

**Une Nouvelle Approche de Mise en Correspondance
Stéréoscopique Dense utilisant des Méthodes
d'Inférences Imprécises**

Christophe Rabaud, Olivier Strauss, Frédéric Comby

► **To cite this version:**

Christophe Rabaud, Olivier Strauss, Frédéric Comby. Une Nouvelle Approche de Mise en Correspondance Stéréoscopique Dense utilisant des Méthodes d'Inférences Imprécises. LFA: Logique Floue et ses Applications, Nov 2004, Nantes, France. pp.197-204. lirmm-00108808

HAL Id: lirmm-00108808

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00108808>

Submitted on 23 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une nouvelle approche de mise en correspondance stéréoscopique dense utilisant des méthodes d'inférences imprécises

Fuzzy dense stereoscopic matching

Christophe RABAUD

Olivier STRAUSS

Frédéric COMBY

LIRMM, Université Montpellier II, 161, rue Ada

34392 Montpellier Cedex 5, France

e-mail : rabaud,strauss,comby@lirmm.fr

Résumé :

Cet article présente une nouvelle méthode de mise en correspondance dense de pixels dans des images stéréoscopiques. Elle s'appuie sur une modélisation sémantique graduée des niveaux de gris de l'image (via les sous-ensembles flous).

Mots-clés :

images stéréoscopiques, mise en correspondance dense, théorie des possibilités.

Abstract:

This paper deals with a new fuzzy-based method allowing a dense stereoscopic matching.

Keywords:

Stereoscopic images, dense matching, possibility theory.

1 Introduction

Depuis que D. Marr [1] a proposé une théorie de la vision stéréoscopique humaine, la communauté scientifique a porté un intérêt croissant aux systèmes binoculaires qui sont devenus très répandus. Le principe de la stéréovision consiste à utiliser plusieurs images d'une même scène prises sous différents points de vue. Ceci permet de déduire des informations 3D sur la scène à partir de ses projections 2D sur les images. Dans la plupart des cas, seules 2 caméras sont utilisées.

Les applications de la stéréovision sont diverses : construction de carte et reconnaissance aérienne [2], navigation pour la robotique [3], reconstruction d'objets [4], etc. Néanmoins, l'estimation de la position d'un point dans l'espace n'est possible que si l'on est capable de déterminer la position de sa projection dans chaque image. Il n'existe pas de solution globale au problème de mise en correspondance en raison principalement des ambiguïtés de mise en correspondance (occultations, réflexions des surfaces, manque de textures, etc.). Cepen-

dant, un certain nombre de contraintes (par exemple la géométrie épipolaire) et d'hypothèses (illumination globale constante, surface uniforme) sont communément employées pour résoudre ce problème.

1.1 Rappel sur la stéréovision

Sur la figure 1, un système stéréoscopique est modélisé en utilisant le modèle sténopé pour représenter les caméras [5]. Ce modèle de caméra est l'un des plus utilisés. Il est caractérisé d'une part par une projection qui transforme un point de l'espace 3D en un point image 2D et, d'autre part, par une transformation d'un repère métrique lié à la caméra vers un repère pixel lié à l'image [6]. Nous en rappelons ici les notations et les différents points principaux. Le système stéréoscopique que nous allons étudier se compose de deux caméras. Soit M , un point de l'espace 3D repéré par ses coordonnées dans un repère fixe lié ou non à la caméra. Soient m_1 sa projection sur la 1^{ère} image et m_2 sa projection dans la 2^{ème} image. La nature du système contraint m_2 à se trouver sur une droite qui est la projection, sur l'image 2, de la droite (c_1M) ; c_1 étant le centre de projection sténopé de la caméra 1. Cette contrainte porte le nom de contrainte épipolaire. La matrice F , appelée matrice fonda-

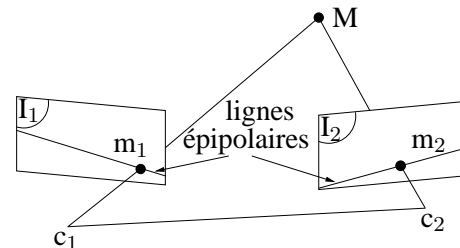


Figure 1 – La géométrie d'un capteur stéréoscopique

mentale [7], décrit cette contrainte dans le repère image. Elle est définie d’une part par la projection du point M sur les rétines de chaque caméra (paramètres intrinsèques de chaque caméra), et d’autre part par la transformation géométrique entre les deux caméras (paramètres extrinsèques inhérents au capteur stéréoscopique ou paramètres épipolaires). La contrainte épipolaire peut être exprimée par la relation :

$$m_2^t F m_1 = 0 \quad (1)$$

1.2 Mise en correspondance

La mise en correspondance est une étape essentielle pour la reconstruction de la scène 3D. Elle consiste à appairer, dans les différentes vues de la même scène, les projections d’un même objet. Il existe deux grandes classes de méthodes d’appariement. La première regroupe les méthodes utilisant des informations locales telles que les mises en correspondances de motifs [8], de primitives (coins, lignes) [9] ainsi que les méthodes utilisant le flot optique [10]. L’inconvénient des deux premières méthodes est le faible nombre de points mis en correspondance. De plus, les méthodes basées sur des informations locales sont peu robustes vis à vis des scènes ayant beaucoup de disparité (profondeur) et aux régions peu texturées. La deuxième classe regroupe les méthodes utilisant des informations globales. Appartiennent à ce groupe les méthodes utilisant la programmation dynamique [11], celles s’appuyant sur des caractéristiques invariantes aux transformations rigides [12] ou encore des techniques utilisant la propagation de croyance [12] et la diffusion non linéaire [13]. Ces méthodes sont généralement moins sensibles aux problèmes des régions ambiguës mais elles sont plus coûteuses en temps de calcul et l’élaboration d’un modèle globale n’est pas toujours possible.

La majeure partie des méthodes classiques de mise en correspondance utilisent une représentation statistique de l’erreur d’appariement. On est en droit de se poser des questions sur l’utilisation des outils statistiques lorsque le nombre d’informations à mettre en correspondance ne dépasse pas deux. Les techniques de régularisation s’appuient sur l’hypothèse d’ergodicité du bruit de mise en correspondance et utilisent le voisinage de chaque pixel pour évaluer une représentation de la statistique des variations de niveaux de gris. Plus le voisinage uti-

lisé est étendu, plus la statistique sera précise, mais moins elle sera fiable. Le manque de fiabilité étant lié, comme nous l’avons précédemment vu (dans l’étude des méthodes de mise en correspondance locales) aux régions ambiguës.

Dans cet article, nous proposons de remplacer les outils statistiques de représentation des variations par une approche ensembliste floue. L’utilisation de la théorie des sous-ensembles flous nous permet de raisonner de façon pondérée sur le pire cas. Cette méthode utilise une formulation sémantique des contraintes stéréoscopiques afin d’augmenter la robustesse de la mise en correspondance vis à vis des ambiguïtés d’appariements.

La suite de cet article est organisée comme suit : après ce paragraphe introductif, le deuxième paragraphe présente notre approche de mise en correspondance dense de pixels. Elle utilise une formulation sémantique des contraintes stéréoscopiques utilisant une modélisation floue des niveaux de gris des pixels. Dans cette même partie, nous présentons les méthodes retenues pour fusionner l’ensemble des informations afin de déterminer les mise en correspondance. Le troisième paragraphe illustre les résultats de notre approche. Enfin, le dernier paragraphe conclut sur l’apport de ce travail et ouvre un certain nombre de perspectives.

2 Notre approche

2.1 Concept de base

Plutôt que de modéliser la mise en correspondance d’un pixel m_1 de l’image 1 avec un pixel m_2 de l’image 2 par une mesure de probabilité (trop normative) ou un domaine ensembliste (pas assez nuancé), nous proposons ici d’utiliser la théorie des possibilités [14] pour associer à chaque pixel m_1 de l’image 1 un domaine flou $\mathcal{D}_2(m_1)$ de l’image 2 des pixels pouvant être mis en correspondance avec m_1 . Pour définir ce domaine, nous proposons de remplacer les règles classiques de restriction d’appariement par leur équivalent flou. Ceci peut s’énoncer de la façon suivante : la mise en correspondance potentielle entre un pixel m_1 de l’image 1 et un pixel m_2 de l’image 2 est mesurée par la possibilité $\Pi(\mathcal{D}_2(m_1); m_2) = \Pi(\mathcal{D}_1(m_2); m_1)$. Cette possibilité définit de façon symétrique aussi bien $\mathcal{D}_2(m_1)$ que $\mathcal{D}_1(m_2)$. Un domaine flou peut être vu comme un emboîtement de domaines classiques affectés de

leur mesure de confiance. Cette représentation des ensembles flous par α -coupe [15] permet de raisonner plus facilement avec des concepts ensemblistes ordinaires.

Comme les mesures de possibilités ne respectent ni le tiers exclus, ni la non-contradiction, il est nécessaire de baser la mise en correspondance sur un processus bipolaire de décision [16]. Ainsi, un même critère peut, pour un niveau de confiance donné, à la fois accepter et refuser un appariement sans pour cela contredire l'axiomatique des possibilités. C'est la prise en compte de cette propriété que l'on nomme bipolarité. Cette représentation permet de modéliser des informations incomplètes aussi bien négatives (restrictions, impossibilités, rejets) que positives (confirmations, renforcement).

2.2 Contrainte chromatique

La mise en correspondance de deux pixels à partir de leur niveau de gris s'appuie sur l'hypothèse que les projections d'un même point M de l'espace tridimensionnel ont des intensités lumineuses comparable. Les mesures de corrélations de motifs ou de distances numériques sur les intensités sont particulièrement perturbées par des variations non ergodiques provoquées par des phénomènes tel que le changement de point de vue, les occultations partielles, l'échantillonnage, la numérisation,... qui peuvent difficilement être modélisés par de simples lois normales. Nous avons proposé, au cours de précédent travaux, d'utiliser une classification floue des niveaux de gris pour rendre plus robuste l'analyse des appariements. Dans [17], la modélisation que nous avons proposé repose sur une représentation par classe des phénomènes de saturation. On modélise le niveau de gris d'un pixel par sa compatibilité avec une des deux classes que sont le noir (niveau d'excitation du pixel inférieur au seuil de saturation bas) et blanc (niveau de gris du pixel supérieur au seuil de saturation haut). Ces compatibilités sont représentées par deux classes flous duales que nous appelons "pixels blanc" et "pixels noirs". Nous avons montré que cette modélisation permet de minimiser l'a-priori dans l'utilisation des informations de niveaux de gris pour la mise en correspondance et de représenter l'imprécision des valeurs de niveaux de gris.

Par contre, les outils développés (histogrammes quasi-continus [18]) s'appuient sur une version ro-

buste des statistiques de modes qui ne peuvent être utilisés ici en raison du peu d'informations disponibles (voisinage considéré restreint). Nous proposons de reprendre cette modélisation en remplaçant les outils d'analyse statistiques par une analyse par règles. La possibilité de mettre en correspondance deux pixels en s'appuyant sur les informations chromatiques peut être mesurée par :

Prop. 1 *La mise en correspondance de deux pixels m_1 et m_2 projections du même point M est envisageable si ces deux pixels appartiennent à la même classe. C'est à dire (m_1 est blanc et m_2 est blanc) ou (m_1 est noir et m_2 est noir).*

L'utilisation de cette proposition en logique floue s'exprime simplement par la formule :

$$\Pi_{C(m_1, m_2)} = \max \left(\begin{array}{l} \min [\mu_b(m_1), \mu_b(m_2)], \\ \min [\mu_n(m_1), \mu_n(m_2)] \end{array} \right) \quad (2)$$

ou $\Pi_{C(m_1, m_2)}$ est une mesure de possibilité et $\mu_b(m)$ (resp. $\mu_n(m)$) représente le degré d'appartenance du pixel m à la classe des pixels blancs (resp. noirs). Ce n'est pas une mesure de similarité mais une mesure de "co-appartenance" à une même classe. Elle est beaucoup moins sensible à une mauvaise modélisation et à des bruits non-ergodiques [17]. Par contre, elle nécessite l'emploi d'une représentation bipolaire. La "non co-appartenance" est mesurée par :

$$\overline{\Pi_{C(m_1, m_2)}} = \max \left(\begin{array}{l} \min [\mu_b(m_1), \mu_n(m_2)], \\ \min [\mu_n(m_1), \mu_b(m_2)] \end{array} \right) \quad (3)$$

2.3 La contrainte épipolaire

Lorsque l'on dispose d'un système stéréoscopique étalonné, les positions de m_1 et m_2 sont contraintes par la géométrie épipolaire. Cette contrainte symétrique est explicitée par l'équation 1. Projetée sur l'image 2, cette contrainte exprime le fait que tout pixel de l'image 2 pouvant être mis en correspondance avec un pixel m_1 appartient à la droite l_2 définie par :

$$l_2 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = Fm_1 = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} m_1 \quad (4)$$

Ce cas idéal n'est pas réaliste. En effet, l'échantillonnage des images et la mauvaise connaissance de la matrice fondamentale F nous conduit à considérer cette relation sous un aspect ensembliste. Pour être cohérent avec notre proposition initiale, nous représentons la distribution de chaque paramètre de la matrice F par un ensemble flou triangulaire $\mathcal{F}_{i,j}$ dont le mode est défini par la valeur la plus probable produite par l'étalonnage du capteur stéréoscopique et dont le support est calculé par une analyse "pire cas" des erreurs d'étalonnage [19]. La relation (4) ne traduit plus une ligne épipolaire sur l'image 2 mais un bandeau flou épipolaire dont une représentation naïve est proposée sur la figure 2-a. Nous représentons l'imprécision liée à l'échantillonnage de l'image par une partition floue forte de \mathbb{R}^2 . Chaque pixel est alors associé à un nombre flou \mathcal{M} de distribution triangulaire dont le noyau est centré sur le pixel m , représentation modélisée sur la figure 2-b.

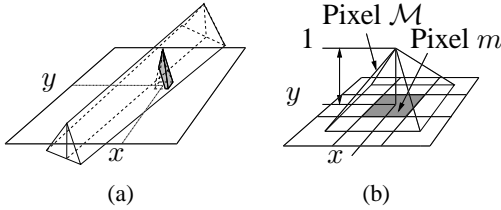


Figure 2 – Intersection entre le bandeau épipolaire et un pixel flou (a). Représentation floue d'un pixel (b).

En se servant des paramètres de la ligne épipolaire définis par l'équation 4 et en remplaçant le pixel m_1 par son modèle flou $\mathcal{M}_1 = (\mathcal{X}_1, \mathcal{Y}_1, 1)$, l'équation de la droite épipolaire peut être réécrite sous la forme ensembliste par :

$$\begin{pmatrix} \mathcal{A} \\ \mathcal{B} \\ \mathcal{C} \end{pmatrix} = \mathcal{F} \odot \mathcal{M}_1$$

$$\begin{cases} \mathcal{A} = \mathcal{F}_{11} \odot \mathcal{X}_1 \oplus \mathcal{F}_{12} \odot \mathcal{X}_2 \oplus \mathcal{F}_{13} \\ \mathcal{B} = \mathcal{F}_{21} \odot \mathcal{X}_1 \oplus \mathcal{F}_{22} \odot \mathcal{X}_2 \oplus \mathcal{F}_{23} \\ \mathcal{C} = \mathcal{F}_{31} \odot \mathcal{X}_1 \oplus \mathcal{F}_{32} \odot \mathcal{X}_2 \oplus \mathcal{F}_{33} \end{cases} \quad (5)$$

Les notations \odot et \oplus représentent respectivement les opérateurs arithmétiques de la multiplication et de l'addition appliqués aux ensembles flous [20]. De même que la contrainte épipolaire met en relation un sous-ensemble de pixels de l'image 2 avec chaque

pixel de l'image 1, la contrainte épipolaire floue ainsi définie permet de mettre en correspondance un sous-ensemble flou de l'image 2 avec chaque pixel de l'image 1. La possibilité $\Pi_{E(m_1, m_2)}$ de mettre en correspondance géométriquement un pixel m_1 de l'image 1 avec un pixel m_2 de l'image 2 s'obtient naturellement en utilisant la relation épipolaire (1) et le principe d'extension de la façon suivante :

$$\Pi_{E(m_1, m_2)} = \text{Sup}_{\substack{f_{ij, m_k} \\ i, j=1..3 \\ k=1..2}} \left\{ \min \left(\begin{array}{l} \mu_{\mathcal{F}_{ij}}(f_{ij}), \\ \mu_{\mathcal{M}_k}(m_k) \end{array} \right) \right\} \quad (6)$$

Cette écriture étant peu pratique à mettre en oeuvre, nous pouvons la modifier en nous plaçant toujours dans le "pire cas". En se servant des paramètres de la ligne épipolaire définis par l'équation (5) et en remplaçant le pixel m_2 par son modèle flou avec pour $\mathcal{M}_2 = (\mathcal{X}_2, \mathcal{Y}_2, 1)^t$, la contrainte épipolaire peut être réécrite sous la forme ensembliste par :

$$(\mathcal{A} \odot \mathcal{X}_2) \oplus (\mathcal{B} \odot \mathcal{Y}_2) \oplus \mathcal{C} \ni 0 \quad (7)$$

\mathcal{X}_2 et \mathcal{Y}_2 étant définies strictement positives, nous pouvons poser que $\mathcal{ABC} = \mathcal{A} \odot \mathcal{X}_2 \oplus \mathcal{B} \odot \mathcal{Y}_2 \oplus \mathcal{C}$. L'équation (7) revient tout simplement à chercher à quel point $0 \in \mathcal{ABC}$. Nous pouvons alors réécrire la possibilité de mise en correspondance géométrique (6) :

$$\Pi_{E(m_1, m_2)} = \mu_{\mathcal{ABC}}(0) \quad (8)$$

Notre processus étant basé sur une représentation bipolaire des informations, nous pouvons établir une mesure de non compatibilité géométrique entre deux pixels $\Pi_{\overline{E(m_1, m_2)}}$. Cela revient à chercher à quel point $0 \notin \mathcal{ABC}$. On peut alors calculer à quel point il est possible de ne pas mettre en correspondance géométriquement nos deux pixels par :

$$\Pi_{\overline{E(m_1, m_2)}} = \mu_{\overline{\mathcal{ABC}}}(0) \quad (9)$$

2.4 La contrainte d'unicité

Si on se limite à observer des objets opaques dont la disparité n'est pas très forte, alors un objet dont la projection est un pixel sur la première image a une projection qui est aussi un pixel sur la seconde image. L'expression d'une telle règle permet de réduire le domaine de mise en correspondance de chaque pixel de l'image 1 en prenant en compte les domaines de mise en correspondance des autres

pixels. On ne peut utiliser cette contrainte pour créer une mise en correspondance, mais uniquement pour modifier une distribution de mise en correspondance de façon à ce que la distribution modifiée viole moins la contrainte d'unicité que la distribution initiale. Soit $\mathcal{D}_2(m_1)$, le domaine de la seconde image mis en correspondance à priori avec un pixel m_1 de la première image. Une distribution qui viole la contrainte d'unicité est une distribution telle qu'il existe un pixel m'_1 dont le domaine $\mathcal{D}_2(m'_1)$ vérifie $\mathcal{D}_2(m_1) \cap \mathcal{D}_2(m'_1) \neq \emptyset$.

Soit $\Pi_{(m_1, m_2)}$ la possibilité de mettre en correspondance les pixels m_1 et m_2 et dont le calcul sera explicité au paragraphe 2.6. L'appariement de m_1 avec m_2 viole la contrainte d'unicité si il existe un pixel $m'_1 \neq m_1$ dans l'image 1 tel que :

$$\Pi_{(m'_1, m_2)} > \Pi_{(m_1, m_2)}$$

Dans ce cas, la possibilité de ne pas mettre en correspondance m_1 avec m_2 du point de vue de la contrainte d'unicité sachant m'_1 s'écrit :

$$\overline{\Pi_{(m_1, m_2)}} = 1 - \underset{m'_1}{Sup} \left\{ \begin{array}{l} \min(1 - \overline{\Pi_{(m_1, m_2)}}), \\ 1 - \Pi_{(m'_1, m_2)} \end{array} \right\} \quad (10)$$

2.5 La contrainte d'ordre

Sous certaines conditions [6], l'ordre des projections est le même dans les deux images. On peut établir la proposition suivante :

Prop. 2 *si m'_1 et m'_2 sont les projections d'un même point M alors tous les pixels à droite (resp. gauche) de m'_1 sont à droite (resp. gauche) du pixel m'_2 .*

Par extension, nous pouvons exprimer la proposition négative duale :

Prop. 3 *si m'_1 et m'_2 sont les projections d'un même point M alors tous les pixels à droite (resp. gauche) de m'_1 ne peuvent pas être à gauche (resp. droite) du pixel m'_2 .*

L'appariement de m_1 avec m_2 viole la contrainte d'ordre si $[(m_1 \text{ est à droite de } m'_1) \text{ ET } (m_2 \text{ est à$

$\text{gauche de } m'_2)] \text{ OU } [(m_1 \text{ est à gauche de } m'_1) \text{ ET } (m_2 \text{ est à droite de } m'_2)] \text{ ET } [\Pi_{(m'_1, m'_2)} > \Pi_{(m_1, m_2)}]$.

Dans ce cas, la possibilité de ne pas mettre en correspondance m_1 avec m_2 du point de vue de la contrainte d'ordre, connaissant l'appariement de m_1 avec m_2 , s'écrit :

$$\overline{\Pi_{(m_1, m_2)}} = 1 - \underset{m'_1, m'_2}{Sup} \left\{ \begin{array}{l} \min(1 - \overline{\Pi_{(m_1, m_2)}}), \\ 1 - \Pi_{(m'_1, m'_2)} \end{array} \right\} \quad (11)$$

2.6 Fusion des contraintes

En considérant uniquement la mise en correspondance des pixels m_1 et m_2 , alors seules sont disponibles les informations portant sur les contraintes chromatiques et géométriques. On peut établir que, pour que ces deux pixels puissent être en correspondance, il faut qu'ils soient compatible aussi bien chromatiquement que géométriquement. Cela se traduit, en choisissant min comme t -norme, par la formule suivante :

$$\Pi_{(m_1, m_2)} = \min(\Pi_{C(m_1, m_2)}, \Pi_{E(m_1, m_2)}) \quad (12)$$

avec $\Pi_{(m_1, m_2)}$ la possibilité de mise en correspondance du pixel m_1 avec le pixel m_2 . De la même façon, nous pouvons établir une mesure de possibilité de ne pas mettre en correspondance les pixels m_1 et m_2 par :

$$\overline{\Pi_{(m_1, m_2)}} = \max(\overline{\Pi_{C(m_1, m_2)}}, \overline{\Pi_{E(m_1, m_2)}}) \quad (13)$$

Ces mesures sont les états initiaux qui vont être utilisés pour calculer les premières mises en correspondance de pixels. A chaque mise en correspondance de pixels, on est capable de trouver des pixels qui enfreignent les contraintes d'ordre ou d'unicité. Ce que l'on souhaite maintenant, c'est de re-calculer les possibilités de non mise en correspondance de façon à obtenir un domaine de possibilité qui viole moins les contraintes d'ordre et d'unicité.

2.7 Cas des images stéréoscopiques redressées

Dans ce paragraphe, nous proposons de restreindre notre algorithme sur le cas particulier des images redressées [6]. Lorsqu'un capteur stéréoscopique est

correctement étalonné, c'est à dire lorsqu'on connaît la géométrie relative des deux caméras et les paramètres qui gouvernent la projection des images, il est possible de "redresser" les images. Redresser les images consiste en la création, par interpolation, des deux images qu'aurait acquies le même capteur stéréoscopique si les deux caméras étaient co-planaires, d'axes optiques confondus et de même orientation.

Lorsqu'une paire d'images est redressée, la contrainte épipolaire se simplifie : à toute ligne x_1 de l'image 1 correspond une ligne x_2 de l'image 2. Il est même possible de redresser les images de façon à ce que les deux lignes soient identiques sur les deux images.

Enfin, si m_1 est un pixel de l'image 1 de coordonnées (x_1, y_1) , m_2 appartient forcément à un segment $[a_2, m_{21}]$ de la ligne épipolaire de coordonnée x_2 . a_2 est l'intersection de cette ligne avec le bord gauche de l'image et m_{21} est le point de coordonnées (x_2, y_1) (figure 3). Cette contrainte, que

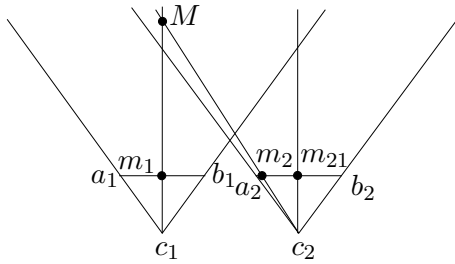


Figure 3 – Un cas particulier de stéréovision

nous avons appelé contrainte du point de visée, nous permet d'établir des mesures de possibilité de mise en correspondance et de possibilité de non mise en correspondance par :

$$\forall m_2 \notin [a_2, m_{21}] \Rightarrow \begin{cases} \Pi_{(m_1, m_2)} = 0 \\ \Pi_{(m_1, m_2)} = 1 \end{cases}$$

Ce même raisonnement doit être effectué en considérant cette fois-ci l'image 2.

3 Algorithme et étude des résultats

Nous proposons, pour simplifier l'interprétation des résultats, d'illustrer notre méthode sur une paire d'images stéréoscopiques redressées (figure 4). On considère les informations d'étalonnage comme

fiable, c'est à dire qu'à un pixel m_1 de l'image 1 appartenant à la ligne x on associe une ligne (et non plus un bandeau) épipolaire horizontale ayant comme ordonnée x dans l'image 2. Ce choix nous permet de nous concentrer uniquement sur les problèmes d'échantillonnage et de variation de luminosité. La ligne de pixels x de l'image 1 est

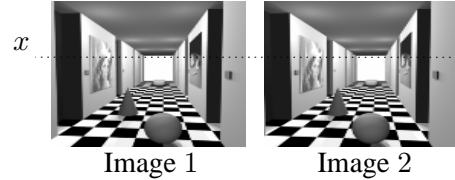


Figure 4 – Images stéréoscopique de synthèse

l'épipolaire de la ligne de pixel x de l'image 2. Nous avons représenté la répartition des possibilités de mise en correspondance (P.M.E.C.) et des possibilités de non mise en correspondance (P.N.M.E.C.) initiales tel que définis par les équations (12) et (13) sur la figure 5. Sont représentés sur l'axe des ordonnées

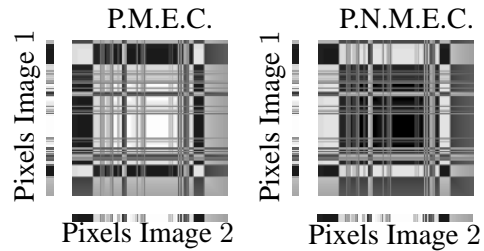


Figure 5 – Répartition des possibilités de mise en correspondance (P.M.E.C.) et de non possibilités de mise en correspondance (P.N.M.E.C.) initiales

données les pixels de la ligne x de l'image 1 et sur l'axe des abscisses les pixels de la ligne x de l'image 2. Le domaine de répartition des mesures d'appariements étant $[0, 1]$ nous proposons de les représenter sous la forme d'une image avec un pixel blanc pour une valeur de possibilité égale à 1 et un pixel noir pour une valeur de possibilité égale à 0. On peut noter que, bien que leur intensité ne soient pas égale, la possibilité de mettre en correspondance des pixels appartenant à la même classe est assez forte. Par contre, on constate que la possibilité de mettre en correspondance un pixel (ou un bloc de pixels) blanc avec un pixel (ou un bloc de pixels noirs) est faible. Inversement, la possibilité de ne pas les mettre en correspondance est forte.

En reprenant la théorie des α -coupes, on se place au niveau de confiance α_1 et on calcule un sous-

ensemble classique Π^{α_1} contenant l'ensemble des appariements tel que $\Pi_{(m_1, m_2)} \geq \alpha_1$ (à l'initialisation de l'algorithme $\alpha_1 = 1$). Pour chaque niveau de coupe, on détermine l'ensemble des couples de pixels mis en correspondance ne répondant pas à ce niveau de confiance d'appariement tel que $\Pi_{(m'_1, m'_2)} \notin \Pi^{\alpha_1}$. Pour tous ces pixels, on calcule une nouvelle valeur de possibilité de mise en correspondance et de possibilité de non mise en correspondance grâce aux contraintes d'ordre, d'unicité et de point de visée.

Sur la figure 6, on peut voir la répartition fi-

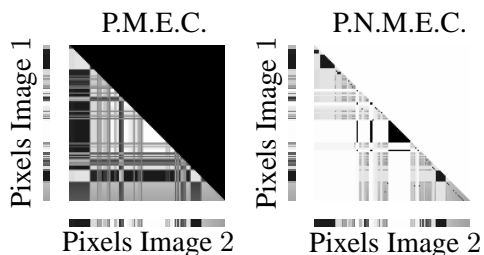


Figure 6 – Répartition des possibilités de mise en correspondance (P.M.E.C.) et de non possibilités de mise en correspondance (P.N.M.E.C.) finales

nale des possibilités de mise en correspondance et les possibilités de non mise en correspondance des pixels de nos deux lignes. On constate tout d'abord les effets de la contrainte de point de visée. Cette contrainte définit, dans la représentation que nous avons choisi, une zone triangulaire d'impossibilité stricte de mise en correspondance (possibilité=0 ; non possibilité=1).

On constate les effets de la contrainte d'unicité : aux pixels blancs (resp. noir) de l'image 1 on leur a uniquement attribuer les pixels blancs (resp. noir) de l'image 2 en rendant non possible les autres appariements. La contrainte d'ordre, un peu de la même manière que la contrainte de point de visée mais de façon plus nuancée, définit des zones de non possibilité de mise en correspondance.

Pour appairier deux pixels, il faut que leur mise en correspondance soit possible et ne soit pas non possible. Nous estimons le degré de vérité de la proposition "le pixel m_1 de l'image 1 est le correspondant du pixel m_2 dans l'image 2" par $V_{(m_1, m_2)} = \Pi_{(m_1, m_2)} - \Pi_{\overline{(m_1, m_2)}}$. Cette mesure, définie sur $[-1; 1]$, nous permet d'établir une nouvelle carte d'appariement. En utilisant un algorithme de recherche d'appariement par maximisation de fonction de ressemblance [21] basée sur cette mesure

de vérité, on peut mettre en correspondance l'ensemble des pixels des deux images. On peut voir sur la figure 7 la carte des disparités. Un pixel blanc représente un point proche, un pixel noir représente un point éloigné. L'utilisation de ce genre d'algorithme semble permettre de gérer les occultations ainsi que les effets de bords d'image. Ce travail fait l'objet de nos expérimentations actuelles.

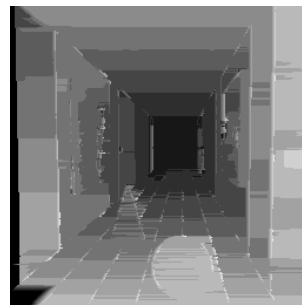


Figure 7 – Carte de disparité

4 Conclusions et perspectives

Nous avons présenté dans cet article une nouvelle approche de mise en correspondance dense. Les principaux objectifs de cette méthode sont de s'affranchir des problèmes liés aux variations de luminosité entre les vues ainsi que des problèmes d'échantillonnage liées à la nature discrète du signal délivré par le capteur. Nous avons constaté, au cours de nos expérimentations, que notre algorithme permet de mettre en correspondance des pixels qui n'ont pas exactement les mêmes niveaux de gris. Il permet aussi de mettre en correspondance des blocs de pixels ayant des niveaux de gris approchant bien que le nombre composant ces blocs dans les deux images soient différents. De plus, notre méthode permet de prendre en compte l'imprécision de la phase d'étalonnage du capteur stéréoscopique. Néanmoins, la formulation que nous proposons pour le calcul de la possibilité de non mise en correspondance ne rend pas compte de l'aspect nuancé de la distribution originale de confiance.

Nous travaillons actuellement sur une nouvelle méthode de calcul des possibilités ainsi que sur une nouvelle méthode de fusion des contraintes. Nous pensons, par ailleurs, pouvoir lever les quelques incertitudes en utilisant les informations du voisinage. Enfin, nous travaillons sur une nouvelle méthode pour le calcul final d'appariement basée sur les

méthodes de recherche de chemin par maximisation de critères.

Références

- [1] D. Marr et T. Poggio. A computational theory of human stereo vision. *Royal Society of London*, B204, 1979.
- [2] M. Courtois et G. Weill. The spot satellite remote sensing mission. *Photogrametric Engineering and Remote Sensing*, pages 1163–1171, 1981.
- [3] M. Kumano, A. Ohya, et S. Yuta. Obstacle avoidance of autonomus mobile robot using stereo sensor. *Second International Symposium on Robotic and Automation ISRA*, pages 497–502, 2000.
- [4] P. Fua et Y. Leclerc. Object-centered surface reconstruction. *International Journal on Computer Vision*, 16 :35–56, 1995.
- [5] R. Hartley et A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [6] O. Faugeras. *Three-Dimensional Computer Vision, a Geometric Viewpoint*. MIT Press, 1993.
- [7] G. Xu et Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recongnition*, volume 6. Kluwer Academic, 1996.
- [8] D. N. Bhat et S. K. Nayar. Ordinal measures for visual correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 :415–423, 1998.
- [9] S. Birchfield et C. Tomasi. Multiway cut fro stereo and motion with slanted surfaces. *Int'l Joint Conf. Artificial Intelligence*, 1 :489–495, 1999.
- [10] B. D. Lucas et T. Kanade. An iterative image registration technique with an application to stereovision. *Int'l Joint Conf. Artificial Intelligence*, pages 674–679, 1981.
- [11] Y. Ohta et T. Kanade. Stereo by intra- and inter-scanline search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7 :139–154, 1985.
- [12] C. Tomasi et R. Manduchi. Stereo matching as a nearest-neighbor problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 :330–340, 1998.
- [13] D. Scharstein et R. Szeliski. Stereo matching with non-linear diffusion. *Int'l Joint Conf. Artificial Intelligence*, 28 :155–174, 1998.
- [14] D. Dubois et H. Prade. *Possibility Theory : An Approach to Computerized Processing of Uncertainty*. Plenum Press, 1988.
- [15] D. Dubois et H. Prade. Fuzzy sets, a convenient fiction for modeling vagness and possibility. *IEEE Transactions on Fuzzy Systems*, pages 16–21, 1994.
- [16] S. Benferhat, D. Dubois, S. Kaci, et H. Prade. Bipolar possibilistic representations. *Uncertainty in Artificial Intelligence*, 2002.
- [17] F. Comby. *Estimation du Mouvement Apparent Majoritaire dans une Sequence d'Images Video par Accumulation de Votes Bimodaux sur un Histogramme Approche*. PhD thesis, Université de Montpellier II, 2001.
- [18] O. Strauss et F. Comby. Modal estimation using rough histograms. *LFA*, pages 35–42, 2002.
- [19] B. Telle. *Methode Ensembliste pour une Reconstruction Tridimensionnelle Garantie par Stereovision*. PhD thesis, Université de Montpellier II, 2003.
- [20] D. Dubois et H. Prade. *Fuzzy Sets and Systems*. Academic Press, 1980.
- [21] I. J. Cox, S. L. Hingorani, et S. B. Rao. A maximun likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3) :542–567, 1996.