HAL
open science

Contrôle de la direction privilégiée dans deux variantes de la méthode de reconstruction stéréoscopique SGM

Mathias Paget, Jean Philippe Tarel, Pascal Monasse

To cite this version:

Mathias Paget, Jean Philippe Tarel, Pascal Monasse. Contrôle de la direction privilégiée dans deux variantes de la méthode de reconstruction stéréoscopique SGM. RFIAP 2018, Reconnaissance des Formes, Image, Apprentissage et Perception et Conférence Française de Photogrammétrie et de Télédétection, Jun 2018, Marne-la-Vallée, France. 8 p. hal-01822857

HAL Id: hal-01822857

https://enpc.hal.science/hal-01822857v1

Submitted on 25 Jun 2018

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.

Contrôle de la direction privilégiée dans deux variantes de la méthode de reconstruction stéréoscopique SGM

M. Paget[1] J-P. Tarel[1] P. Monasse[2]

[1] Université Paris-Est, LEPSIS, COSYS, IFSTTAR, F-77447 Marne-la-Vallée, France
[2] LIGM (UMR 8049), École des Ponts, UPE, Champs-sur-Marne, France

mathias.paget@ifsttar.fr, jean-philippe.tarel@ifsttar.fr, monasse@imagine.enpc.fr

Résumé

La méthode SGM (Semi-Global Matching) se situe dans les meilleures méthodes dans les bases d'évaluation de reconstruction stéréoscopique. Elle comporte en outre l'avantage de s'exécuter assez rapidement pour être intégrée à des applications temps réel. De ce fait, cette méthode a été très étudiée et des variantes ont été proposées pour corriger certains défauts de la méthode, à savoir des cartes de disparité présentant des artefacts en tramage du fait d'une prise en compte partielle de la régularisation. Parmi les variantes de SGM, les méthodes MGM (More Global Matching) et CAT (Cost Aggregation Table) prennent en compte simultanément deux directions lors de l'étape d'accumulation afin de les intégrer à la régularisation. Les solutions obtenues sont plus lisses, avec moins d'artefacts. Toutefois, nous avons remarqué que ces deux méthodes ont tendance à privilégier l'information qui se trouve sur la diagonale des deux directions choisies. Nous proposons donc d'ajouter un paramètre à ces deux méthodes pour contrôler la direction privilégiée. Nous montrons expérimentalement l'influence de ce paramètre sur un extrait des bases Middlebury et KITTI.

Mots Clef

SGM, reconstruction stéréoscopique, optimisation.

Abstract

The SGM method provided some of the best results over most of stereo-vision benchmarks. Moreover, SGM is fast enough to be used in real time applications. This method have thus been studied a lot and variants in the method had been proposed in order to correct defaults, namely dithering artifacts due to partial inclusion of the regularization. Among SGM variants, the MGM (More Global Matching) and CAT (Cost Aggregation Table) methods take two directions into account simultaneously during the aggregation step, in order to better include the regularization. Obtained solutions are smoother with fewer artifacts. However we notice that information located in the diagonal of the two chosen directions is privileged. We thus propose to add a parameter to the two variants in order to control the privileged direction. We show experimentaly the influence of this parameter on an extract of Middlebury and KITTI databases.

Key words

SGM, stereo reconstruction, optimization.

1 Introduction

La résolution d'un problème de reconstruction stéréoscopique binoculaire passe par la recherche de correspondants entre les images du couple stéréoscopique. Dans [14], on distingue les méthodes dites "locales" des méthodes dites "globales". Les méthodes locales consistent à apparier un ensemble de points en comparant une information locale entre les deux images du couple stéréoscopique. Souvent rapides d'exécution, elles sont mises en défaut par les zones ambiguës des images (zones uniformes ou répétitives). Les méthodes globales posent le problème sous la forme de la minimisation d'une énergie calculée sur toute l'image, permettant de gérer les zones ambiguës. Elles sont souvent formulées comme un champs de Markov 2D, or la résolution d'un tel problème est difficile. Différentes méthodes approchées ont été proposées, par propagation de croyance [3] ou par coupure de graphe [17] pour optimiser efficacement une telle énergie.

Une de ces méthodes se démarque des autres : la méthode SGM [7] où la régularisation 2D est approximée efficacement par l'optimisation selon des alignements de pixels. Cette approximation entraîne un manque de cohérence entre deux alignements contigus donnant lieu à des artefacts en tramage dans la carte de disparité. Toutefois le coût de calculs de cette approximation rend cette méthode compatible avec les applications temps réel. De ce fait, cette méthode a beaucoup été étudiée, par exemple dans [1] SGM est mis en lien avec d'autre méthodes d'optimisation discrète. Ainsi, de nombreuses évolutions et variations de la méthode ont été proposées visant à corriger ses défauts. Dans [6], SGM est étendu à une régularisation du deuxième ordre. Dans [5], une approche itérative multi-résolution est mise en œuvre. D'autres approches adaptent le modèle de la reconstruction à l'aide d'indices qui visent à détecter les

valeurs aberrantes. Cet indice peut être construit, comme dans [11], ou appris, comme dans [12], avec dernièrement des approches par apprentissage dit profond, comme dans [16]. Enfin dans [2], MGM et dans [4] CAT, deux directions sont prises en compte, au lieu d'un simple alignement de pixels. Il en résulte des solutions plus lisses grâce à une meilleure uniformité de la régularisation.
En analysant ces deux dernières méthodes, nous avons remarqué qu'elles ont tendance à privilégier l'information se trouvant dans une direction particulière. Nous proposons d'ajouter un paramètre pour contrôler cette direction. Dans la section 2, nous présentons les deux méthodes évoquées précédemment et nous introduisons un paramètre de contrôle de la direction privilégiée pour ces deux méthodes. Dans la section 3, nous montrons expérimentalement l'influence de ce paramètre sur un extrait des bases d'évaluation Middlebury et KITTI. Les principaux résultats de cet article sont tirés de [10].

2 Méthodes

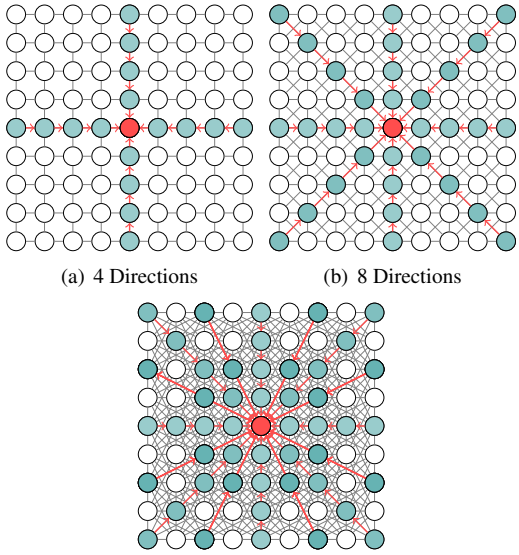
Le problème de reconstruction stéréoscopique en géométrie rectifiée peut se ramener à la recherche d'une carte de profondeur. Chaque pixel de la carte de profondeur se voit attribuer un label de profondeur, l_p étant le label associé au pixel p prenant ses valeurs dans $\mathcal{L}$. A l'aide de l'interprétation bayésienne, la solution du problème est formulée comme l'affectation minimale d'une énergie. Cette énergie est composée de deux termes. Le premier est un terme de vraisemblance, qui est une somme de fonctions de dissimilarité entre deux positions de pixels gauche-droit. Le second est un terme de régularisation qui est une somme de fonctions mettant en jeu un couple de pixels voisins,

$$E(\boldsymbol{l}) = \sum_{p \in \boldsymbol{I}} D_p(l_p) + \sum_{(p,q) \in \boldsymbol{N}} R_{pq}(l_p, l_q) \tag{1}$$

où $\boldsymbol{l}$ est le vecteur des labels, D_p est la fonction de dissimilarité associée au pixel p, par exemple la différence absolue de l'intensité entre les deux pixels appariés. R_{pq} est la fonction de régularisation entre les pixels p et q, souvent une fonction de la différence de leur label. $\boldsymbol{I}$ est la carte de disparité et $\boldsymbol{N}$ un ensemble de pixels voisins. Ici nous prendrons en compte le voisinage en 4-connexité, D_p et R_{pq} non-négatifs avec $R_{pq}(l,l) = 0,\ \forall l \in \mathcal{L}$. Optimiser une telle énergie revient à résoudre un champs de Markov 2D du premier ordre. Il s'agit d'un problème, en général, NP-difficile. Ainsi, une résolution approchée est mise en œuvre.

2.1 SGM

La méthode SGM, proposée dans [7], consiste à décomposer la résolution du champs de Markov 2D en problèmes d'optimisation 1D, en prenant en compte l'information de plusieurs alignements de pixels formant une étoile autour du pixel considéré. Soit $\boldsymbol{\varrho}$ un ensemble des directions dans l'image tel que si $\varrho \in \boldsymbol{\varrho}$ alors $-\varrho \in \boldsymbol{\varrho}$, de sorte que toutes



(a) 4 Directions (b) 8 Directions

(c) 16 Directions

FIGURE 1 – Chemins pris en compte dans l'étape d'accumulation de SGM pour un pixel donné (en rouge) pour 4, 8 et 16 directions. Les pixels pris en compte, l'ensemble noté $\boldsymbol{\chi}_p^{\boldsymbol{\varrho}}$, sont représentés en bleu-vert. Plus le nombre de directions est grand, plus la couverture de l'image est grande.

les directions viennent par couple. La méthode SGM se décompose en deux étapes. La première étape consiste en une accumulation des coûts selon les directions $\boldsymbol{\varrho}$: si la position p est en dehors de l'image, $L_p^{\varrho}(l) = 0$. Sinon $L_p^{\varrho}(l)$ est défini de manière récursive en fonction des valeurs obtenues à la position suivante dans la direction ϱ, $L_{p+\varrho}^{\varrho}(l)$:

$$\begin{aligned} L_p^{\varrho}(l) =& D_p(l) + \min_k R_{p,p+\varrho}(l,k) + L_{p+\varrho}^{\varrho}(k) \\ &- \min_k L_{p+\varrho}^{\varrho}(k) \end{aligned} \tag{2}$$

Le terme $\min_k L_{p+\varrho}^{\varrho}(k)$ est constant et ne modifie pas la solution du problème mais permet de garder les coûts dans un intervalle. Dans la deuxième étape, les coûts calculés pour chaque direction sont sommés,

$$SGM_p(l) = \sum_{\varrho} L_p^{\varrho}(l) \tag{3}$$

et, pour chaque pixel p, le label qui correspond au coût minimal est sélectionné,

$$l_p^{SGM} = \operatorname*{argmin}_k SGM_p(k). \tag{4}$$

Soit $\boldsymbol{\chi}_p^{\boldsymbol{\varrho}}$ l'ensemble des pixels dans les directions $\boldsymbol{\varrho}$ par rapport à p, les pixels de $\boldsymbol{\chi}_p^{\boldsymbol{\varrho}}$ sont représentés en bleu dans la Figure 1 pour 4, 8 et 16 directions par rapport au pixel central.
Dans [1], un terme correctif est appliqué :

$$SGM_p^*(l) = \sum_{\varrho} L_p^{\varrho}(l) - (|\varrho| - 1) D_p(l) \tag{5}$$

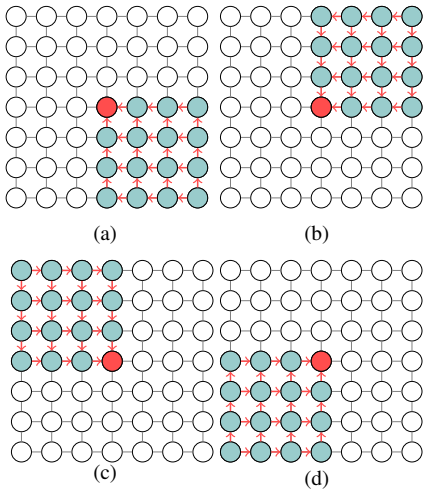


FIGURE 2 – Représentation des parties de l'image mises en jeu lorsque deux directions d'accumulation sont prises en compte simultanément comme dans les méthodes CAT et MGM.

où $|\varrho|$ est le nombre de directions prises en compte. Par construction, les valeurs $SGM_p^*(l)$ calculées peuvent s'écrire comme le minimum d'une énergie sous la contrainte $l_p = l$:

$$SGM_p^*(l) = \min_{\boldsymbol{l}|l_p=l} E^{\boldsymbol{\chi}_p^{\varrho}}(\boldsymbol{l}) \tag{6}$$

avec $E^{\boldsymbol{\chi}_p^{\varrho}}$ l'énergie restreinte aux pixels de $\boldsymbol{\chi}_p^{\varrho}$,

$$E^{\boldsymbol{\chi}_p^{\varrho}}(\boldsymbol{l}) = \sum_{p \in \boldsymbol{I} \cap \boldsymbol{\chi}_p^{\varrho}} D_p(l_p) + \sum_{(p,q) \in \boldsymbol{N} \cap (\boldsymbol{\chi}_p^{\varrho})^2} R_{pq}(l_p, l_q) \tag{7}$$

Dans la suite de l'article, lorsqu'il est question de SGM, nous nous référons à SGM dans 4 directions avec la correction montrée dans (6).

2.2 CAT et MGM

La méthode CAT, introduite dans [4], est une variante de SGM qui prend en compte deux directions lors de l'étape d'accumulation des coûts : la direction ϱ et la direction perpendiculaire $\varrho_\perp$. Cette étape est réalisée de manière analogue à (2) à la différence que le minimum des valeurs obtenues dans chacune des directions est sélectionné :

$$\begin{aligned} L_p^{\varrho}(l) =& D_p(l) \\ &+ \min_{\rho \in \{\varrho, \varrho_\perp\}} \min_k R_{p,p+\rho}(l,k) + L_{p+\rho}^{\varrho}(k) \end{aligned} \tag{8}$$

En analysant le coût L^{ϱ}, nous remarquons qu'il correspond au chemin de coût minimal par déplacement de ϱ ou $\varrho_\perp$ depuis p, de sorte que les plus courts chemins en distance de Manhattan au bord de l'image sont avantagés. Il est à noter que la solution issue de CAT est sensible à une constante ajoutée sur une attache aux données. Cette dernière peut modifier la sélection faite par le minimum selon deux directions. Nous proposons de retrancher une constantes aux coûts accumulés, de manière analogue à [7]. L'expression (8) devient :

$$\begin{aligned} L_p^{\varrho}(l) =& D_p(l) \\ &+ \min_{\rho \in \{\varrho, \varrho_\perp\}} \Big\{ \min_k R_{p,p+\rho}(l,k) + L_{p+\rho}^{\varrho}(k) \\ &\qquad - \min_k L_{p+\rho}^{\varrho}(k) \Big\} \end{aligned} \tag{9}$$

Nous appellerons cette méthode CAT_0. L'interprétation est alors plus difficile. Toutefois, en considérant que la sélection par le minimum entre les deux directions suit une loi de Bernoulli de paramètre 0.5, la probabilité $P^{\varrho}(p,q)$ d'arriver en p en passant par q (q dans le cadran défini par ϱ et $\varrho_\perp$), suit une loi binomiale $B^{\varrho}(p,q)$ de paramètre $(v_i + v_j, 0.5)$:

$$P^{\varrho}(p,q) = B^{\varrho}(p,q) = \frac{1}{2^{v_i+v_j}} \binom{v_i + v_j}{v_i} \tag{10}$$

avec (v_i, v_j) les coordonnées du vecteur entre p et q dans la base $(\varrho, \varrho_\perp)$. D'une part, la probabilité décroît en fonction de la distance de Manhattan entre les deux pixels. D'autre part, la probabilité est maximale lorsque $v_i = v_j$, de sorte que les pixels qui ont le plus de chance d'être sélectionnés sont dans la direction de la diagonale entre ϱ et $\varrho_\perp$ par rapport au pixel p.

La méthode MGM, proposée dans [2], est aussi une variante de SGM qui prend en compte deux directions lors de l'étape d'accumulation des coûts. A la différence de CAT, c'est la moyenne des valeurs obtenues dans chacune des directions qui est sélectionnée :

$$\begin{aligned} L_p^{\varrho}(l) =& D_p(l) \\ &+ \frac{1}{2} \sum_{\rho \in \{\varrho, \varrho_\perp\}} \min_k R_{p,p+\rho}(l,k) + L_{p+\rho}^{\varrho}(k) \end{aligned} \tag{11}$$

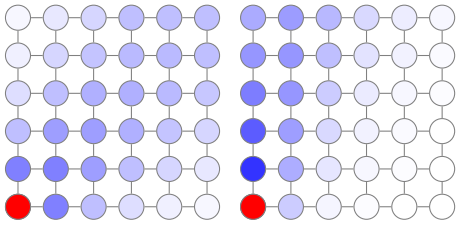
La valeur de L^{ϱ} dépend de tous les pixels compris dans le cadran défini par les vecteurs ϱ et $\varrho_\perp$ comme illustré Figure 2. Afin de comprendre comment se traduit la modification de la récursion de SGM dans MGM sur l'énergie, nous définissons la contribution $C^{\varrho}(p,q)$ comme la sommes des coefficients devant les termes en D_q dans L_p^{ϱ}. En observant la récursion (11), on note que $C^{\varrho}(p,p) = 1$ et que $C^{\varrho}(p,q) = \frac{1}{2}C^{\varrho}(p+\varrho, q) + \frac{1}{2}C^{\varrho}(p+\varrho_\perp, q)$, avec $p \neq q$. On en déduit que la contribution d'un pixel à un autre suit une loi binomiale de paramètre 0.5 qui dépend de leur position relative,

$$C^{\varrho}(p,q) = B^{\varrho}(p,q) \tag{12}$$

Aussi, plus les chemins "zigzaguent", plus ils ont d'importance dans le calcul du coût MGM.

2.3 Ajout d'un paramètre à CAT et MGM

Nous remarquons que les deux variantes qui tentent d'exploiter deux directions ont tendance à peu prendre en



(a) $a = 0.5$ (b) $a = 0.8$

FIGURE 3 – Contributions des pixels au pixel rouge pour deux valeurs de paramètre a dans (13). Une contribution de 1 correspond à un bleu profond. Les valeurs successives selon la perpendiculaire à la première diagonale suivent une loi binomiale. Dans le cas où le paramètre a est proche de 0.5, les pixels situés sur la première diagonale ont une plus forte contribution. Lorsque le paramètre est proche de 1, ce sont les pixels à la verticale qui ont une plus forte contribution.

compte les chemins droits. Cet effet s'explique par le fait que les chemins droits sont moins fréquents que ceux qui zigzaguent. Nous avons vu que ces deux méthodes font intervenir une loi de Bernoulli de paramètre 0.5. Dans la méthode MGM, le paramètre 0.5 vient de l'opérateur moyenne, mais il est possible d'influer sur ce paramètre en modifiant la pondération selon les directions ϱ et $\varrho_\perp$. Ainsi la règle de récursion peut être modifiée en :

$$\begin{aligned} L_p^\varrho(l) =& D_p(l) \\ &+ (1-a)\min_k R_{p,p+\varrho}(l,k) + L^\varrho_{p+\varrho}(k) \\ &+ a \min_k R_{p,p+\varrho_\perp}(l,k) + L^\varrho_{p+\varrho_\perp}(k) \end{aligned} \tag{13}$$

La Figure 3 illustre la modification des contributions en fonction du paramètre a. Pour rendre le problème symétrique, l'accumulation doit être réalisée pour chaque cadran avec les paramètres a et $1-a$. Ainsi pour modifier le MGM, le nombre d'accumulation à réaliser est double par rapport à SGM. Cette méthode est désignée par la suite par MGM_a. Il est à noter que MGM_1 est la méthode SGM et que $\text{MGM}_{0.5}$ est le MGM original.
Dans la méthode CAT_0, le paramètre 0.5 vient du fait que l'on considère la sélection des deux directions comme équiprobable. Cette méthode peut elle-aussi être modifiée pour favoriser la direction ϱ par rapport à la direction $\rho_\perp$ en ajoutant une constante K aux coûts issus de la direction perpendiculaire :

$$\begin{aligned} L_p^\varrho(l) &= D_p(l) \\ &+ \min \begin{cases} \min_k R_{p,p+\varrho}(l,k) + L^\varrho_{p+\varrho}(l) \\ \qquad - \min_k L^\varrho_{p+\varrho}(l) \\ \min_k R_{p,p+\varrho_\perp}(l,k) + L^\varrho_{p+\varrho_\perp}(l) \\ \qquad - \min_k L^\varrho_{p+\varrho_\perp}(l) + K \end{cases} \end{aligned} \tag{14}$$

Cette méthode est désignée par la suite comme CAT_K. Il est à noter que CAT_∞ correspond à la méthode SGM et nous retrouvons la méthode CAT_0 lorsque $K = 0$.

3 Expérimentations

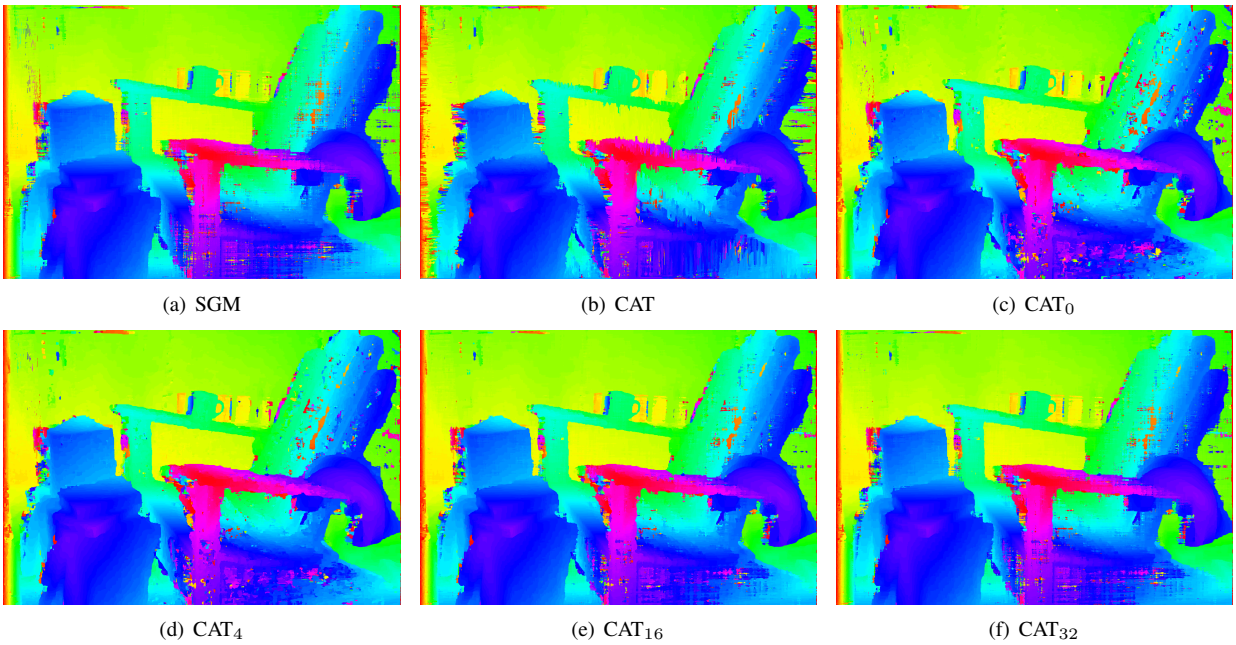
Dans cette section, nous présentons différents résultats de reconstruction stéréoscopique par des variantes de SGM. Nous allons évaluer sur un ensemble de 10 couples stéréoscopiques pris parmi les 12 couples de la base Middleburry que nous nommons "de 2014" : huit couples de la base de 2014 [13] (nommés Adirondack, Motorcycle, Piano, Pipes, Playroom, Playtable, Recycle, Shelves), un de la base de 2005 [8] (ArtL) et un de la base de 2003 [15] (Teddy). Nous avons retiré les couples possédant plus de 256 valeurs de disparité. L'attache aux données du modèle est le Census [18] calculé sur les intensités d'une fenêtre 5 par 5 pixels et agrégé 2 fois par CBCA (*Cross Based Cost Agregation*, [19]). La régularisation est la fonction de Potts lissée,

$$R(l,k) = \begin{cases} 0, & \text{si } l = k \\ P1, & \text{si } |l-k| = 1 \\ P2, & \text{si } |l-k| \geq 1 \end{cases} \tag{15}$$

de paramètres $(P1, P2) = (24, 96)$. Le critère d'évaluation, noté err.M, est le pourcentage de pixels dont la disparité est à une différence de plus de un de la référence. L'évaluation est réalisée sur les points non-occultés de la vérité terrain.
Pour compléter ces résultats, une évaluation sur les 10 premiers couples de la base de test de KITTI de 2015 [9] est réalisée. Le modèle est le même que précédemment avec les paramètres $(P1, P2) = (1.2, 23)$. L'énergie est multipliée par 10 pour ne manipuler que des valeurs entières. Le seuil du critère d'évaluation, noté err.K, est fixé à 3 pixels de différence. Nous utilisons pour cela la vérité terrain "non-occ" qui ne contient que les points visibles dans l'autre image du couple stéréoscopique.

3.1 CAT

Nous comparons la méthode SGM avec les méthodes CAT et les variantes CAT_K pour différentes valeurs de K : 0, 4, 16 et 32. Nous voulons observer l'influence du paramètre K sur les résultats. Les cartes des disparités estimées pour le couple Adirondack sont présentées dans la Figure 4. Nous observons que CAT réduit une partie des artefacts de tramage de la méthode SGM mais des artefacts en ligne apparaissent perpendiculairement aux contours des objets. La variante CAT_0 ne produit pas de tramage comme SGM ni de ligne comme CAT, mais plutôt des tâches d'erreur. Lorsque le paramètre K augmente, les solutions forment un continuum entre la solution de la méthode CAT_0 et celle de la méthode SGM.
Les taux d'erreur pour les couples de la base KITTI et Middleburry sont présentés Figure 5. Sur la base Middleburry, la méthode CAT donne systématiquement un nombre de pixels faux plus élevé que pour la méthode SGM, supérieur en moyenne de +3.05%. Parmi les va-



(a) SGM (b) CAT (c) CAT_0

(d) CAT_4 (e) CAT_{16} (f) CAT_{32}

FIGURE 4 – Résultats de reconstruction pour les méthodes SGM, CAT et CAT_0, CAT_4,CAT_{16} et CAT_{32}. La solution CAT comporte des artefacts en ligne suivant la perpendiculaire aux contours. La méthode CAT_0 présente moins d'artefact au niveau des contours, mais des taches d'erreur supplémentaires. À mesure que le paramètre K augmente, la solution tend vers la solution de la méthode SGM.

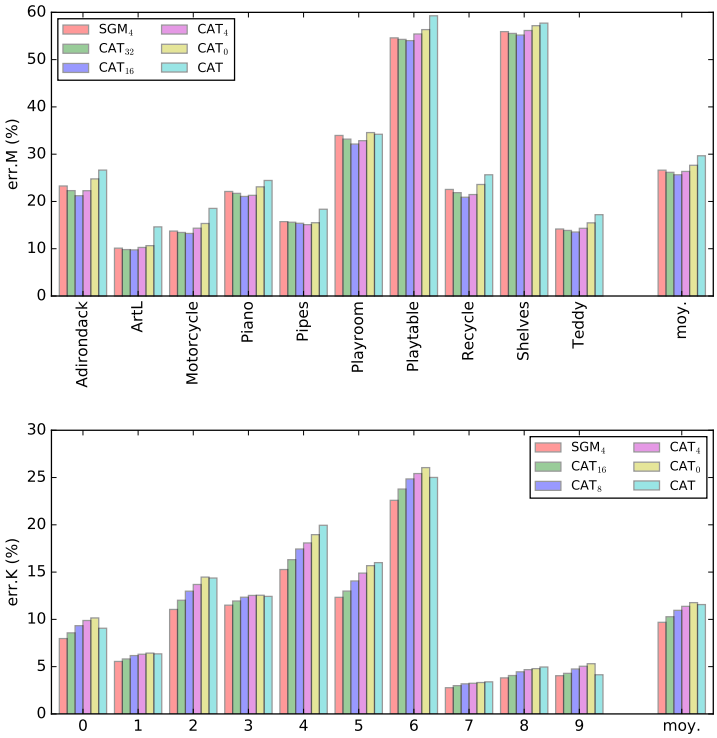


FIGURE 5 – Taux d'erreur des méthodes SGM dans 4 directions, CAT, CAT_0, CAT_4, CAT_{16}, CAT_{32} et sur 10 couples de la base Middlebury (err.M) en haut, et sur les 10 premiers couples de la base KITTI15 (err.K) en bas. Sur la base Middlebury, la méthode CAT détériore les résultats par rapport à SGM. Les variantes de CAT donnent les meilleurs résultats pour $K = 8$ et surpasse ceux la méthode SGM. Sur la base KITTI, toutes les méthodes fondées sur CAT donnent de moins bons résultats que SGM. Les variantes de CAT donnent alors de meilleures résultats à mesure qu'elle s'approche de la méthode SGM.

riantes, les meilleurs taux sont obtenus pour $K = 16$, avec en moyenne -0.98% de pixels faux par rapport à la méthode SGM. Sur la base KITTI, la méthode CAT détériore les résultats de -0.1% à -4.68%, en moyenne de -1.87%. Les variantes de CAT détériorent elles-aussi les résultats. Cette détérioration est plus forte pour les faibles valeurs du paramètre K que pour les grandes valeurs : le nombre de pixels faux augmente en moyenne de +0.59% entre SGM et CAT_{16} et de +2.08% entre SGM et CAT_0.

3.2 MGM

Nous comparons la méthode SGM avec les méthodes MGM et les variantes de MGM_a pour différentes valeurs du paramètres a : 0.9, 0.8, 0.7 et 0.6. Nous voulons observer l'influence de la pondération sur les résultats. Les images de disparité pour le couple Adirondack sont montrées dans la Figure 6. Les artefacts de tramage de la méthode SGM prennent la forme de taches dans la méthode MGM. Nous remarquons que les solutions forment un continuum entre SGM et MGM au fur et à mesure que le paramètre varie entre 1 (SGM) et 0.5 (MGM).
Les taux d'erreur pour les couples de la base KITTI et Middleburry sont présentés Figure 7. Sur la base Middlebury, sur les couples ArtL et Motocycle, les résultats de la méthode MGM sont moins bons que ceux de SGM (+2.64% et +1.10% de pixels faux). Néanmoins sur les dix couples, la méthode MGM donne en moyenne -0.54% de pixel faux que SGM. Si l'on regarde les résultats des variantes de MGM, elles donnent presque les mêmes taux d'erreur que MGM pour les couples Pipes, Playtable et Shelves. Hormis pour le couple ArtL, les variantes de MGM donnent de meilleurs taux d'erreur que SGM. Le meilleur taux moyen est obtenu par $\text{MGM}_{0.8}$, dont le nombre de pixels faux est en moyenne de -1.56% par rapport à SGM. Sur la base KITTI, MGM et ses variantes dégradent les résultats. Cette dégradation se réduit à mesure que les variantes de MGM s'approchent de la méthode SGM. En moyenne, les nombres de pixels faux de SGM, $\text{MGM}_{0.9}$, $\text{MGM}_{0.8}$ sont très proches (moins de 0.2% de différence).

3.3 Discussion

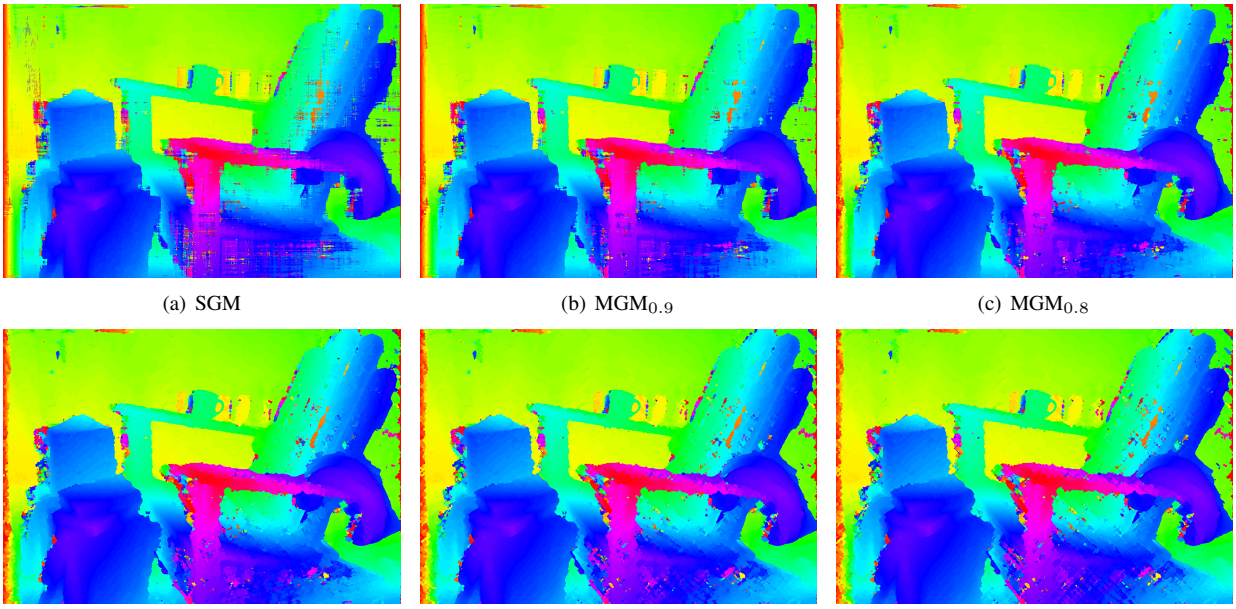
Nous avons observé que les méthodes CAT et MGM ainsi que leur variantes présentent un intérêt relativement limité sur Midllebury, et peuvent détériorer les résultats sur la base KITTI. Cette absence d'amélioration sur la base KTTI peut s'expliquer par la structure de la scène où les directions verticales et horizontales de l'image sont privilégiées (scène routière). Or les variantes de SGM ont tendance à diminuer l'importance des a priori verticaux et horizontaux par rapport aux autres directions. Sur la base Middlebury, les conclusions sont plus nuancées : la méthode CAT détériore les résultats par rapport à SGM. Les variantes proposées de CAT donnent de meilleurs résultats et l'une d'elle, CAT_{16}, donne même de meilleurs résultats que SGM. La méthode MGM et ses variantes donnent des taux d'erreur plus bas que SGM, le meilleur taux étant obtenu pour $\text{MGM}_{0.8}$. Ainsi, parmi les variantes de CAT et MGM, celles qui donnent les meilleurs résultats sont celles qui, de par leur valeur de paramètre, donnent les solutions à mi-chemin entre SGM et la méthode dont elles sont la variante. Cette amélioration reste marginale compte tenu du fait que le temps d'exécution est doublé par rapport à CAT et MGM.

Dans [4], la méthode CAT permet des gains importants sur la base KITTI de 2012 par rapport à SGM. Il en est de même, dans [2], pour MGM sur la base Middleburry alors que le modèle utilisé est très proche du notre. La principale différence entre notre modèle et celui ces articles est l'ajout de l'étape d'agrégation des coûts par CBCA. Cette opération lisse localement les coûts d'attache aux données à disparité constante. Ainsi, les coûts d'attache aux données de pixels voisins ont tendance à être corrélés ce qui limite les incohérences entre les lignes dans la solution SGM. Il apparaît donc que les méthodes CAT et MGM introduisent un lissage local qui est redondant avec une opération d'agrégation des coûts telle que CBCA. Une autre différence est que nous prenons en compte 4 directions dans MGM contre 8 dans [2].

4 Conclusion

Partant du constat que deux variantes de la méthode SGM, MGM et CAT ont tendance à privilégier l'information se trouvant sur la diagonale des deux directions prises en compte dans ces deux méthodes, nous avons proposé un paramètre additionnel pour ces deux méthodes permettant de contrôler la direction privilégiée. Les expérimentations ont été réalisées sur les bases d'évaluation Middlebury et KITTI. En faisant varier le paramètre additionnel de CAT et MGM, nous obtenons un continuum entre les solutions obtenues par ces deux méthodes et la solution de SGM. La modification de la direction privilégiée apporte une amélioration sur la base Middlebury par rapport aux méthodes SGM, MGM et CAT. Cependant, une absence d'amélioration voire une dégradation des résultats par rapport à SGM sont observées sur la base KITTI. Cela peut s'expliquer par la prépondérance des directions verticales et horizontales dans les scènes routières. L'importance de ces directions est de fait moindre dans toutes les variantes de SGM.

De plus, les solutions des méthodes MGM et CAT ont tendance à avoir une énergie plus basse que celle de SGM et donc de se rapprocher du minimum de l'énergie. L'a priori de régularisation utilisé favorise les zones à disparité constante. Il en résulte des erreurs de disparité par taches homogènes, qui sont plus difficiles à corriger en post-traitement que les artefacts de la méthode SGM, avec un filtrage médian de la carte de disparité par exemple. Cela constitue une limite de ces deux méthodes et donc des variantes proposées.



(a) SGM (b) $MGM_{0.9}$ (c) $MGM_{0.8}$

(d) $MGM_{0.7}$ (e) $MGM_{0.6}$ (f) MGM

FIGURE 6 – Résultats de reconstruction des méthodes SGM, $MGM_{0.9}$, $MGM_{0.8}$, $MGM_{0.7}$, $MGM_{0.6}$ et MGM. Les différents solutions forment un continuum entre la solution de la méthode SGM avec un tramage et MGM qui présente des erreurs plus localisées.

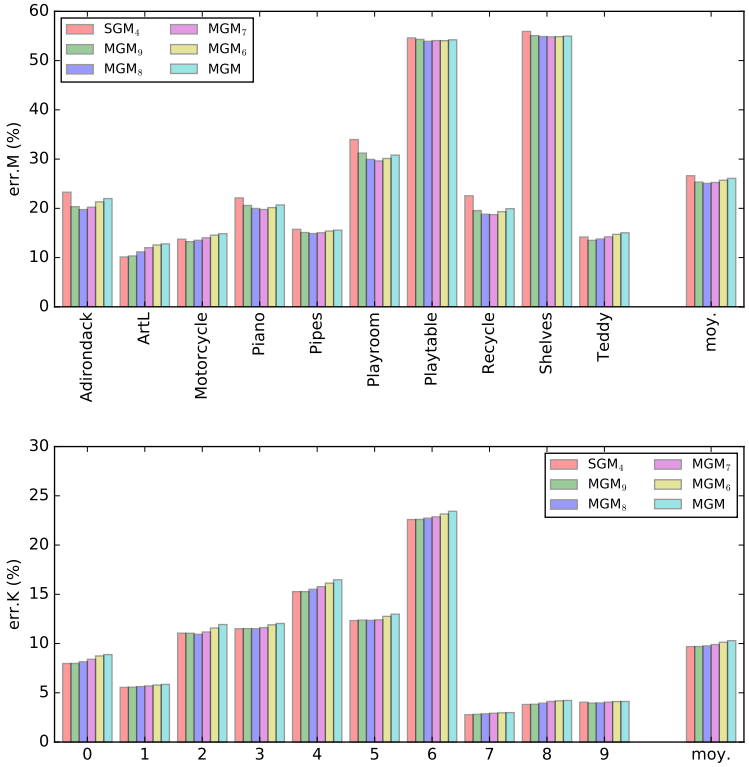


FIGURE 7 – Taux d'erreur des méthodes SGM, $MGM_{0.9}$, $MGM_{0.8}$, $MGM_{0.7}$, $MGM_{0.6}$ et MGM sur 10 couples de la base Middlebury (err.M) en haut, sur les 10 premiers couples de la base KITTI15 (err.K) en bas. Sur la base Middleburry, la méthode MGM donne de meilleurs résultats que SGM sauf pour les couples ArtL et Motorcycle. Le meilleur résultat moyen est obtenu pour $MGM_{0.8}$

Références

[1] DRORY, A., HAUBOLD, C., AVIDAN, S., AND HAMPRECHT, F. A. Semi-global matching : a principled derivation in terms of message passing. In *German Conference on Pattern Recognition* (2014), Springer, pp. 43–53.

[2] FACCIOLO, G., FRANCHIS, C. D., AND MEINHARDT, E. MGM : A Significantly More Global Matching for Stereovision. In *BMVC 2015* (Swansea, United Kingdom, 2015), B. Press, Ed.

[3] FELZENSZWALB, P. F., AND HUTTENLOCHER, D. P. Efficient belief propagation for early vision. *International Journal of Computer Vision 70*, 1 (Oct 2006), 41–54.

[4] HA, J. M., JEON, B., JEON, J., JO, S. Y., AND JEONG, H. Cost aggregation table : a theoretic derivation on the markov random field and its relation to message passing. In *Image Processing (ICIP), 2015 IEEE International Conference on* (2015), IEEE, pp. 2224–2228.

[5] HERMANN, S., AND KLETTE, R. Iterative semi-global matching for robust driver assistance systems. In *Asian Conference on Computer Vision* (2012), Springer, pp. 465–478.

[6] HERMANN, S., KLETTE, R., AND DESTEFANIS, E. Inclusion of a second-order prior into semi-global matching. In *Advances in Image and Video Technology* (Berlin, Heidelberg, 2009), T. Wada, F. Huang, and S. Lin, Eds., Springer Berlin Heidelberg, pp. 633–644.

[7] HIRSCHMÜLLER, H. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence 30*, 2 (2008), 328–341.

[8] HIRSCHMÜLLER, H., AND SCHARSTEIN, D. Evaluation of cost functions for stereo matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (2007), IEEE, pp. 1–8.

[9] MENZE, M., AND GEIGER, A. Object scene flow for autonomous vehicles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 3061–3070.

[10] PAGET, M. *Optimisation discrète et indices de stabilité appliqués à la stéréoscopie en contexte routier*. PhD thesis, Université Paris-Est, 2017.

[11] PAGET, M., TAREL, J.-P., AND MONASSE, P. Stereo ambiguity index for semi-global matching. In *Proceedings of IEEE International Conference on Image Processing (ICIP'17)* (Beijing, China, 2017), pp. 2513–2517. http ://perso.lcpc.fr/tarel.jean-philippe/publis/icip17.html.

[12] PARK, M.-G., AND YOON, K.-J. Leveraging stereo matching with learning-based confidence measures. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 101–109.

[13] SCHARSTEIN, D., HIRSCHMÜLLER, H., KITAJIMA, Y., KRATHWOHL, G., NEŠIĆ, N., WANG, X., AND WESTLING, P. High-resolution stereo datasets with subpixel-accurate ground truth. In *German Conference on Pattern Recognition* (2014), Springer, pp. 31–42.

[14] SCHARSTEIN, D., AND SZELISKI, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision 47*, 1 (Apr 2002), 7–42.

[15] SCHARSTEIN, D., AND SZELISKI, R. High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* (2003), vol. 1, IEEE, pp. I–I.

[16] SEKI, A., AND POLLEFEYS, M. Patch based confidence prediction for dense disparity map. In *British Machine Vision Conference (BMVC)* (2016), vol. 10.

[17] VEKSLER, O. *Efficient graph-based energy minimization methods in computer vision*. PhD thesis, Cornell University, 1999.

[18] ZABIH, R., AND WOODFILL, J. Non-parametric local transforms for computing visual correspondence. In *European conference on computer vision* (1994), Springer, pp. 151–158.

[19] ZHANG, K., LU, J., AND LAFRUIT, G. Cross-based local stereo matching using orthogonal integral images. *IEEE transactions on circuits and systems for video technology 19*, 7 (2009), 1073–1079.