Méthodes proximales pour l'estimation du champ de disparité à partir d'une paire d'images stéréoscopiques en présence de variations d'illumination

Mireille El Gheche, Caroline Chaux, Jean-Christophe Pesquet, Béatrice Pesquet-Popescu, Youmana Farah

▶ **To cite this version:**

Mireille El Gheche, Caroline Chaux, Jean-Christophe Pesquet, Béatrice Pesquet-Popescu, Youmana Farah. Méthodes proximales pour l'estimation du champ de disparité à partir d'une paire d'images stéréoscopiques en présence de variations d'illumination. GRETSI 2011, Sep 2011, Bordeaux, France. pp.ID454. hal-00733649

HAL Id: hal-00733649
https://hal.science/hal-00733649

Submitted on 19 Sep 2012

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.

Méthodes proximales pour l'estimation du champ de disparité à partir d'une paire d'images stéréoscopiques en présence de variations d'illumination

Mireille EL GHECHE[1], Caroline CHAUX[1], Jean-Christophe PESQUET[1], Béatrice PESQUET-POPESCU[2], Joumana FARAH[3]

[1]Université Paris-Est, LIGM and UMR-CNRS 8049, 77454 Marne-la-Vallée cedex, France

[2]Telecom ParisTech, Sign. and Im. Proc. Dept., 37-39, rue Dareau, 75014 Paris, France

[3]Department of Telecommunications, Faculty of Engineering, Holy-Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon

{el.gheche,caroline.chaux,jean-christophe.pesquet}@univ-paris-est.fr,
beatrice.pesquet@telecom-paristech.fr, joumanafarah@usek.edu.lb

Résumé – La présence de variations d'illumination dans les problèmes de mise en correspondance stéréoscopique introduit une difficulté supplémentaire importante. Dans cet article est proposée une nouvelle méthode proximale pour résoudre le problème d'estimation de la disparité en présence de changements d'illumination dans la scène observée. L'approche de mise en correspondance est formulée comme un problème d'optimisation convexe sous contrainte. Une fonction objectif convexe est alors minimisée sur l'intersection d'ensembles convexes. Ces ensembles sont associés à différentes contraintes modélisant des informations *a priori* sur les propriétés des champs à estimer. Pour résoudre ce problème d'estimation sous contraintes, l'algorithme parallèle proximal (PPXA+) a été adopté afin d'estimer conjointement les deux champs de variables qui représentent la disparité et les changements d'illumination. Des résultats espérimentaux obtenus sur des images synthétiques et réelles montrent l'efficacité et la robustesse de la méthode proposée.

Abstract – Illumination variations in stereo mapping make the disparity map estimation a very difficult problem. In this paper, we represent a new proximal approach to solve the problem of disparity estimation in the presence of illumination changes in the observed scene. The matching approach is formulated as a convex constrained optimization problem. A convex objective function is then minimized over the intersection of convex sets. These sets are associated with various constraints modeling prior knowledge and observed information of the fields to be estimated. To solve this estimation problem subject to constraints, a parallel proximal algorithm (PPXA+) is employed to jointly estimate the two variable fields representing the disparity and illumination changes. Experimental results on synthetic and real images demonstrate the robustness and efficiency of our method.

1 Introduction

La stéréovision vise à retrouver la profondeur d'une scène à partir d'images stéréoscopiques. Cette technique connaît un vif succès lié aux nombreuses applications de reconstruction 3D qui requièrent cette connaissance de profondeur. Grâce aux avancées technologiques, les recherches actuelles portent sur l'acquisition de modèles tridimensionnels de très haute qualité et de plus en plus précis. L'information de distance est primordiale pour, par exemple, localiser les objets qui représentent un obstacle potentiel pour un véhicule dans le domaine de la route intelligente. Des enjeux similaires existent dans le domaine du multimédia et notamment de la télévision 3D et de la compression stéréoscopique [1]. Les approches basées sur la vision stéréoscopiques permettent de retrouver l'information de profondeur par un calcul des coordonnées 3D d'un point à partir de ses deux images projetées (appelées vues gauche et droite). Résoudre le problème de mise en correspondance consiste donc à trouver la carte de disparité minimisant une fonction de coût global représentant une certaine distance. Plusieurs méthodes d'optimisation ont été proposées dans la littérature. Nous pouvons distinguer les méthodes à variables discrètes, qui effectuent un appariement de couples de pixels, comme la programmation dynamique [2] ou les coupures de graphes [3], des méthodes à variables continues, où la disparité estimée varie dans $\mathbb{R}$, comme dans les approches variationnelles [4]. Dans ce qui suit, nous nous plaçons dans ce dernier cadre en employant des méthodes d'optimisation parallèles reposant sur de récents développements sur les opérateurs proximaux.

La plupart des méthodes de mise en correspondance reposent sur l'hypothèse de conservation de la luminance. Cependant cette hypothèse physique n'est pas toujours vérifiée en pratique. Notre approche proximale aborde ce problème en modélisant l'illumination avec un facteur multiplicatif variant spatialement qui est une variable de la fonction objectif globale. L'estimation de profondeur se traduit ainsi par une estimation conjointe de la disparité et de la variation d'illumination. L'al-

gorithme considéré offre une grande flexibilité quant au choix du critère à minimiser et à l'incorporation de contraintes multiples [5]. Par ailleurs, il faut souligner que la méthode peut être facilement implantée sur une architecture parallèle (multi-coeurs, multi-processeurs, GPU, ...).

La suite de cet article est organisée comme suit. Dans la Section 2, nous présentons le problème de la correspondance stéréoscopique reformulé comme un problème d'optimisation convexe sous contraintes. Puis, dans la Section 3, une solution efficace itérative basée sur un algorithme parallèle proximal est proposée. Les résultats expérimentaux, montrant la précision des cartes de disparité obtenues sont donnés dans la Section 4. Enfin, la Section 5 conclut ce travail.

2 Formulation convexe du problème

On représente chaque point des images de gauche et de droite par son intensité lumineuse (I_g et I_d resp.). Pour décrire la relation existant entre les luminances des deux projections d'un même point de l'espace, on utilise la formule introduite par Gennert, $I_d(x-u(x,y),y) \simeq v(x,y)I_g(x,y)$ où u est le champ de disparité et v un champ de variation d'illumination éventuelle. La disparité u et la variation d'illumination v peuvent être calculées en minimisant une fonction d'erreur J globale d'appariement

$$J(u,v) = \sum_{s\in D} \phi(v(\mathbf{s})I_g(\mathbf{s}) - I_d(x-u(\mathbf{s}),y))$$

où ϕ est une fonction convexe, semi-continue inférieurement et propre, $\mathbf{s}=(x,y)$ et D est le support de l'image.
J n'étant pas convexe par rapport à u, on suppose alors que I_d est différentiable et on considère le développement de Taylor de $I_d(x-u(\mathbf{s}),y)$ autour d'une disparité initiale $\bar{u}$ au premier ordre :

$$I_d(x-u(\mathbf{s}),y) = I_d(x-\bar{u}(\mathbf{s}),y) - (u(\mathbf{s})-\bar{u}(\mathbf{s}))\nabla I_d^x(x-\bar{u}(\mathbf{s}),y)$$

où ∇I_d^x est le gradient horizontal de l'intensité de l'image droite. On obtient alors

$$J(u,v) = \sum_{s\in D} \phi(T_1(\mathbf{s})u(\mathbf{s}) + T_2(\mathbf{s})v(\mathbf{s}) - r(\mathbf{s})) \quad (1)$$

avec $T_1(\mathbf{s}) = \nabla I_d^x(x-\bar{u}(\mathbf{s}),y)$, $T_2(\mathbf{s}) = I_g(\mathbf{s})$ et $r(\mathbf{s}) = I_d(x-\bar{u}(\mathbf{s}),y) + \bar{u}(\mathbf{s})T_1(\mathbf{s})$.
Soient $\mathbf{w}(\mathbf{s}) = [u(\mathbf{s}),v(\mathbf{s})]^\top$ et $\mathbf{T}(\mathbf{s}) = [T_1(\mathbf{s}),T_2(\mathbf{s})]$, le problème revient alors à minimiser :

$$J(\mathbf{w}) = \sum_{s\in D} \phi\big(\mathbf{T}(\mathbf{s})\mathbf{w}(\mathbf{s}) - r(\mathbf{s})\big). \quad (2)$$

L'optimisation du critère permet d'obtenir des estimés $\widehat{\mathbf{w}}$ du vecteur de paramètres $\mathbf{w}$ sachant que $\mathbf{T}$ et r sont connues.

Le problème d'estimation convexe résultant est un problème inverse mal posé de par les composantes de $T_1(\mathbf{s})$ qui peuvent s'annuler mais aussi par le fait que l'on a deux variables à déterminer en chaque point. Ainsi, pour résoudre le problème d'une manière fiable, il est nécessaire d'intégrer des contraintes supplémentaires. Ceci peut être réalisé en formulant le problème dans un cadre ensembliste [6].

2.1 Formulation ensembliste

Le principal intérêt des approche ensemblistes [6, 7] est leur simplicité dans l'introduction d'informations *a priori* de natures très variées sur la solution désirée. On définit alors individuellement des ensembles $(S_i)_{1\leq i\leq m}$ pour chaque contrainte disponible, et leur intersection S constitue l'ensemble convexe des solutions admissibles sur lequel on doit projeter [8]. En s'inspirant de ce point de vue, notre problème s'écrit alors :

$$\text{trouver } \hat{\mathbf{w}} \in S = \bigcap_{i=1}^{m} S_i, \text{ tel que } J(\hat{\mathbf{w}}) = \inf\ J(S). \quad (3)$$

Pour limiter l'influence des zones d'occlusions où un point n'est visible que dans l'une des deux images, on les exclut du critère minimisé. Notons O l'ensemble des points de l'image occultés, la sommation est alors limitée aux pixels visibles appartenant à $D \setminus O$. Une comparaison des méthodes pour la détection des zones d'occlusion est menée dans [9].

2.2 Contraintes convexes

Comme indiqué précédemment, l'objectif est de formuler des contraintes à partir des observations et des informations *a priori* que l'on peut se donner sur les champs à estimer. Le premier *a priori* qu'on peut imposer porte sur les plages de valeurs prises par u et v :

$$S_1 = \big\{(u,v)\in\mathbb{R}^N\times\mathbb{R}^N |\ u_{\min}\leq u\leq u_{\max} \text{ et } v_{\min}\leq v\leq v_{\max}\big\}.$$

Le second *a priori* est basé sur une estimation en accord avec la régularité des champs de disparité et d'illumination :

$$S_2 = \big\{(u,v)\in\mathbb{R}^N\times\mathbb{R}^N \mid \mathrm{TV}(u)\leq\tau_u \text{ et } \|\nabla v\|_{\ell^2}\leq\tau_v\big\}.$$

On remarque que le choix d'une borne sur la variation totale discrétisée permet, pour la disparité, de respecter les discontinuités aux frontières des objets.

3 Algorithme PPXA+ pour l'estimation de la disparité

La distance de fidélité aux données considérée dans ce travail peut être choisie de manière très générale et peut être non différentiable (une norme ℓ_1 par exemple). Le problème d'estimation résultant est résolu en utilisant un algorithme parallèle proximal (e.g. PPXA+) [10].
On peut remarquer que $S_1 = C_1^u \times C_1^v$ et que S_2 peut être exprimé comme le produit de l'image réciproque de C_2^u par L_2^u et de l'image réciproque de C_2^v par L_2^v où C_1^u, C_1^v (resp. C_2^u, C_2^v) sont des ensembles convexes fermés de $\mathbb{R}^N$ (resp. $\mathbb{R}^{2N}$) et L_2^u, L_2^v sont des matrices de $\mathbb{R}^{2N\times N}$. Le problème qu'on se propose de résoudre peut ainsi se formuler de la manière suivante :

$$\begin{array}{l}\text{minimiser} \\ u\in C_1^u, v\in C_1^v \\ L_2^u u\in C_2^u, L_2^v v\in C_2^v\end{array} \quad J(\mathbf{w}). \quad (4)$$

Cette minimisation peut être réalisée, de manière efficace, en employant l'algorithme PPXA+ [11] décrit par l'Algo. 1.

$(\omega_1, \omega_2) \in]0, +\infty[^2, \gamma > 0$
$(z^u_{i,0})_{1\leq i\leq 3} \in \mathbb{R}^N \times \mathbb{R}^{2N} \times \mathbb{R}^N$
et $(z^v_{i,0})_{1\leq i\leq 3} \in \mathbb{R}^N \times \mathbb{R}^{2N} \times \mathbb{R}^N$
$Q^u = \big(\omega_2 (L^u_2)^\top L^u_2 + (\gamma + \omega_1)\mathbb{I}\big)^{-1}$
et $Q^v = \big(\omega_2 (L^v_2)^\top L^v_2 + (\gamma + \omega_1)\mathbb{I}\big)^{-1}$
$u_0 = Q^u\big(\omega_1 z^u_{1,0} + \omega_2 (L^u_2)^\top z^u_{2,0} + \gamma z^u_{3,0}\big)$
et $v_0 = Q^v\big(\omega_1 z^v_{1,0} + \omega_2 (L^v_2)^\top z^v_{2,0} + \gamma z^v_{3,0}\big)$
Pour $n = 0, 1, \ldots$ **faire**
  Pour $i = 1, 2$ **faire**
    $p^u_{i,n} = P_{C^u_i}(z^u_{i,n})$ et $p^v_{i,n} = P_{C^v_i}(z^v_{i,n})$
  Fin Pour
  $(p^u_{3,n}, p^v_{3,n}) = \mathrm{prox}_{\frac{J}{\gamma}}(z^u_{3,n}, z^v_{3,n})$
  $c^u_n = Q^u\big(\omega_1 p^u_{1,n} + \omega_2 (L^u_2)^\top p^u_{2,n} + \gamma p^u_{3,n}\big)$
  et $c^v_n = Q^v\big(\omega_1 p^v_{1,n} + \omega_2 (L^v_2)^\top p^v_{2,n} + \gamma p^v_{3,n}\big)$
  $z^u_{1,n+1} = z^u_{1,n} + \lambda_n(2c^u_n - u_n - p^u_{1,n})$
  et $z^v_{1,n+1} = z^v_{1,n} + \lambda_n(2c^v_n - v_n - p^v_{2,n})$
  $z^u_{2,n+1} = z^u_{2,n} + \lambda_n(L^u_2(2c^u_n - u_n) - p^u_{2,n})$
  et $z^v_{2,n+1} = z^v_{2,n} + \lambda_n(L^v_2(2c^v_n - v_n) - p^v_{2,n})$
  $z^u_{3,n+1} = z_{3,n} + \lambda_n(2c^u_n - u_n - p^u_{3,n})$
  et $z^v_{3,n+1} = z^v_{3,n} + \lambda_n(2c^v_n - v_n - p^v_{3,n})$
  $u_{n+1} = u_n + \lambda_n(c^u_n - u_n)$
  et $v_{n+1} = v_n + \lambda_n(c^v_n - v_n)$
Fin Pour

Algorithme 1: PPXA+ pour l'estimation conjointe de (u, v).

Cet algorithme consiste à calculer successivement : 1) des projections directes (la projection liée à S_1 est triviale et celle liée à S_2 est réalisée en utilisant l'algorithme proposé dans [12]), 2) des opérateurs proximaux pour lesquels nous avons des expressions explicites [10].

4 Résultats

Nous considérons ici des résultats expérimentaux [1] permettant d'évaluer la robustesse de notre approche vis-à-vis des changements d'illumination. Dans un premier temps, nous considérons un exemple où un changement de luminosité artificiel a été simulé. Dans un second temps, des paires d'images ayant subi de réels changements d'illumination sont traitées.
Le premier exemple que nous avons utilisé pour valider quantitativement notre méthode est l'image synthétique "Dolls" où un changement de luminosité est simulé en multipliant l'image de droite par un profil gaussien, tandis que l'image gauche reste inchangée. Le profil de changement d'illumination gaussien est défini par :

$$g(x, y) = c_1 \exp\Big(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\xi^2}\Big) + c_2 \qquad (5)$$

où (x_0, y_0) est le centre, $\xi = 256$ est l'écart type de l'effet d'illumination et c_1 et c_2 sont deux constantes fixées à 1.8 et -0.6 respectivement.
Pour ces résultats, les paramètres des contraintes ont été fixés à : $[u_{\min}, u_{\max}] = [2\ 74]$, $[v_{\min}, v_{\max}] = [0.8\ 1.2]$, $\tau_u = 92469$ et $\tau_v = 0.293$.

Le deuxième test a été réalisé sur la paire d'images "Book". Dans ce cas, les paramètres des contraintes ont été fixés à : $[u_{\min}, u_{\max}] = [2\ 75]$, $[v_{\min}, v_{\max}] = [0.1\ 1.1]$, $\tau_u = 73644$ et $\tau_v = 399$.

Ces paramètres ont été évalués sur les champs de disparité et d'illumination réelle (vérité terrain), mais lorsque celle ci n'est pas disponible, on peut les fixer par des méthodes *ad-hoc*, par exemple, sur les champs de disparité et d'illumination initiaux.

Par ailleurs, la fonction ϕ a été choisie égale à une norme ℓ_1 et les paramètres de l'algorithme ont été fixés à $\omega_1 = 10$, $\omega_2 = 100$, $\gamma = 200$ et $\lambda_n \equiv 1.5$.
Notons que l'initialisation de notre algorithme est obtenue par la méthode locale de mise en correspondance par corrélation "NCC" (*Normalized Cross-Correlation*). Par ailleurs, afin d'affiner les résultats, l'algorithme PPXA+ est relancé plusieurs fois (3 dans notre cas) et à chaque fois, l'estimée initiale $\bar{u}$ est réinitialisée par la valeur du champ estimé à l'itération précédente. Cette stratégie améliore considérablement la qualité de la solution recherchée et diminue sa dépendance vis-à-vis de l'initialisation.

Les résultats visuels, les EAM (Erreur Absolue Moyenne) ainsi que les pourcentages de pixels erronés Err obtenus sont donnés dans la Fig. 1 où nous nous comparons à la méthode de référence proposée par W. Miled *et al.* [4] qui a démontré d'excellentes performances par rapport aux méthodes de l'état de l'art. On peut remarquer que la méthode proposée ici permet d'obtenir une estimation améliorée des champs de disparité et d'illumination.

5 Conclusion

Nous avons proposé un modèle permettant de compenser localement les changements d'illumination pour des paires d'images stéréoscopiques. Le problème de mise en correspondance est alors approché par un problème d'optimisation convexe où une fonction objectif convexe est minimisée sur l'intersection d'ensembles convexes construits à partir des connaissances *a priori* sur les champs de disparité et d'illumination. Il en résulte une grande flexibilité dans le choix du critère à minimiser, ce dernier n'étant pas restreint à une mesure d'erreur strictement convexe et quadratique.

Références

[1] M. Kaaniche, W. Miled, B. Pesquet-Popescu, A. Benazza-Benyahia, and J.-C. Pesquet. Dense disparity map representations for stereo image coding. *IEEE Interna-*

1. http ://cat.middlebury.edu/stereo/scenes2005/

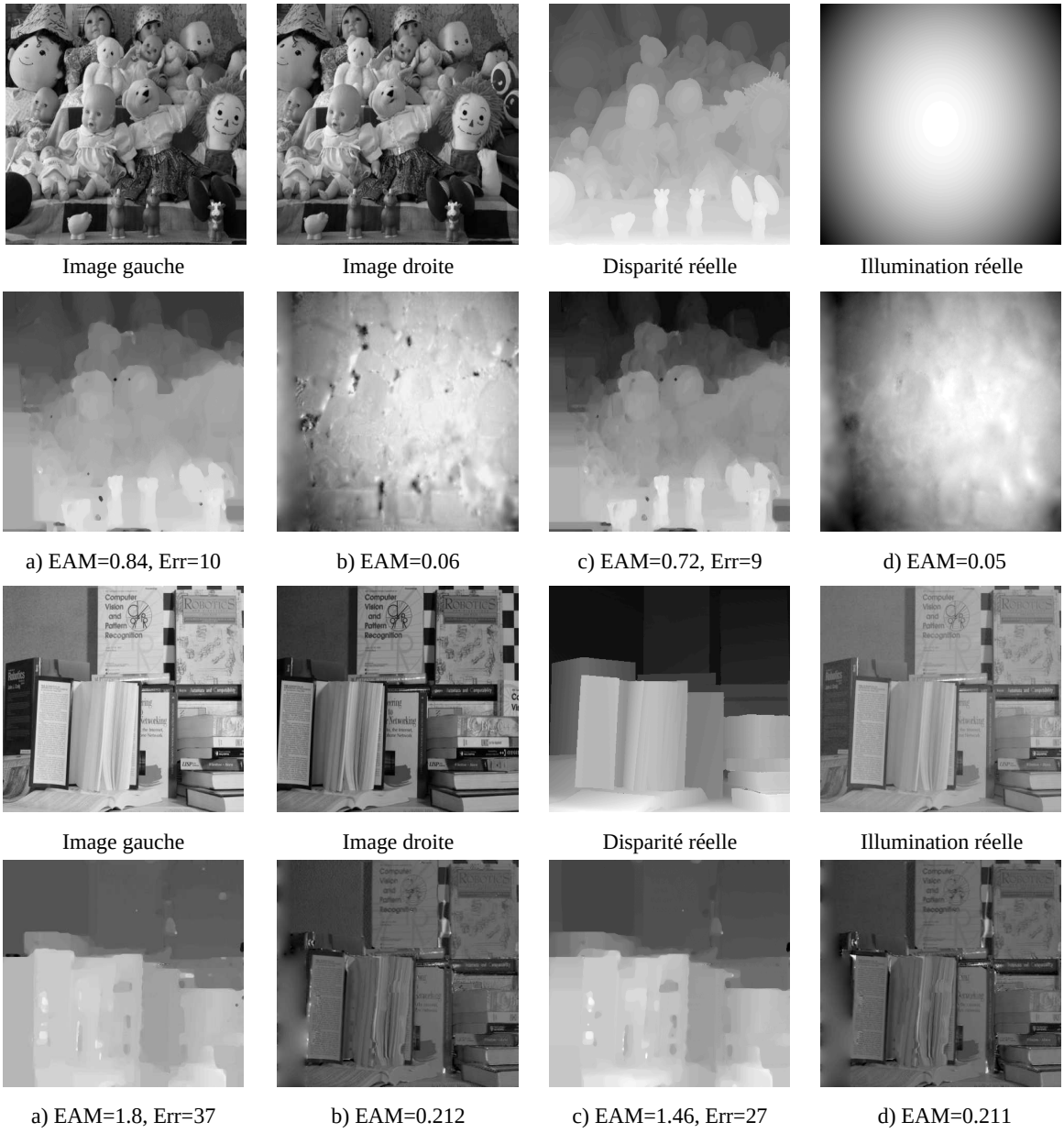

FIGURE 1 – Résultats sur les images “Dolls” et “Book”. Lignes 1 et 3 : les données originales. Lignes 2 et 4 : (a,b) restauration des champs de disparité et d’illumination par la méthode de projection sous différentielle [4] et (c,d) par la méthode proposée.

tional Conference on Image Processing, Nov. 7–10, 2009.

[2] O. Veksler. Stereo correspondence by dynamic programming on a tree. In *IEEE Conference Proceedings of Computer Vision and Pattern Recognition*, volume 2, pages 384–390, San Diego, United States, 2005.

[3] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *IEEE International Conference on Computer Vision*, volume 2, pages 508–515, Vancouver, BC, Canada, July 2001.

[4] W. Miled, J.-C. Pesquet, and M. Parent. A convex optimisation approach for depth estimation under illumination variation. *IEEE Transactions on Image Processing*, 18(4) :813–830, April 2009.

[5] M. El Gheche, J.-C. Pesquet, J. Farah, M. Kaaniche, and B. Pesquet-Popescu. Proximal splitting methods for depth estimation. Prague, Czech Republic, May 2011.

[6] P. L. Combettes. The foundations of set theoretic estimation. *Proceedings of the IEEE*, 81(2) :182–208, Feb 1993.

[7] P. L. Combettes. The convex feasibility problem in image recovery. *Advances in Imaging and Electron Physics*, 95 :155–270, New York : Academic Press, 1996.

[8] D. C. Youla and H. Webb. Image restoration by the method of convex projections. *IEEE Trans. Medical Imaging*, MI-1 :81–94, Oct 1982.

[9] G. Egnal and R.P. Wildes. Detecting binocular half-occlusions : Empirical comparisons of five approaches. *IEEE Trans. Pattern Anal. Mach. Intell*, 24(8) :1127–1133, Aug 2002.

[10] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer-Verlag, New York, 2010.

[11] J.-C. Pesquet. A parallel inertial proximal optimization method. http ://www.optimization-online.org/DB_FILE/2010/11/2825.pdf, 2010.

[12] E. van den Berg and M. P. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM J. on Scientific Computing*, 31(2) :890–912, Nov. 2008.