E-04
In = 243.03
En = 8.60
Ho = 0.09

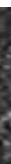


N = 2.82E-04
In = 220.38
En = 8.45
Ho = 0.10

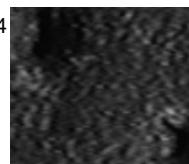


N = 2.74E-04
In = 256.85
En = 8.63
Ho = 0.08

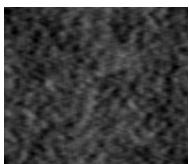
Cracker_b (#60)



N = 8.60E-04
In = 63.99
En = 7.46
Ho = 0.18

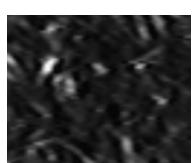


N = 7.94E-04
In = 58.31
En = 7.50
Ho = 0.19

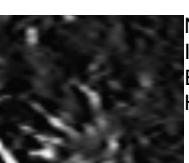


N = 6.19E-04
In = 71.13
En = 7.66
Ho = 0.14

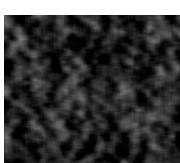
Mousse végétale (#61)



N = 1.31E-03
In = 71.01
En = 7.27
Ho = 0.24



N = 1.16E-03
In = 103.60
En = 7.43
Ho = 0.23



N = 3.77E-03
In = 86.29
En = 7.68
Ho = 0.17

