

Perception du

mouvement propre

Nous avons vu dans le précédent chapitre que la quasi-totalité des modalités sensorielles humaines (hormis le goût et l'odorat) peuvent jouer un rôle dans la perception du mouvement. Nous avons vu également les principes de fonctionnement des systèmes visuel, vestibulaire, somatosensoriel... ainsi que les types d'informations renvoyées par ces modalités sensorielles. Ce chapitre est consacré à étudier comment le système nerveux central extrait de ces signaux des informations sur le déplacement dans l'espace afin de créer la sensation de mouvement.

Nous verrons dans un premier temps que de nombreuses informations peuvent être issues du système visuel seul, notamment à travers le flux optique qui permet d'obtenir à la fois la direction du mouvement et sa vitesse. La perception de la vitesse étant liée à la perception des distances, nous détaillerons également comment l'être humain perçoit la profondeur. Dans une deuxième partie, nous verrons quelles sont les interactions multisensorielles qui existent et qui permettent au cerveau de faire émerger une sensation de mouvement unique et cohérente. Nous détaillerons plus particulièrement le concept de lavection et les théories existantes sur la perception d'un monde stable.

3.1. Perception visuelle du mouvement

La littérature concernant la perception visuelle du mouvement est très fournie car cette modalité visuelle est considérée comme principale pour la perception du mouvement. Le flux optique, notion apparue dans les années 1950, a été largement étudiée car elle permet d'obtenir des informations aussi bien sur la direction du mouvement que sur la vitesse. Après avoir défini les notions de flux optique et de flux rétinien, nous verrons comment ils peuvent être utilisés pour la perception de la vitesse, en relation avec les informations de profondeur.

3.1.1. Flux optique, flux rétinien

Définitions

L'image du monde qui se forme sur nos rétines est en constante évolution. En effet, dès que nous nous déplaçons dans notre environnement, dès que l'œil tourne dans son orbite ou dès qu'un objet se déplace par rapport à nous, l'image formée sur la rétine change. Or, à l'exception près des mouvements de saccade des yeux qui sont des rotations brutales survenant entre $10^\circ/s$ et $1000^\circ/s$, ce changement est continu et est donc descriptible à tout instant par la vitesse instantanée de déplacement d'un point lumineux constituant l'image. On appelle flux optique le champ de ces vecteurs vitesses lorsque l'œil est fixe par rapport à

la tête. La Figure 29 donne un exemple de représentation d'un flux optique radial correspondant à un mouvement vers l'avant.

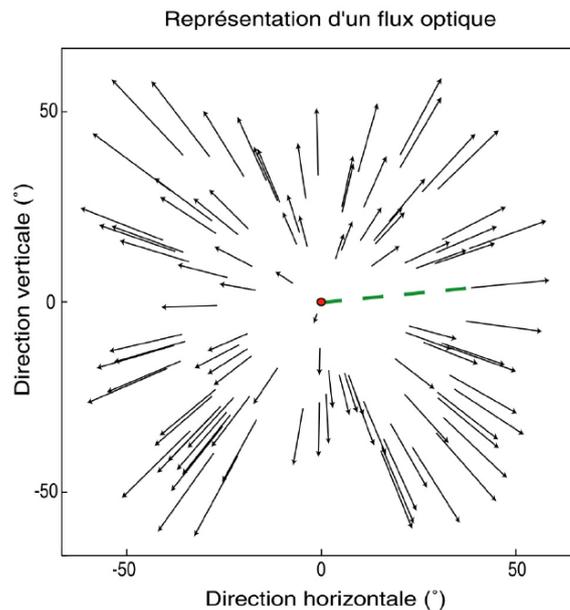
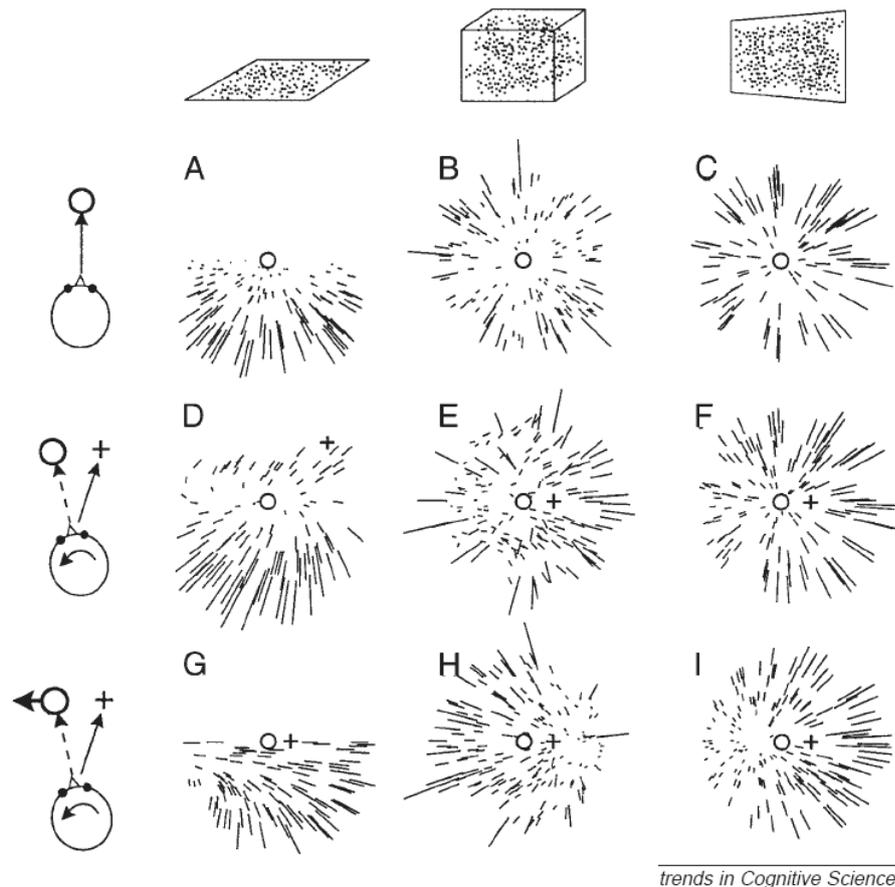


Figure 29 – Exemple de représentation d'un flux optique simulant un mouvement vers l'avant, d'après (Klam 2003).

Au centre (en rouge), le foyer d'expansion. Chaque vecteur de mouvement (lignes fléchées, en noir) est caractérisé par sa ligne de direction (ligne pointillée, en vert) le reliant au foyer d'expansion (en rouge au centre). Chacun de ces vecteurs représente le mouvement visuel provoqué par le déplacement d'un objet par rapport à l'observateur.

C'est Gibson (Gibson 1950) qui, le premier, remarque que le flux optique contient des informations pertinentes sur la locomotion. Sa structure dépend en effet du type de mouvement relatif entre l'observateur et l'environnement. Gibson montre alors que l'homme utilise le flux optique pour obtenir des informations comme la direction de son déplacement, sa vitesse propre ou encore le temps restant avant le contact avec un obstacle. Par exemple dans le cas simple d'un déplacement vers l'avant (flux radial Figure 29), les directions de tous les vecteurs de mouvement se croisent en un point remarquable appelé foyer d'expansion, que l'on interprète dans ce cas comme la direction du déplacement.

Ces indices de mouvement sont faciles à interpréter lorsque les yeux sont immobiles. Toutefois, cette situation reste très rare. Et le champ de vitesses formé sur la rétine devient alors la composition du champ de vitesses dû au mouvement de la tête relativement à l'environnement et du champ de vitesses induit par le mouvement de l'œil par rapport à la tête. Le champ alors obtenu est appelé flux rétinien, dont l'interprétation est beaucoup plus complexe que celle du flux optique. La Figure 30 donne des exemples de structures de flux rétinien.



trends in Cognitive Sciences

Figure 30 – Structures possibles de flux rétiniens, d'après (Lappe, Bremmer et van den Berg 1999).

Le flux rétinien perçu par un observateur en mouvement dépend de la direction du déplacement, la rotation de l'œil, et de la composition de l'environnement.

Les colonnes représentent différents environnements :

(A, D, G) un plan horizontal (« sol ») ;

(B, E, H) un volume de points aléatoires ;

(C, F, I) un plan vertical (« mur »).

Les lignes représentent différentes combinaisons de déplacement de l'observateur et de rotation de l'œil.

(A, B, C) Mouvement vers l'avant en l'absence de mouvement de l'œil.

(D, E, F) Mouvement vers l'avant (en direction de la croix) pendant que le regard (indiqué par un cercle) est dirigé vers un élément de l'environnement ;

(G, H, I) Mouvement vers l'avant (en direction de la croix) pendant que le regard suit une cible se déplaçant horizontalement (indiquée par un cercle). Cette cible n'est pas attachée à l'environnement, si bien que le mouvement de l'œil est découplé du mouvement de l'observateur. En G et I, le mouvement de l'œil s'effectue vers la gauche, en H vers la droite.

On peut observer dans la Figure 30 que l'interprétation du flux rétinien est en effet plus simple pour la première ligne car l'œil reste immobile par rapport à la tête. Les flux rétiniens A, B et C sont donc des flux optiques et on observe dans les 3 cas que le point singulier (foyer d'expansion) correspond à la direction du déplacement. Pour la deuxième ligne (D-F) et la troisième ligne (G-I), l'œil ne reste plus immobile pendant le mouvement. On s'aperçoit alors que lorsqu'un objet de l'environnement est fixé (D-F), le point singulier ne correspond plus à la direction du déplacement mais à la cible car c'est ce point qui est stabilisé sur la rétine. Si le point fixé n'est plus attaché à l'environnement (G-I), le point singulier peut même disparaître et le flux rétinien obtenu peut facilement être confondu avec celui d'une autre configuration. En G par exemple, le flux rétinien pourrait être facilement confondu avec celui perçu dans un virage.

Structure du flux optique

Pour mieux comprendre la structure du flux optique, et donc les informations qu'il est possible d'en retirer, nous allons voir qu'il est possible de le décomposer en « éléments simples ». Comme un mouvement dans l'espace peut se décomposer en un ensemble de translations et de rotations, on peut considérer le flux optique correspondant à ce mouvement comme la composition des flux optiques résultants des translations et rotations composant le mouvement (Longuet-Higgins et Prazdny 1980). De ce fait, en analysant les caractéristiques des flux optiques correspondant à une translation pure ou à une rotation pure, on peut alors déterminer les caractéristiques du mouvement global.

Dans le cas d'une rotation pure (Figure 31.a), l'image rétinienne tourne autour d'un point (R) correspondant à l'intersection de l'axe de rotation avec la rétine. La vitesse angulaire de cette rotation de l'image rétinienne correspond alors simplement à la vitesse de rotation du mouvement. Dans le cas particulier où l'axe de rotation est vertical, le mouvement sur la rétine est uniforme, c'est-à-dire que tous les points de l'image ont la même vitesse angulaire.

Dans le cas d'une translation pure (Figure 31.b), le flux optique possède une disposition radiale (comme dans la Figure 29) dont le pôle (Q) correspond à l'intersection de la direction de translation avec la rétine. Les vitesses des points de l'image rétinienne « fuient » ce pôle, qu'on appelle foyer d'expansion. De plus, la distribution des vitesses autour de ce foyer d'expansion va dépendre de la structure de l'environnement. De façon générale, plus un objet sera proche et plus sa vitesse sur l'image rétinienne sera importante. Et pour des objets situés à la même distance, plus leur excentricité par rapport à l'axe optique sera grande, et plus leur vitesse sur l'image rétinienne sera grande également.

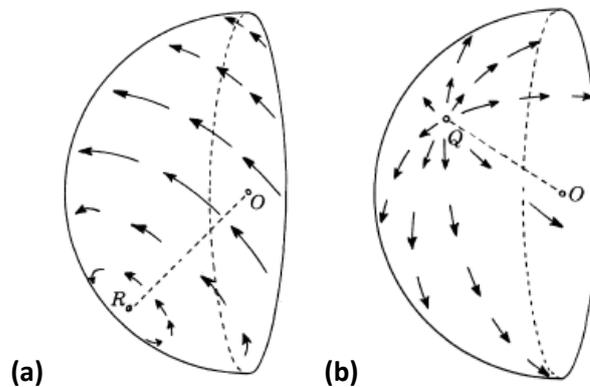


Figure 31 – Représentation du flux optique sur une rétine hémisphérique, d'après (Longuet-Higgins et Prazdny 1980).

O représente le centre optique de la rétine.

- (a) Flux optique dû à une rotation pure. R est l'intersection de l'axe de rotation avec la rétine.*
- (b) Flux d'expansion dû à une translation pure. Le foyer d'expansion Q est l'intersection de la direction de la translation avec la rétine.*

Enfin, les mouvements oculaires vont rajouter au flux rétinien une dernière composante qui est du même type que la composante correspondant à une rotation pure. Les mouvements oculaires sont en effet des mouvements de rotation de la rétine. Ainsi, le flux rétinien global correspondra à la composition des flux de rotation et de translation correspondants respectivement à la rotation et à la translation de la tête dans l'environnement, et d'un flux de rotation correspondant au mouvement oculaire. On notera

toutefois qu'on fait ici l'hypothèse que le monde est rigide, c'est-à-dire que les éléments de l'environnement ne se déplacent pas les uns par rapport aux autres.

3.1.2. Flux optique et vitesse perçue

Ambiguïtés du flux optique

Avant d'étudier la façon dont l'être humain traite le flux optique pour en tirer une information de vitesse, il convient de faire état des ambiguïtés qu'il présente. Tout d'abord, le flux optique ne permet pas de savoir si c'est le sujet qui se déplace dans l'environnement, ou si c'est l'environnement qui se déplace par rapport au sujet. On voit donc qu'il ne permet pas de répondre à la question pourtant simple : « Suis-je en mouvement ? ». Cette ambiguïté a souvent été mise en avant pour apporter une forme de validité aux simulateurs de conduite statiques.

Une autre ambiguïté du flux optique est celle concernant l'indétermination de la composante de la vitesse en profondeur. En effet, le flux optique est uniquement bidimensionnel car il est le résultat d'une projection sur une surface (la rétine). Ainsi, comme le montre la Figure 32, plusieurs vitesses peuvent avoir la même projection sur la rétine et ainsi être interprétées de la même manière.

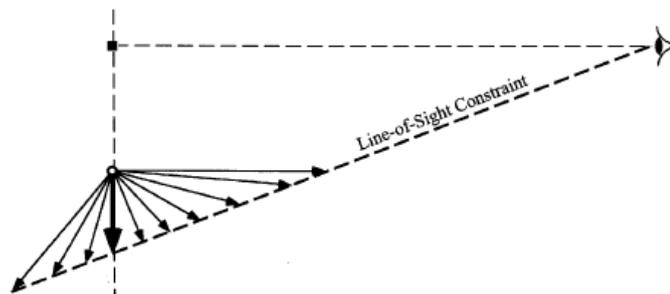


Figure 32 – Indétermination par le flux optique de la composante de mouvement en profondeur, d'après (Clifford, Beardsley et Vaina 1999).

Le point de l'espace représenté par un disque blanc a une vitesse représentée par la flèche en gras. Toutefois, l'observateur peut interpréter ce mouvement par n'importe quelle autre flèche dont l'extrémité est sur la droite appelée « contrainte de ligne de vue » (« Line-of-Sight Constraint »).

Enfin, le flux optique présente également une ambiguïté d'échelle. Comme le montre la Figure 33, les vitesses apparentes d'objets ayant des vitesses pourtant différentes peuvent être identiques. Tout comme l'indétermination de la composante en profondeur de la vitesse, l'ambiguïté d'échelle est due au fait que le flux optique est une projection en 2 dimensions d'un monde en 3 dimensions.

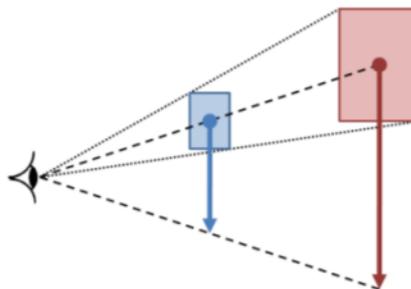


Figure 33 – Illustration de l'ambiguïté d'échelle du flux optique.

Le flux optique seul n'est pas suffisant pour distinguer les mouvements des 2 objets, pourtant de taille et de vitesses différentes.

On verra plus tard que d'autres indices visuels (notamment la parallaxe de mouvement) ainsi que le système vestibulaire sont à-même de lever les ambiguïtés du flux optique. De plus, (Gibson 1950) montre que le cerveau utilise des hypothèses fortes pour l'interprétation du flux optique, notamment celles de la rigidité et de la stabilité de l'environnement. En effet, en conditions écologiques (c'est-à-dire en conditions naturelles), l'environnement ne se déforme pas et c'est le sujet qui se déplace dans l'environnement, et non l'environnement qui se déplace par rapport au sujet. Ce sont donc prioritairement ces hypothèses qui sont utilisées en cas de doute.

Dans l'expérience de (Clifford, Beardsley et Vaina 1999) par exemple, les sujets sont soumis à un flux uniquement radial engendré par un plan face à l'observateur en translation vers l'observateur. De par les ambiguïtés du flux optique, le système visuel peut interpréter ce flux radial d'une infinité de manières, en jouant sur les combinaisons entre déplacement et dilatation du plan. A l'extrême, le plan peut être perçu comme fixe et se dilatant. A l'opposé, il peut être perçu comme rigide et se déplaçant vers le sujet. C'est cette deuxième interprétation qui a été le plus souvent effectuée par les sujets, montrant ainsi l'hypothèse faite a priori par le cerveau sur la rigidité de l'environnement.

Flux optique et vitesse en profondeur

Nous avons vu que flux optique présente plusieurs ambiguïtés, dont notamment une concernant la composante en profondeur de la vitesse d'un objet. Cependant, (Berthoz et Droulez 1982) montrent qu'avec certaines hypothèses sur la structure de l'environnement (notamment sa rigidité), il est possible d'extraire à partir d'un flux radial des informations sur la vitesse de l'observateur relativement à l'environnement. Ils considèrent l'exemple d'un flux radial généré par la translation à la vitesse V d'un observateur relativement à un environnement rigide (Figure 34). Si la projection sur la rétine d'un point situé à la profondeur D se situe à la distance ρ du foyer d'expansion, elle subit alors une dilatation $\dot{\rho}$ (dérivée temporelle de ρ) gouvernée par la loi $\frac{V}{D} \approx \frac{\dot{\rho}}{\rho}$.

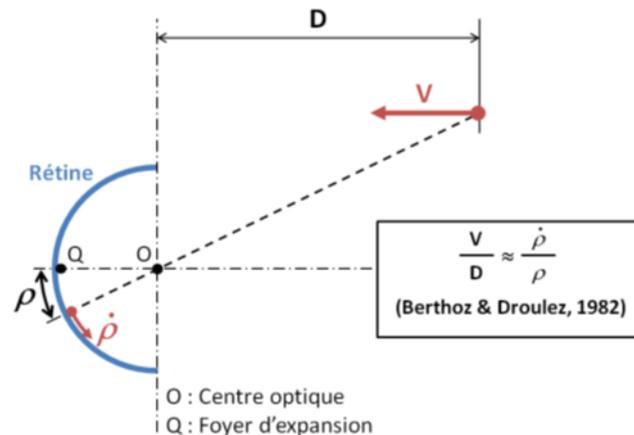


Figure 34 – Illustration de la loi proposée par (Berthoz et Droulez 1982) sur la perception de la vitesse à partir des données du flux optique.

Cette loi est indépendante de la position du point choisi. Le rapport $\frac{\dot{\rho}}{\rho}$ peut donc être considéré comme un invariant du flux optique. Lee (Lee 1976) avait par ailleurs été le premier à proposer que le conducteur puisse utiliser le « temps avant contact » (*TTC Time To Contact* ou *Time To Collision*) comme variable de régulation pour une tâche de freinage ou pour suivre un autre véhicule. Si on considère un point appartenant au véhicule précédant le

véhicule simulé ou plus simplement l'endroit auquel on veut stopper son véhicule, alors le « temps avant contact » n'est autre que $\frac{D}{v}$, en d'autres termes l'inverse du rapport $\frac{\dot{\rho}}{\rho}$ utilisé par (Berthoz et Droulez 1982).

On s'aperçoit donc qu'il est théoriquement possible d'extraire des informations de vitesse en profondeur à partir du flux optique. Cependant, il est pour cela nécessaire de résoudre les ambiguïtés qui lui sont inhérentes : deux éléments permettent d'interpréter correctement le flux optique. Premièrement, la préférence pour la perception d'un environnement rigide (hypothèse de constance de taille) permet de lever l'ambiguïté sur la composante de profondeur. Deuxièmement, la perception correcte de la profondeur de chaque point de l'environnement permet de déterminer la vitesse d'un point grâce à la relation donnée par (Berthoz et Droulez 1982). Comme nous le verrons par la suite, la vision binoculaire ainsi que la parallaxe de mouvement permettent de remplir cette deuxième condition sur la perception en profondeur.

Toutefois, on pourra remarquer que rien jusqu'à présent ne permet de lever l'ambiguïté d'échelle du flux optique. Et elle ne peut être résolue que par l'interaction avec d'autres indices fournissant l'échelle de l'environnement. Typiquement, en conditions réelles, ces indices peuvent être fournis par les systèmes vestibulaires et/ou proprioceptifs (Harris, Jenkin et Zikovitz 2000). D'autres travaux ont également montré qu'avec une simulation uniquement visuelle, l'humain est capable de discerner avec précision deux distances parcourues ((Bremmer et Lappe 1999), (Lappe, Jenkin et Harris 2007)). Or l'estimation de distance reposant sur une certaine intégration de la vitesse (Lappe, Jenkin et Harris 2007), l'humain est donc capable de calibrer l'information de vitesse qu'il perçoit en fonction de la taille supposée de l'environnement. L'ambiguïté d'échelle peut donc être a priori résolue uniquement avec des informations visuelles.

Hauteur du point de vue

Pour la même vitesse de déplacement dans le même environnement, il a été montré que la hauteur du point de vue fait varier le flux optique. La vitesse perçue est donc également impactée par ce paramètre (Larish et Flach 1990). Plus le point de vue est haut, plus l'intensité du flux optique diminue et donc plus la vitesse perçue sera faible. Pour quantifier ce phénomène, un indice : la « vitesse du flux optique global » (*GOFR*, *Global Optical Flow Rate*) est utilisé. Il correspond au rapport de la vitesse par la hauteur du point de vue.

Par exemple, un avion se déplaçant à 250 m/s à une altitude de 9000 m génèrera un *GOFR* de $0,028 \text{ s}^{-1}$. Un piéton se déplaçant à 0,8 m/s (environ 2,88 km/h) dont les yeux seraient à 1,6 m du sol génèrerait un *GOFR* de $0,5 \text{ s}^{-1}$, et aurait donc une sensation de vitesse plus importante que le pilote de l'avion de l'exemple précédent.

Dans le cadre de la simulation de conduite, une étude de (Panerai, et al. 2001) menée avec des conducteurs de poids lourds, montre un effet significatif de la hauteur du point de vue. Plus le point de vue est haut et plus les conducteurs roulent vite, tout en réduisant les distances de sécurité. La hauteur du point de vue a donc également une influence sur la perception des distances, comme nous le verrons par la suite.

« *Vitesse de bord* » (*Edge Rate*)

En conduisant, il est facilement remarquable que l'environnement lui-même peut également avoir une influence sur la perception de la vitesse. Par exemple, il n'est pas rare d'avoir une sensation de vitesse qui augmente lorsqu'on roule sur une route bordée d'arbres après avoir traversé une zone de rase campagne.

La « vitesse de bord » (*ER, Edge Rate*) a été définie par (Denton 1980) comme le nombre d'éléments texturés croisant un point de référence dans le champ visuel de l'observateur (comme par exemple le bord du pare-brise), par unité de temps. Son étude consistait à faire conduire les sujets sur une route possédant des marquages au sol de type « bandes transversales ». En faisant diminuer de façon exponentielle la distance séparant les bandes, Denton a mesuré une réduction significative de la vitesse.

Influences de la luminosité et du contraste

Enfin, la luminosité et le contraste ont également une influence sur le flux optique ((Snowden, Stimpson et Ruddle 1998), (Blakemore et Snowden 1999)), entraînant donc un changement de vitesse perçue. Dans leur étude, (Snowden, Stimpson et Ruddle 1998) soutiennent qu'en cas de brouillard (qu'ils simulent par une baisse de contraste), la vitesse perçue est plus faible, ce qui explique pourquoi les automobilistes rouleraient plus vite, malgré le danger plus important dû au manque de visibilité. Plus récemment, (Pretto, Vidal et Chatziastros 2008) remettent en cause cette explication. Ils soutiennent que la diminution de vitesse perçue en cas de brouillard ne proviendrait pas de la baisse de contraste, mais du fait que le brouillard réduit la visibilité de façon plus forte pour les éléments éloignés (dans la vision centrale) que pour les éléments proches (dans la vision périphérique).

3.1.3. Perception des distances

Nous avons vu qu'une des conditions nécessaires à l'estimation de la vitesse est la bonne perception des distances. La perception de la vitesse et la perception des distances sont donc deux notions liées. Nous allons voir que l'être humain possède plusieurs moyens de percevoir la profondeur et les distances égocentriques (Paillé 2005). Ces informations de profondeur peuvent se classer en 3 catégories : informations monoculaires, informations binoculaires et informations apportées par le mouvement (notamment la parallaxe de mouvement).

Informations monoculaires

Les informations monoculaires peuvent permettre d'avoir une rapide estimation de la profondeur d'un objet. Par exemple, lorsqu'un objet plus proche cache complètement ou en partie un objet plus éloigné, on parle alors d'occlusion. Ce phénomène ne permet pas d'obtenir une information de profondeur absolue, mais plutôt une information qualitative sur la profondeur relative des objets constituant la scène.

Les effets de perspectives (perspective géométrique, perspective linéaire, perspective aérienne, perspective/gradient de texture...) sont également des indices monoculaires permettant d'apprécier la profondeur. Ces effets sont associés au phénomène de constance de taille et de taille relative qui font que plus un objet est loin, plus sa taille

apparente est petite. L'effet du phénomène de constance de taille est par ailleurs d'autant plus marqué lorsque l'objet est connu.

Enfin, la hauteur du point de vue joue également un rôle important dans la perception de la profondeur. C'est Gibson (Gibson 1950) qui le premier rappelle que pour des objets situés sur un plan horizontal, plus leur distance est faible et plus ils sont perçus bas par rapport à l'horizon. Dans leur expérimentation, (Ooi, Wu et He 2001) montrent que le système visuel utilise la déclinaison angulaire par rapport à l'horizon (Figure 35) pour les jugements de distance : lorsqu'ils utilisent des prismes biaisant cette déclinaison angulaire, les sujets sous-estiment la distance.

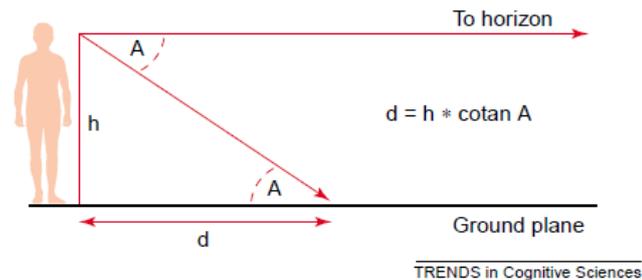


Figure 35 – Illustration de la déclinaison angulaire, d'après (Kemeny et Panerai 2003).

Le système visuel peut calculer la distance absolue (d) à un objet placé sur le sol, à partir de la hauteur des yeux (h) et la déclinaison angulaire par rapport à l'horizon (A). Mathématiquement, la distance est donnée par le produit de (h) par la cotangente de (A).

Les informations monoculaires sont souvent restituées en simulation de conduite et de façon plus générale en réalité virtuelle. Paillé (Paillé 2005) montre cependant qu'une limitation du champ de vision entraîne une sous-estimation de la perception des distances égocentriques. Cette sous-estimation est toutefois résorbée lorsque les mouvements de tête sont autorisés.

Informations binoculaires

En plus des informations monoculaires, l'être humain utilise des informations binoculaires pour la perception de la profondeur. L'accommodation en est une. Elle correspond à la « mise au point » de l'œil pour obtenir une image nette sur la rétine. Cette « mise au point » est obtenue par une déformation du cristallin, qui devra être plus ou moins importante en fonction de la distance de l'objet fixé. De plus, l'accommodation est faite de façon couplée entre les deux yeux. C'est pourquoi elle est généralement classée dans les informations binoculaires.

L'accommodation est accompagnée d'un mouvement de rotation des yeux qu'on appelle vergence, générant lui aussi des informations proprioceptives relatives à la distance de l'objet fixé. Ce mouvement de vergence des yeux permet que l'image de l'objet observé se projette sur les fovéas des deux yeux. Ainsi, plus l'objet fixé sera proche et plus la vergence des yeux sera importante.

Enfin, la disparité binoculaire (ou parallaxe binoculaire) est un des indices binoculaires les plus importants pour la perception de distances moyennes (inférieures à 20m (Cutting et Vishton 1995)). Elle correspond à la différence qui existe entre les images formées sur les rétines des deux yeux. Il est théoriquement possible de percevoir la profondeur uniquement avec cette information : on parle alors de stéréopsie. Cependant, son efficacité diminue avec la distance. En effet, plus un objet est éloigné et plus les axes de

vision deviennent parallèles et il n'y a alors plus de disparité. Toutefois en environnement virtuel, la construction des images stéréoscopiques se base essentiellement sur les indices de parallaxe binoculaire.

Parallaxe de mouvement

Les informations apportées par le mouvement constituent la dernière catégorie d'indices permettant à un sujet d'apprécier la profondeur. Nous avons déjà évoqué les informations amenées par le flux optique lors du déplacement longitudinal ou lors d'une rotation du sujet dans son environnement. Toutefois, lorsqu'un observateur effectue des mouvements de tête latéraux (vers la gauche ou la droite), il perçoit également des indices de parallaxe de mouvement. Ces indices contribuent à la perception de la structure tridimensionnelle de l'environnement (Figure 36).

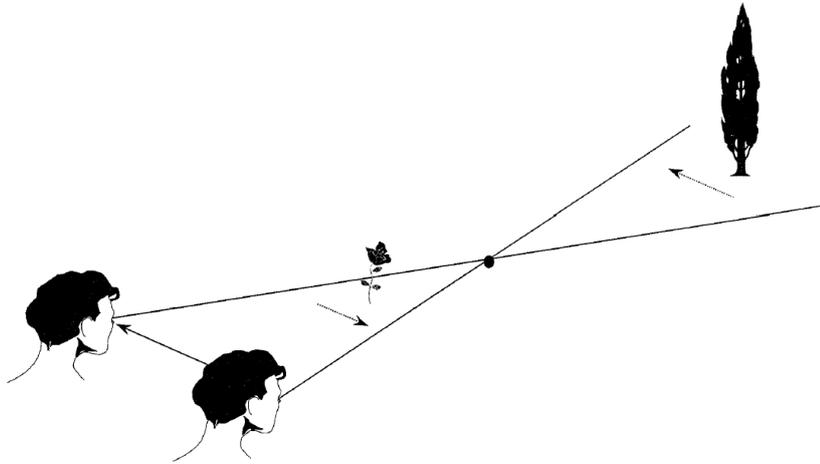


Figure 36 – Illustration du phénomène de parallaxe de mouvement, d'après (Guerraz, et al. 2001).

Lorsque le sujet se déplace latéralement dans une scène stationnaire tout en fixant un point de l'espace, les objets situés en avant de ce point se déplacent par rapport au sujet dans la direction opposée à son déplacement, et les objets situés en arrière se déplacent dans la même direction. L'amplitude de ce déplacement relatif dépend en particulier de l'éloignement des objets par rapport au point de fixation. Ce phénomène, appelé parallaxe de mouvement, constitue un indice important dans la perception des profondeurs et des distances.

Les mouvements de tête de l'observateur entraînent un déplacement relatif apparent des objets constituant l'environnement, qui est fonction de leur distance relative au point de fixation du regard. Plus un objet sera éloigné de ce point de fixation et plus son déplacement apparent sera ample. De plus, ce déplacement apparaîtra dans le sens opposé au mouvement de l'observateur si l'objet correspondant est situé en avant du point de fixation et dans le même sens si l'objet est situé en arrière de ce point.

La parallaxe de mouvement est ainsi un indice fort pour la perception des profondeurs relatives des différents éléments de l'environnement. Des études de (Rogers et Graham 1979) et de (Ono, Rivest et Ono 1986) ont également montré que si les informations de parallaxe de mouvement étaient couplées aux informations sensorielles de mouvement propre et de distance du point de fixation, elles pouvaient alors constituer un indice de perception de profondeur absolu.

Les informations procurées par la parallaxe de mouvement sont utilisées par le cerveau pour reconstruire la forme des objets observés et la structure tridimensionnelle d'une scène visuelle. On appelle ce phénomène le *structure from motion* (Longuet-Higgins et Prazdny 1980). Cependant, comme pour le flux optique cette information n'est pas

suffisante. Elle nécessite encore une fois de faire des hypothèses supplémentaires sur la rigidité de l'environnement.

Une étude de (Wexler, et al. 2001) montre que le système visuel utiliserait plutôt une hypothèse de stationnarité des objets de l'environnement lorsque le flux optique de parallaxe est généré par le mouvement de l'observateur. Dans son expérience les sujets passifs et actifs, même en voyant le même stimulus visuel, perçoivent la structure tridimensionnelle de façon différente. Cela tend à montrer la dépendance de cette perception à des signaux extrarétiniens.

3.2. Interactions multisensorielles pour la perception du mouvement propre

3.2.1. La vection

La vection, une illusion sensorielle ?

La vection est un puissant phénomène d'induction d'un mouvement propre par une stimulation visuelle. L'illusion du train est l'exemple le plus connu et le plus représentatif de ce phénomène. Il s'agit du cas où nous nous trouvons arrêtés dans un train à quai et que le train voisin démarre : nous pouvons alors avoir l'impression que c'est notre train qui est en mouvement. Il est souvent considéré que la vection est une illusion perceptive, car un mouvement propre est perçu alors que le sujet est objectivement immobile. La vection serait alors une erreur de perception commise par le cerveau. Toutefois, ce point ne fait pas l'unanimité.

Gibson a été le premier à remettre en cause le côté illusoire de la vection. Dans son approche écologique de la perception visuelle (Gibson 1979), il explique qu'un observateur mesurant par sa rétine un mouvement de l'environnement relativement à soi se perçoit nécessairement en mouvement, car en conditions écologiques (c'est-à-dire naturelles) l'environnement est toujours stationnaire. L'illusion produite sur un simulateur de vol ou de conduite n'est donc jamais une illusion en situation écologique.

De plus, selon (Dichgans et Brandt 1978), les noyaux vestibulaires sont également activés pendant la vection (voir 2.5.1), suivant la même évolution temporelle que la sensation de mouvement propre. Comme nous avons également vu qu'en vitesse établie le signal renvoyé par le système vestibulaire s'éteint (voir 2.3), ce constat sur l'activation des noyaux vestibulaires a amené l'hypothèse que la vection joue un rôle complémentaire à l'information vestibulaire. Comme nous le verrons par la suite, la vection n'apparaît pas de façon instantanée : le temps de latence de son apparition vient alors renforcer cette hypothèse de complémentarité fréquentielle entre la vision et le système vestibulaire.

Etablissement de la vection

La vection n'est pas un phénomène instantané : elle apparaît avec une latence pouvant varier de 1 à 10 secondes. Puis elle devient exclusive au bout de 5 à 15 secondes : on parle alors de vection saturée. (Dichgans et Brandt 1978) rapportent même que la vection peut durer jusqu'à 30 secondes après la disparition du stimulus par extinction de la lumière. Ces valeurs ont été obtenues en vection circulaire (mouvement de rotation) en utilisant notamment des cylindres texturés en rotation autour du sujet. Néanmoins, (Berthoz, Pavard et Young 1975) trouvent des résultats similaires en vection linéaire (mouvement de translation). La Figure 37 donne plusieurs exemples d'appareils pour l'étude des vections linéaires et circulaires.

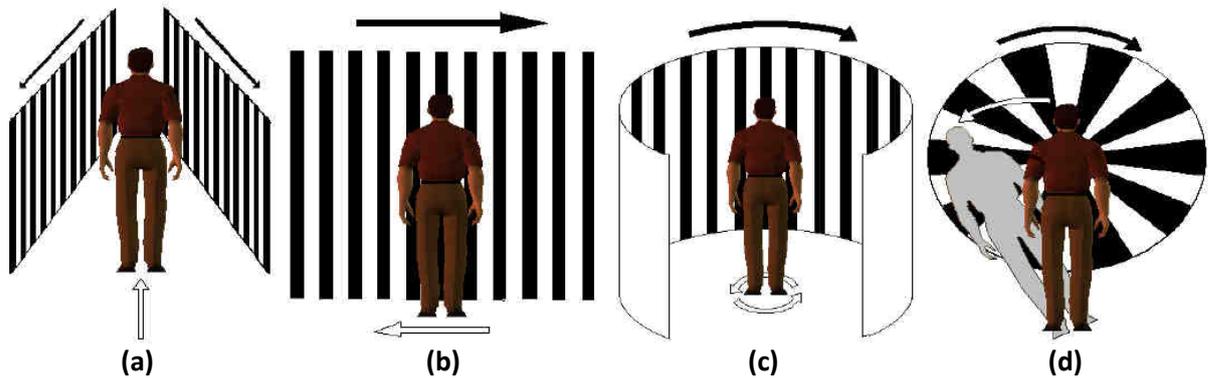


Figure 37 – Exemples d'appareils pour l'étude des vections linéaires (a et b) et circulaires (c et d), d'après (Neimer 2002).

Ces vections sont induites par des flux visuels en translation longitudinale (a) et latérale (b), ainsi qu'en rotation en lacet (c) et en roulis (d). La flèche pleine indique la direction des mouvements réels de la scène visuelle et la flèche évidée la sensation de mouvement du sujet. Dans chaque cas, le sujet reste immobile.

En 1978, (Borah, Young et Curry 1978) résumant les principales caractéristiques de la vection :

- L'établissement de la vection est dominé par la vision périphérique. Nous reviendrons sur ce point par la suite.
- La présence d'un objet stationnaire situé à l'avant du stimulus n'affecte pas la vection et peut même la renforcer. A l'inverse, si l'objet stationnaire est situé en arrière-plan du mouvement visuel, la vection peut être annihilée.
- La latence du phénomène de vection dépend aussi bien du contexte que de l'individu.
- La sensation de mouvement propre induite par la vection se rapproche de celle induite par une stimulation vestibulaire à la différence qu'aucune accélération n'est perçue lors d'un changement de vitesse.
- La vection est fortement influencée par la présence d'informations vestibulaires.

Concernant ce dernier point, il a été montré que lorsque la stimulation vestibulaire est cohérente avec le mouvement visuel, le temps de latence de la vection ainsi que le temps d'établissement de la vection saturée sont réduits. Ces résultats ont été obtenus aussi bien en vection circulaire (Groen, Howard et Cheung 1999) qu'en vection linéaire (Berthoz, Pavard et Young 1975). Ils montrent ainsi l'importance de la restitution inertielle en simulation de conduite. En revanche, lorsque la stimulation vestibulaire est en conflit avec le mouvement visuel, la vection circulaire ne semble peu ou pas affectée (Groen, Howard et Cheung 1999) alors que la détection et l'identification de la direction du mouvement semblent fortement dégradés en vection linéaire (Berthoz, Pavard et Young 1975). Par ailleurs, des études ((Pavard et Berthoz 1977), (N. Filliard 2009)) montrent que la présentation d'une stimulation vestibulaire peut entraîner l'illusion d'un ralentissement (voir même d'un arrêt) du flux visuel.

La sensation de mouvement propre résulte donc de l'interaction entre un stimulus visuel et un stimulus inertiel. De plus, elle présente des caractéristiques optimales lorsque ces deux stimuli sont en accord. Un conflit résulte soit en une dégradation de la vection, soit en un choix d'une modalité dominante. Nous verrons par la suite que les conflits peuvent également provoquer l'apparition du mal du simulateur.

Concernant l'établissement de la vection, des études récentes ((Palmisano et Chan 2004), (Wright, DiZio et Lackner 2006)) montrent l'influence de facteurs cognitifs. Dans ces études, les résultats indiquent qu'une connaissance a priori de la possibilité d'un déplacement physique cohérent avec le mouvement visuel améliore la crédibilité du mouvement, mais n'a cependant aucun effet sur la latence de la vection. Dans le même ordre d'idées, (Riecke, et al. 2006) montrent que l'utilisation d'un environnement photo-réaliste favorise la diminution de la latence.

Influences de la vision centrale et de la vision périphérique

De nombreuses études ont été menées sur les influences de la vision centrale et de la vision périphérique sur la vection. Plusieurs théories émanent de ces études, que (Bardy, Warren et Kay 1999) proposent de classer en 3 catégories : hypothèse de dominance périphérique, hypothèse de sensibilité fonctionnelle et hypothèse d'invariance rétinienne.

La première hypothèse, l'hypothèse de dominance périphérique, est celle de (Dichgans et Brandt 1978), qui postule simplement que les stimulations en vision périphérique seraient prépondérantes dans la perception du mouvement de soi, par opposition aux informations visuelles centrales impliquées majoritairement dans la perception du mouvement d'objet. Cette hypothèse est la plus répandue de nos jours, et trouve également des fondements dans l'étude anatomique de l'œil (voir 2.2).

La deuxième hypothèse, défendue par (Warren et Kurtz 1992) postule que l'influence des visions centrale et périphérique sur le mouvement de soi dépende en fait de leur sensibilité fonctionnelle, à savoir que la rétine périphérique serait plutôt sensible au flux lamellaire alors que la vision centrale serait préférentiellement sensible au flux radial.

Enfin, la troisième hypothèse va à l'encontre des deux premières en postulant quant à elle que la vision centrale et la vision périphérique ont un rôle équivalent dans la perception du mouvement. Par exemple, (Crowell et Banks 1993) montrent à l'aide de flux radiaux et lamellaires présentés à différentes excentricités, que les visions centrales et périphériques ont la même sensibilité à la perception du mouvement. (Post 1988) montre quant à lui que pour de la vection circulaire, l'excentricité n'a pas d'influence sur la vection tant que la surface du stimulus reste constante. (Tarita-Nistor, et al. 2006) trouvent des résultats similaires pour de la vection linéaire, mais seulement si un point de fixation est rajouté sur le stimulus. En l'absence de ce point de fixation, la vision périphérique a, selon les auteurs, plus d'effet que la vision centrale pour induire le phénomène de vection.

Au vu de ces différentes hypothèses sur les influences respectives de la vision centrale et de la vision périphérique sur la vection, il semble peu évident de conclure. De plus, le manque de consensus aussi bien au niveau de la définition même de ces champs de vision en termes d'amplitude, qu'au niveau des paradigmes expérimentaux utilisés ne permet pas de comparer facilement les différents résultats expérimentaux. Toutefois, on pourra retenir qu'une stimulation de la vision centrale sous certaines conditions peut engendrer le phénomène de vection et que la stimulation de la vision périphérique semblerait favoriser l'émergence de la sensation de mouvement propre.

3.2.2. Distinction entre mouvement propre et mouvement d'objets

Nous avons évoqué précédemment le fait que l'œil ne serait pas capable à lui seul de faire la distinction entre le déplacement dans un environnement stable et le déplacement de tout l'environnement par rapport à lui. L'œil ne véhicule en effet qu'une information de déplacement relatif. Pourtant, comme nous venons de le voir avec la vection, cette distinction est faite et le cerveau parvient à faire émerger la sensation d'un environnement stable pendant le mouvement propre. Comment y parvient-il ? La littérature propose plusieurs théories pour répondre à cette question centrale pour la simulation de conduite.

Théorie de la perception directe

Selon cette approche théorique, proposée initialement par Gibson ((Gibson 1966), (Gibson 1979)), la séparation des composantes de mouvement propre et du mouvement des objets dans le flux rétinien peut être réalisée à partir des seules informations que ce flux contient. C'est l'hypothèse dominante sur laquelle se fondent la majorité des études sur la perception visuelle du mouvement que nous avons vues dans la partie précédente.

La théorie de la perception directe fait partie de la théorie écologique de la perception que nous avons déjà évoquée précédemment, qui explique la perception dans les conditions les plus naturelles possibles. En particulier, une hypothèse forte de la théorie écologique est la stationnarité de l'environnement. Le monde environnant ne peut donc jamais, par hypothèse, être perçu comme étant en mouvement.

Selon cette théorie, le flux optique suffit pour percevoir le mouvement. Des développements mathématiques ont montré qu'il était en effet possible de construire un algorithme analysant le flux optique afin d'en déduire le mouvement propre ou le mouvement des objets, en ajoutant quelques hypothèses simplificatrices (Longuet-Higgins et Prazdny 1980). Les informations vestibulaires et les informations proprioceptives des mouvements oculaires ne sont pas utilisées dans cette théorie. Cet aspect multi-sensoriel n'a toutefois pas été le plus investigué, les recherches qui ont suivi se focalisant principalement sur l'analyse du flux optique.

On notera que cette théorie confère une certaine forme de validité aux simulateurs de conduite statiques (voir Chapitre 4), les conducteurs étant effectivement parfaitement capables de piloter leur véhicule virtuel uniquement à partir des indices visuels du mouvement.

Théories inférentielles

Il existe en fait deux théories inférentielles : la théorie afférente et la théorie efférente. Cependant, toutes deux se basent sur le postulat que l'information visuelle n'est pas suffisante à elle seule pour résoudre les questions liées à la perception du mouvement. Le cerveau utiliserait donc toutes les informations disponibles pour les utiliser dans un processus semblable à un raisonnement permettant de construire la sensation de mouvement correspondant le plus vraisemblablement à la réalité physique.

Dans la théorie afférente, le cerveau utilise, en plus des informations rétinienne, les informations proprioceptives des muscles oculaires. Ces dernières correspondent à l'étirement des muscles et permettent donc en théorie de connaître la position de l'œil dans son orbite. Cependant, le décalage temporel des informations proprioceptives par rapport

aux informations rétinienne font qu'elles sont souvent considérées comme inefficaces. Il faut en effet attendre que les commandes motrices des mouvements oculaires parviennent aux muscles, que le mouvement soit effectué, que ce mouvement soit mesuré par les capteurs proprioceptifs qui finalement renvoient l'information au cerveau.

Cependant les résultats expérimentaux supportent mieux la théorie efférente, dans laquelle l'information sur le mouvement des yeux n'est pas extraite des informations proprioceptives des muscles oculaires, mais de la copie des commandes motrices (ou copie d'efférence) de ces muscles (von Holst et Mittelstaedt 1950). Ce signal permet alors de différencier facilement les composantes du flux rétinien dues au mouvement de l'œil de celles dues aux mouvements de l'environnement. Une expérience simple est classiquement citée afin de prouver l'existence d'un tel signal. En fermant un œil et en appuyant doucement sur le blanc de l'autre œil, le monde paraît soudainement instable, car le mouvement de translation de l'œil (sur lequel on appuie) n'est pas un mouvement volontaire.

Théorie du signal de référence

Bien que la théorie de la perception directe et les théories inférentielles soient en conflit, notamment sur la manière dont le cerveau fait émerger la sensation de mouvement des différentes informations sensorielles, on trouve pourtant des résultats expérimentaux supportant chacune de ces théories. Un consensus est en général trouvé en considérant que toutes les informations sensorielles sont utilisées, mais différemment suivant le contexte. La théorie du signal de référence (Wertheim 1994) est une théorie permettant d'unifier les théories de la perception directe et inférentielles.

La Figure 38 donne le modèle fonctionnel de la théorie du signal de référence. Les flèches noires épaisses correspondent au processus proposé par la théorie inférentielle expliquant la perception du mouvement des objets. Les flèches plus fines correspondent aux interactions visio-vestibulaires (vection, fusion visio-vestibulaire) typiques considérées par les modèles de perception du mouvement propre. Enfin, les flèches grises sont spécifiques au modèle proposé par Wertheim, qui tente ainsi de connecter ces deux aspects de la perception du mouvement, le plus souvent étudiés séparément.

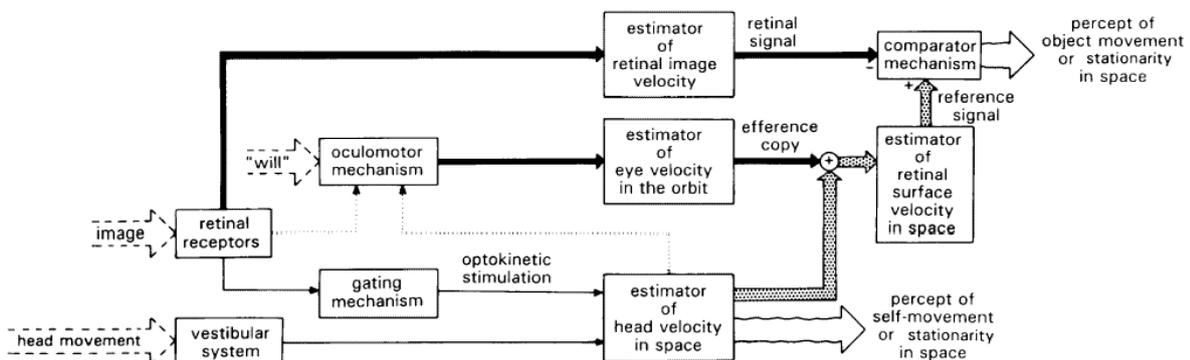


Figure 38 – Modèle fonctionnel décrivant la construction du signal de référence et des percepts de mouvement propre et de mouvements d'objet dans l'espace, d'après (Wertheim 1994).

3.2.3. Conflits sensoriels et mal du simulateur

En simulation de conduite, les conflits sensoriels (entre par exemple le mouvement visuel et la stimulation vestibulaire) sont nombreux. On peut même les considérer comme inévitables car un simulateur n'a pas la possibilité physique de déplacer le conducteur de la même façon que le ferait un véhicule en conditions réelles. Ces conflits sensoriels peuvent provoquer dans certains cas le mal du simulateur, qui est un cas particulier de cinétose.

Les cinétoses sont définies par un ensemble de réactions physiologiques se présentant sous la forme de symptômes cliniques (maux de têtes, nausées, pâleur etc.) largement ressentie dans la population générale soumise à des mouvements, physiques ou virtuels (Reason et Brand 1975). Il existe plusieurs variantes de cinétoses, comme le mal de mer, de l'air ou en voiture. Le mal du simulateur est une de ces variantes, résultant d'une exposition à un mouvement virtuel.

Les symptômes du mal du simulateur recourent en partie ceux identifiés dans le mal des transports classiques. (Kennedy, et al. 1993) ont développé le questionnaire *SSQ (Simulator Sickness Questionnaire)* largement utilisé dans les études relatives au mal du simulateur. Ce questionnaire liste exactement 16 symptômes : malaise général, fatigue, mal de tête, fatigue oculaire, difficulté d'accommodation, hyper salivation, sueur, nausée, difficulté de concentration, sensation de tête lourde, vision floue, étourdissement yeux ouverts, étourdissement yeux fermés, vertige, réveil de la zone gastrique et vomissement.

Théorie explicative du mal du simulateur

La théorie explicative du mal du simulateur la plus souvent rencontrée est celle du conflit sensoriel (Reason et Brand 1975). Selon cette théorie, la cause du mal du simulateur serait le conflit perceptif dû aux différences entre les informations fournies par les systèmes sensoriels relatifs à la perception du mouvement. Le mal du simulateur serait donc le résultat d'un conflit que le cerveau ne peut pas résoudre.

Les conflits sont généralement classés en 2 catégories : les conflits intermodaux (entre le système visuel et le système vestibulaire par exemple) et les conflits intra-modaux (entre les canaux semi-circulaires et les organes otolithiques du système vestibulaire par exemple). Ces catégories se décomposent elles-mêmes en 2 types de conflits : soit les deux signaux sont présents et fournissent des informations contradictoires, soit un des deux signaux est absent. Dans ce deuxième cas, c'est principalement l'attente d'un signal qui provoque le conflit et donc la cinétose.

Cette théorie des attentes a une forte valeur explicative et peut rendre compte du lien étroit qui existe entre mal du transport et mal du simulateur. Ainsi dans le cas d'un simulateur de conduite statique par exemple (voir Chapitre 4), les informations de mouvement fournies par le système visuel entreraient en conflit avec l'absence d'informations vestibulaires concordantes, puisque le conducteur ne bouge pas réellement. Or compte-tenu du fait qu'en conduite réelle, le conducteur a accès conjointement à ces deux sources d'informations, il s'attend naturellement à cette perception, dès lors qu'il s'assoit derrière un volant pour conduire. Et c'est cette différence entre ce qui est attendu et ce qui est perçu qui créerait le conflit contribuant à induire le mal du simulateur.

Ce modèle des attentes a évolué, comme par exemple avec les modèles de (Oman 1990) et de (Droulez et Darlot 1989), afin de prendre en compte la spécificité des différents

capteurs sensoriels. D'autres études plus récentes, comme celle de (Groen et Bos 2008), cherchent à déterminer l'influence de la fréquence du conflit, afin notamment de pouvoir prédire de façon plus certaine l'apparition du mal du simulateur.

3.3. Conclusion

La perception du mouvement est un phénomène complexe qui émane de plusieurs modalités sensorielles (modalités visuelle, vestibulaire et proprioceptive entre autres) dont les signaux renvoyés ont été étudiés au Chapitre 2. La modalité visuelle semble au premier abord être celle fournissant le plus d'informations, aussi bien sur la perception de la direction du mouvement que sur la perception de la vitesse de déplacement ou encore sur la perception de la profondeur.

Le flux optique, correspondant au changement de l'image rétinienne au cours du temps, semble notamment être une information permettant de déterminer la direction du déplacement ainsi que sa vitesse. Certains auteurs, suivant la théorie de la perception directe (ou théorie écologique de la perception) proposée par Gibson (Gibson 1979) soutiennent même que le flux visuel peut suffire pour percevoir à la fois son déplacement dans l'environnement et le déplacement d'objets dans ce même environnement.

Un mouvement visuel peut en effet provoquer à lui seul une sensation de mouvement propre : on appelle ce phénomène la vection. Cependant, des études montrent que la vection n'émanerait pas de la seule modalité visuelle, mais plutôt d'une interaction multisensorielle, notamment entre les modalités visuelle et vestibulaire. La vection viendrait alors compenser les déficiences du système vestibulaire, ce dernier étant inefficace à vitesse constante.

Cet aspect multisensoriel de la perception du mouvement n'est pas pris en compte dans la théorie de la perception directe. Les ambiguïtés du flux optique, notamment concernant la perception des profondeurs et la stabilité de l'environnement, seraient résolues par le cerveau grâce à des hypothèses simplificatrices. D'autres théories, comme les théories inférentielles ou la théorie du signal de référence, présentent une approche multisensorielle prenant en compte les informations provenant de tous les capteurs de mouvement du corps.