

Déconvolution aveugle de séquences vidéo

Ferréol SOULEZ^{1,2,3}, Éric THIÉBAUT¹, Yves TOURNEUR⁴, Alain GRESSARD⁵, Raphael DAUPHIN⁵

¹Centre de Recherche Astrophysique de Lyon
Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval cedex, 69561, France;

²Laboratoire de Contrôle Non Destructif par Rayonnements Ionisants
INSA de Lyon, 25 avenue Jean Capelle, 69621 Villeurbanne, France

³Laboratoire Hubert Curien
18 rue Pr Benoît Lauras, 42000 Saint-Etienne, France

⁴Centre Commun de Quantimétrie, Université Lyon 1,
8 avenue Rockefeller - 69373 Lyon cedex 08, France

⁵Service de Cardiologie de l'hôpital de la Croix Rousse, Hospices Civiles de Lyon,
103, Grande Rue de la Croix-Rousse - 69317 Lyon cedex 04, France
ferreol.soulez@obs.univ-lyon1.fr, thiebaut@obs.univ-lyon1.fr

Résumé – Nous proposons ici une méthode pour la déconvolution aveugle de séquences d'images. Dans une approche maximum *a posteriori* (MAP), nous proposons une régularisation séparable pour prendre en compte les corrélations suivant les différentes dimensions tant pour la PSF que pour l'objet. La combinaison d'une régularisation spatiale préservant les bords avec une régularisation temporelle sur l'objet permet de contraindre efficacement le problème de la déconvolution aveugle. Cela aboutit à une restauration importante des détails de l'objet observé. Cette méthode a été testée avec succès pour différentes applications bio-médicales : la coronarographie, la microscopie confocale et la microscopie classique.

Abstract – We present a new blind deconvolution method for video sequence. It is derived following an inverse problem approach in a maximum *a posteriori* (MAP) framework. This method exploits the temporal continuity of both object and PSF. Combined with edge-preserving spatial regularization, a temporal regularization constrains the blind deconvolution problem, improving its effectiveness and its robustness. We demonstrate these improvements by processing various real video sequences obtained by different imaging techniques.

1 Introduction

Le problème de la restauration d'un signal flouté par un processus pas ou peu connu est appelé « déconvolution aveugle ». La déconvolution aveugle d'image est un sujet de recherche très actif depuis une vingtaine d'années (lire [2] pour une revue). Cependant, si plusieurs auteurs comme Schultz[10] et Farsiu[5] utilisent des séquences temporelles d'image pour des méthodes de super-résolution (plusieurs images floues d'une même scène), peu de travaux ont étudié le problème de la déconvolution aveugle de séquence vidéo[6]. Dans ce papier, nous proposons d'utiliser les corrélations temporelles des séquences vidéo pour définir un algorithme de déconvolution aveugle efficace.

2 Description du modèle

Pour une séquence de T images y_t de N pixels chacune, le modèle de formation des données \mathbf{y} est défini par :

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x} + \mathbf{n}, \quad (1)$$

où $\mathbf{y} = (\mathbf{y}_{t_1}^T, \dots, \mathbf{y}_{t_T}^T)^T$ and $\mathbf{x} = (\mathbf{x}_{t_1}^T, \dots, \mathbf{x}_{t_T}^T)^T$ sont des vecteurs de taille $N \cdot T$, \mathbf{n} est vecteur aléatoire correspondant à un bruit

additif et \mathbf{H} est une matrice de taille $N \cdot T \times N \cdot T$.

Nous supposons ici qu'il n'y a pas de rémanence temporelle du capteur et que le flou est invariant par translation. Ceci permet d'exprimer l'Eq. (1) à l'aide d'un produit de convolution spatial $*$ et la PSF \mathbf{h}_t :

$$\mathbf{y}_t = \mathbf{h}_t * \mathbf{x}_t + \mathbf{n}_t. \quad (2)$$

Sous l'approximation circulante, cette expression peut être calculée rapidement par transformée de Fourier rapide.

3 Algorithme

Ce problème de la déconvolution aveugle est abordé par l'approche inverse. La solution au sens du maximum *a posteriori* est le couple objet/PSF $\{\mathbf{x}, \mathbf{h}\}^{\text{MAP}}$ qui minimisent la fonctionnelle $\Phi(\mathbf{h}, \mathbf{x})$:

$$\Phi(\mathbf{h}, \mathbf{x}) = \Phi_{\text{MV}}(\mathbf{h}, \mathbf{x} | \mathbf{y}) + \Psi_{\text{regul}}(\mathbf{h}) + \Theta_{\text{regul}}(\mathbf{x}), \quad (3)$$

où $\Phi_{\text{MV}}(\mathbf{h}, \mathbf{x} | \mathbf{y})$ est un terme d'attache aux données \mathbf{y} . $\Psi_{\text{regul}}(\mathbf{h})$ (resp. $\Theta_{\text{regul}}(\mathbf{x})$) est la fonction de régularisation sur la PSF (resp. l'objet).

Cette fonctionnelle est minimisée de manière alternée[1, 13] en résolvant itérativement $\Phi(\mathbf{h}, \mathbf{x})$ par rapport à \mathbf{x} uniquement puis par rapport à \mathbf{h} . Ces minimisations sont effectuées par la méthode d'optimisation continue à grand nombre de paramètres VMLMB[12] qui est une variante de la méthode à métrique variable avec mise à jour BFGS[8].

3.1 Terme d'attache aux données

Pour un bruit gaussien, le terme d'attache aux données s'écrit :

$$\Phi_{\text{kl}}(\mathbf{x}, \mathbf{h}; \mathbf{g}) = [\mathbf{g} - \mathbf{m}(\mathbf{x}, \mathbf{h})]^T \cdot \mathbf{C}_{\text{noise}}^{-1} \cdot [\mathbf{g} - \mathbf{m}(\mathbf{x}, \mathbf{h})], \quad (4)$$

où $\mathbf{C}_{\text{noise}}$ est la matrice de covariance du bruit et $\mathbf{m}(\mathbf{x}, \mathbf{h})$ le modèle défini d'après l'équation Eq. (2).

Si le bruit n'est pas corrélé, $\mathbf{C}_{\text{noise}}$ est diagonal et Eq. (4) se simplifie en :

$$\Phi_{\text{kl}}(\mathbf{x}, \mathbf{h}, \mathbf{g}) = \sum_t \sum_k w_{t,k} ((\mathbf{H}_t \cdot \mathbf{x}_t)_k - g_{t,k})^2, \quad (5)$$

où $1/w_{t,k}$ est la variance du bruit au pixel k de l'image t . Un tel modèle peut être utilisé avec un bruit non stationnaire et permet ainsi de prendre en compte rigoureusement les bords et les données manquantes ($1/\sigma_{t,k}^2 = 0$ pour les pixels non mesurés).

3.2 Termes de régularisation

Plutôt que de traiter les vidéos image par image, nous proposons de considérer l'ensemble des données dans une perspective unifiée. Pour cela, nous avons développé une régularisation séparable en décomposant les fonctions d'*a priori* sur l'objet et la PSF suivant leurs dimensions. Ces régularisations séparable permettent de prendre en compte facilement les corrélations présentes *a priori* suivant chacune des dimensions du signal.

On a ainsi la régularisation sur l'objet :

$$\Phi_{\text{Regul}}(\mathbf{x}) = \alpha^S \Theta_S(\mathbf{x}) + \alpha^T \Theta_T(\mathbf{x}), \quad (6)$$

avec un terme de régularisation spécifique le long de chaque dimension : (i) Θ_S la régularisation spatiale et (ii) Θ_T la régularisation temporelle. α^S et α^T sont des hyper-paramètres permettant de régler l'importance des *a priori* injectés.

Et de même pour la régularisation sur la PSF :

$$\Psi_{\text{regul}}(\mathbf{h}) = \beta^S \Psi_S(\mathbf{h}) + \beta^T \Psi_T(\mathbf{h}), \quad (7)$$

avec (i) Ψ_S et β^S la régularisation et l'hyper-paramètre spatial (ii) Ψ_T et β^T la régularisation et l'hyper-paramètre temporel.

3.2.1 *a priori* sur l'objet

Régularisation spatiale : Le bruit corrompant principalement les hautes fréquences, une contrainte de lissage est utilisée pour éviter l'amplification du bruit :

$$\Theta_S(\mathbf{x}_t) = \sum_k \sum_{v \in V_k} \varphi \left(\frac{x_{t,k} - x_v}{d_v} \right), \quad (8)$$

où x_v est la valeur du pixel v dans le voisinage du pixel k (ici un voisinage V_8) et d_v sa distance à ce pixel. Pour ne pas lisser

les bords francs présents dans l'image, nous utilisons comme fonction φ une norme non quadratique de type l_1-l_2 [3, 7] :

$$\varphi(u; \eta) = 2\eta^2 [|u|/\eta - \log(1 + |u|/\eta)], \quad (9)$$

Cette norme est asymptotiquement quadratique (resp. linéaire) pour de petites différences (resp. grande) comparées à η . La valeur de η n'est pas critique et elle peut être fixée une fois pour toute à la valeur d'un niveau de quantification. Dans ce cas, cette régularisation est très proche d'une régularisation purement ℓ_1 qui favorise la parcimonie du gradient[4], mais elle peut être minimisée plus rapidement par notre algorithme d'optimisation.

Régularisation temporelle : Si l'évolution de l'objet est lente comparée à la vitesse d'échantillonnage — ce qui n'est pas le cas dans les résultats présentés — une régularisation selon l'axe temporel similaire à la régularisation spatiale peut être définie.

3.2.2 *a priori* sur la PSF

Les causes de flou étant très variées (bougé, défocus, turbulence...), le choix de la régularisation dépendra des conditions expérimentales. Ce choix sera principalement entre une contrainte de lissage identique à celle décrite pour l'objet ou une contrainte de forme si l'on dispose d'une forme paramétrique *a priori* de la PSF[11].

Comme pour l'objet, si l'évolution de la PSF est lente devant la vitesse d'échantillonnage, une régularisation lissante peut être utilisée :

$$\Psi_T(\mathbf{h}) = \sum_{t=1}^{T-1} (\mathbf{h}_t - \mathbf{h}_{t-1})^2. \quad (10)$$

En plus, de ces *a priori* des contraintes strictes de positivité et de normalisation sont utilisées sur la PSF.

4 Résultats

4.1 Coronarographie

La coronarographie est un examen radiologique des artères du cœur (les coronaires). Cela consiste à filmer l'écoulement dans l'arbre coronaire d'un liquide iodé opaque aux rayons X. L'examen de ces courtes séquences vidéo permet aux cardiologues de localiser un obstacle dans les artères pouvant être responsable d'infarctus. La diffusion des rayons X au cours de la traversée du corps du patient provoque un flou sur ces données dont le contraste est de plus très faible en raison de la puissance limitée des rayons X émis.

Dix séquences de plusieurs dizaines d'images (environ 70) de taille 512×512 ont été acquises. Une image de la première séquence est présentée sur la figure 1. Notons en particulier que la dynamique de l'image est très faible; bien qu'il y ait 128 niveaux possibles, la différence de niveaux entre un pixel situé dans une coronaire et un pixel du fond est de quelques niveaux (< 10).

Une image de la reconstruction de la première séquence est affichée sur la figure 2 avec la PSF reconstruite figure 3. Bien

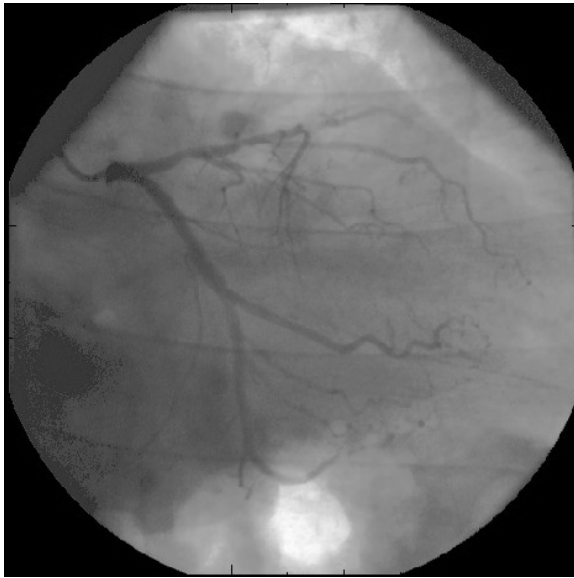


FIG. 1 – Trame 38 de la séquence coronarographique numéro 1.

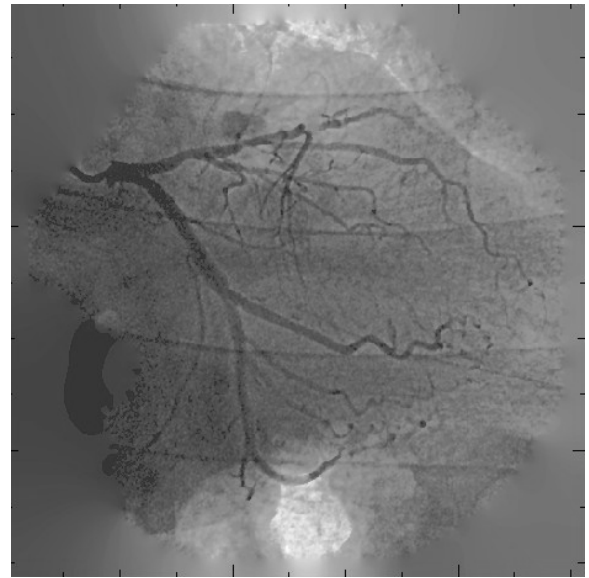


FIG. 2 – Déconvolution aveugle correspondante.

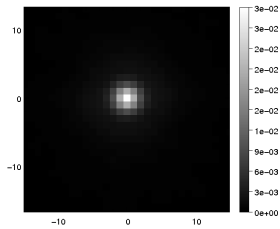


FIG. 3 – PSF reconstruite pour la première séquence.

que la quantification utile soit faible, une inspection visuelle permet de constater l'efficacité de la méthode proposée. Comme la séparation des artères coronaires et du fond est plus nette, la perception du mouvement sur la séquence vidéo est très largement améliorée. La justesse de cette amélioration a été confirmée par les cardiologues à qui ces résultats ont été présentés.

4.2 Microscopie confocale

La microscopie confocale en fluorescence permet de réaliser une carte de la distribution de fluorophores dans les trois dimensions (x,y,z) avec une très faible profondeur de champ. Les fluorophores marquant des protéines particulières, cette technique permet de visualiser des organites à l'intérieur des cellules.

Nous avons testé notre méthode sur une séquence temporelle d'images en microscopie confocale prise toujours dans le même plan. Cette séquence (x,y,t) est issue d'une partie des données utilisées dans le cadre de l'expérience menée par S. Pelloux[9] qui étudie les mitochondries de cellules cardiaques nommées "non-beating HL-1".

Seulement deux itérations de minimisation alternée ont été suffisantes pour aboutir à un résultat satisfaisant. La déconvolution aveugle de cette séquence de 64 images a été effectuée en 24 minutes soit 22 secondes par image. Le résultats de cette

déconvolution aveugle est affichée sur la figure 5 et est à comparer avec les données affichées avec la même palette de couleurs Fig. 4. Si peu de détails supplémentaires ont été restaurés, on peut constater que le bruit a été largement supprimé.

4.3 Microscopie classique

Nous avons aussi appliqué notre méthode sur des données prises en microscopie classique. Ces données sont des séquences d'images de cellules ciliées épithéliales prises au microscope à transmission. Le but des biologistes est ici de caractériser le mouvement des cils de ces cellules pour étudier les dysfonctionnements de ces mouvements dans certaines maladies génétiques. Pour cela ils utilisent des séquences d'images prise au microscope à transmission avec un fort grossissement pour distinguer ces cils d'environ 70 nm de diamètre.

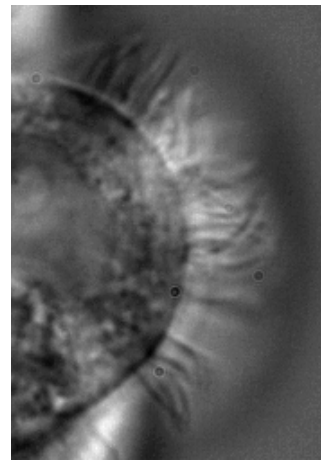


FIG. 6 – Première image de la séquence sur la cellule ciliée.

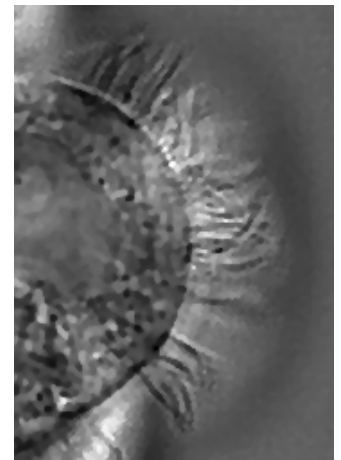


FIG. 7 – Déconvolution aveugle correspondante.

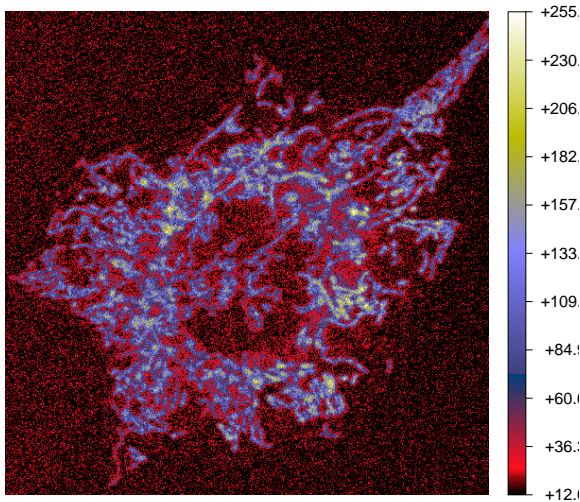


FIG. 4 – Première image de la séquence sur les mitochondries.

Nous avons déconvolué en aveugle une séquence expérimentale de 400 images de taille 253×178 pixels. La première image de la séquence déconvoluée est affichée sur la figure 7. On peut constater sur cette reconstruction que dans même ce cas “limite” ou le flou n’est pas tout à fait isoplanétique, l’algorithme permet une amélioration de la résolution et, par exemple, permet de mieux séparer certains cils.

5 Conclusion

Nous avons présenté une méthode de déconvolution aveugle adaptée pour la restauration de séquence vidéo. L’utilisation de régularisations, sur l’objet ou la PSF, séparables suivant les dimensions a permis d’exploiter avec succès les corrélations spatiales comme temporelles pour contraindre efficacement le problème. En particulier, l’utilisation de contraintes temporelles sur la PSF a abouti à une bonne caractérisation de celle-ci et ainsi une bonne déconvolution de la séquence vidéo. Enfin, l’efficacité et la robustesse de cette méthode ont été illustrées sur des séquences expérimentales provenant de différentes techniques d’imagerie dynamique.

Remerciement

Les auteurs tiennent particulièrement à remercier les personnes ayant effectué les préparations biologiques et acquis les vidéos : Sophie Pelloux pour les mitochondries et Brigitte Chhin pour les cellules épithéliales.

Références

- [1] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3) :259–302, 1986.
- [2] P. Campizi and K. Egiazarian, editors. *blind image deconvolution : theory and application*. CRC Press, 2007.
- [3] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in

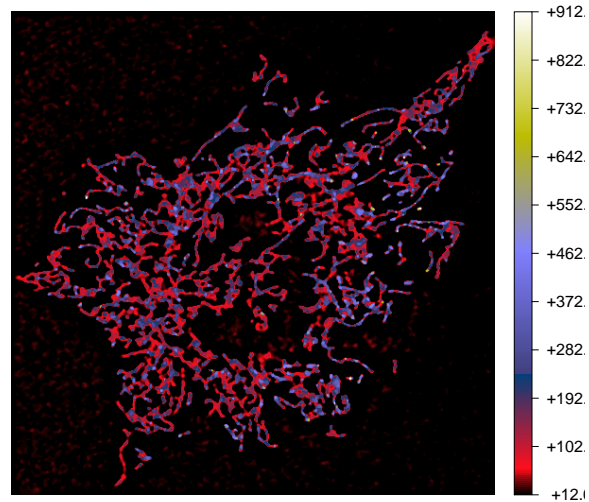


FIG. 5 – Déconvolution aveugle correspondante.

computed imaging. *Image Processing, IEEE Transactions on*, 6(2) :298–311, 1997.

- [4] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4) :1289–1306, 2006.
- [5] S. Farsiu, MD Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *Image Processing, IEEE Transactions on*, 13(10) :1327–1344, 2004.
- [6] D. Li, R.M. Mersereau, D.H. Frakes, and M.J.T. Smith. New method for suppressing optical turbulence in video. In *Proc. European Signal Processing Conference (EUSIPCO)*, 2005.
- [7] M. Nikolova. Regularisation functions and estimators. In *International Conference on Image Processing*, volume 1, pages 457–460, 1996.
- [8] Jorge Nocedal. Theory of algorithms for unconstrained optimization. *Acta Numerica*, 1 :199–242, 1992.
- [9] S. Pelloux, J. Robillard, R. Ferrera, A. Bilbaut, C. Ojeda, V. Saks, M. Ovize, and Y. Tourneur. Non-beating h1-1 cells for confocal microscopy : Application to mitochondrial functions during cardiac preconditioning. *Progress in Biophysics and Molecular Biology*, 90(1-3) :270–298, 2006.
- [10] T. J. Schulz. Multifram blind deconvolution of astronomical images. *JOSA-A*, 10(5) :1064–1073, may 1993.
- [11] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial Mathematics, 2005.
- [12] E. Thiébaud. Optimization issues in blind deconvolution algorithms. In Jean-Luc Starck, editor, *Astronomical Data Analysis II*, volume 4847, pages 174–183, dec 2002.
- [13] Y.L. You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *Image Processing, IEEE Transactions on*, 5(3) :416–428, 1996.