

Approche Variationnelle pour la fusion d'images multifocus bruitées.

COSMIN LUDUSAN^{1,2}, OLIVIER LAVIALLE¹

¹Université de Bordeaux, UMR CNRS-UB1 IMS 5218, ENITA de Bordeaux, Talence, France.

²Technical University of Cluj-Napoca, The Faculty of Electronics, Telecommunications and Information Technology, Cluj-Napoca, Roumanie

Cosmin.ludusan@ims-bordeaux.fr, olivier.lavialle@ims-bordeaux.fr

Résumé – Nous proposons une approche variationnelle permettant la fusion d'images multifocus corrompues par du bruit. Notre objectif est donc de mener en parallèle une opération rehaussement des contours et de débruitage anisotropique. Nous nous appuyons pour cela sur la théorie de l'estimation robuste et sur un formalisme fondé sur les équations aux dérivées partielles. Notre méthode est guidée par la géométrie locale de l'image en utilisant un schéma d'interpolation subpixelique qui couplé à un processus de fusion fondé sur une équation aux dérivées partielles permet de répondre au double problème du rehaussement et du débruitage. Nous démontrons dans cet article l'efficacité de notre approche pour une problématique de fusion d'images bruitées.

Abstract – we propose a variational approach for image fusion with concurrent denoising of multifocus images using robust error estimation theory and Partial Differential Equations (PDEs). The novelty of our approach consists in the development of an image geometry-driven, anisotropic fusion model, described using an efficient pixel interpolation scheme and coupled with a PDE-based, weighted fusion process, capable of concurrent fusion and denoising. The experimental analysis shows that robust denoising can be attained in parallel with efficient image fusion, thus bringing two paramount image processing tasks into complete synergy. We demonstrate the efficiency of our approach in case of unsupervised image fusion of multifocus noise-corrupted images.

1. Introduction

La fusion d'image est un processus qui consiste à combiner différentes sources dans le but d'améliorer la qualité de l'image résultante. Les problématiques de fusion d'image sont très diverses selon que l'on se positionne au niveau du pixel (bas niveau), de l'objet (niveau intermédiaire) ou de la décision (haut niveau). L'approche que nous proposons se situe entre une approche bas-niveau à l'échelle du pixel et le niveau intermédiaire à travers l'utilisation de la carte des contours dans le cadre du processus de rehaussement.

Dans la classification des méthodes de fusion proposée par Blum et al.[1], notre approche appartient au groupe des méthodes non fondées sur la décomposition multi-échelles. En cela, les méthodes concurrentes présentées récemment dans la littérature sont celles de Socolinsky [8], Socolinsky et Wolff [9], John et Vorontsov[3], Wang et Ye[10] et Piella[5]. Ces méthodes sont toutes fondées sur l'hypothèse que les images initiales ne sont pas corrompues par du bruit, ce qui peut apparaître très limitant pour un grand nombre d'applications en fusion d'image.

Dans la pratique, les images à fusionner sont bien souvent perturbées par du bruit que ce soit durant la

phase d'acquisition, de transmission ou de stockage. Quelques auteurs ont proposé des approches fondées sur des équations aux dérivées partielles permettant la prise en compte du bruit présent dans les images. (Pop et al [6]).

Dans la section suivante nous allons décrire le concept de théorie de l'estimation robuste de l'erreur dans le cadre de la fusion d'image par approche variationnelle en liant quelques concepts théoriques aux modèles de fusion proposés dans la littérature. Le modèle que nous proposons sera décrit en section 3. Enfin, quelques résultats sont présentés en section 4.

2. Théorie de l'estimation robuste de l'erreur et approche variationnelle pour la fusion.

Soit une image dégradée I avec $I(p)$ la valeur de I au pixel p et $I_r(p)$ la version restaurée de $I(p)$. $I_r(p)$ peut être estimée à travers une fonctionnelle $E(I_r)$:

$$E(I_r) = \int_{\Omega} \rho(p, I_r(p), |\nabla I_r(p)|) dp \quad (1)$$

Ω est le domaine de l'image et $|\nabla I_r(p)|$ est le gradient de l'image. La norme de l'erreur $\rho(\cdot)$ est définie en fonction du type d'application et du type de dégradation. Par exemple, dans le cas d'un bruit additif blanc gaussien, un choix pertinent est de considérer pour $\rho(\cdot)$ le critère des moindres carrés.

Pour estimer un extrémum de (1), nous pouvons utiliser l'équation d'Euler-Lagrange :

$$\frac{\partial \rho}{\partial I_r} - \nabla \left(\frac{1}{|\nabla I_r|} \frac{\partial \rho}{\partial |\nabla I_r|} \nabla I_r(p) \right) = 0 \quad (2)$$

La solution de (2) n'est pas immédiate et réclame l'utilisation d'une méthode de résolution numérique comme une optimisation par descente de gradient. Ainsi I_r peut être estimé itérativement en utilisant :

$$I_r(p, t + 1) = I_r(p, t) - \tau \frac{\partial I_r(p, t)}{\partial t} \quad (3)$$

avec

$$\frac{\partial I_r(p, t)}{\partial t} = -\frac{\partial \rho}{\partial I_r} + \nabla \left(\frac{1}{|\nabla I_r|} \frac{\partial \rho}{\partial |\nabla I_r|} \nabla I_r(p, t) \right) \quad (4)$$

et les conditions initiales $I_r(p, 0) = I(p)$.

L'évolution décrite par (4) est stoppée lorsqu'un critère défini a priori est atteint. En pratique, un nombre fini d'itérations est nécessaire pour obtenir des résultats satisfaisants [3].

3. Fusion et débruitage d'image anisotrope

Nous avons vu dans la section précédente qu'un choix possible pour la norme de l'erreur $\rho(\cdot)$ dans (1) est le critère des moindres carrés :

$$\rho(p, I_r(p), |\nabla I_r(p)|) = \frac{1}{2} |\nabla I_r(p)|^2 \quad (5)$$

(5) permet de filtrer le bruit de manière isotrope dans une image. Un choix plus pertinent doit être obtenu pour satisfaire à la fois à l'objectif de débruitage et de rehaussement de contours dans la mesure où l'objectif de la fusion est de transférer toute l'information pertinente au sein de l'image résultat, i.e. l'information définie par les contours des objets. Un exemple de norme d'erreur ayant de bonne propriété a été défini par John et Vorontsov [3] :

$$\rho(p, I_r(p), |\nabla I_r(p)|) = \frac{\alpha}{2} |\nabla I_r(p)|^2 + \frac{\beta}{2} J_1(p) (I_r(p, t) - I(p))^2 \quad (6)$$

où α et β permettent de contrôler le poids relatif du terme de filtrage et de rehaussement. J_1 est une carte des contours lissée par une gaussienne :

$$J_1(p) = \int |\nabla I(q)|^2 G(p - q, \sigma) dq \quad (7)$$

où $G(\cdot)$ est un noyau gaussien d'écart type σ .

Dans la pratique, le fait d'utiliser un terme de filtrage isotrope n'est pas pertinent dans la mesure où ce processus tend à diminuer l'information pertinente. Par suite, les deux termes de (6) ont un effet contradictoire et le résultat n'est jamais complètement satisfaisant pour les deux objectifs recherchés. C'est pourquoi nous proposons d'introduire une norme de l'erreur de nature anisotrope qui apporte une réponse plus complète aux objectifs de la fusion d'image à la fois concernant le terme de filtrage et le terme de rehaussement :

$$\rho(p, I_r(p), |\nabla I_r(p)|) = -\alpha \int \mathcal{D}_a(I_r) dI_r + \frac{\beta}{2} J(p, t) (I_r(p, t) - I(p))^2 \quad (8)$$

$J(p, t) = J_1(p) - J_r(p, t)$ est une fonction anisotrope inspirée des travaux de John et Vorontsov [3] :

$$J(p, t) = \int |\nabla^* I(q)|^2 G(p - q, \sigma) dq - \int |\nabla^* I_r(q, t)|^2 G(p - q, \sigma) dq \quad (9)$$

où ∇^* est le gradient calculé le long de la direction principale du contraste obtenue à travers l'utilisation d'une analyse en composantes principales (direction orthogonale au structure). En utilisant la norme de l'erreur définie en (9) dans l'équation (4), nous obtenons l'expression de l'équation d'évolution :

$$\frac{\partial I_r(p, t)}{\partial t} = \alpha \mathcal{D}_a(I_r) - \beta \Theta(J(p, t)) J(p, t) (I_r(p, t) - I(p)) \quad (10)$$

$\mathcal{D}_a(\cdot)$ est une fonction de diffusion et Θ autorise l'injection de l'information pertinente (i.e. correspondant à des gradients élevés)

$$\Theta(J) = \begin{cases} 1 & \text{si } J \geq 0 \\ 0 & \text{si } J < 0 \end{cases} \quad (11)$$

Finalement, le modèle de fusion que nous proposons dans cet article est établi en étendant la notion d'image reconstruite I_r à la notion d'image fusionnée I_f . Chaque image d'entrée $I_n(p)$ avec $n=1 \dots N$ est vue comme une image dégradée I reconstruite en utilisant (10). L'image I_f est alors le résultat d'une combinaison linéaire des N entrées au sein de laquelle l'information utile est transférée par l'intermédiaire de poids $w_n(p, t)$: $I_f(p, t) = \sum_{n=1}^N w_n(p, t) \cdot I_n(p)$.

L'ensemble des poids peut être estimé conjointement conformément à (3) et (10) [4]:

$$w_n(p, t + 1) = w_n(p, t) - \tau \frac{\partial I_r(p, t)}{\partial t} \quad (12)$$

avec

$$\frac{\partial w_n(p, t)}{\partial t} = \frac{1}{I_n(p)} [\alpha \Delta_a(I_n) - \beta \theta(J(p, t)) J(p, t) (I_n(p, t) - I(p))] \quad (13)$$

Les poids sont initiés à $1/N$ en début de processus, conduisant à une simple moyenne des images. Les poids évoluent ensuite pour permettre la mise en évidence de l'information pertinente.

4. Résultats

Nous montrons dans un premier temps, à la figure 1 (images de la base de Lehigh University [2]), les résultats de notre approche concernant la fusion de deux images que nous avons perturbées par un bruit important d'écart-type 20 (sur une dynamique classique de 256 niveaux de gris). La première image présente une zone floue à l'arrière plan, la seconde étant, à l'inverse, nette sur l'arrière plan et floue au premier plan. Nous remarquons la qualité visuelle de l'image fusionnée avec un traitement simultané du flou et du bruit.

Dans un deuxième temps, nous nous proposons au tableau 1 d'analyser l'évolution de l'index de similarité structurale (SSIM) de l'image fusionnée suivant le niveau de bruit des images d'entrée. Le SSIM mesure la pertinence du rehaussement et la préservation des structures dans l'image résultante par comparaison à l'image de départ non floutée et non bruitée. La décroissance du SSIM est lente avec des valeurs qui restent relativement fortes pour des niveaux de bruit extrêmement élevés ($\sigma=50$). Ces résultats illustrent la grande robustesse au bruit de l'approche proposée avec un maintien de la qualité de rehaussement de

l'information pertinente en présence d'un faible rapport signal sur bruit.

Nous proposons enfin de comparer les résultats de notre approche avec des approches concurrentes de fusion permettant de traiter le cas d'images bruitées. La figure 2 présente les images fusionnées obtenues à travers l'utilisation de 3 méthodes comparées à l'approche proposée dans cet article : la simple moyenne entre les deux images de départ est ici donnée à titre de référence sachant qu'elle n'est évidemment pas pertinente pour traiter le cas bruité. La seconde méthode est fondée sur la Shift Invariant Discrete Wavelet Transform (SIDWT) proposé par Rockinger dans le cadre d'une problématique de fusion [7]. Enfin, la troisième méthode développée en 2007 par Pop et al [6] relève également d'une démarche fondée sur des équations aux dérivées partielles en proposant l'évolution conjointe des images à fusionner dans le but de les faire converger.

Trois critères sont utilisés pour évaluer les performances des méthodes. Ces critères faisant appel à une image de référence, le résultat d'une approche classique de fusion par pyramide laplacienne effectuée sur les images non bruitées a été choisi pour garantir une évaluation sans biais des méthodes testées (tableau 2). Les trois critères sont l'erreur quadratique entre l'image fusionnée et l'image référence (RMSE), le PSNR et à nouveau le SSIM qui mesure spécifiquement la qualité du rehaussement. Pour ces trois critères, l'ordre des performances est identique entre les méthodes. Si les trois méthodes dédiées (SIDWT, Pop, approche proposée dans cet article) surpassent largement les performances de la simple moyenne, nous pouvons remarquer la plus grande efficacité de notre approche, ce qui est confirmé par un jugement visuel des résultats.

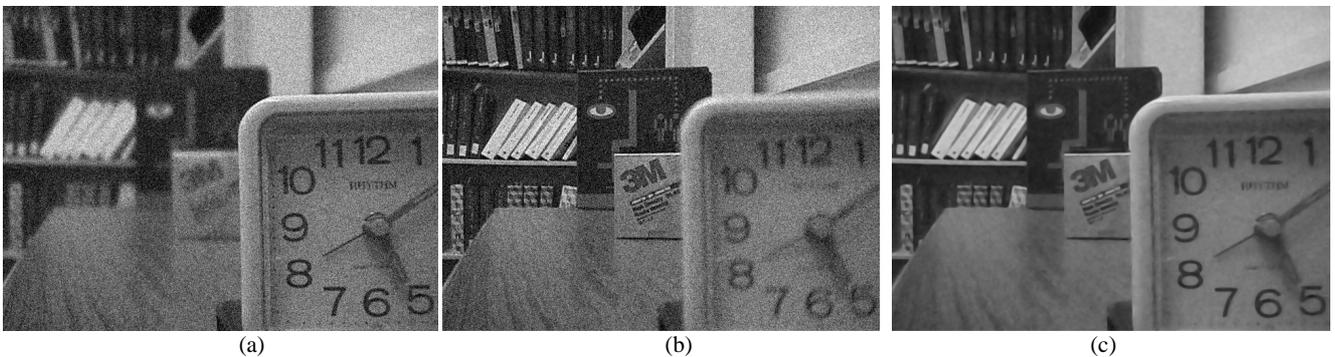


Figure 1 : (a) et (b) images d'entrée à fusionner (bruit d'écart-type 20) (c) résultat de fusion.

Tableau 1 : Evolution du niveau de SSIM en fonction du bruit

σ bruit	0	1	2	4	10	20	30	50
SSIM	1	0,993	0,991	0,987	0,96	0,924	0,903	0,857

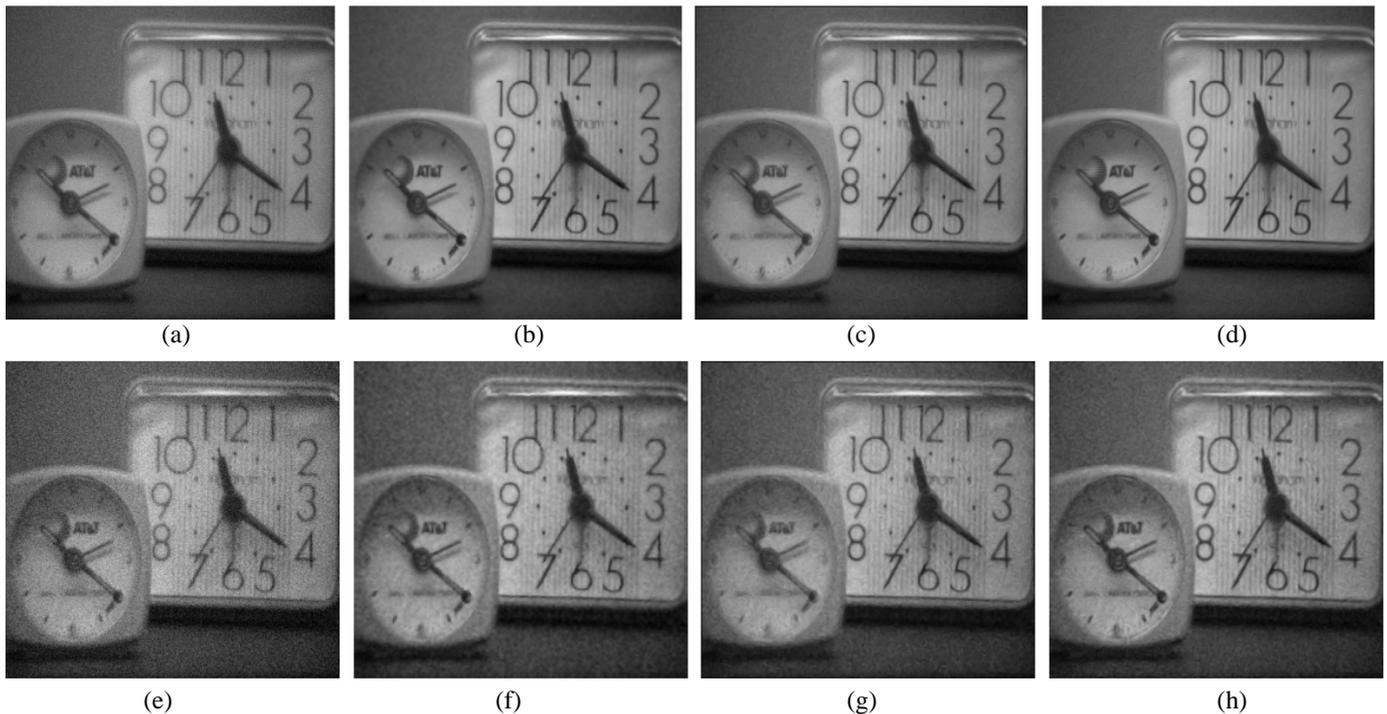


Figure 2 : Résultats obtenus par les différentes approches de fusion. $\sigma=13$ (en haut) et $\sigma=30$ (en bas) ; (a)et (e) Image moyenne, (b) et (f) SIWDT, (c) et (g) Pop [6], (d) et (h) Méthode proposée dans cet article.

Tableau 2 : Comparaison des performances de fusion sur images bruitées pour 2 niveaux de bruits à partir de trois indicateurs

niveau de bruit	méthode de fusion	RMSE	PSNR	SSIM
$\sigma=13$	moyenne	10,9	21,4	0,25
	SIDWT-Hard	5,8	32,9	0,91
	Pop	9,6	30,2	0,88
	Approche proposée	5,2	33,8	0,89
$\sigma=30$	moyenne	21,7	21,4	0,25
	SIDWT-Hard	8,8	29,2	0,77
	Pop	10	28,2	0,79
	Approche proposée	7,4	30,7	0,82

REFERENCES

[1] Blum, R.S., Xue, Z., Zhang, Z., 2006. Multi-Sensor Image Fusion and Its Applications. CRC Press Taylor and Francis Group, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742. chapter An Overview of Image Fusion. First edition. pp. 1–36.

[2] ECE, Lehigh University, 2010. Investigations of image fusion. <http://www.ece.lehigh.edu/SPCRL/IF/disk.htm>.

[3] John, S., Vorontsov, M., 2005. Multiframe selective information fusion from robust error estimation theory. IEEE Trans. Image Process. 14, 577–584.

[4] Mitianoudis, N., Stathaki, T., 2008. Image Fusion: Algorithms and Applications. Academic Press London. Chapitre “Enhancement of multiple sensor

images using joint image fusion and blind restoration.” First edition. pp. 299–326.

[5] Piella, G., 2009. Image fusion for enhanced visualization: A variational approach. Int. J. Comput. Vision 83, 1–11.

[6] Pop, S., Lavielle, O., Terebes, R., Borda, M., 2007. Low-level fusion: a PDE-based approach, 10th International Conference on Information Fusion, pp. 1–8.

[7] Rockinger O., 1999. Image fusion toolbox <http://www.metapix.detoobox.htm>

[8] Socolinsky, D.A., 2000. A Variational Approach To Image Fusion. Ph.D. thesis. The Johns Hopkins University.

[9] Socolinsky, D., Wolff, L., 2002. Multispectral image visualization through first order fusion. IEEE Trans. on Image Process. 11, 923–931.

[10] Wang, C., Ye, Z., 2006. First-order fusion of volumetric medical imagery. Vision, Image and Signal Processing, IEE Proceedings 153, 191–198.