Complémentarité de mesures de corrélation pour la mise en correspondance

S. CHAMBON[1] A. CROUZIL[2] M. EL MIZIANI[3] G. LEMARIE[3] P. LE CALLET[3]

[1] Laboratoire Central des Ponts et Chaussées, LCPC, Nantes
[2] Institut de Recherche en Informatique de Toulouse, IRIT
[3] Institut de Recherche en Communications et en Cybernétique de Nantes, IRCCyN

chambon@lcpc.fr,crouzil@irit.fr, patrick.lecallet@univ-nantes.fr

Résumé

Dans le cadre de la mise en correspondance stéréoscopique de pixels, nous présentons une étude des mesures de corrélation. Plus particulièrement, nous évaluons le degré de complémentarité de huit mesures différentes. Ces mesures sont les plus représentatives de chacun des types de mesures considérées : classiques, croisées, fondées sur le calcul des dérivées, non-paramétriques et robustes. Cette étude met en évidence les deux mesures les plus complémentaires : la corrélation des gradients et une variance robuste. Enfin, nous proposons deux approches permettant de combiner automatiquement ces deux mesures : une méthode qui fusionne les critères de corrélation et une seconde méthode itérative qui, grâce à notre protocole d'évaluation, apparaît comme étant la plus performante.

Mots Clef

Stéréovision, mise en correspondance, corrélation, fusion.

Abstract

We present a study of correlation measures for pixel matching. We particularly evaluate how much eight measures are complementary. These measures are the most significant of five categories: classic, cross-correlation, derivative-based, non-parametric and robust. We highlight the two most complementary measures: gradient correlation and robust variance. We also introduce two algorithms to combine these measures: one based on merging the criteria and one iterative which obtains the best results.

Keywords

Stereovision, matching, correlation, fusion.

1 Introduction

Une des étapes fondamentales pour retrouver le relief d'une scène à partir d'un couple stéréoscopique consiste à retrouver les primitives homologues entre les deux images. Ces primitives peuvent être des pixels ou des éléments plus complexes comme des droites, des contours, des régions. Dans le cadre de notre étude, nous allons nous préoccuper de la mise en correspondance de pixels.

Il existe de nombreuses techniques de mise en correspondance. Dans la littérature, l'habitude est de distinguer les méthodes locales des méthodes globales [3]. Une autre façon de procéder est de décomposer les méthodes de mise en correspondance en éléments constituants [5]. Cette description permet alors d'affiner la séparation en quatre catégories : les méthodes locales, les méthodes globales ne faisant pas intervenir de critère local, les méthodes mixtes (avec critère local qui intervient dans un processus global) et les méthodes à passages multiples (qui font intervenir de manière séquentielle et complémentaire différentes méthodes locales et/ou globales). Enfin, un travail conséquent à permis de mettre en concurrence et de distinguer les méthodes les plus performantes [11]. Les auteurs de cet article ont aussi mis un protocole d'évaluation à la disponibilité de la communauté [1]. Cependant, cette évaluation permet surtout de comparer les performances des méthodes globales. Dans le cadre des méthodes locales, il est encore difficile de mesurer les qualités respectives de chacune des méthodes proposées. Pourtant, ces méthodes restent actuellement les plus rapides et les plus utilisées pour des applications assez diverses, comme la robotique, l'étude de déformations de surfaces [14]. Récemment, une première étude a été menée [4, chap. 3 à 5] afin d'établir les qualités de chacune des méthodes locales proposées, et en particulier, l'influence des différents critères de similarité sur la qualité des résultats dans différentes zones de l'image ou encore l'influence de la couleur. Dans cette première étude, une mesure robuste a notamment été mise en avant ainsi que des algorithmes permettant de combiner une mesure classique, dans les zones sans occultation, avec une mesure robuste dans les zones présentant des occultations. Les résultats encourageants on permis de dégager trois nouvelles questions :

- Quelles sont les mesures de corrélation qui ont une réelle « complémentarité » ?
- Serait-il judicieux de combiner plus de critères et jusqu'à combien ?
- Si on arrive à répondre à ces deux questions, peut-on proposer de nouveaux algorithmes plus performants que ceux initialement proposés ?

[1] http://vision.middlebury.edu/stereo/eval/

Ainsi, dans un premier temps, nous allons présenter l'état de l'art en ce qui concerne les méthodes locales, puis, nous proposerons un nouvel algorithme de fusion de critères de corrélation. Ensuite, afin d'illustrer les résultats obtenus, nous commencerons par détailler le protocole d'évaluation, avant de présenter notre étude de complémentarité et le gain de performance obtenu en utilisant la fusion des mesures de corrélation retenues.

2 Mise en correspondance locale

L'hypothèse des méthodes locales est que deux pixels qui se correspondent, ainsi que leurs voisinages respectifs, se ressemblent d'un point de vue photométrique. De nombreuses difficultés se posent : les changements de luminosité, les zones uniformes, les occultations. Ainsi, un très grand nombre de mesures ont été proposées afin de répondre à ces difficultés et nous avons choisi de les classer en cinq grandes familles [4, chap. 3] :

(1) *Croisée* – Il s'agit de toutes les mesures qui utilisent un produit scalaire.

(2) *Classique* – Cela concerne toutes les mesures qui s'appuient sur des outils des statistiques classiques.

(3) *Dérivée* – Dans cette catégorie, à la place des niveaux de gris, les auteurs utilisent les dérivées des images.

(4) *Non-paramétrique* – Ce type de mesures, comme les mesures *dérivées*, ne s'appuient plus sur les niveaux de gris, mais, sur l'ordre des niveaux de gris dans le voisinage pris en compte.

(5) *Robuste* – Afin d'être robuste aux problèmes liés aux occultations, ces mesures font intervenir des outils des statistiques robustes.

Au cours des différentes études que nous avons menées [4, 6], nous avons mis en évidence les mesures les plus performantes dans chacune des familles. Nous allons donc nous limiter à la présentation de ces huit mesures.

2.1 Critères de similarité

Avant de présenter les mesures de corrélation retenues pour notre étude de complémentarité, nous présentons les notations utilisées :

- La taille des fenêtres de corrélation est $(2N_v + 1) \times (2N_h + 1)$ et le nombre de pixels dans la fenêtre de corrélation est noté $N_f = (2N_v + 1)(2N_h + 1)$.
- Les niveaux de gris des pixels des fenêtres de corrélation (matrices) sont stockés dans les vecteurs $\mathbf{f}_l$:

$$\mathbf{f}_l = (\cdots I_l^{i+p,j+q} \cdots)^T = (\cdots f_l^k \cdots)^T$$

  où f_l^k est l'élément k du vecteur $\mathbf{f}_l$, $p \in [-N_v; N_v]$, $q \in [-N_h; N_h]$ et $k \in [0; N_f - 1]$.
- Nous notons $\mathrm{I}_{\mathrm{max}}$ le niveau de gris maximal dans les images.
- Le vecteur gradient au pixel $(i\ j)^T$ de l'image I_l est représenté par $\nabla I_l^{i,j}$.
- Les P-normes, ou distances L_P, $P \in \mathbb{N}^*$, sont notées :

$$\|\mathbf{f}_l\|_P = \left(\sum_{k=0}^{N_f-1} |f_l^k|^P \right)^{1/P}.$$

  Pour la norme euclidienne, nous notons : $\|\mathbf{f}_l\| = \|\mathbf{f}_l\|_2$.
- Le produit scalaire est donné par :

$$\mathbf{f}_g \cdot \mathbf{f}_d = \sum_{k=0}^{N_f-1} f_g^k f_d^k.$$

- Les moyennes sont notées :

$$\mathrm{moy}(\mathbf{f}_l) = \frac{1}{N_f} \sum_{k=0}^{N_f-1} f_l^k.$$

  Et nous utilisons le vecteur des moyennes défini par :

$$\overline{\mathbf{f}_l} = \underbrace{(\mathrm{moy}(\mathbf{f}_l)\ \cdots\ \mathrm{moy}(\mathbf{f}_l))}_{N_f \text{ colonnes}}{}^T.$$

- Les valeurs ordonnées du vecteur $\mathbf{f}_l$ sont notées :

$$(f_l)_0 \leq \ldots \leq (f_l)_k \leq \ldots \leq (f_l)_{N_f-1},$$

  où k indique le rang de la donnée considérée dans le vecteur $\mathbf{f}_l$.

Famille croisée – Lors de notre première étude [4, chap. 3], nous avons trouvé que, dans cette famille, la mesure qui obtenait les meilleurs pourcentages d'appariements corrects correspondait à la corrélation normalisée ou *Normalized Cross Correlation* :

$$\mathrm{NCC}(\mathbf{f}_g, \mathbf{f}_d) = \frac{\mathbf{f}_g \cdot \mathbf{f}_d}{\|\mathbf{f}_g\| \|\mathbf{f}_d\|}. \tag{1}$$

Les valeurs de cette mesure de similarité appartiennent à l'intervalle [0;1]. Elle est très utilisée dans la littérature et c'est la raison pour laquelle nous avons choisi de la conserver dans notre étude, même si notre seconde évaluation [6] a aussi permis de souligner les bonnes performances de la mesure de Moravec [9]. Notée MOR, cette mesure de similarité est centrée et utilise une normalisation différente de celle utilisée par la mesure NCC (cette normalisation est moins coûteuse en temps de calculs d'après les auteurs) :

$$\mathrm{MOR}(\mathbf{f}_g, \mathbf{f}_d) = \frac{2(\mathbf{f}_g - \overline{\mathbf{f}_g}) \cdot (\mathbf{f}_d - \overline{\mathbf{f}_d})}{\|\mathbf{f}_g - \overline{\mathbf{f}_g}\|^2 + \|\mathbf{f}_d - \overline{\mathbf{f}_d}\|^2}. \tag{2}$$

Les valeurs de $\mathrm{MOR}(\mathbf{f}_g, \mathbf{f}_d)$ appartiennent à l'intervalle $[-1; 1]$. D'après Aschwanden et Guggenbül [1], la mesure de Moravec est robuste face à un bruit impulsionnel.

Famille classique – Dans cette famille, nous avons considéré quatre groupes de mesures : celles s'appuyant sur une distance, celles localement centrées, les mesures utilisant une variance et celle utilisant le kurtosis. Les deux études réalisées [4, 6] ont mis en évidence deux mesures relativement performantes :

- *La somme des valeurs absolues des différences* – Il s'agit de la norme L_1, notée SAD (*Sum of Absolute Differences*) :

$$\mathrm{SAD}(\mathbf{f}_g, \mathbf{f}_d) = \|\mathbf{f}_g - \mathbf{f}_d\|_1. \qquad (3)$$

  Cette mesure de dissimilarité a pour intervalle de variation $[0; \mathrm{I}_{\max}^P N_f]$.
- *La somme des valeurs absolues des différences localement centrée* – L'objectif de cette mesure est de faire en sorte que la moyenne des niveaux de gris sur la fenêtre de corrélation de droite soit la même que celle de gauche afin d'être invariant aux changements de luminosité. La mesure de dissimilarité, notée LSAD (*Locally scaled Sum of Absolute Differences*), est donnée par :

$$\mathrm{LSAD}(\mathbf{f}_g, \mathbf{f}_d) = \left\| \mathbf{f}_g - \frac{\overline{\mathbf{f}_g}}{\overline{\mathbf{f}_d}} \mathbf{f}_d \right\|_1. \qquad (4)$$

  D'après Aschwanden et Guggenbül [1], LSAD est résistante au bruit impulsionnel.

Famille dérivée – Pour la plupart des mesures exploitant les dérivées des images, seule la direction du gradient est utilisée. Or cette information, seule, peut entraîner des erreurs, surtout dans le cas de gradients de faible norme. Une autre mesure, la plus performante de cette famille, appelée corrélation des champs de gradients, notée GC (*Gradient field Correlation*) [7], utilise la similarité de vecteurs gradients et est donnée par :

$$\mathrm{GC}(\mathbf{f}_g, \mathbf{f}_d) = \frac{\sum_A \|\nabla I_g^{i+p,j+q} - \nabla I_d^{i+p,v+q}\|}{\sum_A (\|\nabla I_g^{i+p,j+q}\| + \|\nabla I_d^{i+p,v+q}\|)}$$
$$\text{avec } \sum_A = \sum_{p=-N_v}^{N_v} \sum_{q=-N_h}^{N_h} \qquad (5)$$

Il s'agit d'une mesure de dissimilarité dont les valeurs appartiennent à l'intervalle $[0; +\infty[$. Pour calculer le gradient, deux solutions ont été proposées : le filtre de Shen et Castan [13] ou l'opérateur de Sobel.

Famille non paramétrique – Trois grands types de mesures ont été distingués, les mesures qui comparent les niveaux de gris de la fenêtre de corrélation avec celui du pixel étudié [8], les mesures non-paramétriques de Zabih et Woodfill [15] et les mesures ordinales de Bhat et Nayar [2]. Le principe des mesures ordinales est d'établir un classement des niveaux de gris des pixels au sein de la fenêtre de corrélation. Ceci permet d'être plus robuste aux bruits et aux occultations, mais rend ces mesures trop ambiguës pour permettre de bonnes performances, c'est-à-dire que dans certains cas, un mauvais correspondant peut obtenir un meilleur score de corrélation que le bon correspondant. Ainsi, les deux mesures les plus performantes sont :

- *Increment Sign Correlation* – Elle s'appuie sur les vecteurs $\mathbf{b}_l$ suivants :

$$\mathbf{b}_l = \left(\dots b_l^k \dots\right)^T \text{ pour } k = 0 \cdot \cdot N_f - 2$$
$$\text{avec } b_l^k = \begin{cases} 1 & \text{si } f_l^{k+1} \geq f_l^k \\ 0 & \text{sinon.} \end{cases}$$

  Si le niveau de gris augmente entre f_l^k et f_l^{k+1}, alors b_l^k vaut 1 ou 0 dans le cas contraire. La mesure ISC compare les vecteurs $\mathbf{b}_g$ et $\mathbf{b}_d$. Elle détermine si les niveaux de gris varient dans le même sens. Elle est donnée par :

$$\mathrm{ISC}(\mathbf{f}_g, \mathbf{f}_d) = \frac{1}{N_f - 1}(\mathbf{b}_g \cdot \mathbf{b}_d + (\mathbf{1} - \mathbf{b}_g) \cdot (\mathbf{1} - \mathbf{b}_d)). \quad (6)$$

  C'est une mesure de similarité et ses valeurs appartiennent à l'intervalle $[0; 1]$.
- *Corrélation de rang* – La transformation de rang correspond au nombre de pixels dans la fenêtre de corrélation dont le niveau de gris est inférieur au niveau de gris du pixel central. Elle est donnée par :

$$\mathrm{R}_{rank}(\mathbf{f}_l) = \left(\dots \quad rank(\mathbf{p}_l^{i+p,j+q}) \quad \dots\right)$$
$$\text{avec } rank(\mathbf{p}_l^{i,j}) = \#(\{I_l^{i+p,j+q} \mid I_l^{i+p,j+q} < I_l^{i,j}$$
$$p \in [-N_v; N_v], q \in [-N_h; N_h], \}).$$

  Le terme $\#\{\mathbf{x}\}$ correspond au cardinal de $\mathbf{x}$. La corrélation de rang, notée RANK, est donnée par :

$$\mathrm{RANK}(\mathbf{f}_g, \mathbf{f}_d) = \|\mathrm{R}_{rank}(\mathbf{f}_g) - \mathrm{R}_{rank}(\mathbf{f}_d)\|_1. \qquad (7)$$

  C'est une mesure de dissimilarité et les valeurs de $\mathrm{RANK}_1(\mathbf{f}_g, \mathbf{f}_d)$ appartiennent à l'intervalle $[0; N_f^2]$.

Famille robuste – Dans [4, chap. 3], nous avons considéré les mesures partielles, les pseudo-normes et proposé des mesures s'appuyant sur des variances robustes, des M-estimateurs et des R-estimateurs. La mesure qui a obtenu les meilleures performances s'appuie sur l'estimateur SMAD (*Smooth Median Absolute Deviation*), qui correspond à une estimation robuste de la variance, proposée dans [10] et définie par :

$$\mathrm{SMAD}(\mathbf{f}_g, \mathbf{f}_d) = \sum_{k=0}^{h-1} \left(|\mathbf{f}_g - \mathbf{f}_d - \mathrm{med}(\mathbf{f}_g - \mathbf{f}_d)|^2\right)_k,$$

où k indique l'ordre. Ainsi, nous définissons la somme des puissances des h premiers écarts à la médiane, notée SMPD_P (*Smooth Median Powered Deviation*), et donnée par :

$$\mathrm{SMPD}_P(\mathbf{f}_g, \mathbf{f}_d) = \sum_{k=0}^{h-1} \left(|\mathbf{f}_g - \mathbf{f}_d - \mathrm{med}(\mathbf{f}_g - \mathbf{f}_d)|^P\right)_k. \qquad (8)$$

Les valeurs de cette mesure de dissimilarité appartiennent à l'intervalle $[0; {\mathrm{I}_{\max}}^P h]$[2].

[2]Nous avons choisi $h = N_f/2$ et $P = 2$.

2.2 Algorithme de fusion

Le bilan des travaux réalisés dans [4] est que la combinaison par fusion de cartes de disparités[3] obtenues par deux mesures de corrélation différentes : ZNCC (*Zero-mean* NCC) et SMPD, donnent des résultats performants aussi bien dans les zones sans occultation que dans les zones avec occultations. Toutefois, les performances peuvent encore être améliorées. Nous pensons qu'une étude plus approfondie de la complémentarité des différentes mesures utilisées est une première étape importante pour l'amélioration de ce travail. De plus, nous pensons que d'autres techniques de fusion sont aussi envisageables. Ainsi, le but de ce paragraphe est de présenter des techniques de fusion de critères de corrélation pour l'appariement afin de combiner les avantages de chaque critère et afin de tenter d'améliorer les performances des méthodes par corrélation.
Nous proposons deux méthodes :

(1) La méthode SCORE qui consiste à proposer un critère prenant en compte différents critères. Plus précisément, en supposant que nous n'avons que des critères de dissimilarité, nous allons chercher à minimiser :

$$M_{\text{Fusion}}(\mathbf{f}_g, \mathbf{f}_d) = \underset{i \in [1..N_c]}{\text{Fusion}}\ M_i(\mathbf{f}_g, \mathbf{f}_d), \tag{9}$$

où N_c est le nombre de critères fusionnés, M_i est la $i^{\text{ème}}$ mesure prise en compte. Nous pouvons envisager différentes manières de fusionner les critères, mais, dans le cas de cette première étude, nous avons choisi d'effectuer une moyenne. Dans ce cas, il est nécessaire d'avoir des critères qui ont le même intervalle de variation. Pour cela, dans la suite de nos expérimentations, nous avons choisi de normaliser toutes les mesures (qui ne le sont pas) afin que les intervalles de variations soient identiques. Ainsi, plus précisément, nous avons choisi le critère suivant :

$$M_{\text{Fusion}}(\mathbf{f}_g, \mathbf{f}_d) = \sum_{i \in [1..N_c]} \frac{M_i(\mathbf{f}_g, \mathbf{f}_d)}{\max(M_i)}, \tag{10}$$

où $\max(M_i)$ est la valeur maximale obtenue par le critère de corrélation sur l'image étudiée.

(2) La méthode ITÉRATIVE qui consiste à calculer la carte de disparités avec chacune des mesures utilisées, puis à effectuer une fusion de tous les résultats. Le principe de la fusion repose sur deux aspects : d'une part, le fait que lorsqu'on a au moins deux cartes qui donnent la même disparité, alors, le résultat est considéré comme fiable et est validé, et, d'autre part, lorsqu'on est dans une zone indéterminée (huit cartes avec huit disparités différentes), on va propager les disparités voisines tout en validant le choix par les possibilités de disparités données par chaque carte. Ces deux règles se formalisent de la manière suivante :

[3] Une carte de disparités représente, pour chaque pixel, l'amplitude du déplacement de ce pixel entre les deux images. Dans nos exemples, plus le pixel est clair et plus la disparité est importante. Les pixels noirs correspondent aux pixels occultés entre l'image de gauche et de droite.

(a) *Initialisation pour chaque pixel* $\mathbf{p}_g$
Si $\exists\ d$ tel que $\#\{f_i(\mathbf{p}_g) = d\} \geq 2$ alors on conserve d, sinon la disparité est indéterminée.

(b) *Itération*
Tant que la carte de disparités n'est pas dense, pour chaque pixel, on calcule la disparité moyenne d_m dans son voisinage (8-connexe). Si $\exists\ d$ tel que $f_i(\mathbf{p}_g) = d\ \&\ |d - d_m| < \epsilon$ alors on conserve d[4]. Le terme f_i est la fonction de disparités associée à la mesure M_i.

3 Protocole d'évaluation

L'objectif de ce protocole est, d'une part, de visualiser les zones de performances de chacune des mesures, et, d'autre part, d'évaluer le gain de performances obtenu par les deux variantes de fusion proposées. Pour cela, nous avons besoin de déterminer les images à évaluer et les critères à utiliser.

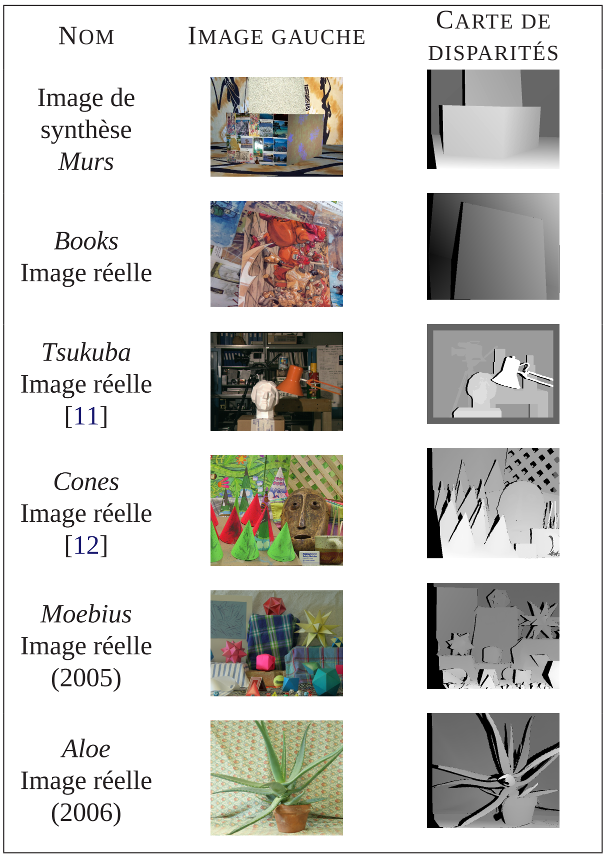


FIG. 1 – Exemples d'images utilisées pour les tests.

3.1 Images

Quarante-deux images ont été testées (voir les exemples sur la figure 1) : un stéréogramme aléatoire, deux paires d'images synthétiques, une paire d'images réelles réalisées par Benoît Bocquillon[5] et trente-huit paires d'images

[4] Par la suite, nous avons choisi $\epsilon = 1$.
[5] http://www.irit.fr/~Alain.Crouzil/matching.html

réelles introduites par Scharstein et Szeliski (huit ont été proposées en 2001 [11], deux en 2003 [12], six en 2005 vingt-deux en 2006). Les dernières images proposées sont les images les plus complexes. Comparé au protocole d'évaluation proposé par Scharstein et Szeliski, notre protocole a l'avantage de résumer les résultats sur toutes les images et non sur quatre comme il est proposé sur leur site.

3.2 Critères

Un grand nombre de critères peuvent être pris en compte [4, chap. 2], comme les pourcentages d'appariements corrects, d'appariements erronés, de faux négatifs, de faux positifs. On peut aussi distinguer différentes zones dans l'image comme les zones avec ou sans occultation, les zones texturées et non texturées. Dans le cadre de ce travail, dans un premier temps nous avons voulu évaluer les pourcentages d'appariements erronés. Nous réalisons seulement une visualisation de la qualité des résultats en fonction des zones avec ou sans occultation, cf. § 4.1.

4 Résultats expérimentaux

4.1 Étude de la complémentarité

Pour calculer la complémentarité de ces mesures de similarité, nous nous appuyons uniquement sur les pourcentages de correspondances erronées. On souhaite minimiser ce critère et on va chercher la combinaison de mesures qui, théoriquement, devrait le minimiser. Pour cela on calcule ce pourcentage pour chaque mesure séparément, d'une part, et pour chaque combinaison de mesures en supposant que l'on conserve toujours le bon correspondant, cf. table 1. Une première analyse de ces résultats montrent clairement qu'au delà de 2 mesures, le gain de performances n'est plus aussi important. Ce qui ressort de cette première évaluation est que la combinaison des mesures GC et SMPD semble la plus complémentaire, contrairement à l'hypothèse faite dans [4, chap. 5]. Nous proposons aussi une évaluation visuelle de la complémentarité des mesures grâce à trois cartes de visualisation :

(1) *Visualisation des mesures seules* – On affiche sur l'image les zones où seule une mesure (sur les huit testées) a fourni la bonne réponse. Une illustration dans la figure 2.(a) montre que la mesure qui se distingue le plus des autres, c'est-à-dire qui est la plus complémentaire aux autres, est SMPD.

(2) *Visualisation 2 à 2* – On affiche les différences de résultats obtenues entre deux mesures. Sur la figure 2.(b), on peut voir, comme espéré, que SMPD permet de compléter les zones proches des discontinuités alors que GC permet de compléter les zones difficiles à apparier hors occultations.

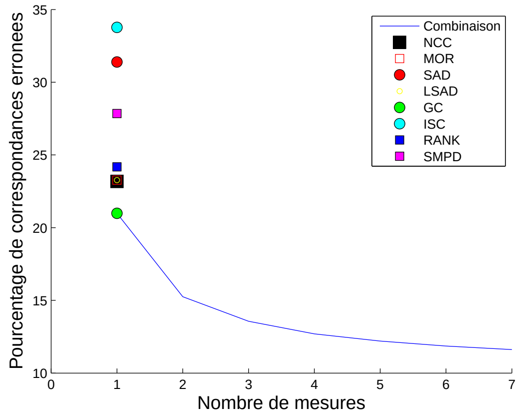
(3) *Visualisation des zones sensibles* – Afin de visualiser les zones difficiles à apparier, nous affichons par pixel le nombre de mesures qui ont trouvé le bon correspondant. La figure 2.(c) met en évidence les zones homogènes, de discontinuités ou d'occultations.

(a)

NB	MESURES	ER	NB	MESURES	ER
1	NCC	23.2	1	MOR	23.3
1	SAD	31.4	1	LSAD	23.3
1	GC	**21**	1	ISC	33.8
1	RANK	24.2	1	SMPD	27.9

NB	MESURES RETENUES	ER
2	GC, SMPD	15.3
3	GC, ISC, SMPD	13.6
4	NCC, GC, RANK, SMPD	12.7
5	NCC, GC, ISC, RANK, SMPD	12.2
6	NCC, SAD, GC, ISC, RANK, SMPD	11.9
7	MOR, SAD, LSAD, GC, ISC, RANK, SMPD	11.6

(b)



TAB. 1 – Pourcentage de correspondances erronées (ER) suivant les différentes combinaisons – La courbe (b) est l'illustration graphique des résultats donnés en (a) où les mesures retenues correspondent à la combinaison qui donne le pourcentage moyen de correspondances erronées le plus faible (NB, nombre de mesures combinées). Nous affichons ER pour chaque mesure puis pour chaque meilleure combinaison.

En conclusion de ces premiers résultats, nous avons choisi d'étudier les résultats de mise en correspondance en combinant deux mesures de corrélation : GC et SMPD.

4.2 Résultats de mise en correspondance

Nous avons analysé les améliorations sur le pourcentage d'appariements erronés avec les deux nouveaux algorithmes proposés, cf. § 2.2. De plus, pour tenter de détecter les pixels occultés, nous utilisons la contrainte de vérification bidirectionnelle : calcul des appariements de la gauche vers la droite puis de la droite vers la gauche et élimination de toutes les correspondances non cohérentes (ces pixels sont marqués en noir dans les figures 1 et 3). Le tableau 2 résume l'ensemble des résultats obtenus sur la totalité des images testées. La méthode SCORE ne permet pas d'améliorer les performances de la mesure GC seule alors que la méthode ITÉRATIVE améliore les résultats : le pourcentage de réponses erronées diminue. Cette diminution est comprise entre 2.47 et 4.08 et, en particulier, elle est importante pour les images difficiles à apparier. Toutefois, la méthode n'atteint pas les performances théoriques d'une méthode combinant parfaitement les deux mesures. En effectuant cette combinaison de mesures, nous espérons, dans les cas où il y a ambiguïté, c'est-à-dire, où les disparités estimées sont différentes, que nous allons conserver la réponse correcte. Pour cela, nous avons tenté de proposer une méthode qui favorise la réponse correcte. Dans le cas de la méthode SCORE, en normalisant la somme des scores, nous donnons autant de chance à chacune des mesures, quelque soit le contexte, alors que, dans la méthode ITÉRATIVE, nous favorisons celle qui semble la plus cohérente avec le contexte, c'est-à-dire, avec les résultats de disparités obtenus dans le voisinage. Cette différence de comportement est illustré par la différence de performances entre ces deux méthodes et explique en partie pourquoi la méthode SCORE est parfois moins performante que les méthodes classiques. Une autre manière de visualiser les résultats est de regarder les cartes de disparités obtenues comme dans la figure 3. Cela montre que la carte de disparités obtenue avec la méthode ITÉRATIVE est plus claire que celles obtenues avec les autres méthodes, c'est-à-dire, qu'elle possède moins de faux négatifs. Les zones d'occultations sont aussi mieux délimitées (les contours sont nets et il n'y a pas de « trous » dans ces zones).

Méthode	H+O	O	H	R	Total
Mesure GC	25.68	17.54	19.64	15.93	20.99
SCORE	25.52	17.14	20.18	17.14	20.92
ITÉRATIVE	**22.09**	**13.46**	**16.86**	**13.46**	**17.45**
Théorique	19.43	10.81	15.44	10.81	15.25

TAB. 2 – Résultats des deux méthodes proposées – Nous donnons les pourcentages moyens d'appariements erronés que nous comparons à ceux obtenus avec GC seule et dans le cas d'une combinaison théorique avec SMPD. Nous avons distingué 4 catégories d'images : H signifie que la scène contient de nombreuses zones homogènes, O de nombreuses zones occultées et R aucune région difficile. Ce classement a été réalisé de manière visuelle. Il apparaît que la méthode SCORE ne parvient pas à améliorer les résultats et même, dans le cas des images les moins difficiles à apparier (R), elle dégrade la qualité des résultats obtenus avec GC. En revanche, la méthode ITÉRATIVE permet d'améliorer les résultats, particulièrement pour les images plus délicates à apparier (H et/ou O).

Conclusion

Ainsi, nous avons proposé, d'une part, une étude de la complémentarité de mesures de corrélation, illustrée par trois cartes de visualisation différentes, et, d'autre part, deux méthodes de combinaison de mesures de corrélation. Les tests sur quarante-deux images nous ont permis, tout d'abord, de déterminer le couple de mesures les plus complémentaires, et, ensuite, de valider la méthode ITÉRATIVE qui consiste à conserver les correspondances validées par les deux mesures et à les propager dans le voisinage.
Les premiers résultats obtenus sont encourageants mais cependant la méthode ne permet pas d'obtenir d'aussi bons résultats que si on combinait parfaitement les deux mesures, c'est-à-dire que parfois, on ne conserve pas la bonne disparité parmi les deux calculées. C'est pourquoi, nous envisageons d'étudier des méthodes de fusion s'appuyant sur la notion de vote. De plus, nous souhaitons tester des variantes faisant intervenir les huit mesures sélectionnées.

Références

[1] P. ASCHWANDEN et W. GUGGENBÜL. Experimental results from a comparative study on correlation type registration algorithms. Dans W. FÖRSTNER et S. RUWIEDEL, éditeurs, *Robust computer vision: Quality of Vision Algorithms*, pages 268–282. 1992.

[2] D. N. BHAT et S. K. NAYAR. « Ordinal Measures for Image Correspondence ». *PAMI*, 20(4):415–423, 1998.

[3] M. Z. BROWN, D. BURSCHKA et G. D. HAGER. « Advances in Computational Stereo ». *PAMI*, 25(8):993–1008, 2003.

[4] S. CHAMBON. « *Mise en correspondance stéréoscopique d'images couleur en présence d'occultations* ». Thèse de doctorat, Université de Toulouse, 2005.

[5] S. CHAMBON et A. CROUZIL. « Mise en correspondance par corrélation avec prise en compte des occultations ». *Trait. du signal*, 24(6):429–446, 2007.

[6] S. CHAMBON et A. CROUZIL. « Occlusions handling in dense stereo matching ». *PR*, 2009. Soumis.

[7] A. CROUZIL, L. MASSIP-PAILHES et S. CASTAN. « A New Correlation Criterion Based on Gradient Fields Similarity ». Dans *ICPR*, volume 1, pages 632–636, 1996.

[8] S. KANEKO, I. MURASE et S. IGARASHI. « Robust Image registration by Increment Sign Correlation ». *PR*, 35(10):2223–2234, 2002.

[9] H. MORAVEC. « *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover* ». PhD thesis, Université de Carnegie Mellon, CMU, 1980.

[10] P. J. ROUSSEEUW et C. CROUX. L_1-Statistical Analysis and Related Methods. Dans Y. DODGE, éditeur, *Explicit Scale Estimators with High Breakdown Point*, pages 77–92. Elsevier, 1992.

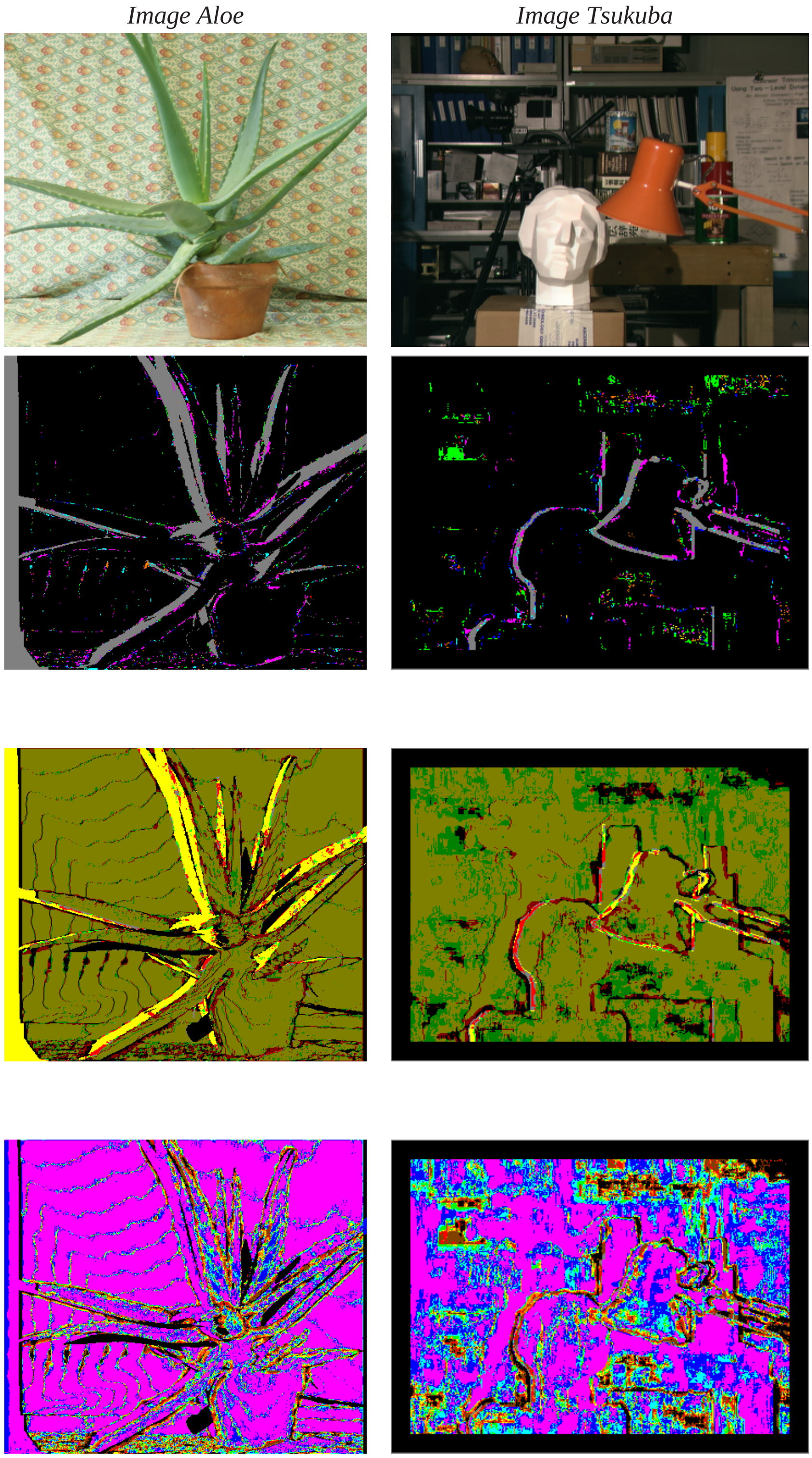
[11] D. SCHARSTEIN et R. SZELISKI. « A Taxomomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms ». *IJCV*, 47(1):7–42, 2002.

[12] D. SCHARSTEIN et R. SZELISKI. « High-Accuracy Stereo Depth Maps Using Structured Light ». Dans *CVPR*, volume 1, pages 195–202, 2003.

[13] J. SHEN et S. CASTAN. « An optimal linear operator for step edge detection ». *Computer Vision, Graphics, and Image Processing*, 24(2):112–133, 1992.

[14] M. A. SUTTON, J. H. YAN, V. TIWARI, H. W. SCHREIER et J. J. ORTEU. « The effect of out-of-plane motion on 2D and 3D digital image correlation measurements ». *Optics and Lasers in Engineering*, 46:746–757, 2008.

[15] R. ZABIH et J. WOODFILL. « Non-parametric Local Transforms for Computing Visual Correspondence ». Dans *ECCV*, pages 151–158, 1994.



(a) *Visualisation des mesures les plus complémentaires* – Les codes sont les suivants : NCC : marron, MOR : rouge, SAD : orange, LSAD : jaune, GC : vert, ISC : cyan, RANK : bleu, SMPD2 : violet. Les zones noires sont celles où plus d'une mesure a fourni le bon correspondant.

(b) *Comparaison de mise en correspondance entre GC et SMPD* – Les codes sont les suivants : jaune foncé/clair : les 2 mesures donnent la bonne réponse, respectivement dans une zone sans/avec occultations, noir : les 2 mesures donnent la mauvaise réponse (toutes zones confondues), rouge foncé/clair : SMPD donne la bonne réponse seule respectivement dans une zone sans/avec occultations, vert foncé/clair : GC donne la bonne réponse seule respectivement dans une zone sans/avec occultations.

(c) *Visualisation du degré de difficulté des zones à apparier* – Les codes couleur représentent le nombre de mesures (de 0 à 8) qui ont donné le bon correspondant : 0 : noir, 1 : marron, 2 : rouge, 3 : orange, 4 : jaune, 5 : vert, 6 : cyan, 7 : bleu, 8 : violet.

FIG. 2 – Comparaison des résultats obtenus avec les différentes mesures – En (a), on voit que la mesure robuste se distingue des autres, particulièrement pour l'image de la plante. Les zones où elle est plus performante que les autres correspondent aux zones de discontinuités. La deuxième image met en évidence les performances de GC dans les zones faiblement texturées, cf. coin supérieur droit de l'image. Les exemples (b) montrent la bonne complémentarité de ces deux mesures : SMPD est performante près des occultations et discontinuités et GC comble les difficultés de SMPD dans les zones faiblement texturées. Sur les images (c), il apparaît clairement que les zones difficiles à apparier correspondent aux discontinuités de profondeur et aux zones faiblement texturées.

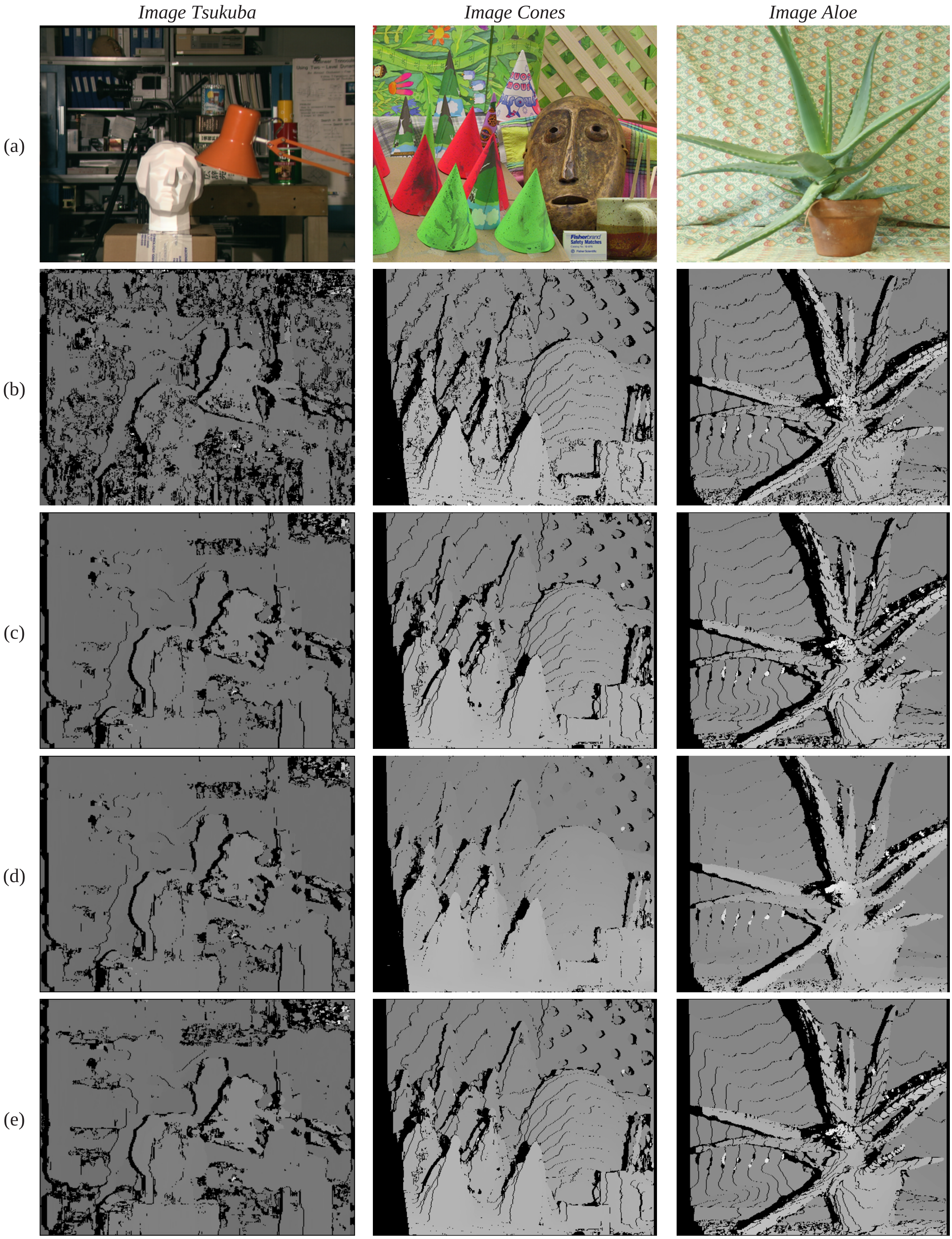


FIG. 3 – Cartes de disparités des méthodes comparées – De haut en bas, nous illustrons le travail en affichant les images de gauche, (a), les cartes de disparités obtenues avec la mesure SMPD, (b), avec la mesure GC, (c), avec la méthode ITÉRATIVE, (d), et la méthode SCORE, (e). La quatrième ligne expose les meilleurs résultats obtenus avec la méthode ITERATIVE proposée.