

DÉTECTION

Le but de l'extraction de primitives est de passer de l'espace initial des données à un espace de représentation compacte et riche en information. Des moyens comme la morphologie mathématique, la détection des contours et de coins, ou les courbures sont utilisés pour caractériser les propriétés locales de l'espace des données initiales. Selon l'application finale, la fiabilité d'une représentation de primitives est fortement rattachée à sa stabilité pour des variations liées aux transformations de l'espace, à la résolution, au bruit et aux occultations. En particulier, les courbures de la surface renseignent sur la forme de l'objet et sont robustes à des variations d'angles de vues. Dans cette thèse, pour caractériser la forme des objets dans des vues partielles, nous adoptons une approche de détection de points d'intérêt 3D basée sur les courbures de la surface. Dans ce chapitre, nous exposons, dans une première partie, une sélection de méthodes de détecteurs inspirés de la littérature et nous introduisons notre principale contribution. Dans une deuxième partie, nous mettons l'accent sur notre apport à travers les comparaisons avec les détecteurs présentés. Les notions de bases utilisées dans ce chapitre ont été présentées dans les deux sections précédentes.

3.1. Principes des détecteurs testés

3.1.1. Seuillage des valeurs de SI

Le détecteur, que nous avons nommé « SI », est un détecteur à échelle fixe proposé par (Chen, et al., 2007) basé sur un seuillage des valeurs de l'indice de forme (Équation 1.13) des points. Les points d'intérêt sont sélectionnés aux points ayant une valeur SI localement extremum par rapport aux SI de leur voisinage. Un point p est considéré comme point d'intérêt si son SI_p est un extremum local des SI de son voisinage et est significativement plus grand ou plus petit que la moyenne des SI sur le voisinage. Mathématiquement, SI_p doit satisfaire l'Équation 3.1:

$$\begin{cases} SI_p = \min(SI_k); k \in \text{voisinage} \text{ et } SI_p \leq (1 - \beta) \times \mu \\ SI_p = \max(SI_k); k \in \text{voisinage} \text{ et } SI_p \geq (1 + \alpha) \times \mu \end{cases} \quad 0 \leq \alpha, \beta \leq 1$$

Équation 3.1

Dans cette équation, μ est la moyenne des indices de formes sur les valeurs voisines. Notons que dans l'Équation 3.1, les paramètres α et β contrôlent le nombre de points d'intérêt sélectionnés. La stabilité au bruit et aux transformations de la surface des résultats d'une classification basée sur SI, est notable par rapport à une classification basée uniquement sur les courbures gaussiennes et moyennes. De plus, contrairement à ces deux derniers estimateurs dépendant de seuils pour détecter les valeurs de courbure proches de zéro, la formulation de SI sépare les types de surfaces par des seuils fixes. Avec

ce détecteur, la sélection de PIs se fait sur les zones convexes (les maximums en SI) et concaves de la forme (les minimums en SI). Dans Figure 3-1, les pixels clairs représentent les formes convexes et les pixels sombres indiquent les formes concaves.

Dans le travail proposé par (Chen, et al., 2007), le calcul des valeurs de SI se fait sur l'image de profondeur. Nous proposons de faire ce calcul directement sur le nuage de point en passant par un maillage qui s'oriente par l'organisation 2D de l'image de profondeur.

Dans la sélection de PIs, les valeurs de SI du voisinage sont pris en compte ce qui permet de faire une meilleure sélection basée sur la géométrie locale de la forme. La taille du voisinage reste un paramètre décisif pour cette approche créant un compromis entre l'efficacité de la détection et le temps de calcul.

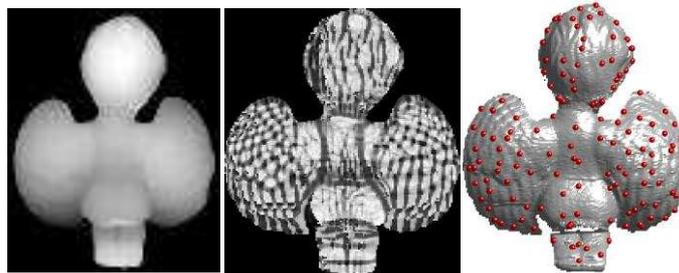


Figure 3-1-Gauche-Image de profondeur du modèle d'un ange - Milieu- Image de l'indice de forme en niveau de gris-Droite- Les PIs détectés.

3.1.2 Sélection des Points Harris 3D

Après avoir associé une valeur de l'opérateur de Harris à chaque sommet, les auteurs (Sipiran, et al., 2010) proposent deux façons de sélectionner les points d'intérêt. En premier, les maximums locaux des réponses de Harris h sont extraits, ce qui correspond aux sommets v vérifiant la condition suivante :

$$h(v) > h(w), \forall w \in \text{ring}_1(v)$$

Le ring_1 correspond aux sommets directement liés au sommet v dans le maillage. Ensuite, deux approches sont proposées pour sélectionner l'ensemble des PIs finaux :

- **Sélection des points ayant la plus grande réponse de Harris:** Où il est question d'extraction d'une portion constante de points d'intérêt. L'extraction, seulement, des points de grande saillance peut pénaliser certaines parties de l'objet.

- **Regroupement des PIs (Clusters).** Cette approche permet d'obtenir une distribution des PIs sur toute la surface de l'objet. L'ensemble des points présélectionnés selon la valeur de l'opérateur de Harris sont ordonnés dans un ordre décroissant puis, l'algorithme de regroupement (clustering) suivant est leur appliqué :

Algorithm 1 Interest Points Clustering

Require: Set P of pre-selected interest points in decreasing order of Harris operator value

Ensure: Final set of interest points

```

1: Let  $Q$  be a set of points
2:  $Q \leftarrow \emptyset$ 
3: for  $i \leftarrow 1$  to  $|P|$  do
4:   if  $\min_{j \in [1, |Q|]} \|P_i - Q_j\|_2 > \rho$  then
5:      $Q \leftarrow Q \cup \{P_i\}$ 
6:   end if
7: end for
8: Return  $Q$ 
  
```

Figure 3-2 Algorithme de regroupement des points d'intérêt (clustering)

La valeur de ρ peut être une fraction de la diagonal de la boîte englobant de l'objet et sa valeur détermine le nombre de PI final. Nous avons utilisé le code de Harris 3D donné par l'auteur pour notre comparaison des performances des détecteurs. Les paramètres réglables par l'utilisateur de cet algorithme concernent la taille du voisinage (nombre de rings), le paramètre K pour le calcul de la réponse de Harris des points, le nombre de rings pour la détection des maxima et le type de l'algorithme de sélection final (Fraction ou Clustering) avec la valeur du paramètre de sélection. Pour la suite, nous allons appeler le Harris qui utilise la sélection par Fraction « Harris_fract » et celui qui utilise le regroupement « Harris_clust ».

La Figure 3-3 représente un exemple du résultat de la sélection des points d'intérêts avec les deux détecteurs Harris_fract et Harris_clust.

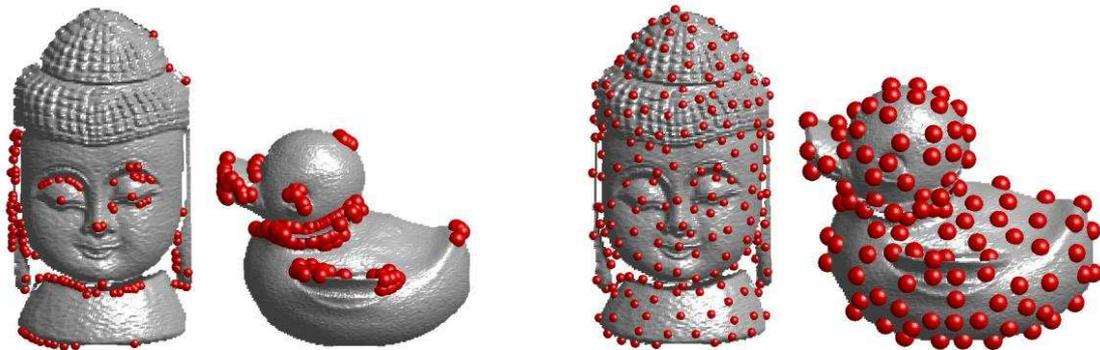


Figure 3-3- A gauche -Sélection des points avec la plus grande réponse d'Harris et - A droite - Sélection des points Harris avec regroupement.

3.1.3 SURF 3D

Un autre détecteur de l'état de l'art faisant une extension d'une méthode connue en 2D est le SURF 3D introduit par (Knopp, et al., 2010). Différemment aux autres approches multi-échelles, le 3D SURF construit des espaces d'échelle à partir de la version **voxélisée** du maillage original. Une mesure de saillance S est définie pour chaque bin de la grille cubique sur plusieurs échelles (trois octaves). S

correspond à la valeur absolue du déterminant de la matrice Hessienne H des dérivées gaussiennes secondes L , déduites des moyennes des box de filtrage.

$$S(\mathbf{x}, \sigma) = |H(\mathbf{x}, \sigma)| = \left| \begin{pmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) & L_{xz}(\mathbf{x}, \sigma) \\ L_{yx}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) & L_{yz}(\mathbf{x}, \sigma) \\ L_{zx}(\mathbf{x}, \sigma) & L_{zy}(\mathbf{x}, \sigma) & L_{zz}(\mathbf{x}, \sigma) \end{pmatrix} \right|$$

Équation 3.2

Les valeurs propres de la matrice H ne sont tous de même signe, ce qui correspond à des points détectés sur des zones sous forme de blobs et de selles. La sélection finale des PIs est faite avec un algorithme de suppression non-maximale. Chaque point extrait lui est associée une échelle.

Dans notre étude comparative, nous avons utilisée la version (exécutable) fournie par les auteurs et qui calcule le détecteur et le descripteur à la fois. Les paramètres fixés par l'utilisateur sont : la dimension du cube englobant l'objet, la distance entre les bords du cube et le modèle, et le seuil utilisé par la suppression non-maximale. Nous remarquons dans l'exemple donné dans la Figure 3-4, que les positions des PIs SURF ne sont pas forcément sur la surface de l'objet, étant donné qu'elles correspondent à des centres de voxels.

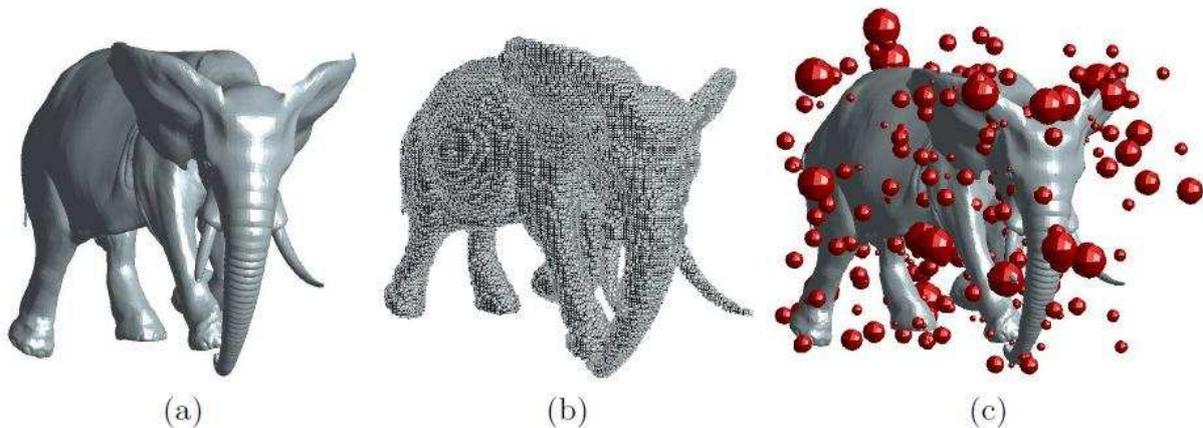


Figure 3-4 Illustration des PIs détectés par le SURF 3D. La forme (a) est voxélisée en une grille cubique (b). Les PIs détectés (c), et représentés par des sphères dont le rayon illustre l'échelle. (Knopp, et al., 2010)

3.1.4 Facteur de qualité

Le troisième détecteur implémenté est basé sur une mesure de qualité du PI introduite par (Mian, et al., 2009) qui l'ont utilisée pour classer les PIs détectés. En chaque point, nous associons la mesure de qualité Q_k que nous rappelons sa formulation:

$$Q_k = \frac{1000}{r^2} \sum |K| + \max(100K) + |\min(100K)| + \max(10k_p^1) + |\min(10k_p^2)|; \quad K = k_p^1 k_p^2$$

Équation 3.3

Dans cette équation, $K = \kappa_1 \kappa_2$ est la courbure gaussienne et $k_p^1 = \kappa_1$ et $k_p^2 = \kappa_2$ sont les courbures principales.

Les fonctions de sommation, maximum et minimum sont calculées sur le voisinage du point. Le pouvoir descriptif des valeurs de courbure étant indépendant de leur signe, ce sont les valeurs absolues qui sont utilisées dans la formule, afin d'éviter que des contributions de signes opposés ne s'annulent. Les valeurs des constantes multiplicatives sont choisies empiriquement pour donner le poids approprié à chaque terme. Pour des surfaces lissées, les valeurs des courbures principales sont en général plus petites que 1 et comme la courbure gaussienne est le produit des courbures principales, elle est alors d'ordre 10 fois plus petite que les courbures principales. Cette relation a été confirmée expérimentalement par (Mian, et al., 2009). Ainsi, le terme multiplicatif pour K est choisi 10 fois plus grand que les courbures principales. Les données d'apprentissage ont montré que le premier terme de l'équation est 100 fois plus petit que les courbures principales. En effet, la valeur de la courbure gaussienne, pour la plupart des points sur la surface lissée, est faible et qui une fois moyenné sur la surface donne une valeur qui est 10 fois plus petite que les valeurs maximum et minimum absolu de la courbure gaussienne. Pour finir, tous les termes de l'équation sont multipliés par 10 pour avoir une mesure de qualité Q toujours supérieure à 1.

Après le calcul des valeurs du facteur qualité de tous les points du nuage, nous les ordonnons selon leur valeur Q et les groupons à l'aide de l'algorithme 1 de clustering (Figure 3-2). L'image de droite dans la Figure 3-5 illustre la carte des valeurs du facteur qualité des points du modèle de l'ange. Les pixels les plus clairs correspondent aux valeurs de Q les plus importantes et sont, en effet, situés aux zones les plus descriptives (à fortes variations locale de forme par rapport au voisinage).

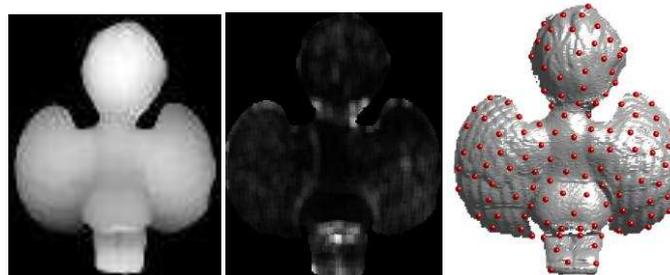


Figure 3-5-A gauche-Image de profondeur d'un modèle d'ange –Au milieu- L'image correspondant aux valeurs de la mesure facteur de qualité, en niveau de gris- A droite- Les PIs détectés.

3.1.5 La classification HK et SC

Rappelons que l'idée ici est de construire un espace de classification de formes en utilisant la paire des courbures Moyenne-Gaussienne (HK) ou le couple Indice de forme-intensité de Courbure (SC). Le type d'une surface locale est reconnu par l'intermédiaire de la valeur de la paire considérée. Ces couples sont géométriquement invariants et renseignent sur l'ordre second de l'information géométrique de la forme.

Typiquement, pour la classification HK, la fonction type T_p (Equation 1.9) associe à chaque valeur du couple H et K un type unique correspondant à une forme particulière. Neuf valeurs de type sont considérées. De plus, deux seuils sur les valeurs de H et K permettent de classer les surfaces planes. Dans l'espace Indice de forme-Intensité de Courbure (SC), l'indice de forme SI renseigne sur la forme et l'intensité de Courbure C définit le degré de courbure. Comme l'illustre bien la Figure 3-6, les détails de la forme sont visuellement accentués.

Pareillement, que la représentation HK, SI permet d'avoir une graduation continue et l'espace SC subdivise les formes de surface en 9 catégories. Dans la Figure 3-7 illustrant cette catégorisation, les droites séparatrices sont des valeurs constantes de S: $S = \pm 5/8, \pm 3/8$ et $\pm 3/16$. Le cercle autour de l'origine est le seuil pour les valeurs de C à zéro et d'équation: $\sqrt{(\kappa_1^2 + \kappa_2^2)} = C_{zéro}$.

Nous utilisons ces deux classifieurs pour extraire les régions saillantes qui sont sélectionnées au niveau des surfaces en forme de dôme, de cuvette, de sphère, de selle ornière et de selle crête, et qui correspondent aux primitives primaires que l'humain peut observer. Un grand nombre de PIs résulte suite à cette extraction. Nous proposons dans notre implémentation de ces détecteurs d'appliquer un processus de filtrage des points selon leur distance relative par l'algorithme 1 de clustering. Nous désignons ces deux détecteurs par « HK » et « SC ».

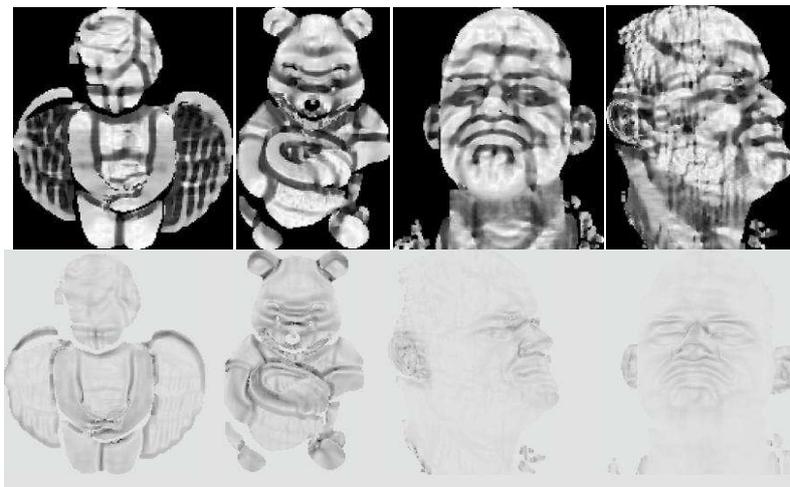


Figure 3-6 Images correspondant à -en haut- mesures de SI et -en bas- mesures de C

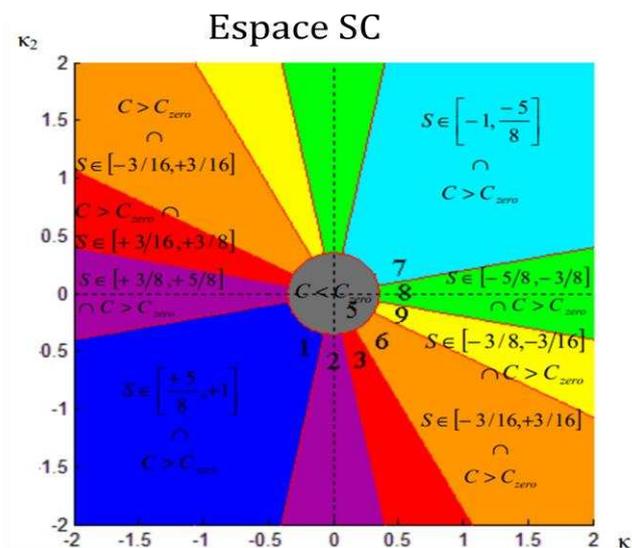


Figure 3-7 Classification SC dans le plan (κ_1, κ_2) (Akagündüz, 2011)

3.1.6 Combinaison des critères

a) Version mono-échelle

Théoriquement, les deux classifications SC et HK convergent vers le même résultat. Cependant, comme nous pouvons le remarquer sur les graphes de la Figure 3-7 de l'espace SC et ceux de l'espace HK, des différences dans la classification de quelques régions subsistent. Par exemple, la région plane, définie dans l'espace HK (Figure 1-10), est caractérisée autrement dans l'espace SC (régions 1, 2, 5, 9, 10, 11, 14). Pareillement, les régions elliptiques et paraboliques se diversifient dans les deux espaces quand les valeurs de courbures principales augmentent. Une confusion entre les deux types de surface est remarquée pour les régions 4, 6, 13, 15. Et finalement, la définition des régions 3, 7, 8, 12, 16, 17 restent ambigus.

De ce fait, nous suggérons, dans un premier temps, de combiner les deux espaces en un espace couplé HK&SC pour augmenter l'efficacité du résultat de classification (

Figure 3-8 et Figure 3-9). Seulement les régions communes sont classées. En d'autres termes, nous ne sélectionnons un PI que s'il est saillant pour les deux classifications SC et HK. L'intersection garde les points dont les deux labels du type correspondent à une région saillante (dôme, sphérique ou selle). Au final, notre résultat est validé par deux mesures de détection de PIs et nous réduisons le nombre de point représentatifs de la forme initiale (Figure 3-10). Quant au type de régions choisies comme saillantes, nous les définirons selon la nature des objets qui constituent la base de données. Par exemple, si des objets dont les surfaces sont planes existent, nous allons sélectionnés, outre les points déjà mentionnés, les points de courbure nulle.

Convex (Elliptic) ⁽¹⁾	$H < - H_{zero} \cap K > + K_{zero} \cap S \in [+5/8, 1] \cap C > C_{zero}$
Convex (Parabolic) ⁽²⁾	$H < - H_{zero} \cap K < K_{zero} \cap S \in [+3/8, +5/8] \cap C > C_{zero}$
Saddle Ridge ⁽³⁾	$H < - H_{zero} \cap K < - K_{zero} \cap S \in [+3/16, +3/8] \cap C > C_{zero}$
Planar ⁽⁵⁾	$H < H_{zero} \cap K < K_{zero} \cap C < C_{zero}$
Hyperbola ⁽⁶⁾	$H < H_{zero} \cap K < - K_{zero} \cap S \in [-3/16, +3/16] \cap C > C_{zero}$
Concave (Elliptic) ⁽⁷⁾	$H > + H_{zero} \cap K > + K_{zero} \cap S \in [-1, -5/8] \cap C > C_{zero}$
Concave (Parabolic) ⁽⁸⁾	$H > + H_{zero} \cap K < K_{zero} \cap S \in [-5/8, -3/8] \cap C > C_{zero}$
Saddle Valley ⁽⁶⁾	$H > + H_{zero} \cap K < - K_{zero} \cap S \in [-3/16, +3/16] \cap C > C_{zero}$

Figure 3-8 Formulation mathématique de la classification du couple HK&SC (Akagündüz, 2011)

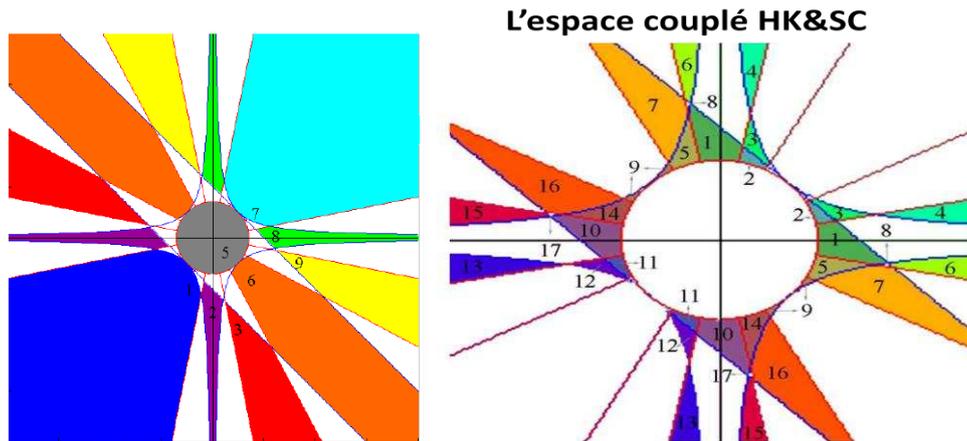


Figure 3-9 L'espace couplé HK&SC de classification dans le plan (k_1, k_2) (Akagündüz, 2011)

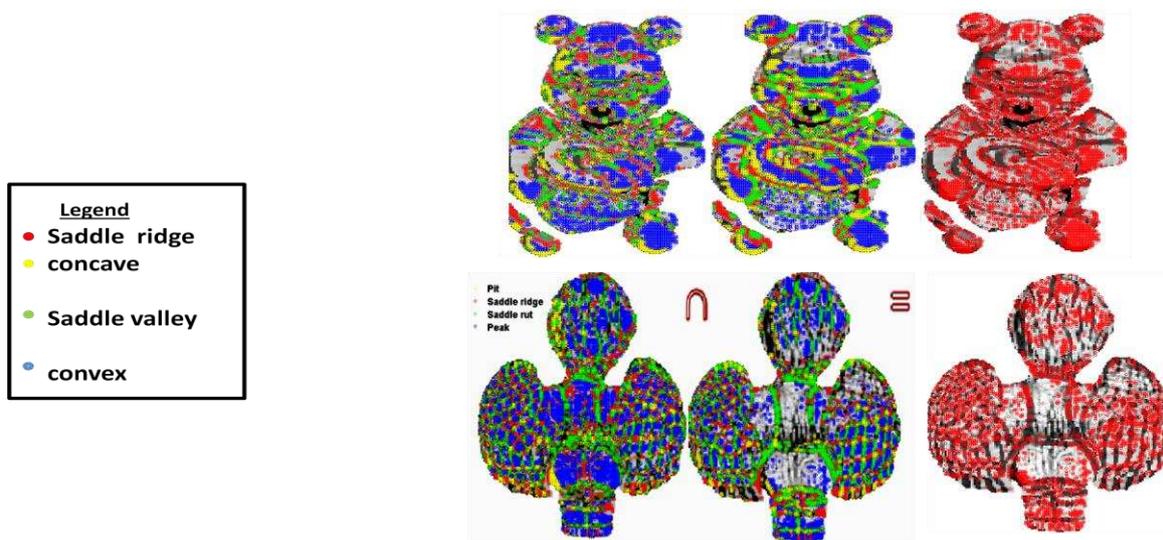


Figure 3-10 Résultat du processus de croisement de deux espaces SC et HK sur deux modèles d'objets.

Nous proposons, dans un deuxième temps, de rajouter aux critères de sélection SC et HK, un critère de tri. Nous parcourons le nuage point par point et sélectionnons les points répondant au critère de saillance de la paire (SC, HK). La sélection finale des PIs pour le détecteur SC_HK se fait de deux façons. :

- La première façon de faire est de trier l'ensemble des PIs trouvé selon une valeur de mesure de tri puis d'appliquer le principe de filtrage des PIs basé sur l'écart spatial des coordonnées 3D (algorithme 1de clustering Harris) sur l'ensemble des points ordonnés. Nous avons utilisé un premier critère de tri basé sur l'intensité de courbure C pour sélectionner les valeurs maximales. Nous appelons la combinaison SC_HK avec le critère de tri sur C , détecteur « SC_HK_C ». Nous introduisons un deuxième détecteur qui classe les PIs selon la valeur du facteur qualité FQ des points. Nous notons ce dernier détecteur « SC_HK_FQ ». Un troisième détecteur basé sur un tri avec une valeur de confiance γ utilisée par Ho et Gibbins (Ho, et al., 2009) et qui mesure la déviation de leur intensité de courbure

par rapport à celle du voisinage. La formule de cet indice est donnée dans la Figure 3-11. Nous notons ce dernier détecteur « SC_HK_Conf » et appelons la mesure γ par indice « Conf ».

$$\gamma(p, r_k) = \frac{|c_p - \mu_{N_p}|}{\sigma_{N_p}} \quad \begin{aligned} \mu_{N_p} &= \frac{\sum_{p_j \in N_p} c_{p_j}}{n} \\ \sigma_{N_p} &= \sqrt{\frac{\sum_{p_j \in N_p} (c_{p_j} - \mu_{N_p})^2}{n-1}} \end{aligned}$$

Figure 3-11 Formule de l'indice de confiance γ basée sur l'intensité de courbure C (Ho, et al., 2009)

- Pour la deuxième sélection, nous procédons à un étiquetage de chaque point du nuage de points selon le critère considéré, puis nous extrayons toutes les composantes connexes. La sélection finale se fait en prenant le centre des composantes qui répondent au critère de tri final. Nous allons considérer uniquement le critère basé sur la valeur maximale en C et qui donne le détecteur « SC_HK_C_Con ». Nous introduisons la notion de labellisation des composantes connexes comme suit :

- **Labellisation des composantes connexes**

L'étiquetage des composantes connexes est utilisé pour exprimer la connexité ou grouper les pixels connectés dans une image. La connexité est la relation entre deux pixels ou plus. Deux pixels sont dits connexes s'ils sont adjacents et de valeur de pixel égale. 4 ou 8- connexités sont les plus utilisées. Pour la connexité d'ordre 4, les 4 voisins (dessus, dessous, gauche et droite) sont vérifiés. Pour la connexité 8, s'ajoute une vérification des voisins dans les 4 diagonales. Le groupement des pixels connexes forme une composante connexe. Chaque composante connexe est étiquetée par un nombre différent. Deux pixels p_1 et p_2 sont connectés s'il existe un chemin les reliant. Durant le processus de parcours des pixels, pour un pixel p_1 et ses quatre voisins, si un des voisins a été déjà étiqueté dans les étapes précédentes, p_1 est étiqueté comme son voisin. Et si aucun pixel voisin n'a été étiqueté, p_1 est marqué par une nouvelle étiquette. Ce processus est répété pour toute l'image. Au final, les composantes connexes sont séparées.

- **Algorithme général de combinaison**

Nous présentons les étapes générales de l'algorithme de combinaison des critères :

- Pour chaque point du nuage, générer le maillage (détection de la boîte englobante + calcul de la distance moyenne entre les points, utilisée pour construire le maillage)
- Détection des points du bord
- Calculer les mesures de saillances sur chaque patch (SC et HK) avec un voisinage N_p proportionnel à la diagonale de la boîte englobante et extraction des points saillants
- Calculer une valeur de mesure (C ou FQ ou Conf) des PIs
- Regrouper les points détectés :
 - sous forme de composantes connexes (label=types selon SC- Type selon HK) : Garder les composantes connexes les plus larges et prendre le PI ayant la plus grande intensité de courbure C comme représentant
 - ordonner et regrouper avec l'algorithme l'ensemble des PIs selon l'un des critères : C ou FQ ou Conf

Contrairement au tri selon le critère C calculé sur points PIs seulement, les mesures FQ et Conf des PIs font intervenir les mesures de courbure du voisinage N_p de ces points. Un paramètre, ParSelect_PIs

qui est un ratio permettant de prendre une proportionnel à la taille de la diagonale de l'objet, est utilisé pour filtrer les PIs selon la distance spatiale entre eux.

Dans la Figure 3-12, le processus d'extraction des PIs en utilisant le détecteur SC_HK_C_Con est exposé. Nous notons que l'étape des composantes connexes permet de réduire considérablement le nombre final de PIs. Les représentants de chaque composante connexe sont situés sur des régions saillantes du visage (nez, front, bouche, menton, joue, oreille) puisqu'ils sont choisis sur les valeurs de courbure maximale.

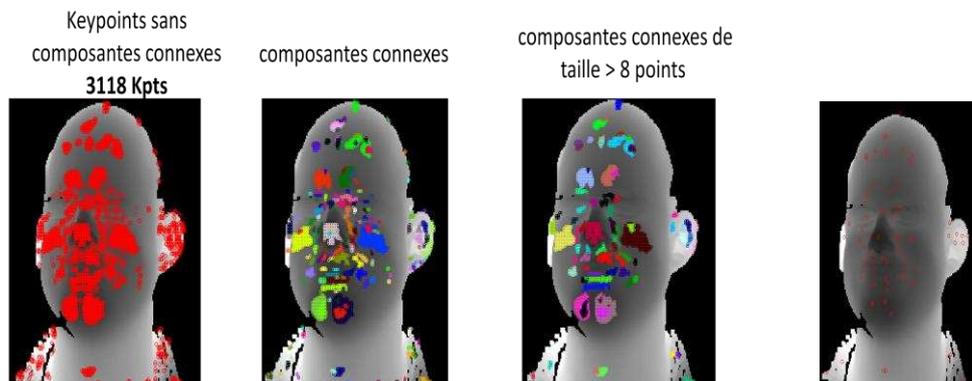


Figure 3-12 Illustration du processus d'extraction des PIs en utilisant le détecteur SC_HK_C_Con

b) Version Multi-échelle

Similairement à l'idée introduite par l'étude de (Ho, et al., 2009), nous proposons, pour une version multi-échelles de notre travail, de s'appuyer sur l'augmentation de la taille du voisinage au lieu d'appliquer un lissage par une pyramide de gaussiennes. Le principe est de calculer la mesure de saillance sur différents niveaux de voisinages (échelles) et de sélectionner le niveau pour lequel la valeur de saillance est extremum par rapport aux valeurs de ses voisins à l'échelle en question et aux valeurs des échelles voisines. L'algorithme adopté est alors le suivant :

- **Algorithme de la version Multi-échelle**

- **Données :**

$P = \{p_i \in \mathbb{R}^3\}$: ensemble des points 3D du modèle.

$R = \{r_k\}$: un ensemble d'échelles.

- **Algorithme :**

1: **for** $r \in \{r_k\}$ **do**

2: **for** $p \in \{p_i\}$ **do**

3: Trouver le voisinage N_r à l'échelle r

4: Calculer les mesures de saillance basées sur k_1, k_2, SI, C, H, K

5: Vérifier si p est saillant avec le critère $(SC \cap HK)$ à l'échelle r

6: **end for**

7: Sélectionner les PIs de position p saillante en termes de (FQ ou C ou Conf) par rapport à son voisinage sur les rayons r_k , $r_k - 1$ et $r_k + 1$ et ayant une mesure de saillance maximale à l'échelle r_k

8: Regrouper les PIs sous forme de composantes connexes et sélectionner pour chaque composante les PIs de courbure maximale C

ou

8: Trier les PIs selon leur saillance (FQ ou C ou Conf) puis appliquer un clustering

9: **end for**

Dans les tests effectués, nous n'avons considéré que la méthode avec les composantes connexes de l'étape 8. Par contrainte de temps, nous n'avons pas pu mener à bout cette version multi-échelles et exploiter ces résultats. Nous nous sommes contentés de quelques résultats visuels qui seront présentés ultérieurement.

3.2. Eléments impactant sur les mesures de courbure

3.2.1 Artefacts et bruit

En optique, une surface claire réfléchit plus de lumière qu'une surface sombre. C'est ce qui fait que pour les scanners optiques certaines zones de la scène comme les cheveux noirs ou des miroirs ne sont pas correctement détectés. Les réflexions multiples de la surface donnent des points aberrants (outliers). De plus, les limitations physiques des capteurs peuvent ajouter du bruit ou des artefacts et affecte l'échantillonnage et les valeurs de profondeurs des points. Le type de bruit des scanners 3D est typiquement gaussien. Similairement au bruit 2D qui est convolué avec l'intensité des pixels, le bruit 3D perturbe les coordonnées des points du nuage. Un exemple de l'impact de bruit sur le calcul des normales est présenté dans la Figure 3-13.

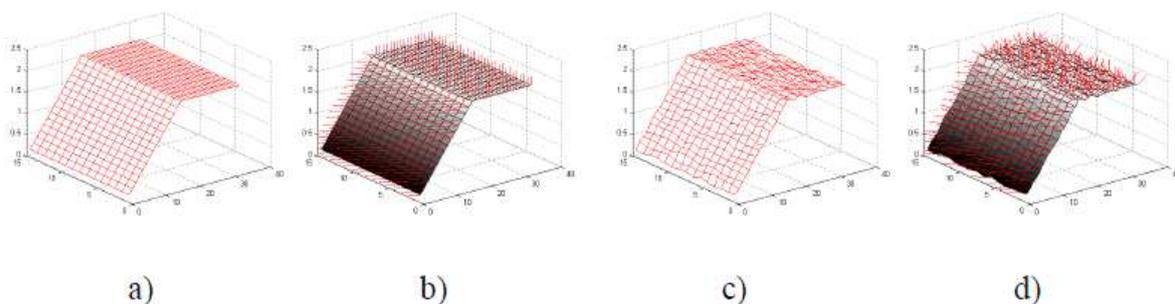


Figure 3-13 a) Surface 3D artificielle sans bruit b) Directions des normales sur la surface non bruitée c) Surface 3D artificielle avec bruit. d) Directions des normales sur la surface bruitée (Bozkurt, et al., 2009)

La surface de l'objet scanné étant inconnue, il est très difficile de décider de la validité ou non-validité des points. Le bruit peut être partiellement enlevé avec par exemple un filtre bilatéral ou un lissage Laplacien. Le filtre bilatéral diminue le bruit tout en préservant les détails de la forme 3D contrairement à un lissage Laplacien qui ne lisse pas seulement le bruit mais aussi les caractéristiques de la forme. Une comparaison entre le lissage Laplacien (implémenté dans la bibliothèque VTK) et le filtrage bilatéral gaussien (Fleishman, et al., 2003) est présenté dans la Figure 3-14. Nous notons la perte des détails de la forme autour de l'œil de bunny sur le modèle lissé par rapport au modèle dé-bruité.

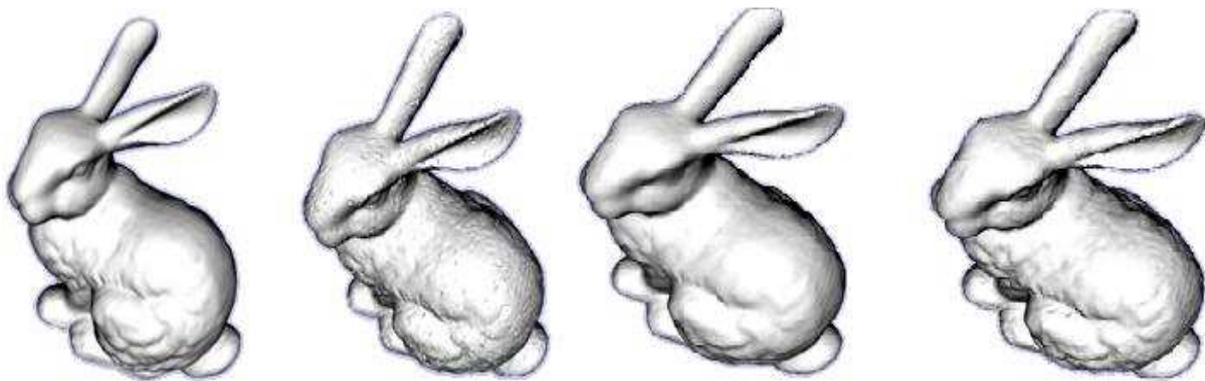


Figure 3-14 De gauche à droite : Modèle Original – Ajout de bruit Artificiel - Modèle lissé par un lissage Laplacien – Modèle dé-bruité par un filtrage bilatéral

Après cette comparaison, nous proposons d'appliquer, en prétraitement, un filtrage bilatéral sur les données des bases **bruitées**. Dans la Figure 3-15, nous présentons le pseudo-code d'une itération de l'algorithme de filtrage bilatéral gaussien de (Fleishman, et al., 2003) que nous allons utiliser dans notre travail. L'idée de ce filtre sélectif consiste à faire une moyenne gaussienne sur les points les proches dans le voisinage. Le choix des paramètres comporte essentiellement le nombre d'itérations et la taille du voisinage.

```

DenoisePoint(Vertex v, Normal n)
  {qi} = neighborhood(v)
  K = |{qi}|
  sum = 0
  normalizer = 0
  for i := 1 to K
    t = ||v - qi||
    h = ⟨n, v - qi⟩
    wc = exp(-t2 / (2σc2))
    ws = exp(-h2 / (2σs2))
    sum += (wc · ws) · h
    normalizer += wc · ws
  end
  return Vertex  $\hat{v} = v + n \cdot (sum / normalizer)$ 

```

Figure 3-15 Pseudo-code de l'algorithme de filtrage bilatéral (Fleishman, et al., 2003)

3.2.2 Invariance aux transformations, échelle et résolution

Les valeurs des courbures principales d'une surface sont indépendantes des rotations 3D. C'est une propriété très importante pour la reconnaissance d'objets quand la capture des objets se fait selon des orientations arbitraires. Les primitives obtenues à partir des courbures principales sont alors invariantes à l'orientation. De ce fait, les valeurs H, K, S et C sont indépendantes de la translation et de la rotation dans l'espace 3D. Cependant, la variation d'échelle modifie la différence finie des valeurs échantillonnées des modèles et impacte le calcul des dérivées partielles discrètes d'une même surface. De ce fait, les courbures principales κ_1 et κ_2 , calculées à partir d'une Hessienne ou des dérivées partielles de deuxième ordre d'une surface, sont dépendantes de l'échelle et varient proportionnellement à un changement d'échelle de la surface dans l'espace 3D. De plus, avec le caractère discret des surfaces obtenues avec les scans 3D, le ratio de l'échantillonnage est un deuxième facteur impactant sur la mesure des courbures. Ainsi les valeurs de H, K et C ne sont pas invariantes par changement d'échelle ou de résolution. Par contre, S est indépendant de l'échelle et de la résolution d'après sa formulation. L'indépendance de l'indice de forme à la résolution est illustrée dans la Figure 3-16, dans laquelle, différents niveaux d'échantillonnage de la même forme sont considérés. La remarque est que la résolution de l'image SI diminue certes, mais les valeurs de SI (intensité en niveau de gris) ne sont pas affectées.

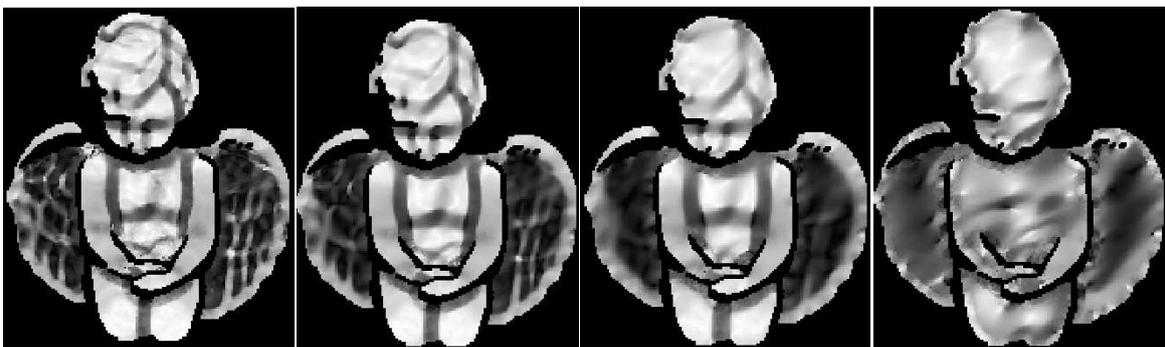


Figure 3-16 Images de l'indice de forme pour différentes résolutions de la forme initiale

3.2.3 Maillage et voisinage

Rappelons que l'objectif de cette thèse est de reconnaître un objet à partir de quelques vues 2.5D et les mettre en correspondance après extraction des primitives.

- Maillage : La subdivision du nuage de points en des sous-régions locales (patches) permet d'avoir des mesures d'une façon indépendante dans chaque portion du modèle. Nous exploitons la structure régulière en (i, j) de l'image 2D pour construire un maillage du nuage de points. Nous générons ce maillage de la façon suivante: nous cherchons les coordonnées des points minimums et maximums sur l'axe des x et des y pour former la boîte englobante de l'objet. Ces coordonnées servent à calculer la taille approximative de la diagonale de l'objet. Ensuite, nous calculons la distance moyenne des points contenus dans cette boîte. Puis, nous parcourons la grille des points dans l'ordre donnée par l'image de profondeur pour former, finalement, des facettes triangulaires à partir de points dont la distance entre eux est inférieure à la distance moyenne du nuage.
- Voisinage : Pour la détection des points d'intérêt, nous mesurons la saillance de chaque point. L'unité de mesure est le couple des courbures principales qui se calcule sur un voisinage d'un point. Le choix de la taille de ce voisinage est critique. Dans notre cas cette

taille est déterminée soit par le nombre d'anneaux autour du point soit par la valeur du rayon r d'une sphère centré au point. Dans le cas où les triangles du maillage ont la même taille (tessellation uniforme), nous pouvons considérer un nombre constant d'anneaux ou un rayon r fixe. Par contre, pour un maillage irrégulier et complexe, ces méthodes n'approximent pas le voisinage adéquatement. Une approche adaptative permet de sélectionner la taille du voisinage selon la tessellation autour du point.

- Considérons l'objet comme un graphe $G(V, E)$, où V est l'ensemble des sommets du maillage et E est l'ensemble des arêtes d'adjacence dans l'objet. Pour un point $v \in V$, un k -ring autour de v correspond à l'ensemble des points dont la longueur du chemin le plus court à partir de v est k . Une façon adaptative de choisir ce nombre est de prendre une proportion du nombre des facettes dans l'objet.
- Dans le cas où le voisinage des points est donné par une région sphérique autour du point, le rayon de cette sphère est pris proportionnellement à la diagonale de la boîte englobante pour assurer une invariance aux changements d'échelles et à l'échantillonnage.
- Nous évitons également dans le calcul des courbures les points situés sur les bords et dans les zones de transition de visibilité. Ces points de discontinuité sont détectés par la recherche des lignes perpendiculaires à l'angle de vue et par leur faible densité (voisinage).

3.3. Résultats expérimentaux

3.3.1 Outils de développements et bases de données

Estimer de manière précise et fiable les courbures principales en tout point est une étape essentielle dans notre détection. Par souci de simplicité, l'approche adoptée est fondée sur l'ajustement local d'une surface polynomiale de second degré 2-jet. L'application d'une Analyse en Composantes Principales (ACP) sur le voisinage des points, suivi d'un ajustement polynomial permet de déduire la base et les coefficients de Monge (Figure 3-17). La bibliothèque CGAL (CGA10) est utilisée pour l'estimation des quantités différentielles d'une surface lissée d'un maillage triangulaire.

Dans le but d'accélérer le temps de calcul, nous avons pu exploiter également, l'implémentation donnée par (Tombari, et al., 2010) pour définir un système de référence local (RF) utilisé pour le calcul du descripteur SHOT. La normale est donnée par le vecteur propre d'une matrice de covariance formée par une combinaison linéaire des distances des points du voisinage. Nous notons ce deuxième algorithme par **RF**.

La bibliothèque VTK (VTK10) est utilisée pour stocker la structure du maillage, et l'application des transformations d'ajout de bruit et de décimation des données.

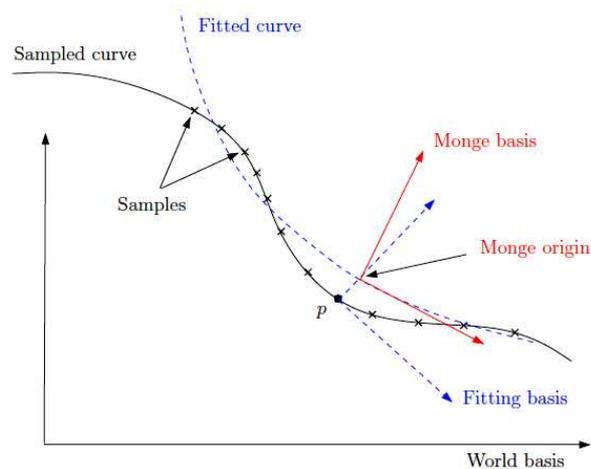


Figure 3-17 Les trois bases orthonormées utilisées pour l'estimation des courbures

Dans nos expérimentations, nous avons utilisé des images de profondeurs de modèles, présentant une variation en angle de prise de vue et d'échelle. Le degré de similarité entre les formes d'objets varie selon le type de la base de données. Des bases différentes en termes de taille, qualité, densité et type d'objets, générées par différents capteurs, sont exploitées dans nos tests. L'évaluation quantitative des détecteurs sera donnée uniquement sur les données de la base Minolta.

L'ensemble de ces bases sera exploité essentiellement dans le chapitre reconnaissance où chaque jeu de données est séparé en deux groupes : un ensemble pour la phase d'apprentissage et un autre pour la requête ou le test.

La base Stuttgart et la base Minolta

La base de Stuttgart University Range Image (Stu10) est composée de 42 objets. Ces images de profondeurs sont générées synthétiquement pour avoir un grand jeu de données offrant une variation de l'angle de vue. A chaque objet correspond 66 poses d'apprentissage et 258 poses pour le test. Ce qui donne un total de $66 \times 42 = 2772$ poses en apprentissage et $258 \times 42 = 10836$ poses en test. Les angles de vues sont pris sur les différents angles de la sphère avec un écart de $23-26^\circ$ dans la base d'apprentissage et un décalage de $11.5-13^\circ$ pour les requêtes de test. L'exemple de la Figure 3-18 représente l'ensemble d'apprentissage pour l'objet "machine".

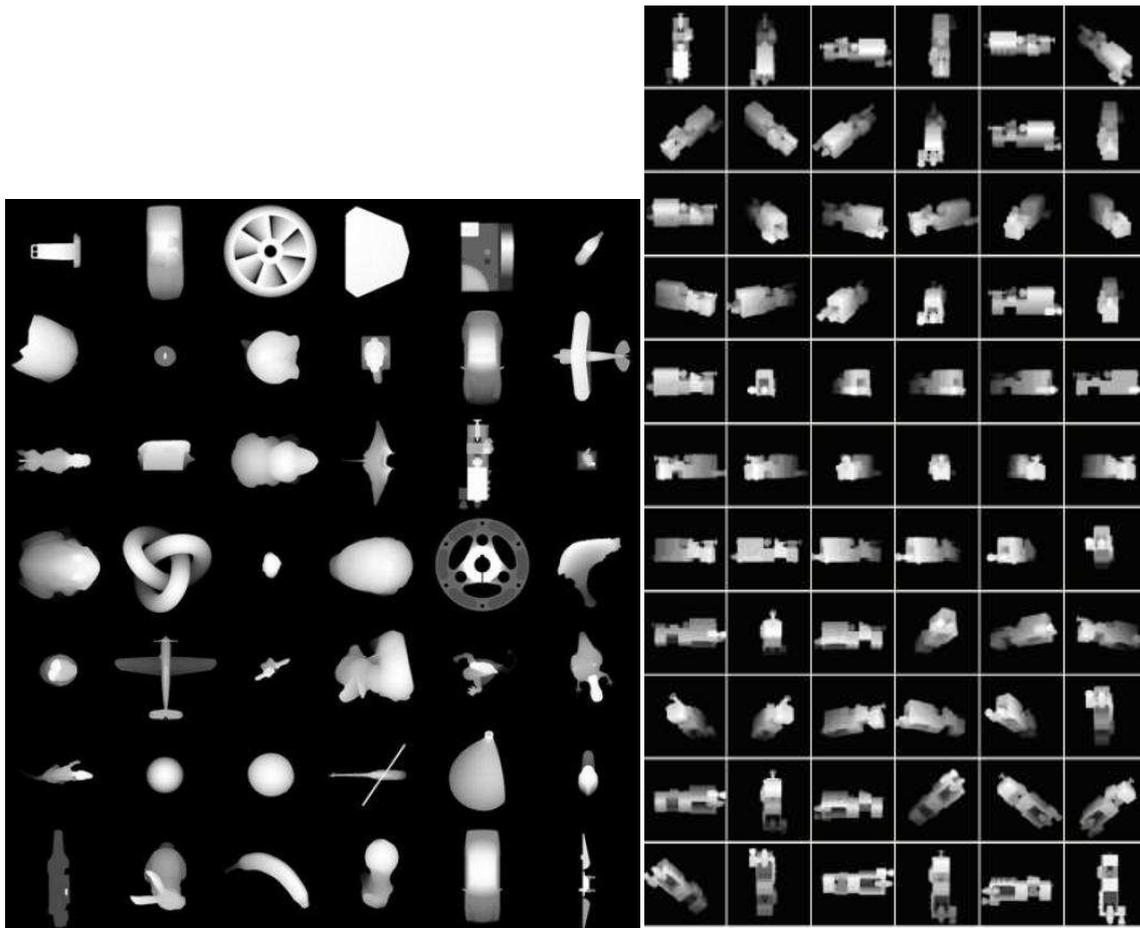


Figure 3-18 A gauche- Les 42 objets de la base de Stuttgart –A droite- L’ensemble des images de la base d’apprentissage pour l’objet « machine »

Une deuxième base d’images de profondeur est la base publique Minolta acquise par un scanner laser (Minolta Vivid 900 laser range scanner) (Min11) par l’université Ohio State University. Un grand jeu de données est fourni avec plus de 30 objets et présente des variations d’angles de 1° , 20° et 36° . Pour les chapitres détection et description, nous avons choisi de travailler avec 9 objets (Figure 3-19) de cette base avec un total de 162 vues d’écart de 20° . Cette sélection d’objet a été proposée par (Chen, et al., 2004). Nous avons remplacé l’objet 9 orangedino de cette base par l’objet bunny.

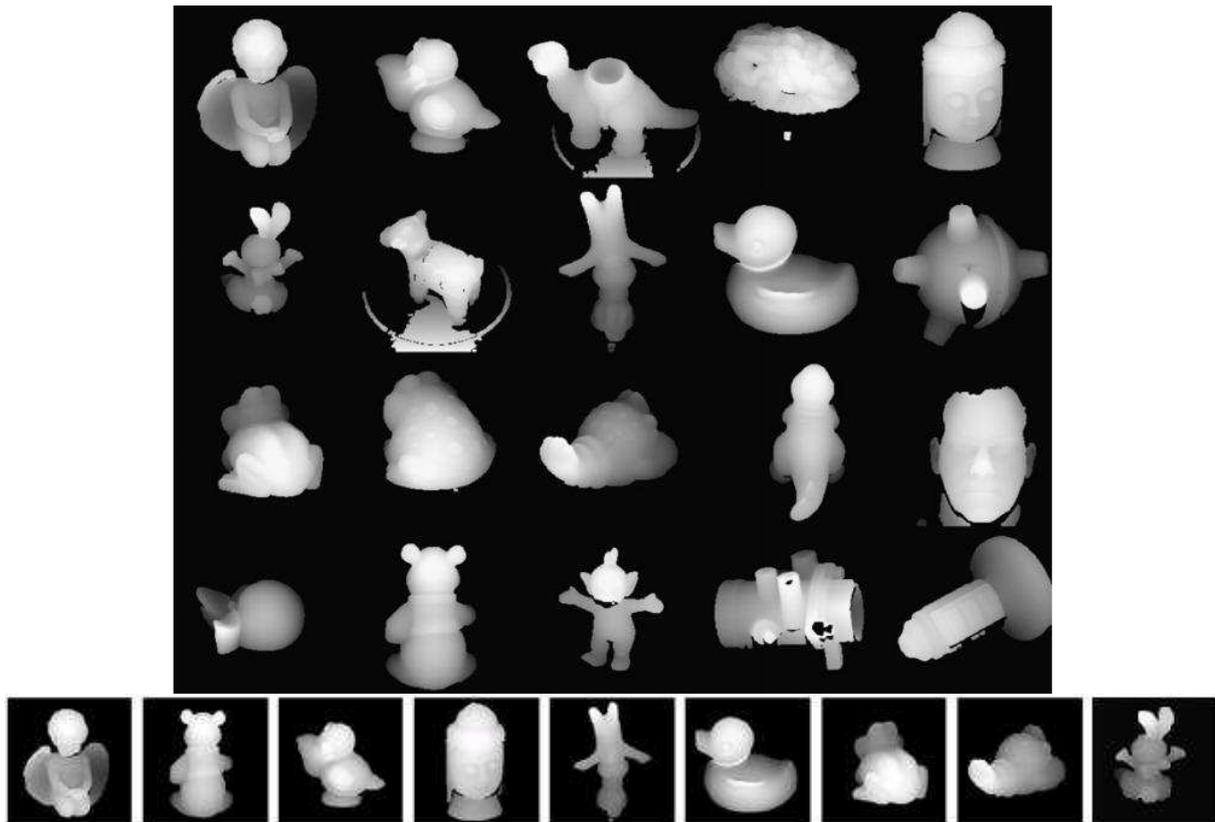


Figure 3-19 En haut- 20 objets de la base Minolta, en bas-Images de profondeurs des 9 objets de la base Minolta utilisés pour notre évaluation (Chen, et al., 2007)

Dans la base Minolta, se présentent, également, des vues prises sous différentes échelles de 2 à 4 échelles pour quelques objets. Dans ce jeu de données issues d'acquisitions réelles de scènes, les objets sont extraits par un processus de segmentation qui laisse des bouts du fond sur quelques vues. Nous avons pu récupérer 7 objets dont les vues ne présentent quasiment pas de problèmes de segmentation sauf pour quelques vues (exemple sur la ligne 1 et 2 de la Figure 3-20). Ces objets sont comme suit avec le nombre d'échelles par objet: bluedino (3 échelles), brain (3 échelles), facesimages (3 échelles), gc_bottle (4 échelles), reddino (3 échelles), valve (4 échelles) et yellowhorn (3 échelles). Pour chaque objet, nous avons sélectionné 11 vues. Le nombre de points dans des vues prises sous le même angle augmente en passant d'une échelle à une échelle plus proche du capteur. Par exemple, pour la vue 0 du modèle Rick des facesimages, nous avons 15523 points et 29920 facettes à l'échelle 0, contre 17883 points et 34813 facettes à l'échelle 1, et 18483 points et 35896 facettes à l'échelle 2.

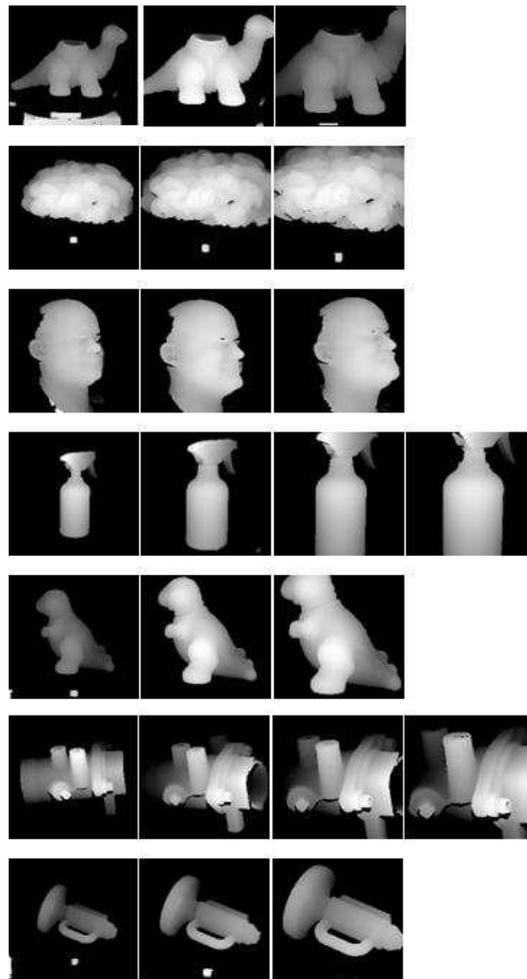


Figure 3-20 Les 7 objets présentant une variation d'échelle. Sur chaque ligne, nous illustrons les vues 0° de chaque échelle

La base RGB-D Object

La base d'objet "RGB-D Object Dataset" (RGB12) est, à notre connaissance, la plus grande base publique de données RGB-D et est composée de 300 objets communs « domestiques » (household). Ces objets sont organisés en 51 catégories rangées selon le système de relations WordNet hypernym-hyponym (similaire à ImageNet). Par exemple la catégorie soda est divisée en des instances comme Pepsi Can et Mountain Dew Can. Les captures sont faites avec la caméra 3D Kinect qui synchronise et aligne des images RGB 640x480 avec des images de profondeurs à une fréquence de 30 HZ. Chaque objet est placé sur un plateau tournant et une capture est prise pour toutes les rotations. La caméra est placée à plusieurs hauteurs pour avoir différents angles de prise de vues.

La Figure 3-21 représente 46 exemples d'objets segmentés de la base RGB-D, qui sont respectivement : apple_1, ball_1, banana_1, bell_peper_1, binder_1, calculator_1, camera_1, cap_1, cell_phone_1, cereal_box_3, coffee_mug_1, comb_1, flashlight_1, food_bag_1, food_box_1, food_can_1, food_cup_1, garlic_1, greens_1, hand_towel_1, instant_noodles_1, keyboard_1, Kleenex_1, lemon_1, lightbulb_1, lime_1, marker_1, mushroom_1, notebook_1, onion_1, orange_1,

peach_1, pear_1, pitcher_1, plate_1, potato_1, rubber_eraser_1, scissors_1, shampoo_1, soda_can_1, sponge_1, stapler_1, tomato_1, toothbrush_1 and watter_bottle_1.



Figure 3-21 Les 46 exemples d'objets segmentés de la base RGB-D (RGB12)



Figure 3-22 Deux vues du même objet pour la base RGB-D Object

Nous énumérons à présent quelques caractéristiques de la Kinect :

- Portée: ~ 50 cm à 5 m.
- Résolution horizontale: 640 x 480. Pour des géométries simples, la résolution est de ~ 0.75 mm par pixel en x sur y à une distance 50 cm, et de ~ 3 mm par pixel en x sur y à une distance de 2 m.
- FOV: 45° en vertical et 58° en horizontal
- Résolution de la profondeur: ~ 1.5 mm pour 50 cm et environ 5 cm à 5 m.
- Bruit: Environ +1 DN pour toutes les distances, mais ce bruit est non-linéaire. Ce qui veut dire qu'il est égale à +1 mm pour des distances faibles, et +- 5 cm pour les plus éloignés.

Base « Carotte » du laboratoire

Cette base a été construite dans notre laboratoire dans le cadre du projet CAROTTE « CARTographie par ROboT d'un Territoire » dans lequel l'une des tâches d'un robot mobile était de reconnaître les

objets de son environnement. La caméra 3D Kinect est utilisée pour capturer quelques prises de vues (3 à 10) de 20 objets différents : ballon rouge, botte fourrée, botte en plastique verte, bouteille mauve, casier de rangement, chaise bleue à roulette, chaise beige à quatre pieds, cylindre vert, disque rose, boîte d'archive bleu, gros bidon, livre, marteau, parallélépipède mauve, boîte d'archive blanc, petit bidon noir, poubelle blanche, prisme jaune, robot Pioneer quatre roues, ventilateur (Figure 3-23). Ces objets sont de tailles différentes allant de 10cm à 1m. La distance moyenne de capture est de 1m50. L'avantage de ce type de caméra est que l'information RGB est stockée en plus de l'information de profondeur.

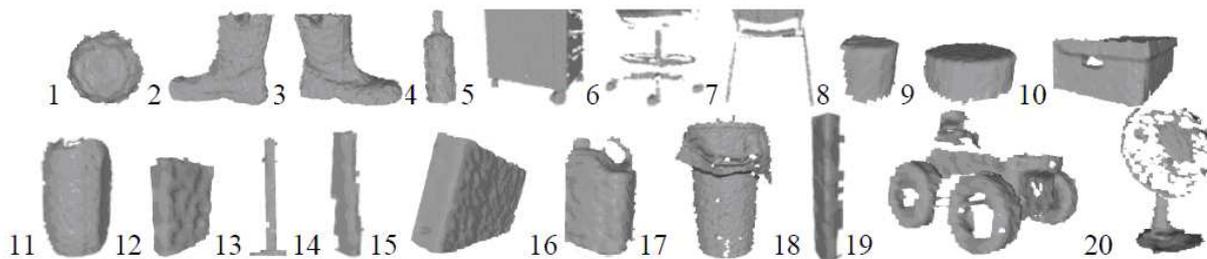


Figure 3-23 Les 20 objets de la base carotte du laboratoire

3.3.2 Evaluation des détecteurs

3.3.2.1 Evaluation visuelle

Nous proposons de commencer cette partie par une brève évaluation visuelle de la performance des 10 détecteurs SC_HK_FQ, SC_HK_C, SC_HK_C_Conf, SC_HK_C_Con, Harris, SI, HK, SC et SURF. Le détecteur FQ, ne fera pas parti des détecteurs testés. Par raison de son temps de calcul et sa faible performance, il a été éliminé.

Nous présentons quelques résultats de ces détecteurs pour les bases: Minolta (Figure 3-24 et Figure 3-25), Carotte (Figure 3-26), Stuttgart (Figure 3-27), et RGB-D (Figure 3-28). Nous allons mettre l'accent sur des aspects particuliers pour classer ces résultats.

❖ **Avantage du processus de combinaison:**

Par le croisement des deux espaces de classification SC et HK, nous profitons des avantages de l'un et de l'autre. En effet, la combinaison des deux estimateurs de courbure SC et HK est robuste aux variations d'angles (Figure 3-25) et élimine les points qui ne sont pas classés pareillement par les deux classificateurs de surface (Figure 3-24). Les points éliminés se caractérisent par une faible variation locale de la forme de l'objet ou peuvent correspondre à du bruit ou à un artefact. En outre, le filtrage basé sur un tri des mesures de courbure (C ou FQ ou Conf) des PIs renforce cette saillance en prenant les points avec la plus forte variation de courbure par rapport à leur voisinage. La dispersion des PIs des détecteurs SC, HK, SC_HK_C, SC_HK_Conf et SC_HK_FQ, sur la surface de l'objet, est due à l'opération de regroupement basée sur l'écart spatial entre les PIs triés.

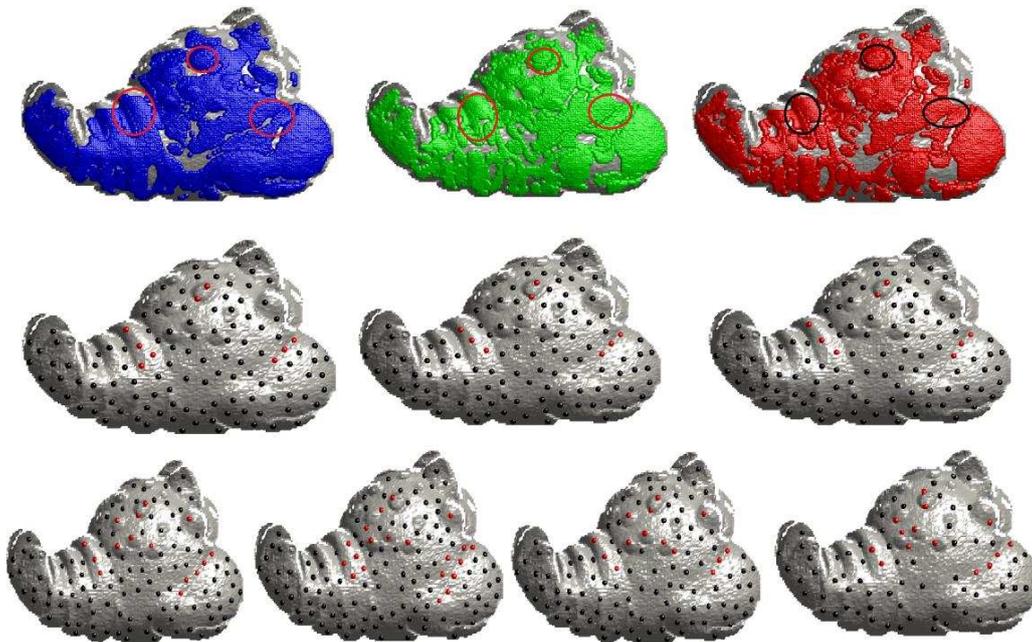


Figure 3-24 Résultat d'extraction des PIs pour le processus de combinaison.

Les détecteurs sont les suivants : -de gauche à droite- Ligne 1: HK sans filtrage (6695), SC sans filtrage (5462), SC_HK sans filtrage (4950). Ligne 2: HK avec filtrage (121), SC avec filtrage (121) et SC_HK_Conf (115). Ligne 3: SC_HK_C (121), SC_HK_FQ (131), SC_HK_Conf (115) et SC_HK_Con (81).

Le nombre de PIs est donné entre parenthèses. Les points colorés en rouge dans la ligne 2 correspondent aux points retenus parmi les points de la zone entourée dans la ligne 1. Dans la ligne 3, les positions en rouge montrent des points extraits sur les mêmes zones de l'objet sur les 4 détecteurs.

Entre les trois méthodes de sélection combinée SC_HK_C, SC_HK_Conf et SC_HK_FQ basées sur un tri final sur les valeurs de C ou FQ ou Conf, la différence n'est pas très notable visuellement (quasiment les même positions de PIs sur la Figure 3-24). L'explication est que ces trois estimateurs mesurent la même information de degré de déviation de la courbure par rapport à un voisinage. Ce voisinage intervient, plus fortement, dans les mesures FQ et Conf puisque des notions d'extremum ou de moyenne, sur les valeurs de courbure des voisins, sont utilisées.

D'autre part, la version combinée SC_HK_C_Con, utilisant le processus de composantes connexes, sélectionne moins de points (81 PIs) que les autres détecteurs combinés (~120 PIs). En effet, la recherche de la connexité, entre les points en se basant sur le couple (SC-HK) de classification de la surface, réduit dans un premier temps l'espace de sélection des PIs. Ensuite dans un deuxième temps, l'étendu de chaque composante connexe sur l'objet est représenté par un seul PI avec la valeur de C maximale. Pour cette raison, les positions des PIs de ce dernier sont plus proches des positions du détecteur SC_HK_C que celles des autres versions combinées.

❖ Comparaison avec des détecteurs de l'état de l'art:

Nous présentons le résultat des détecteurs Harris_fract, Harris_clust, SURF, SI et SC_HK_FQ pour 4 vues de l'objet Ange dans la Figure 3-25. La particularité de cet objet est le fort degré de détails

présents sur sa forme. Pour les deux versions de Harris : avec la sélection d'une fraction de PIs parmi les plus grandes réponses de Harris et avec le regroupement par clustering, le résultat est très différent. Pour le premier, le Harris_fract, les positions sont condensées et situées sur des zones de transitions locales (presque au niveau des contours de la forme). Quant au Harris_clust, les PIs sont répartis sur toute la surface. Néanmoins, la répétabilité de ces deux détecteurs entre les différentes vues est notable et les positions des PIs sont saillantes avec une variation locale distincte de la forme.

Dans notre cas, nous avons privilégié avoir des PIs dispersés sur la surface, que des PIs condensés et collés les uns aux autres. Cette dispersion garantit la prise en compte d'un maximum de détails sur la forme et réduit la redondance dans la représentation à sauvegarder. Il en résulte la nette stabilité des PIs du détecteur combiné SC_HK_FQ sur les différentes vues. En outre, la nature du type de l'objet testé est adéquate pour le critère de notre détection (caractère courbée). En effet, la surface de l'Ange présente des courbures prononcées, ce qui augmente le nombre de PIs détectés.

Par ailleurs, en dépit du faible nombre de PIs détectés par SI, ces positions restent relativement répétables. Pareillement que notre détecteur, ce dernier extrait des extremums en mesure de courbure (indice de forme).

La remarque à donner concernant le détecteur SURF est que quelques positions de PIs se situent à l'extérieur de l'objet. La raison est que ce dernier utilise, contrairement aux autres détecteurs, une information volumique (à base de voxels obtenus par la répartition de l'intersection d'un cube avec la surface de l'objet). Cependant, l'avantage de ce détecteur est l'attribution d'une échelle (représentée par le rayon de la sphère sur la figure) aux PIs.

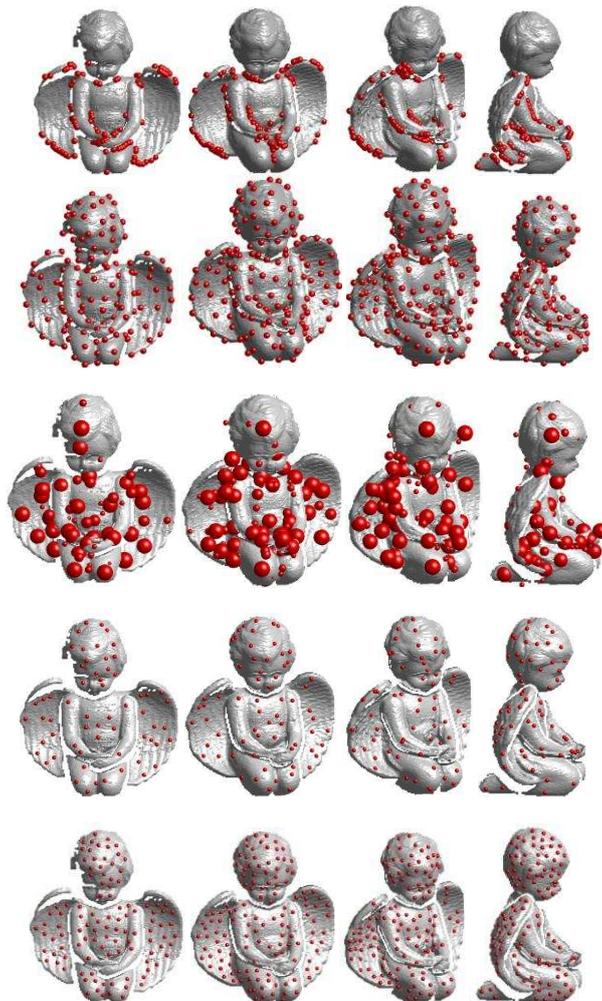


Figure 3-25 Positions des PIs détectés par respectivement de haut vers les bas : Harris_fract, Harris_clust, SURF, SI et SC_HK_FQ pour les angles (0°, 20°, 40° et 80°)

❖ Présence de bruit:

Les résultats sur la base de Carotte présente un autre type d'objets avec des conditions plus réelles. Malgré le bruit existant dans les données de la Kinect, nos algorithmes combinés réussissent à extraire des PIs situés sur des les zones caractéristiques de la forme de l'objet. Par exemple sur la Figure 3-26 des objets de la base Carotte, les PIs sont concentrés sur les plis au niveau du sac de la poubelle, sur les coins, les bords et le bouchon pour l'objet bidon, et sur le support et les pales du ventilateur. Ces positions saillantes sont relativement stables et se répètent sur les différentes vues du ventilateur.

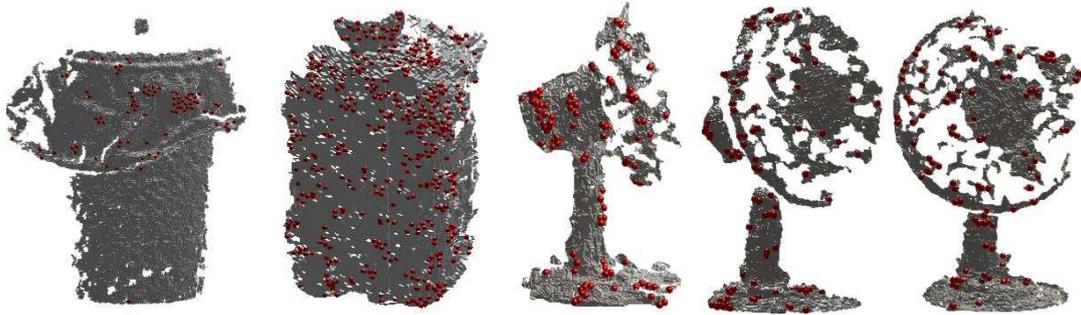


Figure 3-26 Résultat de détection des PIs sur les objets de la base Carotte avec respectivement de gauche à droite: les détecteurs SC_HK_FQ (poubelle), SC_HK_C (bidon) et SC_HK_Conf (ventilateur)

❖ Objets mécaniques:

L'illustration de la Figure 3-27 montre le résultat de détection de l'algorithme combiné SC_HK_C sur des objets de types mécaniques pour les modèles agfa, machine et auto de la base Stuttgart. Ces surfaces sont à dominance de surfaces planes. Comme la sélection basée sur les critères SC et HK ne prend pas en compte ces zones, nous avons rajouté le type plan dans les types de surfaces sélectionnées. N'ayant pas un critère de tri sur ce type de surface, néanmoins, le filtrage utilisant une condition sur le seuil de la distance spatiale entre les PIs permet d'avoir des PIs répartis sur la surface de l'objet. L'inconvénient de cet approche est que le nombre de PIs peut être grand (~400 PIs) pour des modèles à grande densité de points comme le cas des objets de la base Stuttgart. L'information prélevée sur ces PIs est redondante et le risque de confondre ces PIs entre eux augmente, au moment mise en correspondance.



Figure 3-27 Positions des PIs détectés par le détecteur SC_HK_C sur des objets de type mécanique pour les modèles agfa, machine et auto de la base Stuttgart

❖ Similarité entre les objets:

Parmi les contraintes d'une méthode locale, est que l'extraction de PIs sur des objets de forme très similaire donne les mêmes positions physiques. L'illustration de la Figure 3-28 appuie cette remarque. En effet, les vues partielles des objets oignon, tomate et orange se ressemblent fortement. Le bruit des données est réparti de la même façon sur la surface des objets et renforce cette similitude. Cette limitation du détecteur peut être surpassée par une définition d'un descripteur distinguant ces points. Par exemple en utilisant la texture de la surface, nous parvenons à différencier ces objets comme le cas de la couleur des modèles oignon, tomate et orange.

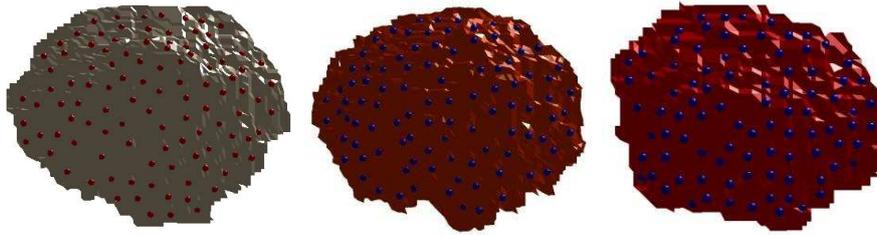


Figure 3-28 Les PIs détectés sur les objets: oignon, orange et tomate de la base RGB-D montrant la forte similarité de la forme. Une couleur de la surface est rajoutée pour différencier ces objets.

❖ **Version multi-échelles :**

Nous avons pu entamer l'implémentation de l'approche multi-échelle décrite précédemment. Nous présentons quelques figures (Figure 3-29, Figure 3-30 et Figure 3-31) illustrant des résultats de ce détecteur. D'après la Figure 3-29, l'étendu des composantes connexes dans la version multi - échelles est réduit par rapport à la version mono-échelle et le nombre des PIs finaux est largement plus petit.

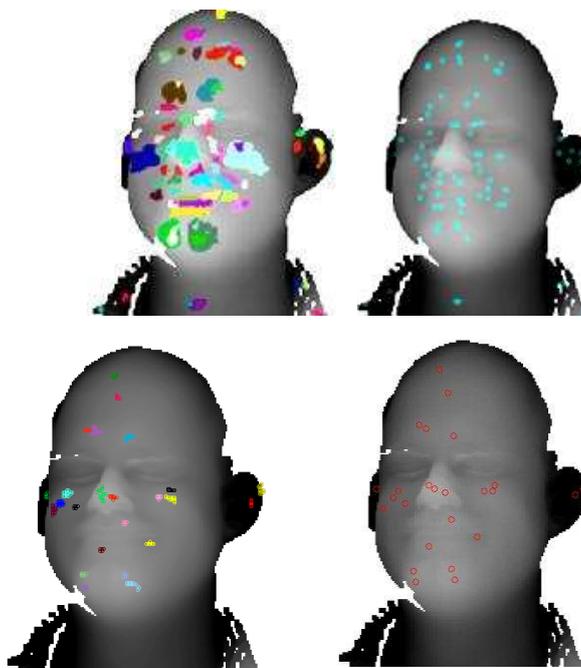


Figure 3-29 Résultat de la détection en échelle fixe (première ligne) et en multi échelles (deuxième ligne) pour le détecteur SC_HK_Con. La première colonne représente l'étendu des composantes connexes et la deuxième les positions des PIs finaux.

D'après la Figure 3-32 illustrant, sur des images de l'indice de forme, le résultat du processus de détection par l'algorithme SC_HK_Con sur trois tailles croissantes de voisinage. Les mesures de l'indice de forme représentées sont calculées sur ces trois niveaux de voisinages. Nous remarquons que les PIs détectés sont concentrés sur les mêmes régions de l'objet de visage (la bouche, le nez, les yeux et les oreilles). Outre cette stabilité à l'échelle, nous avons constaté que les figures (Figure 3-30, Figure 3-31) mettent l'accent sur une forte répétabilité des PIs pour des variations d'angle de vues. Par

exemple, nous avons compté 18 PIs qui sont répétables pour l'approche SC_HK_Con multi-échelle par rapport à 15 PIs répétables pour l'approche de Sparse points introduite par (Castellani, et al., 2008).

Bien que prometteuse pour le peu de tests effectués, l'évaluation de cette méthode n'a pas pu être poussée plus loin, par contrainte de temps.

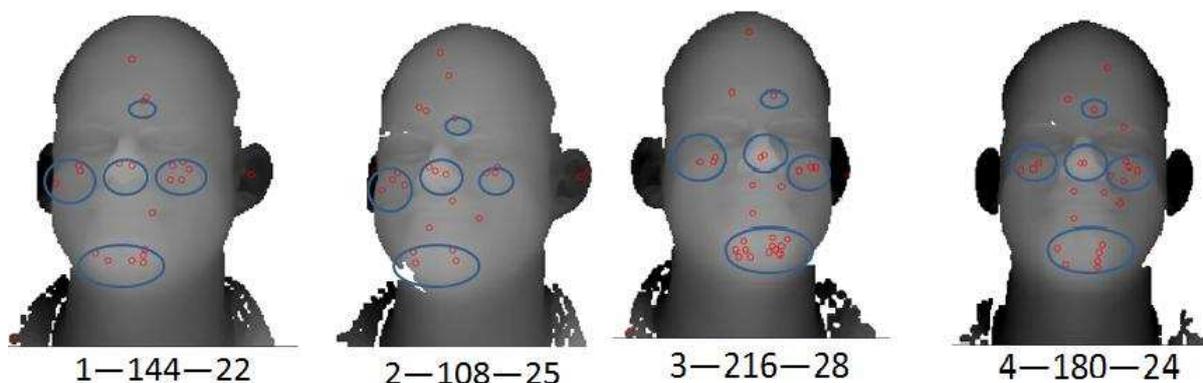


Figure 3-30 Invariance aux angles de vues pour l'approche SC_HK_Con multi-échelle. Les nombres représentent respectivement : l'ordre de la vue, le numéro de la vue dans la base Minolta et le nombre des PIs.

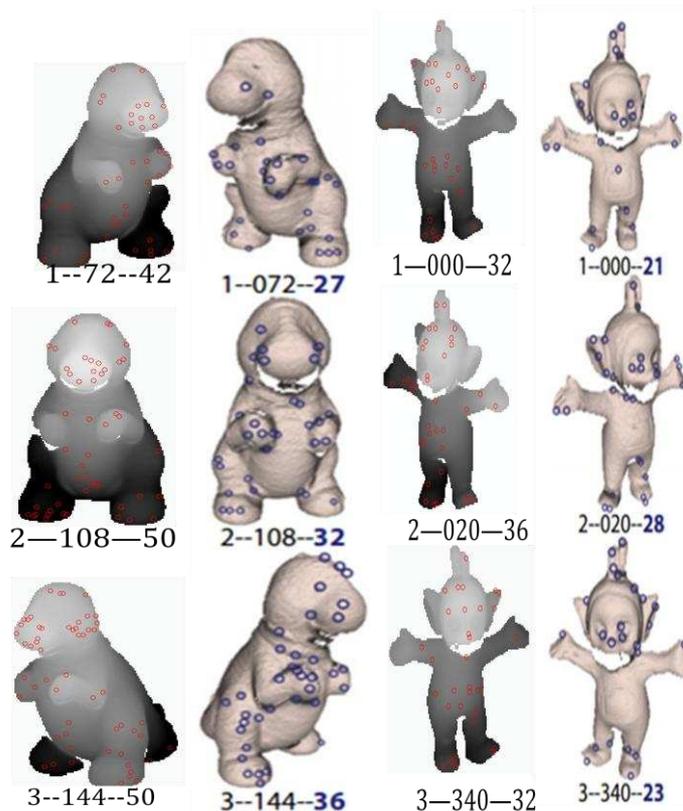


Figure 3-31 18 Résultats de comparaison de l'approche SC_HK_Con multi-échelle par rapport à l'approche de (Castellani, et al., 2008)

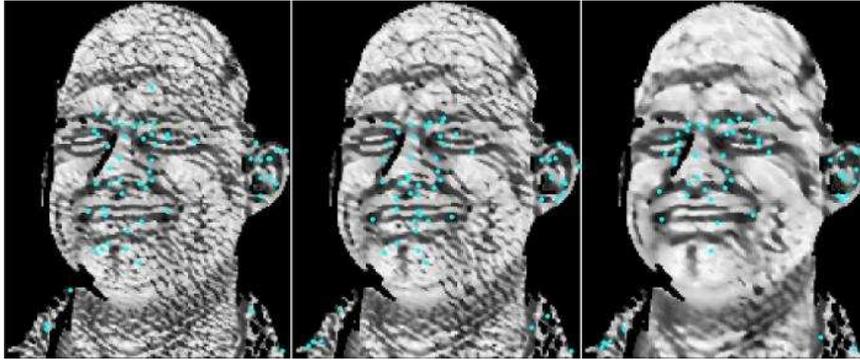


Figure 3-32 Représentation du résultat de détection en utilisant trois niveaux de voisinages sur les images d'indice de forme.

La comparaison visuelle de cette partie reste approximative et n'évalue pas la réelle performance des détecteurs. Pour le paragraphe suivant, nous introduisons une analyse quantitative de la répétabilité de ces détecteurs aux transformations géométriques de rotation et de changement d'échelle, et de leur robustesse au bruit et à l'échantillonnage. Cette étude comparative est faite uniquement sur les objets de la base Minolta.

3.3.2.2 Stabilité aux changements d'angles

❖ Description du protocole de test :

Le critère d'évaluation des expérimentations est la répétabilité des points d'intérêt. Etant donné un objet O et une fonction de transformation T , qui peut être une translation, une rotation, un changement d'échelle ou un ajout de bruit, $T(O)$ est l'objet après l'application de la transformation. Soit aussi P_O l'ensemble des points d'intérêt extraits sur O . Ainsi, la répétabilité est définie par:

$$R_{O,T(O)} = \frac{|P_O \cap P_{T(O)}|}{|P_O|}$$

Un seuil sur la distance entre les points détectés sur $T(O)$ et les points détectés sur O après les avoir transformés avec T permet de compter le nombre de points répétables. La quantification de cette répétabilité se fait en comparant la distance physique des points appariés à un seuil de distance. La définition de ce seuil décide de la correspondance ou pas entre les positions des points: si la distance entre deux points est inférieure au seuil fixé, le nombre de points répétables est incrémenté.

Une première façon d'évaluer la répétabilité est de fixer le seuil de la distance et calculer la répétabilité entre le modèle original et l'objet transformé, et puis faire la moyenne pour chaque objet de la collection. Finalement, la répétabilité correspond à la valeur médiane de toutes les valeurs de répétabilité obtenues pour tous les objets de la base.

Une deuxième évaluation est faite en traçant la courbe médiane des différentes mesures de répétabilité en fonction d'une variation du seuil de distance.

Pour avoir une normalisation dans la comparaison des différents détecteurs, nous pouvons soit faire le rapport entre le premier et le deuxième plus proche voisin, soit, le rapport entre la distance du premier plus proche voisin et une distance maximale calculée sur toute la base et pour tous les détecteurs. Nous avons suggéré que la deuxième normalisation est plus adéquate pour notre comparaison, vu qu'elle permet d'avoir la même normalisation pour tous les détecteurs et surtout ne pénalise pas les détecteurs qui sortent des points dispersés sur la surface de l'objet.

Les paramètres de chaque méthode sont choisis empiriquement d'une façon à avoir des nombres similaires de PIs et pouvoir effectuer des comparaisons de répétabilité au même ordre. Par exemple pour le détecteur SI, nous jouons sur les paramètres alpha et beta pour varier le nombre de PIs. Pareillement, nous ajustons les entrées et les seuils des algorithmes des détecteurs Harris et SURF.

Les seuils H_{zero} , K_{zero} et C_{zero} sont choisis tel que $K_{zero} = H_{zero}^2$ et $C_{zero} = K_{zero}$ comme proposé par (Akagündüz, 2011). Nous avons sélectionné empiriquement pour ces seuils respectivement les valeurs 0.01, 0.0001 et 0.01.

Pour les expérimentations sur le détecteur de Harris, les valeurs du paramètre de Harris est $K = 0.04$ et la valeur du paramètre de sélection finale est égale à 0.01 pour la méthode de Harris_fract et égale à 0.006 pour la méthode de Harris_clust.

Sur les 9 objets de la base Minolta, nous appliquons les détecteurs sur les 18 vues de chaque objet. Nous avons au total 17 transformations. La courbe de répétabilité des PIs entre les différentes vues pour les 10 détecteurs : SC_HK_FQ, SC_HK_C, SC_HK_Conf, SC_HK_Con, Harris_fract, Harris_clust, SI, SC, HK et SURF, est présentée dans la Figure 3-33. Seulement la partie gauche des graphes obtenus est représentée (plus importante car ce qui nous intéresse est l'existence de points homologues très proches entre vue originale et vue transformée).

❖ Résultats :

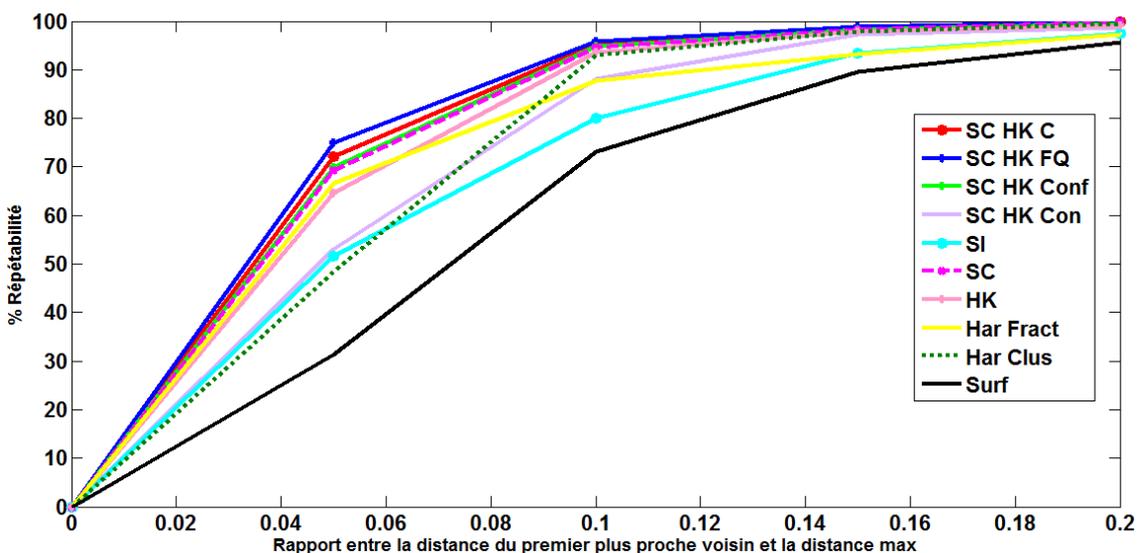


Figure 3-33 Répétabilité des PIs pour les 10 détecteurs : SC_HK_FQ, SC_HK_C, SC_HK_Conf, SC_HK_C_con, Harris_fract, Harris_clust, SI, SC, HK et SURF sur les 9 objets de la base de Minolta.

La performance de la répétabilité des trois détecteurs combinés SC_HK_C, SC_HK_Conf et SC_HK_FQ est comparable pour une variation d'angles de vue. La courbe de SC_HK_FQ est légèrement au dessus de celles des autres détecteurs. Les explications données, dans la partie évaluation visuelle, sont valables pour justifier ce comportement.

L'exception de la courbe du SURF qui est bien au dessous des autres est à noter. Nous pensons que peut être le fait d'ajuster les paramètres pour avoir un nombre de PIs comparables, a pu faire baisser la performance de ce détecteur.

Le Harris dans sa version de sélection par fraction est meilleur que la version Harris_clust pour les faibles valeurs du seuil. Ce comportement s'inverse à partir de la valeur 0.09 du seuil.

D'après le Tableau 3-1, pour une valeur du seuil fixée à 0.16, les cinq premières répétabilités se présentent comme suit : 98.8% pour SC_HK_FQ, 98.5% pour SC_HK_C, 98.2% pour SC, 98.2% pour SC_HK_Conf et 97.9% pour Harris_clust.

L'invariance des mesures S, C, H et K à la rotation explique la haute répétabilité des détecteurs basés sur ces estimateurs. Le processus de combinaison de critères a augmenté considérablement cette répétabilité en comparant aux résultats des détecteurs basés sur uniquement un de ces critères. Les PIs regroupés par composantes connexes sont moins stables que ceux ordonnés et regroupés par clustering. La répartition de ces points sur toute la forme par la méthode de clustering, par rapport à une concentration sur certaines zones obtenue par les composantes connexes, peut expliquer ce léger avantage. Un comportement inverse est observé avec le détecteur Harris. La sélection d'une fraction des points après les avoir ordonnés selon leur réponse d'Harris est plus stable qu'une sélection de PIs par clustering du moins pour les petites valeurs du seuil.

D'autre part, la version combinée SC_HK_C_Con a une répétabilité largement plus faible que les autres détecteurs combinés (pour des seuils de distances petits). Ce détecteur ne sera pas pris en compte dans le reste des évaluations.

Seuil	0.06	0.16
SI	51.7%	93.3%
SC_HK_FQ	74.9%	98.8%
SC_HK_C	72.0%	98.5%
SC_HK_Conf	69.8%	98.2%
SC_HK_Con	54.0%	97.2%
HK	64.6%	97.8%
SC	69.3%	98.2%
Harris_fract	66.6%	93.2%
Harris_clust	48.4%	97.9%
SURF	31.4%	89.6%

Tableau 3-1 Valeurs de répétabilité, à deux seuils de distance, entre les angles de vues pour les 10 détecteurs

3.3.2.3 Stabilité à l'échelle

❖ Description du protocole de test :

Base à variation *artificielle* d'échelles

Dans un premier temps de nos expérimentations, nous proposons de générer des vues avec une variation d'échelles faite manuellement. Cette transformation correspond à une variation de la densité des points sur le maillage. Pour ce faire, nous appliquons une transformation de décimation (qui réduit

le nombre de sommet sur le modèle) avec un facteur de réduction f égal à 0.3 sur les 18 vues de chacun des 9 objets Minolta. Nous disposons alors de 18 nouvelles vues pour chaque objet. Une troisième échelle est obtenue par une réduction de la densité des points dans les modèles d'origines avec un taux de 50% du nombre de points initial. Au final, nous constituons notre base avec trois variations d'échelle : les vues originales (échelle 0), les vues de densité réduite de 30% (échelle 1) et les vues de densité réduite de 50% (échelle 2). Un exemple de décimation est donné dans la Figure 3-34. Pour chaque objet, nous disposons de 54 vues au total.

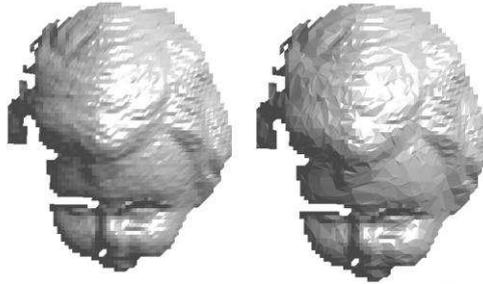


Figure 3-34 Exemple de décimation avec un facteur $f = 50\%$ - à gauche le modèle original et à droite le modèle après décimation

Sur les 9 objets de la base Minolta, nous appliquons les détecteurs sur les 18 vues de chaque objet, et à chaque échelle. Nous avons considéré les 2 transformations de changement d'échelles suivantes : la première, **Ech_1**, s'applique aux 18 vues de l'échelle 1 vers les 18 vues de l'échelle 0 et la deuxième, **Ech_2**, transforme les PIs des 18 vues de l'échelle 2 vers les vues de l'échelle 1. Les courbes de répétabilité des PIs entre ces deux transformations pour les 9 détecteurs : SC_HK_FQ, SC_HK_C, SC_HK_Conf, Harris_fract, Harris_clust, SI, SC, HK et SURF, est présentées dans les figures Figure 3-35 et Figure 3-36. Pour mieux visualiser l'écart entre ces courbes, nous avons normalisé la distance des PIs, cette fois ci, par la distance moyenne de toutes les distances de la base et des détecteurs (au lieu de prendre la distance maximale).

Base à variation réelle d'échelles

Dans un deuxième temps de notre évaluation, les données des 7 objets de la base Minolta, présentant une variation de vues et d'échelles sont utilisées. Pour chaque objet, nous considérons les 11 vues sur 2 échelles choisies parmi les trois ou quatre échelles de cet objet.

Nous appliquons les 9 détecteurs sur les 11 vues de chaque échelle pour tous les objets. La courbe de répétabilité des PIs entre ces échelles pour ces 9 détecteurs est illustrée dans la Figure 3-37.

La Figure 3-38 montre les positions des PIs extraits par les détecteurs Harris_fract, Harris_clust, SC_HK_C et Surf pour la même vue et sur deux échelles différentes de l'objet *gc_bottle*.

❖ **Résultats :**

Les remarques générales sur ces trois courbes sont comme suit :

- Entre les deux transformations artificielles de l'échelle, nous constatons la même allure des courbes des détecteurs. L'échelle 2, où les points sont réduits à 50%, simule une distance d'acquisition plus éloignée que celle de l'échelle 1, dont la décimation est de 30% du nombre initial des points. De ce fait, la répétabilité de tous les détecteurs baisse entre la transformation de l'échelle la plus proche (Ech_1) vers la transformation la plus éloignée (Ech_2).
- Entre les transformations artificielles et la transformation réelle de l'échelle, le comportement des détecteurs SURF et Harris_clust change. Au plus faible seuil de distance, le Surf et le

Harris_fract, suivis par SI ont la meilleure performance sur les transformations artificielles. Sur la variation réelle d'échelle, c'est le Harris_fract, Harris_clust et SI qui l'emportent.

- Étonnamment, sur le changement d'échelle des données réelles, les résultats montrent que le Surf, qui est une approche multi-échelle, a clairement une faible répétabilité. Nous pensons que le faible nombre de points détectés sur les vues peut justifier ce résultat. Un paramétrage adéquat pour augmenter le nombre de PI donnera probablement un meilleur résultat. Cependant, la réelle valeur de ce détecteur peut résider dans le caractère descriptif exprimé par le voisinage des PIs (grâce à l'attribution d'une échelle adéquate à chaque PI), que nous étudierons dans la phase de description et de reconnaissance qui suivent.
- En variation réelle d'échelles, les PIs des détecteurs SC_HK_C, SC_HK_FQ, SC et HK se répètent de la même façon. Nous remarquons une légère supériorité de SC_HK_C par rapport aux deux autres détecteurs combinés pour les plus faibles seuils de distance. Le processus de combinaison, n'améliore pas la performance des détecteurs simples, en changement d'échelles.

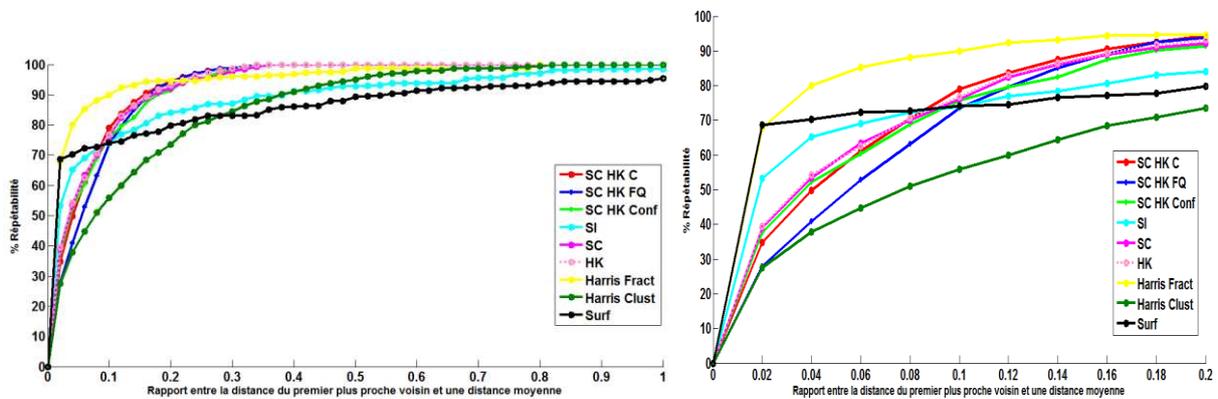


Figure 3-35 Répétabilité des 9 détecteurs pour la transformation d'échelle Ech_1- A droite-Zoom sur la partie gauche de la première courbe

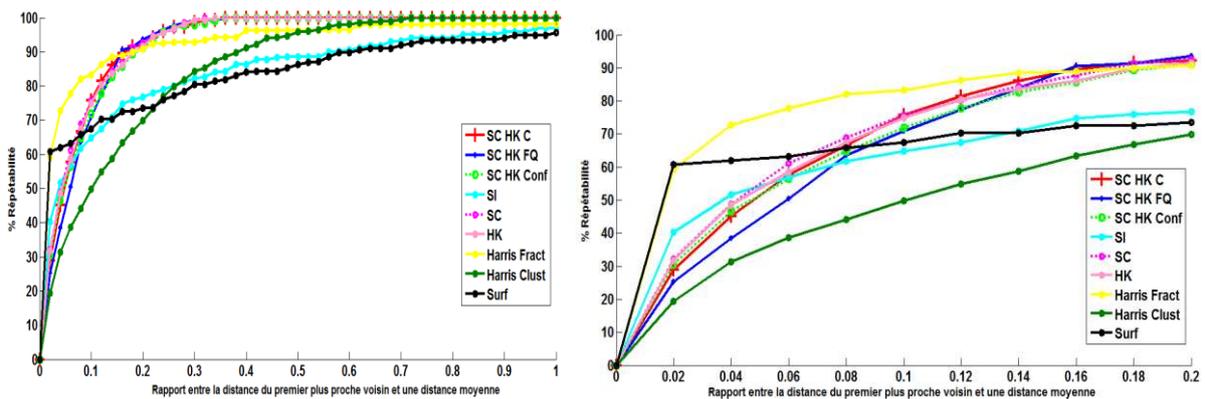


Figure 3-36 Répétabilité des 9 détecteurs pour la transformation d'échelle Ech_2 -A droite-Zoom sur la partie gauche de la première courbe

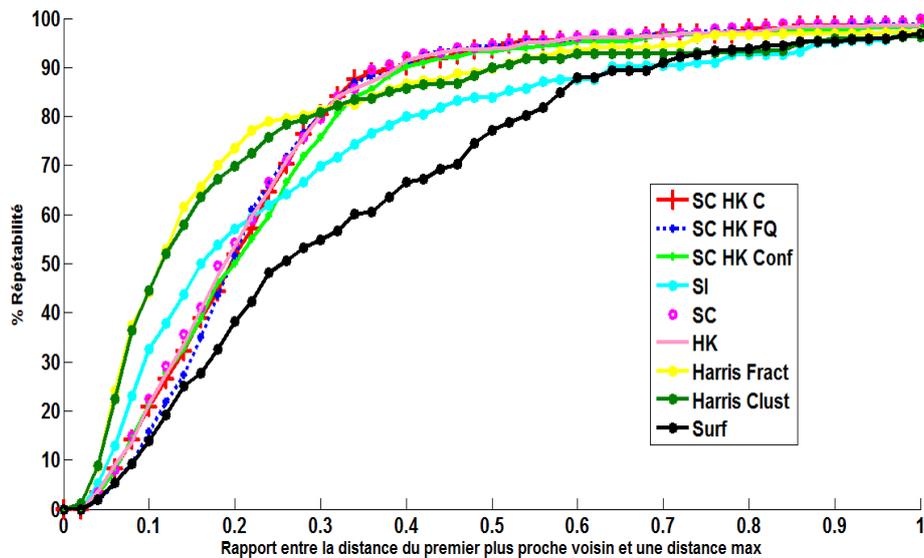


Figure 3-37 Répétabilité des 9 détecteurs SC_HK_FQ, SC_HK_C, SC_HK_Conf, SI, SC, HK, Harris_fract, Harris_clust et SURF sur 11 vues de deux échelles réelles

En variation d'échelles, à un seuil fixé à la valeur 0.4, où la majorité des détecteurs ont une performance supérieure à 80%, nous notons les performances de stabilité suivantes : 89.7% pour SC_HK_FQ, 89.7% pour SC_HK_C, 87.8% pour SC_HK_Conf, 83.9% pour Harris_fract, 81.9% pour Harris_clust, 78.1% pour SI et 62.0% pour SURF. Ces résultats nous emmènent à conclure, concernant l'invariance à l'échelle, sur la stabilité des détecteurs proposés et leur performance qui reste comparable à celle des détecteurs de l'état de l'art.

Le choix adaptatif du voisinage (avec une valeur proportionnelle à la taille de l'objet) nécessaire pour le calcul des courbures et des mesures de saillance est l'une des raisons de cette stabilité. Ce paramètre est le seul facteur, prenant en compte la robustesse aux changements d'échelles dans la version mono-échelle de notre algorithme. Nous pensons que la performance de nos détecteurs pourrait éventuellement s'améliorer avec la version multi-échelle.

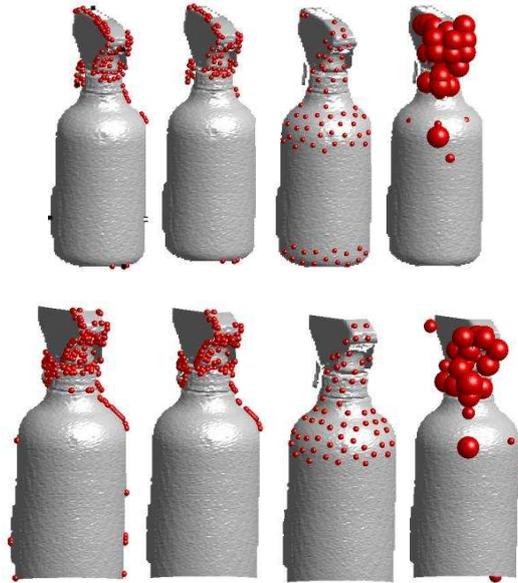


Figure 3-38 Répétabilité des PIs extraits par les détecteurs Harris_fract, Harris_clust, SC_HK_C et Surf dans l'ordre de gauche à droite, pour la même vue sur deux échelles : échelle 1 en première ligne et échelle 2 en deuxième ligne.

Dans notre évaluation de la répétabilité, nous cherchons des homologues des PIs des vues de l'échelle 1 se trouvant dans les vues de l'échelle 2 qui est plus proche du capteur. Or, comme le montre la Figure 3-38, la vue de l'échelle 2 ne présente qu'une portion de la surface de l'objet. Notre détecteur, SC_HK_C, détecte des points sur la vue de l'échelle 1 qui n'ont pas de correspondants réels sur la vue de l'échelle 2. La distance d'appariement de ces points est alors grande, c'est qui justifie l'infériorité de la performance de notre détecteur par rapport à celle du Harris_fract, par exemple. Nous pensons, que la transformation inverse, c'est-à-dire de l'échelle 2 l'échelle 1, fera augmenter la répétabilité pour nos détecteurs. Point de vue pratique, dans notre appariement des PIs pour la reconnaissance d'objet, nous suggérons de prendre en considération cette remarque. Nous proposons de fixer une distance seuil (sur la distance physique des PIs) à partir de laquelle nous décidons de valider ou pas la mise en correspondance d'une paire de PIs.

3.3.2.4 Stabilité au bruit

❖ Description du protocole de test :

L'évaluation de l'impact du bruit sur le résultat de détection se fait par l'ajout d'un bruit gaussien au maillage initial. Puis, nous mesurons la répétabilité des PIs détectés sur le nouveau maillage bruité par rapport à ceux détectés dans le maillage initial. Pareillement qu'en rotation, la comparaison des répétabilités des détecteurs se fait par la variation d'un seuil de la distance des PIs normalisée par une distance maximale calculée sur toute la base. Un bruit gaussien de variance $0.2 \times MR$ (résolution du maillage) est appliqué à 5 vues de chacun des 9 objets Minolta. Nous présentons la Figure 3-39, les courbes de la médiane des valeurs de répétabilité de chacun des 9 détecteurs: SC_HK_FQ, SC_HK_C, SC_HK_Conf, Harris_fract, Harris_clust, SI, SC, HK et SURF. Dans le Tableau 3-2, nous montrons le pourcentage de répétabilité pour quelques détecteurs en fixant deux seuils du rapport des distances entre les PIs.

❖ Résultats :

Le bruit fausse les valeurs initiales des profondeurs des points et par conséquent la mesure de courbure est bruitée, surtout si la taille du voisinage est petite. Pour les deux détecteurs de la Figure 3-40, le nombre de PIs augmente sur les vues bruitées par rapport aux vues initiales.

Le Harris_fract donne nettement la meilleure répétabilité pour les petites valeurs du seuil des distances. Nous expliquons ce résultat par la forte concentration de ses PIs sur certaines zones de la surface. La distance entre ces PIs est, de ce fait, très faible que celle pour un détecteur qui extrait des points éparpillés sur la surface.

En général, pour l'ajout d'un bruit, les détecteurs, utilisant les espaces SC et/ou HK, ont une stabilité comparable. Nous notons, particulièrement, que la répétabilité de nos détecteurs combinés est rapidement au dessus de 90%, dès la valeur 0.06 du seuil correspondant au ratio des distances des PIs sur la distance maximale. A la valeur 0.12% du seuil, les trois détecteurs combinés ont 100% de répétabilité. La performance légèrement supérieure de SC_HK_FQ peut s'expliquer par l'implication des valeurs de courbures de son voisinage dans le calcul de son facteur de qualité dans la phase du tri. Il en résulte que ce facteur renseigne sur une plus large zone spatiale que celle par exemple des PIs SC_HK_C pour lesquels la saillance implique la mesure de courbure C du PI uniquement. La sélection finale par regroupement permet également d'obtenir des positions réparties sur toute la forme de l'objet et augmente la possibilité d'approcher plus de points sur la surface de l'objet.

Par ailleurs, la stabilité de ces détecteurs est fortement liée à la robustesse du calcul des quantités différentielles (normale et courbure) sur des données bruitées. En effet, dans notre cas, pour chaque point du nuage de point, nous calculons ces quantités en appliquant la méthode de Monge sur le voisinage de ce point. Ainsi, la mesure sur le point porte la moyenne des contributions des points qui l'entourent. De ce fait, la taille du voisinage intervient dans cette stabilité. Pour un choix d'une taille faible du voisinage, les courbures seront sensibles aux perturbations des données. Par ailleurs, l'inconvénient majeur de cette méthode est le temps de calcul qui reste linéaire avec le nombre de points dans le nuage.

Pour des bases de données qui sont naturellement bruitées, nous suggérons d'appliquer un prétraitement de dé-bruitage avant d'appliquer nos algorithmes de détection, tel le cas des bases de la Kinect.

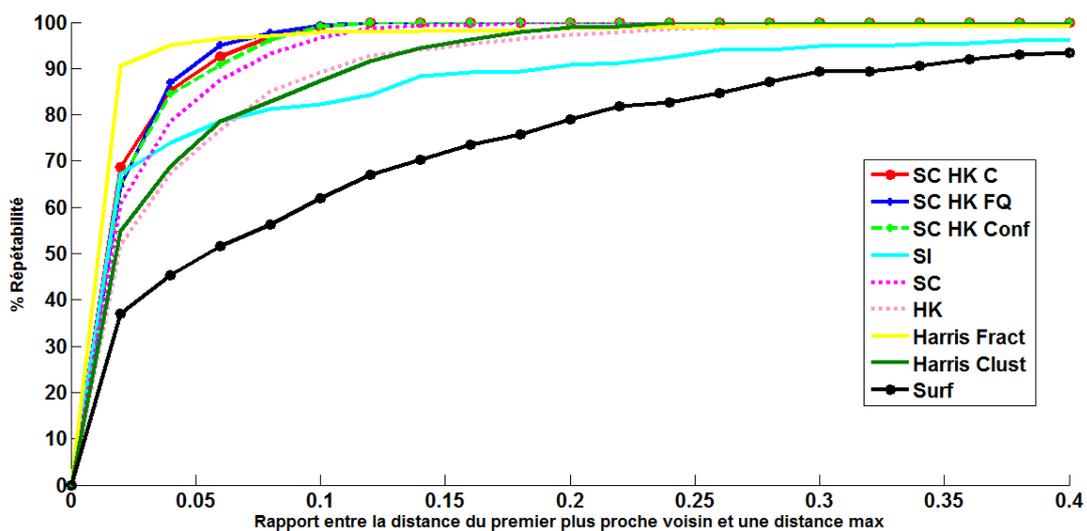


Figure 3-39 Courbes de répétabilité des 9 détecteurs testés sur les 9 objets Minolta et pour un bruit de variance 0.2 MR

Seuil	0.06	0.12
SI	78.7%	84.3%
SC_HK_FQ	94.9%	100%
SC_HK_C	92.6%	100%
SC_HK_Conf	90.8%	100%
Harris_fract	96.4%	97.9%
Harris_clust	78.5%	91.6%
SURF	51.5%	67.0%

Tableau 3-2 Valeurs de répétabilité pour un ajout de bruit gaussien de variance 0.2 MR au maillage

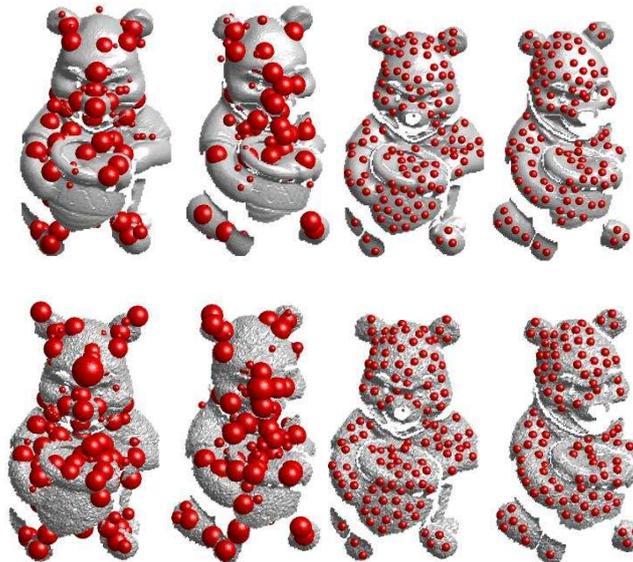


Figure 3-40 Répétabilité des PIs entre les deux vues 0° et 20° initiales (première ligne) et leur version bruitée (deuxième ligne) pour les détecteurs Surf (à gauche) et SC_HK_FQ (à droite)

3.3.2.5 Temps de calcul

Parmi les critères de la faisabilité de ces méthodes de détection et de représentation de la forme pour des applications temps réel, le cout du temps de calcul est principalement décisif. Dans notre cas, nous n'avons pas optimisé au maximum le cout d'exécution de nos algorithmes par contrainte de temps. Cependant, notre performance de calcul reste acceptable pour des applications s'exécutant en temps quasi réel et pour des données 3D pas très denses. Cependant, la précision et les détails de la forme peuvent manquer pour ce genre de données. Nous constatons que la tendance de la plupart des applications du domaine est d'utiliser des caméras 3D de faible résolution comme la Kinect par rapport à des scanners laser plus précis. Dans le Tableau 3-3, nous présentons le temps d'exécution des deux méthodes Monge et RF utilisées pour le calcul de courbures et des normales. L'évaluation est faite sur deux types de données (synthétique et réelle) de densité de points différente.

L'exécutable fournie par les auteurs du SURF 3D exécute la détection et la description en même temps. Dès lors, nous ne pouvons pas estimer le temps de calcul du détecteur SURF. Cette information sera présentée dans le chapitre descripteur. L'exécution du Harris se fait sous le système linux et nos algorithmes tournent sous le système Windows.

Les propriétés système utilisées correspondent à un processeur Intel CORE i7, 2.8GHz et 4Go de mémoire RAM.

❖ **Résultats :**

Il est à expliquer que dans nos algorithmes, nous calculons, dans l'étape de détection toutes les informations de quantités différentielles qui vont nous servir pour l'étape de détection et description à la fois. Les deux détecteurs Harris_Fract et Harris_Clust, fournissent juste les coordonnées des PIs et une information sur la réponse de Harris en ces points.

Densité	Algo	SC_HK_C	SC_HK_FQ	SC_HK_Conf	SI	Harris_Fract	Harris_Clust
Stuttgart	Monge	18.08	32.53	33.9	35.9	0.48	0.61
10948 points	RF	3.5	16.1	16.5	18.06		
Carotte	Monge	9.2	17.2	17	18.05	0.22	0.24
4685 points	RF	1.9	13.8	14.2	16.7		

Tableau 3-3 Temps de calcul des détecteurs en (s)

La remarque est que nos détecteurs sont gourmands en temps de calcul par rapport au Harris. L'exécution de la méthode de Monge (la plus précise en performance) est linéaire avec le nombre de points dans le nuage de points. Le SC_HK_C est le plus rapide des trois détecteurs implémentés puisque les voisins ne sont pas revisités, contrairement au cas de SC_HK_FQ et SC_HK_Conf. L'approche RF est nettement moins coûteuse en temps de calcul et peut, du coup, s'exécuter en temps quasi-réel. Cependant, sa précision dans le calcul des normales et courbures est moins bonne.

Il est clair que la linéarité de l'approche de Monge reste handicapante pour une exécution en temps quasi réel. Néanmoins, nous avons pu tester une alternative à cette contrainte, en sous-échantillonnant le nuage de points en des sous régions (patches). Ainsi, les mesures des courbures se font sur des patches au lieu des points. Le traitement de recherche de PIs saillants est remplacé par une recherche de patches saillants (en considérant la classe de la surface du patch (sur les espaces SC et HK) et un tri sur son indice de courbure (C ou FQ ou Conf)). Un PI est sélectionné à la position du point le plus proche du barycentre du patch. Par conséquent, le temps de calcul devient proportionnel au nombre de patch et non plus au nombre de points. Une légère perte d'information sur les détails de la forme, est détectée avec cette approche puisque nous décimons les points (seulement une proportion sera prise en compte (1 point sur 4)). Cette perte reste négligeable pour des données denses et pour un pas d'échantillonnage pas très grand. En revanche, nous n'avons pas eu le temps de faire une évaluation de la répétabilité et la robustesse des détecteurs avec les résultats de cette décimation.

3.4. Synthèse

Dans ce chapitre, nous avons, tout d'abord, présenté les principes de quelques détecteurs de points d'intérêt 3D sélectionnés de l'état de l'art. Ce choix n'est pas fait au hasard mais selon, d'un côté, l'idée d'avoir des détecteurs connus (du moins, à notre connaissance) pour pouvoir s'y comparer, et d'un autre côté, la considération des travaux qui nous ont inspiré dans notre démarche. Ensuite, nous avons exposé notre contribution, qui propose de combiner des critères de saillance adoptés dans ces méthodes. Avec quatre façons de sélectionner les PIs finaux après une première extraction, nous avons mis en place nos quatre nouveaux détecteurs combinés: SC_HK_C, SC_HK_FQ, SC_HK_Conf et SC_HK_Con. Nous avons enchainé avec une étude comparative de la performance de nos nouveaux détecteurs par rapport aux approches présentées. Suite aux expérimentations menées, le processus de combinaison des critères (SC_HK) a réussi à surpasser la performance des versions séparées SC et HK. Selon le critère d'évaluation, nos méthodes ont un comportement soit supérieure soit assez proche de celui des méthodes de l'état de l'art étudiées (Harris_3D, SI et SURF_3D).

En résumé, les résultats et les conclusions de cette évaluation sont comme suit :

- Invariance à l'angle de vue
 - **Nette supériorité de la répétabilité des détecteurs combinés, en particulier celle de SC_HK_FQ et SC_HK_C**
- Invariance à l'échelle
 - Performance de répétabilité de nos détecteurs supérieure au SURF et au Harris_Clust, mais inférieure au Harris_Fract
 - Proposition de fixer une distance seuil sur la distance physique des PIs mis en correspondance lors du processus de reconnaissance d'objet
- Robustesse au bruit
 - performance légèrement supérieure du Harris_fract et une meilleure stabilité de nos détecteurs combinés par rapport au reste des détecteurs
 - nécessité d'appliquer un prétraitement de dé-bruitage pour une meilleure performance sur les bases de la Kinect qui sont bruitées
- Temps de calcul de nos algorithmes encore à optimiser pour le cas de données très denses, surtout avec la technique la plus précise ; suggestion de l'idée du sous échantillonnage en considérant des patches comme base des mesures de saillance, qui permettrait de trouver un compromis entre perte des détails de la forme et optimisation du temps
- Proposition d'une version multi-échelle des détecteurs proposés. L'évaluation de cette extension est donnée en perspective pour des éventuels prochains travaux.

Il est à noter que ces résultats et conclusions sont donnés sur un ensemble de 9 objets de la base Minolta. Bien que le nombre de vues par objet (18 vues) soit acceptable pour notre évaluation, la taille de cet ensemble de test reste petite pour pouvoir faire des généralisations. En outre, la spécificité des objets de cette base est leur caractère naturel (animaux ou peluches). En effet, ces objets se distinguent par leurs détails de la forme et leurs courbures bien accentuées. Dès lors, le comportement et la performance des différents détecteurs est susceptible de changer pour de objets de type mécanique avec la dominance des formes géométriques planes (comme dans l'exemple de la Figure 3-27 illustrant la détection sur quelques objets de la base Stuttgart). A ce niveau, nous pouvons conclure

que ces résultats ont permis de constituer une première idée sur la performance des détecteurs proposés relativement à quelques détecteurs de l'état de l'art.

Certes, les critères d'évaluation, adoptés dans ce chapitre, ont permis de confirmer la stabilité et la robustesse des détecteurs proposés, néanmoins, la détection n'est que la première étape de la chaîne d'un système de reconnaissance. La réelle valeur de ces détecteurs se confirmera avec le complément d'une description du voisinage des PIs extraits. La continuité de ce traitement de description sera au cœur du prochain chapitre.