

Détection de points d'intérêt pour la mise en correspondance par propagation

Guillaume GALES¹Alain CROUZIL¹Sylvie CHAMBON²¹ Institut de Recherche en Informatique de Toulouse, IRIT² Laboratoire Central des Ponts et Chaussées, LCPC, Nantes¹ Université Paul Sabatier, 118 route de Narbonne, 31062 TOULOUSE cedex 9² route de Bouaye, BP 4129, 44341 BOUGUENNAIS cedex

{gales,crouzil}@irit.fr, chambon@lcpc.fr

Résumé

Dans le contexte de la stéréovision binoculaire, nous nous intéressons aux techniques de mise en correspondance de pixels par propagation de germes. Notre travail se situe plus précisément au niveau de la sélection de ces germes. Parmi les différentes façons de procéder, certaines font appel aux points d'intérêt. L'objectif de ces travaux est de proposer une étude portant sur dix détecteurs de points d'intérêt afin de déterminer lesquels sont les plus appropriés pour sélectionner les germes initiaux de la propagation. Pour cela, nous proposons d'adapter la mesure d'auto-corrélation de Moravec en utilisant une mesure de corrélation différente sur des fenêtres carrées et avec un voisinage plus grand. Différents classements sont établis à partir de mesures de répétabilité et de répartition spatiale effectuées sur des couples d'images de référence. Les tests réalisés permettent de donner des indications sur la stratégie de mise en correspondance qu'il est préférable de mettre en place afin de trouver des germes fiables et judicieusement répartis. Nous examinons ensuite la complémentarité des détecteurs les plus performants selon les critères précédents. Puis, nous regardons l'influence des détecteurs les plus complémentaires dans le contexte de la mise en correspondance stéréoscopique.

Mots Clef

Points d'intérêt, germes, propagation, stéréovision.

Abstract

In a binocular stereovision context, we focus on matching techniques based on seed propagation. Our work is about the seed selection step. Among all the different solutions, some are based on feature points. The goal of this work is to propose a study on ten feature point detectors in order to find out which ones are the most suitable to select the initial seed set for the propagation. In order to do so, we adapt the auto-correlation measure proposed by Moravec by using a different correlation measure over squared windows and by looking over a bigger neighborhood. Several rankings

are established measuring repeatability and spatial distribution computed from a set of reference stereo image pairs. This work gives indications about the most adequate strategy to select reliable and well distributed seeds. Then, we take a look at the complementarity of the most efficient detectors according to our criteria. Then, we take a look at the influence of the most complementary detectors in the stereo matching context.

Keywords

Feature points, seeds, propagation, stereovision.

1 Introduction

La mise en correspondance de pixels est une étape importante de la reconstruction 3D à partir de couples d'images stéréoscopiques. Elle consiste à retrouver pour chaque pixel de l'image de gauche, le pixel qui correspond à la même entité de la scène sur l'image de droite. Le déplacement entre un pixel et son correspondant est représenté par un vecteur de disparité. On distingue deux types de méthodes de mise en correspondance. Les méthodes globales, d'une part, minimisent une fonction de coût sur le déplacement des pixels entre l'image de gauche et celle de droite [22]. Ces méthodes sont complexes et coûteuses en temps de calcul. Les méthodes locales, d'autre part, recherchent les pixels qui se ressemblent le plus en utilisant des mesures de corrélation [3]. La recherche du correspondant est locale et s'effectue dans une zone de recherche. Mais plus celle-ci est grande, plus la méthode est coûteuse et plus la probabilité de faire une erreur d'appariement est élevée. Une méthode alternative a été introduite [20] et est généralement qualifiée de mise en correspondance par propagation de germes (ou *region-growing*). Elle s'appuie sur un ensemble d'appariements fiables et suppose que les voisins d'un germe ont la plupart du temps une disparité proche de celle du germe. La zone de recherche des correspondants est alors réduite à ce petit voisinage augmentant ainsi la fiabilité des appariements et la rapidité d'exécution. Les nouveaux appariements sont ensuite uti-

lisés comme de nouveaux germes et le processus est itéré jusqu'à l'obtention d'une carte de disparité dense ou quasi-dense [4, 6, 12, 14, 20]. La qualité du résultat final est donc conditionnée par le choix des germes initiaux. Ces germes doivent satisfaire au mieux les critères suivants :

- (C1) *bonne répartition* – On souhaite avoir au moins un germe par région de disparités homogènes.
- (C2) *appariements fiables* – On souhaite ne pas avoir des appariements de départ erronés.

Ce dernier critère n'est pas indispensable mais tout de même souhaitable lorsque la technique de propagation est basée sur une stratégie de « meilleur d'abord » qui la rend robuste aux erreurs de départ [12]. Pour trouver ces germes, on procède généralement en deux étapes successives :

1. sélection des pixels ;
2. appariement de ces pixels.

Dans cet article, nous étudions plus particulièrement la première étape. Dans les différents travaux qui proposent une technique de mise en correspondance par propagation de germes, nous pouvons distinguer les méthodes suivantes pour la sélection des germes :

1. sélection et mise en correspondance manuelle [26] ;
2. recherche des appariements les plus fiables sur un ensemble de pixels tirés au sort en divisant l'image en baquets [4] ;
3. mise en correspondance de points d'intérêt [12, 20].

La première méthode n'est pas fiable, n'est pas reproductible et est fastidieuse pour l'opérateur qui n'est pas infailible. La seconde permet une répartition homogène des germes sur l'image mais la recherche des correspondants se fait sur une large zone de recherche ce qui augmente le risque de choisir un mauvais correspondant. De plus, elle dépend du critère choisi pour caractériser cette fiabilité (mesure d'ambiguïté, mesure d'imprécision). La dernière méthode réalise une mise en correspondance entre deux ensembles de points d'intérêt car ils sont faciles à apparier et sont donc de bons candidats pour satisfaire les critères (C1) et (C2). C'est cette technique que nous avons décidé d'étudier plus en détails dans ces travaux.

L'objectif est de proposer des critères d'évaluations qui permettent de déterminer quel détecteur donne un ensemble de points d'intérêt satisfaisant les critères de sélection des candidats pour les germes. Nous proposons les mesures suivantes :

- *répétabilité* – Elle indique si un détecteur donne des points d'intérêt dont les correspondants dans l'autre image sont aussi détectés comme des points d'intérêt.
- *répartition* – Elle indique si la répartition spatiale des points d'intérêt permet une meilleure propagation.
- *complémentarité* – Elle indique si deux ensembles de points d'intérêt peuvent être associés pour satisfaire encore mieux les critères.

Dans un premier temps, nous décrivons brièvement les différents détecteurs de points d'intérêt utilisés dans nos tests

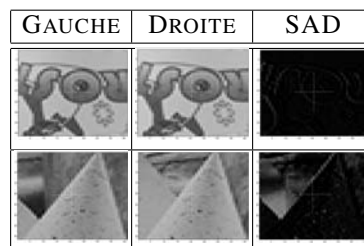


FIGURE 1 – Les imagettes montrées ici sont extraites du couple *cones*, voir figure 3. On montre le voisinage d'un pixel à gauche et celui de son correspondant à droite ainsi que leur différence en valeur absolue (SAD). Dans le premier cas, où il n'y a pas de discontinuité de profondeur, la différence est nulle. Dans le second cas, où il y a des discontinuités de profondeur, la différence n'est pas nulle car les niveaux de gris de l'arrière-plan sont différents.

ainsi que les modifications apportées à la mesure d'Auto-Corrélation proposée par Moravec. Dans un deuxième temps nous décrivons plus en détails les critères d'évaluations proposés. Enfin, nous analysons les résultats obtenus et notamment l'influence des détecteurs complémentaires sur la qualité de la mise en correspondance par propagation.

2 Détecteurs de points d'intérêt

2.1 Étapes de détection

Les points d'intérêt sont des points de l'image qui se distinguent des autres. Ils ont la propriété d'être plus identifiants que les autres pixels de l'image et donc ils sont plus faciles à suivre, à reconnaître ou, dans notre cas, à mettre en correspondance. Les différents détecteurs s'appuient généralement sur les étapes suivantes :

1. calcul pour chaque pixel d'une réponse basée sur les variations locales des niveaux de gris en utilisant directement les valeurs des intensités ou après avoir estimé leurs dérivées ;
2. suppression des non-maxima ;
3. post-traitement optionnel (suppression des fausses réponses, localisation au sous-pixel).

La première étape est celle qui caractérise un détecteur et c'est elle que nous décrivons dans la suite. Depuis 1977, de nombreux détecteurs ont été proposés. Certains ont pour objectif la détection de coins, d'autres sont plus généralement destinés à détecter des structures texturées en niveaux de gris. Nous avons choisi de travailler avec les détecteurs les plus utilisés : Harris, Moravec, Kitchen-Rosenfeld, Beaudet, SUSAN, FAST, SIFT, Harris-Laplace, Hessian-Laplace et Kadir. Nous proposons également une méthode d'auto-corrélation qui recherche les pixels de l'image qui ne ressemblent pas à leurs voisins.

2.2 Méthodes existantes

Nous présentons brièvement les différents détecteurs utilisés. On distingue deux types de détecteurs :

- « à échelle fixe » – La réponse est calculée avec une taille de fenêtre et une force de lissage gaussien prédéfinis.
- « multi-échelle » – La réponse est calculée avec des tailles de fenêtre (ou des résolutions d'image) et des forces de lissage gaussien différentes.

Ces derniers permettent de rendre le détecteur robuste aux changements d'échelle. Les couples stéréoscopiques ne présentent généralement pas de changement d'échelle entre l'image de gauche et l'image de droite. Toutefois, nous pensons que ces méthodes peuvent également permettre de détecter des points d'intérêt sur des structures de différentes tailles et c'est pourquoi nous les utilisons dans notre évaluation.

• Détecteurs à échelle fixe :

- *Harris* (HA) [9] – La réponse est calculée à partir de la matrice des moments du second ordre de chaque pixel. Cette matrice est calculée à partir des dérivées premières [5, 8, 9, 19, 24].
- *Moravec* (MO) [18] – La réponse est donnée par une mesure de la variance des niveaux de gris dans huit directions.
- *Kitchen-Rosenfeld* (KR) [11] – La réponse est déterminée par la courbure du gradient des niveaux de gris.
- *Beaudet* (BE) [2] – La réponse utilise les dérivées secondes (matrice hessienne).
- *SUSAN* (SU) [25] – La réponse mesure la surface des niveaux de gris de même intensité au voisinage du pixel considéré.
- *FAST* (FA) [21] – La réponse étudie la configuration des niveaux de gris des pixels situés sur le périmètre d'un cercle centré sur le point considéré.

• Détecteurs multi-échelle :

- *SIFT* (SI) [13] – La réponse de ce détecteur repose sur le laplacien des niveaux de gris obtenu par une différence de gaussiennes.
- *Harris-Laplace*, *Hessian-Laplace* (HL, HEL) [16] – La réponse est basée sur celle du détecteur de Harris ou Beaudet mais les points détectés doivent également être des maxima dans l'espace-échelle du laplacien.
- *SURF* [1] – La réponse est basée sur le déterminant de la matrice hessienne. Les points détectés doivent également être des maxima dans l'espace-échelle du déterminant de la matrice hessienne.
- *Kadir* (KA) [10] – La réponse repose sur l'entropie de l'histogramme des niveaux de gris au voisinage du point considéré.

2.3 Auto-Corrélation

L'objectif de ce détecteur Auto-Corrélation (A-C) est de sélectionner des germes qui pourront respecter au mieux le critère (C2). Une manière d'avoir de tels appariements consiste à sélectionner des points dans des zones non ambi-

guës, c'est-à-dire des points très différents de leurs voisins. Pour cela, nous proposons de modifier la mesure de Moravec. Pour chaque pixel, la réponse mesure la corrélation entre le voisinage du pixel considéré et celui de ses voisins, voir figure 2.

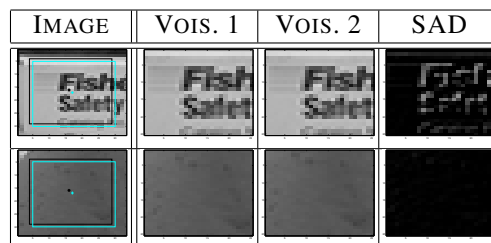


FIGURE 2 – Cette figure illustre le principe du détecteur A-C. L'exemple est montré sur une partie de l'image *cones*, voir figure 3. La première colonne montre la partie de l'image considérée. Le centre de la fenêtre noire est le pixel candidat dont le voisinage est montré dans la seconde colonne (VOIS. 1) et le centre de la fenêtre cyan est un voisin dont le voisinage est montré dans la troisième colonne (VOIS. 2). En pratique, on regarde tous les voisins et on fusionne les résultats. La dernière colonne montre l'image des différences des niveaux de gris (SAD). La première ligne est un exemple illustrant un candidat que l'on peut considérer comme un point d'intérêt contrairement à la seconde.

Formellement, la réponse d'un pixel $\mathbf{p}_{i,j}$ est donnée par :

$$R(\mathbf{p}_{i,j}) = \underset{(k,l) \in V(i,j)}{\text{fusion}} M(\mathbf{f}_1, \mathbf{f}_2) \quad (1)$$

où \mathbf{f}_1 et \mathbf{f}_2 sont des vecteurs qui contiennent respectivement le niveau de gris du pixel étudié $\mathbf{p}_{i,j}$ et celui d'un point $\mathbf{p}_{k,l}$ du voisinage, ainsi que ceux de leurs voisinages respectifs. Le terme M correspond à une mesure de dissimilarité et fusion est un opérateur de fusion des différents résultats de la mesure. Pour ce détecteur, il est nécessaire de choisir : le critère de similarité, la taille du voisinage et l'opérateur de fusion. Nos expérimentations nous ont conduit à choisir respectivement la somme des valeurs absolues des différences (SAD) avec une taille de fenêtre de corrélation 9×9 , la taille 9×9 pour le voisinage et l'opérateur minimum. Nous avons testé deux autres mesures de corrélations : la somme des différences au carré (SSD), comme dans le détecteur proposé par Moravec, et la corrélation croisée centrée normalisée (ZNCC). En effet, la mesure SAD donne une répétabilité supérieure (de l'ordre de 5%) à la mesure SSD. Les mesures normalisées de type ZNCC donnent des taux de répétabilité inférieur à SSD, la normalisation rendant difficile l'extraction des meilleurs maxima locaux. Pour chaque fenêtre les réponses sont normalisées et lors de l'extraction des maxima locaux, la comparaison de deux maxima n'est donc pas équitable. Ainsi, les différences par rapport au détecteur original sont les suivantes : les fenêtres sont carrées au lieu d'être rec-

tangulaires, la mesure de corrélation utilisée est différente et le voisinage examiné est plus important.

2.4 Évaluations existantes

Les critères suivants sont généralement utilisés pour l'évaluation des détecteurs de points d'intérêt :

Répétabilité. Lorsque l'on soumet à un détecteur de points d'intérêt deux images d'une même scène prise sous des angles (ou à des intervalles de temps) différents, il est souhaitable qu'il retourne pour chaque image les points correspondant aux mêmes entités de la scène. Dans ce cas, on dit que ces points sont « répétés ». Le taux de répétabilité d'un détecteur est le pourcentage de points répétés d'une image sur l'autre. Si ce taux est élevé lorsque la transformation entre les deux images est importante (rotation, changement d'échelle, changement de perspective, changement de luminosité), on dit que le détecteur est robuste à cette transformation [7, 8, 23].

Information content. Dans [23], les auteurs mesurent l'*information content*. Il s'agit de calculer dans quelle mesure les points d'intérêt d'une image sont différents les uns des autres. Ce critère est lié au descripteur choisi pour caractériser les points d'intérêt.

Complémentarité. Dans [15], la complémentarité est mesurée dans un contexte de reconnaissance d'objet. Une classification des points d'intérêt est effectuée et le critère que les auteurs proposent mesure le nombre de points d'intérêt issus de deux détecteurs différents que chaque classe possède. Dans l'idéal, chaque classe ne doit contenir que des points d'intérêt issus d'un même détecteur.

Taux de rappel et de précision. Dans [17], ces critères permettent de juger de la qualité de l'appariement des points d'intérêt. Ils sont liés au descripteur choisi pour les caractériser mais également à la stratégie de mise en correspondance utilisée. Le taux de rappel représente le rapport entre le nombre d'appariements corrects effectués et le nombre d'appariements à trouver. La précision représente le rapport entre les appariements corrects et les appariements réalisés.

3 Évaluation proposée

Pour cette évaluation, nous examinons la répétabilité et la répartition spatiale des points d'intérêt. Dans tous les calculs, nous considérons une marge de tolérance de ϵ pixels.

3.1 Répétabilité

Différentes variantes de cette mesure ont été proposées dans [7, 8, 23]. Nous mesurons la répétabilité pour déterminer le pourcentage de points d'intérêt qui ont une chance d'être mis en correspondance en utilisant ces définitions :

- *répétition* – Un point d'intérêt de l'image I_1 est répété si un point d'intérêt a été détecté dans l'image I_2 à une distance inférieure à ϵ pixels de son correspondant théorique.

- *taux de répétabilité* – Le taux de répétabilité d'un détecteur D pour un couple d'images (I_1, I_2) est donné par :

$$\text{rep}(I_1, I_2) = \frac{\text{rep}_D(I_1 \rightarrow I_2) + \text{rep}_D(I_2 \rightarrow I_1)}{2} \quad (2)$$

$$\text{avec } \text{rep}_D(A \rightarrow B) = \frac{\# \text{pdi de } A \text{ répétés dans } B}{\# \text{pdi de } A} \quad (3)$$

où $\# \text{pdi de } A$ est le nombre de points d'intérêt dans l'image A .

- *répétabilité* – La répétabilité rep_D d'un détecteur D est obtenue en effectuant la moyenne des taux de répétabilité trouvés pour ce détecteur sur l'ensemble des couples d'images testés.

3.2 Répartition spatiale des points d'intérêt

On propose de mesurer la répartition spatiale des points d'intérêt. On distingue la répartition des points d'intérêt dans les zones occultées et dans les zones de discontinuité de profondeur, d'une part, et la répartition par région, d'autre part.

Répartition dans les zones occultées et les zones de discontinuité de profondeur. Un détecteur ne peut pas « prédire » *a priori* quelles sont les zones occultées (ZO), par conséquent le pourcentage de points d'intérêt dans les ZO dépend des images et n'est donné qu'à titre indicatif. Le pourcentage de points d'intérêt dans les zones de discontinuité (ZD) est davantage révélateur du comportement du détecteur et constitue un indicateur utile dans le contexte de la mise en correspondance stéréoscopique. En effet, il est préférable de ne pas avoir de points d'intérêt au niveau des zones de discontinuité qui sont plus difficiles à mettre en correspondance. Un détecteur de points d'intérêt est conçu pour détecter les structures locales très texturées, typiquement les coins des objets de la scène. Or, certains détecteurs ont tendance à sélectionner des points d'intérêt sur les contours qui coïncident souvent avec les discontinuités de profondeur.

Répartition par région. Nous souhaitons obtenir au moins un germe dans chaque région de disparités homogènes. Si une région n'a pas de germe, la propagation peut ne pas se faire ou bien elle peut être initiée par une région voisine de profondeur différente, impliquant un risque de propager des erreurs. Les régions sont obtenues par segmentation couleur. En effet, nous supposons que des pixels voisins ayant des couleurs proches correspondent à des points de la scène situés à des profondeurs proches et donc à des disparités proches. Nous utilisons également ces régions lors de l'étape de propagation où nous contraignons les germes à ne se propager qu'à l'intérieur de la région à laquelle ils appartiennent, voir [6]. Dans l'idéal, on souhaite obtenir au moins un point d'intérêt dans chaque région. Pour un détecteur D , on mesure donc le rapport entre le nombre de régions qui contiennent au moins un point d'intérêt sur le nombre total de régions :

$$\text{RS}_D = \frac{\# \text{régions contenant au moins un pdi}}{\# \text{régions}} \quad (4)$$

Un rapport faible peut-être révélateur d'un manque de points d'intérêt ou bien d'une concentration des points d'intérêt sur une même zone.

3.3 Complémentarité

L'objectif de tous les détecteurs de points d'intérêt est commun : rechercher les points les plus singuliers de l'image. Mais nous pouvons nous poser la question suivante : tous les détecteurs de points d'intérêt retournent-ils les mêmes points ? Nous proposons une étude pour savoir dans quelle mesure différents détecteurs peuvent être complémentaires. Nous examinons l'apport de nouveaux points d'intérêt d'un détecteur par rapport à un autre, mais également le gain de répétabilité et de répartition spatiale de l'union de deux détecteurs.

Apport. Soit \mathcal{P}_{D_1} l'ensemble des points d'intérêt retourné par un détecteur D_1 et \mathcal{P}_{D_2} l'ensemble des points d'intérêt retourné par un détecteur D_2 . L'apport du détecteur D_2 sur le détecteur D_1 est mesuré par :

$$\text{apport}_{D_2|D_1} = \frac{\text{card}\{\mathcal{P}_{D_2}\} - \text{card}\{\mathcal{P}_{D_1} \cap \mathcal{P}_{D_2}\}}{\text{card}\{\mathcal{P}_{D_1}\}} \quad (5)$$

où l'intersection $\mathcal{P}_{D_1} \cap \mathcal{P}_{D_2}$ est établie en considérant que deux points d'intérêt sont identiques s'ils sont situés à une distance inférieure de ϵ pixels. Plus l'apport est élevé avec un ϵ important, plus les nouveaux points ajoutés par D_2 sont éloignés des points issus de D_1 .

Gain de répétabilité. La répétabilité est mesurée indépendamment pour les détecteurs D_1 et D_2 , voir équation (3). Le gain de répétabilité de l'union $D_1 \cup D_2$ des deux ensembles est déterminé par :

$$\text{gain}_{\text{rep}} = \text{rep}_{D_1 \cup D_2} - \max(\text{rep}_{D_1}, \text{rep}_{D_2}) \quad (6)$$

L'union est calculée en considérant que deux points d'intérêt sont identiques s'ils sont situés à une distance inférieure de ϵ pixels.

Gain de répartition spatiale. La répartition spatiale est mesurée indépendamment pour les détecteurs D_1 et D_2 , voir équation (4). Le gain de répartition spatiale de l'union $D_1 \cup D_2$ des deux ensembles est donnée par :

$$\text{gain}_{\text{RS}} = \text{RS}_{D_1 \cup D_2} - \max(\text{RS}_{D_1}, \text{RS}_{D_2}) \quad (7)$$

4 Résultats

4.1 Images de référence

Les couples stéréoscopiques utilisés dans nos tests proviennent du site de Middlebury¹. Ces paires sont largement utilisées par la communauté et sont caractéristiques des principaux problèmes rencontrés en stéréovision. Les images ont subi une rectification épipolaire et la vérité terrain, obtenue par une technique de lumière structurée, est fournie [22]. Les quatre couples de référence sur lesquels nous travaillons sont représentatifs de la base de données de Middlebury : *aloe*, *cones*, *dolls* et *teddy*, voir figure 3.

1. vision.middlebury.edu/stereo/data/



FIGURE 3 – Les différentes images utilisées pour nos tests : *aloe*, *cones*, *dolls*, *teddy*. Dans l'ordre, l'image de gauche et la carte de disparité théorique.

4.2 Paramètres des détecteurs

La littérature des points d'intérêt est très abondante et l'exhaustivité n'est pas envisageable. C'est pourquoi notre choix s'est orienté ici vers un représentant caractéristique de chaque grande famille. Nous avons récupéré les programmes des auteurs des détecteurs HL², HEL³, SU⁴, FA⁵, SI⁶ et KA⁷. Nous avons implémenté les détecteurs HA, MO, KR, BE et A-C.

Pour chaque détecteur, différents paramètres influent sur le résultat. Ces paramètres sont les suivants :

- *taille de la fenêtre pour le calcul de la réponse* (dans le cas des détecteurs à échelle fixe) – Plus cette taille est petite, plus le détecteur sélectionne les petites structures de l'image mais plus il est sensible au bruit. Inversement, plus cette taille est grande, plus le détecteur sélectionne les structures les plus importantes de l'image mais plus le lissage est important et plus les points d'intérêt détectés sont délocalisés.
- *taille de la fenêtre pour la suppression des non-maxima* – Plus cette taille est petite, plus on peut retourner des points d'intérêt proches les uns des autres.
- *seuil sur la réponse* – Un seuil sur la réponse du détecteur est souvent appliqué pour ne conserver que les meilleures réponses. Plus ce seuil est strict, moins les points d'intérêt sont nombreux mais ils sont *a priori* « meilleurs » au sens du détecteur. Dans le cas multi-échelle, la suppression des non-maxima prend en compte la réponse « courante » mais aussi les réponses des échelles suivantes et précédentes.

Chaque paramètre a donc son influence sur le résultat. Il serait intéressant de les faire varier mais ils sont trop nombreux. De plus, pour les détecteurs dont nous avons récupéré le code fourni par les auteurs, nous n'avons pas toujours accès à tous les réglages, ces derniers ayant fait des choix. En outre, il faudrait normaliser les valeurs des différents paramètres pour chaque détecteur afin de les rendre comparables. Par conséquent, face à ce problème, nous avons décidé de faire des choix et de fixer ces paramètres

2. robots.ox.ac.uk/~vgg/research/affine/
3. robots.ox.ac.uk/~vgg/research/affine/
4. users.fmrib.ox.ac.uk/~steve/susan/index.html
5. svr-www.eng.cam.ac.uk/~er258/work/fast.html
6. cs.ubc.ca/spider/lowe/keypoints/siftDemoV4.zip
7. robots.ox.ac.uk/~timork/salscale.html

pour les détecteurs que nous avons implémentés. Nous avons sélectionné des valeurs qui donnaient des résultats satisfaisants selon nos nombreuses expérimentations. Dans tous les cas, nous avons cherché les paramètres tels que chaque détecteur retourne le maximum des points d'intérêt bien répartis dans l'image en utilisant un seuil assez permissif sur la réponse. Nous avons ainsi cherché les valeurs des paramètres les plus favorables pour chaque détecteur, voir figure 4. Les cardinalités sont données à titre indicatif dans la table 1.

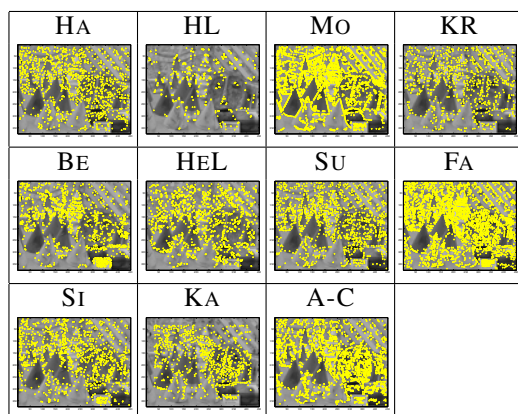


FIGURE 4 – Les différents ensembles de points d'intérêt détectés sur l'image gauche du couple *cones*.

4.3 Répétabilité

La répétabilité est mesurée en prenant $\epsilon = \{0, 1.5, 3\}$. Ces deux dernières valeurs correspondent à des tailles de voisinage 3×3 et 5×5 . Au delà de 5×5 , il n'est pas envisageable de considérer si un point d'intérêt est répété. Les résultats pour la répétabilité sont montrés sur la figure 5.

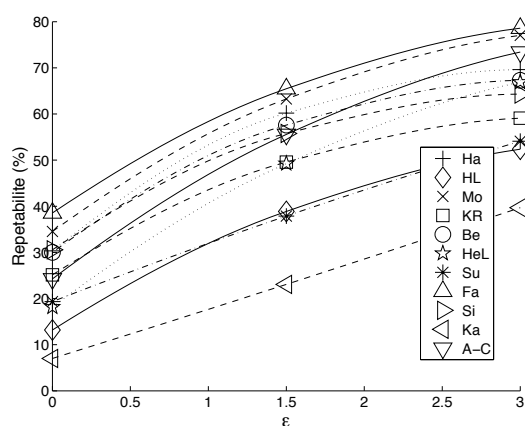


FIGURE 5 – Répétabilité moyennes en fonction de ϵ .

La répétabilité à 0 pixel près est égale au mieux à 39%. Cela signifie que si la mise en correspondance des points d'intérêt se fait entre les points d'intérêt de gauche et les

points d'intérêt de droite, nous aurons un ensemble d'appariements initiaux avec au mieux 39% de correspondances correctes. Par conséquent, nous proposons d'adopter la stratégie suivante : la mise en correspondance des points d'intérêt doit se faire entre les points d'intérêt de gauche et les voisinages des points d'intérêt de droite. La courbe de la figure 5 montre que $\epsilon = 3$ qui correspond à un voisinage de 5×5 pixels semble raisonnable. Le classement CL1 des détecteurs les plus répétables est montré dans la table 2.

4.4 Répartition spatiale

Répartition ZO, ZD. Les ZO sont calculées en effectuant une vérification bidirectionnelle sur les cartes de disparités théoriques. Les ZD sont obtenues en repérant les sauts de disparités sur les cartes théoriques. Ce terme inclus également les zones proches des discontinuités obtenues par dilatation de la zone initialement trouvée, voir figure 6. Les résultats des répartitions dans les ZO et ZD sont donnés dans la table 1. Le classement CL2 des trois détec-

<i>D</i>	CARD.	ZO%	ZD%	RS(1)%	RS(2)%
HA	1011	16	33	76	61
HL	384	23	43	43	24
MO	2369	19	37	79	68
KR	707	18	37	53	45
BE	926	17	36	72	57
HEL	1471	15	40	70	51
SU	1056	20	33	66	60
FA	2611	18	30	89	79
SI	1310	17	33	77	63
KA	680	13	33	55	53
A-C	2089	14	27	78	65

TABLE 1 – Cardinalités moyennes, répartitions moyennes dans les ZO et les ZD et répartitions spatiales par région avec les segmentations S1 (RS(1)) et S2 (RS(2)).

teurs qui retournent le moins de points d'intérêt dans les ZD est montré dans la table 2. Le détecteur A-C donne le meilleur résultat car, contrairement à la plupart des autres détecteurs, il ne cherche pas à sélectionner les coins ou les contours.

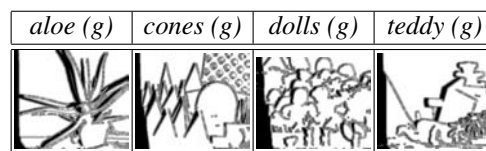


FIGURE 6 – Les zones occultées (ZO) apparaissent en noir et les zones de discontinuité (ZD) apparaissent en gris.

Répartition par région. Le programme EDISON⁸ est utilisé pour réaliser les segmentations en régions. Deux segmentations sont calculées :

8. caip.rutgers.edu/riul/research/code/EDISON/

- une « sous-segmentation » S1 qui donne environ 100 régions sur nos images.
- une « segmentation moyenne » S2 qui donne environ 500 régions sur nos images.

Les résultats sont présentés dans la table 1, et le classement CL3 des meilleurs détecteurs est montré dans la table 2. Nous remarquons également que la répartition n'est jamais parfaite (100%). Certaines régions sont trop homogènes pour que l'on puisse y détecter des points d'intérêt.

RANG	CL1	CL2	CL3
1	FA	A-C	FA
2	MO	FA	MO
3	HA ou A-C	HA, SU, SI ou KA	A-C

TABLE 2 – Classements établis sur les critères : répétabilité (CL1), répartition ZD (CL2), répartition par région (CL3).

4.5 Complémentarité

Nous poursuivons notre étude en faisant l'union de deux détecteurs parmi HA, FA et A-C. La complémentarité est étudiée en prenant les couples Harris+Fast (HA+FA), Harris+Auto-Corrélation (HA+A-C) et Fast+Auto-Corrélation (FA+A-C). Ces détecteurs sont ceux qui obtiennent les meilleurs rangs moyens de répétabilité et de répartition ZD. Il s'agit de deux critères plus discriminants que le critère de répartition par région. Dans la pratique nous avons pu constater que prendre l'union des points d'intérêt issus d'un détecteur répétable et ceux d'un détecteur moins répétable implique un ensemble de points d'intérêt final moins répétés que le meilleur ensemble initial. Une remarque analogue s'applique pour les points d'intérêt situés dans les ZD. En revanche, si on prend un ensemble de points d'intérêt dont la répartition spatiale par région est bonne et qu'on l'associe à un autre ensemble dont la répartition est mauvaise, il est possible que le second ensemble complète le premier et que l'ensemble final ait une répartition meilleure que le premier ensemble.

Apport. Les apports sont montrés dans la figure 7. Ajouter les points d'intérêt de FA ou A-C aux points d'in-

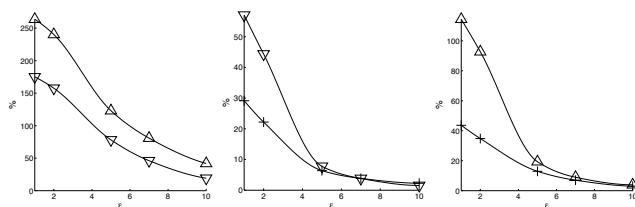


FIGURE 7 – Ces courbes montrent l'apport des détecteurs étudiés sur un détecteur de référence en fonction de ϵ . La légende est la même que pour la figure 5.

térêt issus de HA permet d'apporter beaucoup de nouveaux

points. De plus, une partie de ces nouveaux points sont relativement éloignés des anciens ($\epsilon \geq 5$).

Gain de répétabilité. Le gain de répétabilité est calculé en prenant $\epsilon = 1.5$. Les résultats sont présentés dans la table 3. Le gain (ou la perte) de l'union de deux détecteurs n'est pas significatif puisque qu'il est inférieur à 1% dans tous les cas.

DÉTECTEUR	gain _{rep} %	RS(1)%	RS(2)%
HA+FA	+0.45	93 (+4)	84 (+5)
HA+A-C	-0.99	84 (+6)	76 (+11)
FA+A-C	-0.65	93 (+4)	84 (+5)

TABLE 3 – Gain de répétabilité (gain_{rep}), répartitions (RS) et gains (gain_{RS}), entre parenthèses, calculés avec les segmentations S1 et S2.

Gain de répartition. Les résultats sont présentés dans la table 3. Associer HA et A-C permet d'obtenir un gain de répartition spatiale de +11%. Cela signifie que les deux ensembles de points sont complémentaires dans le sens où ils sont différents et répartis différemment (pas de concentration de points d'intérêt dans une même zone). Le détecteur FA qui donnait à la base une meilleure répartition, donne toujours au final une meilleure répartition lorsqu'il est associé avec HA ou A-C (84%) avec un gain de +5%.

4.6 Application à la propagation

À présent, nous étudions les résultats de propagation que nous obtenons en utilisant les détecteurs étudiés qui permettent de satisfaire le critère (C1), c'est-à-dire, HA, FA, A-C, HA+FA, HA+A-C et FA+A-C. Nous avons commencé par réaliser une étude avec une mise en correspondance théorique des différents points d'intérêt puis nous avons analysé les résultats avec une mise en correspondance calculée. Dans le premier cas, les appariements sont effectués en utilisant les cartes de disparités théoriques et dans le second, nous utilisons une méthode d'appariement par corrélation (SAD) en adoptant la stratégie de mise en correspondance des points d'intérêt de l'image de gauche et du voisinage des points d'intérêt de l'image de droite et en réalisant une vérification bidirectionnelle. Nous proposons ensuite en utilisant la méthode présentée dans [6] avec des seuils stricts. Les disparités estimées sont ensuite évaluées en calculant le pourcentage d'appariements corrects. Le pourcentage d'appariements corrects est similaire lorsqu'on utilise des points d'intérêt d'un même détecteur ou de l'union de deux détecteurs complémentaires mais dans ce dernier cas, la densité du résultat est améliorée, voir la figure 8.

5 Conclusion et perspectives

Nous avons proposé une étude pour la sélection des germes dans le cadre de la mise en correspondance par propagation. Elle s'appuie sur des mesures de répétabilité et de




	FA	A-C	FA+A-C
			
COR%	81	81	82
DEN%	54	46	59

FIGURE 8 – La première ligne montre les cartes de disparité obtenues après propagation sur le couple *dolls*. La seconde ligne indique le pourcentage d'appariements corrects des disparités calculées et la troisième indique la densité. Sur cet exemple, lorsque nous prenons l'union FA+A-C, la densité est améliorée (+5%), de même que le pourcentage d'appariements corrects (+1%).

répartition spatiale des points d'intérêt. Nous avons également proposé une amélioration de la mesure d'Auto-Corrélation proposée par Moravec. Nous avons mis en évidence que les détecteurs qui répondent le mieux aux critères pour la sélection des germes sont Harris, FAST et l'Auto-Corrélation. Nous avons mesuré la complémentarité de ces trois détecteurs pris deux à deux et nous avons montré que prendre l'union de deux détecteurs complémentaires permettait d'améliorer les performances de la propagation. Nous avons proposé une stratégie d'appariement des points d'intérêt qui considère le voisinage des points d'intérêt détectés dans l'autre image.

L'étape de mise en correspondance des germes peut échouer dans certains cas et ainsi nous priver de germes qui auraient pu être bons. Nous envisageons une étude détaillée de cet aspect dans le contexte de la propagation. De plus, nous avons également pu constater que les détecteurs de points d'intérêt ont parfois des limites, en particulier dans les régions trop homogènes. Une étude sur une technique alternative pour la mise en correspondance de ces régions reste à explorer.

Références

- [1] H. BAY, A. ESS, T. TUYTELAARS et L. V. GOOL. « SURF : Speeded Up Robust Features ». Dans *CVIU*, p. 346–359, 2006.
- [2] P. R. BEAUDET. « Rotationally invariant image operators ». Dans *ICPR*, p. 579–583, 1978.
- [3] S. CHAMBON et A. CROUZIL. « Dense matching using correlation : new measures that are robust near occlusions ». Dans *BMVC*, volume 1, p. 143–152, 2003.
- [4] Q. CHEN et G. MEDIONI. « A Volumetric Stereo Matching Method : Application to Image-Based Modeling ». Dans *CVPR*, volume 1, p. 1029–1034, 1999.
- [5] W. FÖRSTNER. « A framework for low level feature extraction ». Dans *ECCV*, p. 383–394, 1994.
- [6] G. GALES, A. CROUZIL et S. CHAMBON. « Approches simultanées et séquentielles de la mise en correspondance par propagation ». Dans *ORASIS*, 2009.
- [7] A. GIL, O. MARTINEZ, M. BALLESTA et O. REINOSO. « A comparative evaluation of interest point detectors and local descriptors for visual SLAM ». *MVA*, 2009.
- [8] V. GOUET, P. MONTESINOS, R. DERICHE et D. PELÉ. « Évaluation des détecteurs de points d'intérêt pour la couleur ». Dans *RFIA*, p. 257–266, 2000.
- [9] C. HARRIS et M. STEPHENS. « A Combined Corner and Edge Detector ». Dans *Alvey Vision Conference*, p. 147–151, 1988.
- [10] T. KADIR, A. ZISSERMAN et M. BRADY. « An affine invariant salient region detector ». Dans *ECCV*, p. 404–416, 2004.
- [11] L. KITCHEN et A. ROSENFELD. « Gray level corner detection ». *PRL*, 1(2) :95–102, 1982.
- [12] M. LHUILLIER et L. QUAN. « Match Propagation for Image-Based Modeling and Rendering ». *PAMI*, 24(8) :1140–1146, 2002.
- [13] D. G. LOWE. « Object Recognition from Local Scale-Invariant Features ». Dans *ICCV*, volume 2, p. 1150–1157, 1999.
- [14] Y. Z. MA et W. J. LIU. « Progressive Matching Based on Segmentation for 3D reconstruction ». Dans *International Conference on Computer and Information Technology*, p. 575–579, 2005.
- [15] K. MIKOLAJCZYK, B. LEIBE et B. SCHIELE. « Local Features for Object Classe Recognition ». Dans *ICCV*, volume 2, p. 1792–1799, 2005.
- [16] K. MIKOLAJCZYK et C. SCHMID. « Scale & Affine Invariant Interest Point Detectors ». *IJCV*, 60(1) :63–86, 2004.
- [17] K. MIKOLAJCZYK et C. SCHMID. « A Performance Evaluation of Local Descriptors ». *PAMI*, 27(10) :1615–1630, 2005.
- [18] H. P. MORAVEC. « Toward automatic visual obstacle avoidance ». Dans *IJCAI*, volume 2, p. 584–584, 1977.
- [19] J. A. NOBLE. « Finding corners ». *IVC*, 6(2) :121–128, 1988.
- [20] G. P. OTTO et T. K. W. CHAU. « “Region-growing” algorithm for matching of terrain images ». *IVC*, 7(2) :83–94, 1989.
- [21] E. ROSTEN et T. DRUMMOND. « Machine learning for high-speed corner detection ». Dans *ECCV*, volume 1, p. 430–443, 2006.
- [22] D. SCHARSTEIN et R. SZELISKI. « A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms ». *IJCV*, 47(1) :7–42, 2002.
- [23] C. SCHMID, R. MOHR et C. BAUCKHAGE. « Evaluation of interest point detectors ». *IJCV*, 37(2) :151–172, 2000.
- [24] J. SHI et C. TOMASI. « Good features to track ». Dans *CVPR*, p. 593–600, 1994.
- [25] S. M. SMITH et J. M. BRADY. « SUSAN – A New Approach to Low Level Image Processing ». *IJCV*, 23(1) :45–78, 1997.
- [26] Z. ZHANG et Y. SHAN. « A Progressive Scheme for Stereo Matching ». Dans *European Workshop on 3D Structure from Multiple Images of Large-Scale Environments*, volume 2018 de *LNCS*, p. 68–85, 2000.