

Filtrage adaptatif multicomposante pour l'amélioration des séquences video couleur

M. El Hassouni

H. Cherifi

LIRSIA, Université de Bourgogne

Faculté des Sciences Mirande

BP 47870, 21078 Dijon cedex FRANCE

{Mohamed.Elhassouni,Hocine.Cherifi}@u-bourgogne.fr

Résumé

Nous présentons dans cet article une méthode robuste de débruitage de séquences vidéo couleur en présence de perturbations de type bruit non gaussien. Pour traiter le problème nous utilisons une approche de filtrage compensée en mouvement basée sur les statistiques d'ordre supérieur (SOS). La compensation du mouvement est assurée par un estimateur robuste. Ensuite, nous appliquons un filtre adaptatif multicomposante sur les images reconstruites dans deux espaces couleur différents (RGB et CIE-LAB). Nous avons validé notre approche sur des séquences vidéo couleur.

Mots clefs

Compensation du mouvement, filtrage multicomposante, espaces couleur, SOS.

1 Introduction

Lors de l'acquisition, la transmission et le stockage, les séquences d'images sont souvent altérées par différents types de bruit. Pour éliminer ce bruit il faut disposer d'un algorithme d'estimation de mouvement robuste ainsi que d'une technique adéquate de filtrage spatio-temporel.

La phase initiale d'estimation du mouvement permet la compensation du mouvement apparent dans la séquence. Pour cela, nous proposons d'utiliser une méthode récursive d'estimation du mouvement global basée sur les statistiques d'ordre supérieur (SOS) [2]. Cette technique se base sur un modèle de mouvement affine simplifié. Cet algorithme a été testé sur des séquences réelles. Les résultats expérimentaux ont mis en évidence sa supériorité en terme de robustesse aux dégradations par rapport aux techniques basées sur les statistiques du second ordre (SS0) dans des situations de bruit impulsionnel et de bruit mixte (gaussien et impulsionnel).

Dans un second temps, une famille de filtres spatio-temporels est appliquée à la séquence d'images compensées. L'approche spatio-temporelle présente l'avantage d'être moins sensible aux non-stationnarités suivant les deux directions. Les techniques de filtrage proposées

permettent d'améliorer la qualité visuelle et de faciliter l'exécution des tâches ultérieures telles que le codage, l'analyse ou l'interprétation.

Plusieurs filtres ont été proposés pour éliminer le bruit dans des séquences d'images à niveau de gris [7]. On peut citer le filtre de Wiener 3-D qui présente des caractéristiques d'optimalité dans le cas de perturbations gaussiennes. Les filtres adaptatifs à moyennes pondérées (3-D AWA, Adaptive Weighted Average) ont été utilisés pour des séquences d'images contaminées par un bruit blanc gaussien additif. Dans le cas du bruit impulsionnel plusieurs filtres non linéaires ont été proposés tels que le médian spatio-temporel et plus particulièrement le médian "3-D multilevel" et le "3-D multistage" [8].

Dans le cas des séquences d'images couleurs, la majorité de ces filtres ont été proposés en utilisant la stratégie marginale [4, 5]. Cette stratégie consiste à traiter chaque composante séparément. Cette approche ignore totalement la corrélation pouvant exister entre les différentes composantes. On peut également noter que cette stratégie demande autant de traitements qu'il y a de composantes, ce qui peut se révéler coûteux en temps de calcul.

Dans cet article, nous développons un filtre d'ordre (L -filter) pour des séquences d'images couleur bruitées. Ce filtre utilise la stratégie vectorielle. Cette stratégie est a priori plus satisfaisante du point de vue de la prise en compte du contexte multicomposante [1]. L'intérêt de cette approche se situe également dans le fait qu'elle ne nécessite qu'un seul traitement, quelque soit le nombre de composantes. Cet avantage doit être pondéré par une complexité accrue de ce seul traitement. La qualité des images obtenues dépend de l'espace couleur utilisé. Pour cela nous avons choisi deux espaces colorimétriques différents, l'espace à composantes corrélées (RGB) et l'espace uniforme (CIELAB) [3]. Ce filtre est basé sur les statistiques d'ordre supérieurs, il utilise comme critère de minimisation le Kurtosis de l'erreur estimée au lieu de l'erreur quadratique moyenne [6]. Ce critère permet de prendre en compte le caractère non-gaussien des perturbations tant dans la phase d'estimation du mouvement, que dans la phase du filtrage.

L'adaptativité de ce filtre doit permettre de prendre en compte les régions mal compensées et n'appartenant pas au mouvement dominant dans l'image.

Cet article est organisé comme suit : la deuxième section est consacrée à décrire l'algorithme d'estimation du mouvement utilisé. Ensuite, la méthode du filtrage multi-composante basée sur les SOS est décrite dans la section 3. Dans la section 4, nous présentons quelques résultats expérimentaux y compris des évaluations qualitatives et quantitatives. Enfin, nous proposons quelques conclusions et remarques dans la section 5.

2 Estimation de mouvement

Dans cette étape nous utilisons une méthode recursive basée sur les statistiques d'ordre supérieur pour l'estimation du mouvement apparent dans les séquences vidéo couleur.

Afin de décrire le mouvement dans chaque région, nous utilisons le modèle du mouvement linéaire simplifié décrit par $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4)^T$ où θ_1 et θ_2 sont respectivement les translations suivant les axes i et j , θ_3 est le facteur de divergence, θ_4 est l'angle de rotation. Le vecteur de déplacement $d(i, j) = (di, dj)^T$ d'un objet est défini par :

$$\begin{cases} di = \theta_1 + \theta_3(i - i_g) - \theta_4(j - j_g) \\ dj = \theta_2 + \theta_3(j - j_g) + \theta_4(i - i_g) \end{cases} \quad (1)$$

(i_g, j_g) sont les coordonnées du centre de la région.

Etant donnée une séquence d'images g_1, \dots, g_N , avec $g_n = \{g_n(s)\}_{s \in S}$, où $s = (i, j)^T$ représente la position du pixel et N le nombre d'images pour chaque composante dans la séquence vidéo. On définit la $DFD(s)$ (Displaced Frame Difference) pour chaque composante par :

$$DFD(s) = g_{n+1}(s + d(s)) - g_n(s) \quad (2)$$

La méthode recursive classique pour calculer le vecteur du mouvement entre deux images utilise la minimisation de l'erreur quadratique moyenne de la $DFD(s)$. Sans information a priori sur la distribution du bruit, il est préférable de minimiser le Kurtosis de la DFD. Ce critère reste optimal lorsque le bruit est gaussien. Le vecteur de déplacement est calculé en minimisant le Kurtosis de la $DFD(s)$:

$$J_4(s) = E\{DFD^4(s)\} - 3E\{DFD^2(s)\}^2 \quad (3)$$

Le champ de mouvement est calculé par optimisation de la DFD des trois composantes couleur :

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \{J_4(w_1 DFD_{comp1}(s) + w_2 DFD_{comp2}(s) + w_3 DFD_{comp3}(s))\} \quad (4)$$

w_i désigne les poids dans chaque composante. La minimisation du Kurtosis est obtenue en utilisant l'algorithme du gradient descendant. Ensuite, le champ de mouvement est alors décomposé en quadtree. Ceci permet une segmentation basique des objets en mouvement et une réduction de complexité de l'algorithme.

3 Filtrage spatio-temporel multi-composante

Une séquence d'images couleur bruitées est modélisée par :

$$G_n(i, j) = F_n(i, j) + N_n(i, j) \quad (5)$$

Avec $G_n(i, j) = (g_n^{comp1}(i, j), g_n^{comp2}(i, j), g_n^{comp3}(i, j))$ est le vecteur intensité du pixel dans la position (i, j) de la $n^{\text{ème}}$ image de la séquence, $F_n(i, j) = (f_n^{comp1}(i, j), f_n^{comp2}(i, j), f_n^{comp3}(i, j))$ est le vecteur intensité du pixel dans la position (i, j) de la $n^{\text{ème}}$ image de la séquence originale et $N_n(i, j) = (\eta_n^{comp1}(i, j), \eta_n^{comp2}(i, j), \eta_n^{comp3}(i, j))$ est un bruit additif.

L'intensité estimée dans la position (i, j) dans la $n^{\text{ème}}$ image est calculée en utilisant un cube de taille $(2p + 1) \times (2q + 1) \times (2l + 1)$ et donnée par :

$$\hat{F}_n(i, j) = \sum_{(p, q, l) \in S} \hat{a}(p, q, l) G_{n-l}^r(i - p, j - q) \quad (6)$$

avec S est le support du filtre, $G_n^r(i, j)$ est le vecteur d'observations ordonnées traitant les trois composantes de la séquence couleur. \hat{a} est le vecteur des coefficients du filtre qui minimisent le Kurtosis de la différence entre l'image estimée et l'image originale :

$$J(s) = E\{e^4(s)\} - 3E\{e^2(s)\}^2 \quad (7)$$

qu'on peut l'exprimer par une notation simple :

$$\hat{a} = \underset{a}{\operatorname{argmin}} [J(e(s))] \quad (8)$$

En tenant compte de la contrainte d'énergie :

$$\sum \hat{a}(s) = 1 \quad (9)$$

$e(s)$ est l'erreur de prédiction sur le pixel s est :

$$e(s) = \hat{F}(s) - F(s) \quad (10)$$

L'équation de la mise à jour des coefficients du filtre du $s^{\text{ème}}$ pixel traité de la $n^{\text{ème}}$ image de la séquence

$$\hat{a}(s + 1) = \hat{a}(s) + \mu \frac{\partial J(s)}{\partial a} \quad (11)$$

la dérivée de $J(s)$ est exprimée par :

$$\frac{\partial J(s)}{\partial a} = E\{4e^3(s) \frac{\partial e(s)}{\partial a}\} - 6E\{e^2(s)\} E\{2e(s) \frac{\partial e(s)}{\partial a}\} \quad (12)$$

L'équation (12) devient :

$$\hat{a}(s + 1) = \hat{a}(s) + \mu [e^3(s) - 3E\{e^2(s)\}e(s)] \frac{\partial e(s)}{\partial a} \quad (13)$$

avec

$$\frac{\partial e(s)}{\partial a} = \frac{\partial \hat{F}(s)}{\partial a} = G_n^r(s) \quad (14)$$

nous obtenons

$$\hat{a}(s+1) = \hat{a}(s) + \mu e^3(s) G_n^r(s) \quad (15)$$

Le pas d'adaptation μ peut être fixé a priori pour toute l'image. Il est préférable d'utiliser un pas adaptatif qui dépend du contenu local de l'image. En pratique le pas d'adaptation adaptatif μ est le produit de μ_0 par un terme de normalisation $\|G_n^r(s)\|^2$. Alors le pas d'adaptation est exprimé par :

$$\mu = \frac{\mu_0}{\|G_n^r(s)\|^2} \quad (16)$$

l'équation de la mise à jour des paramètres devient :

$$\hat{a}(s+1) = \hat{a}(s) + \frac{\mu_0}{\|G_n^r(s)\|^2} e^3(s) G_n^r(s) \quad (17)$$

4 Résultats expérimentaux

Nous avons appliqué la méthode d'estimation du mouvement suivie de la méthode du filtrage spatio-temporel proposée sur plusieurs séquences vidéo couleur. Nous présentons dans cet article quelques résultats appliqués sur quelques images de la séquence **Mobile**¹. Cette séquence de taille 720×576 est classiquement utilisée pour analyser des mouvements peu complexes. Elle présente l'avantage de contenir des objets en mouvement dans l'arrière plan ce qui permet d'évaluer les performances du filtrage.

Dans la figure 1, nous montrons les images résultantes du filtrage spatio-temporel multi-composante basé sur les statistiques d'ordre supérieurs. La figure 1(a) décrit la 3ème image de la séquence originale. Fig 1(b) présente l'image contaminée par un bruit mixte (gaussien ($\sigma=20$) et impulsif ($p=0.05$)). L'image filtrée en utilisant le filtre proposé dans l'espace couleur RGB est représentée dans la figure 1(c). La figure 1(d) montre le résultat du filtrage proposé en utilisant l'espace colorimétrique CIELAB. L'évaluation visuelle montre la supériorité de l'utilisation des espaces colorimétriques uniformes tant en terme d'une bonne estimation du mouvement qu'en terme du filtrage spatio-temporel, tout en préservant les détails de l'image.

La qualité des images filtrées est mesurée par l'erreur quadratique moyenne EQM et NCD (Normalized Color Difference). Le tableau 1 résume les EQM, NCD obtenus lors du filtrage de la 3ème image de la séquence mobile. Ces mesures montrent bien l'efficacité de l'utilisation de l'espace CIELAB en terme de réduction de bruit tout en utilisant une approche de filtrage compensé en mouvement.

Nous avons aussi utilisé d'autres critères de différences² pour évaluer la qualité des images filtrées. Ces critères (luminance, chrominance, émergence, corrélation, contour, fisher) ont été appliqués sur les images filtrées en projetant sur les espaces colorimétriques (RGB, CIELAB). Les résultats de ces évaluations sont résumés dans le tableau 2. Nous pouvons remarquer que la comparaison utilisant ces critères de

différence dans l'espace CIELAB est plus meilleur qu'en utilisant l'espace RGB.

MESURE	NCD	EQM
Image bruitée	0,35	3628,13
Image filtrée (RGB)	0.086	161,89
Image filtrée (LAB)	0.065	105,72

Tableau 1 – Erreur Quadratique Moyenne pour les images filtrées utilisant l'espace RGB et l'espace CIELAB

5 Conclusion

Dans cet article, nous avons présenté un filtre spatio-temporel multicomposante pour la réduction du bruit mixte dans les séquences vidéo. Le filtrage dans l'espace CIELAB se révèle meilleur en terme de qualité visuelle et de certains critères de différences par rapport aux techniques de filtrage dans l'espace classique RGB. A ce niveau, le choix des poids w_i dans chaque composante couleur était arbitraire, un prochain travail sera concerné pour développer cette partie.

Références

- [1] Brian. C. Tom, Aggelos K. Katsaggelos. Resolution enhancement of monochrome and color video using motion compensation, *IEEE transactions on image processing*, Vol. 10, No 2, pp. 278-287, February 2001.
- [2] E. Ibn-elhaj, D. Aboutajdine, S. Pateux and L. Morin. HOS-based method of global motion estimation for noisy image sequences, *Electronics Letters*, vol. 35, no. 16, pp. 1320-1322, August 1999.
- [3] X. Zhang and B. A. Wandell. Color image fidelity metrics evaluated using image distortion maps, *Signal Processing*, vol. 33, no. 3, pp. 201-214, november 1998.
- [4] P.R.Giaccone, G.A. Jones, S. Minelly and A. Curley. Motion-compensated multichannel noise reduction of color film sequences, *Journal of Electronic Imaging*, vol. 8, no. 3, pp. 246-254, June. 1999.
- [5] C. Kotropoulos and I. Pitas. Adaptive Multichannel Marginal L-filters, *SPIE Optical Engineering*, vol. 38, no. 4, vol. 38, no. 4, pp. 688-704, April 1999.
- [6] M. El Hassouni, H. Cherifi et D. Aboutajdine. Débruitage vidéo par statistiques d'ordre supérieurs, *CORESA'2001*, 12-13 novembre 2001. Dijon, France.
- [7] S.A.C. Kokaram. Motion Picture Restoration : Digital Algorithms for Artefacts Suppression in degraded Motion Picture Film and Video. Springer Verlag, 1998.
- [8] J. S. Kim and H. W. Park. Adaptive 3-D median filtering for restoration of an image sequence corrupted by impulsive noise. *Signal Processing : Image Communication* 16 (2001) 657-668.

¹<http://www.cipr.rpi.edu/resource/sequences>

²http://www.univ-reims.fr/Labos/LERI/membre/luc/FRENCH_COLOR_GROUP/



(a)



(b)



(c)



(d)

Figure 1 – Résultats de la 3ème image de la séquence mobile couleur. (a) l'image originale, (b) l'image contaminée par un bruit mixte (gaussien ($\sigma = 20$) + impulsionnel ($p=5\%$)), (c) l'image filtrée en utilisant l'espace colorimétrique RGB et (d) l'image filtrée utilisant l'espace colorimétrique CIELAB

Critères de diff.	Lum		Chro		Emerg		Corr		Cont		Fisher	
	RGB	LAB	RGB	LAB	RGB	LAB	RGB	LAB	RGB	LAB	RGB	LAB
Image bruitée	92,84	67,75	56,22	85,78	21,77	20,77	29,49	23,15	43,40	49,81	29,72	30,20
Image filtrée (RGB)	85,80	86,02	31,09	63,29	20,61	20,45	34,69	36,85	67,24	70,48	23,12	24,29
Image filtrée (LAB)	89,64	90,27	30,10	59,37	24,65	20,82	43,05	38,21	74,30	61,53	39,36	25,79

Tableau 2 – L'évaluation de la qualité de la 3ème image de la séquence mobile filtrées utilisant plusieurs critères de différence en projetant sur deux espaces colorimétriques (RGB, CIELAB)