

# Défloutage de vidéos en temps réel

J. Anger

CMLA, ENS Cachan

Université Paris-Saclay, 94235 Cachan, France

jeremy.anger@cmla.ens-cachan.fr

## Résumé

Nous proposons d'améliorer une méthode de défloutage vidéo récente afin de palier à certaines de ses limitations et de permettre une implémentation temps réel. Pour cela, nous adoptons pour un recalage par translations locales estimées dans le domaine de Fourier. Nous proposons également une gestion simple des objets mobiles présents dans la scène.

## Mots Clef

Défloutage vidéo, recalage, transformée de Fourier

## Abstract

We propose to improve a recent video deblurring method in order to overcome some of its limitations and allow real-time implementations. To this end, we adopt an image registration based on local translations estimated in the Fourier domain. We also propose a simple adaptation to handle object motion in the scene.

## Keywords

Video deblurring, registration, Fourier transform

## 1 Introduction

Lors de l'acquisition de vidéo par une caméra en mouvement, les images subissent des dégradations dues notamment aux tremblements du porteur de la caméra. Ces dégradations peuvent être modélisées par l'équation suivante :

$$u_i = k_i * v_i + n_i,$$

où  $u_i$  est l'image observée,  $k_i$  est le noyau de flou de bougé associé à cette image,  $n_i$  est un bruit Gaussien et  $*$  est l'opérateur de convolution. L'objectif du défloutage vidéo est de reconstruire les images  $v_i$  à partir des images observées. À la différence des méthodes de déconvolution non aveugle, les noyaux  $k_i$  sont ici inconnus et différents pour chaque image. En vision par ordinateur, restaurer les vidéos en retirant le flou permet de meilleures performances lors de traitements d'analyse de scène ou de reconnaissance. Ce problème intervient également pour reconstruire une image nette à partir de plusieurs images acquises par rafale et dont le porteur de la caméra aurait été sujet à des micro tremblements.

## 1.1 Travaux antérieurs

Local Fourier Burst Accumulation (FBA) [1] est une méthode récente de défloutage vidéo qui consiste à faire une agrégation des coefficients de Fourier pour chaque fréquence. La pondération associée à chaque coefficient de Fourier dépend de son module. En effet, l'opération de flou réduisant plus ou moins le module selon la forme du noyau de flou, le coefficient le moins affecté par le flou est celui qui a un module le plus élevé. L'accumulation s'exprime donc par la formule suivante, où  $G_\sigma$  est un opérateur de flou Gaussien permettant de lisser la carte de pondérations fréquentielles, et  $U_i$  et  $V_i$  sont respectivement les transformées de Fourier discrètes de  $u_i$  et  $v_i$  :

$$V_i(\omega) = \frac{\sum_i U_i(\omega)(G_\sigma|U_i(\omega)|)^p}{\sum_i (G_\sigma|U_i(\omega)|)^p}$$

Les images  $u_i$  sont préalablement recalées par flot optique et l'accumulation est effectuée par tuile de taille  $128 \times 128$  avec recouvrement à 50%.

La méthode est capable de reconstruire une image nette dans le cas où les noyaux de flou sous-jacents sont suffisamment différents.

L'une des limitations, en terme de qualité d'image et de temps d'exécution, est l'utilisation d'un algorithme de flot optique. Celui ci doit être calculé à une résolution inférieure afin de ne pas compenser les déformations introduites par le flou ce qui le rend moins précis. Afin de gérer les images mobiles, un schéma *forward-backward* est utilisé afin de ne pas faussement recalcr les images ; ce schéma est la partie la plus couteuse de la méthode.

## 2 Améliorations proposées

### 2.1 Recalage par translation

Afin d'améliorer FBA, nous proposons de remplacer le recalage non rigide par un recalage dans Fourier.

Un modèle de recalage global ne permettant pas une gestion convenable des objets mobiles, nous conservons la notion de tuiles. Celle ci a également l'avantage d'éviter la propagation d'artefacts, et se prête particulièrement bien à une implémentation parallèle de la méthode.

Si les tuiles sont suffisamment petites par rapport à l'intensité des mouvements de la caméra, nous pouvons supposer un modèle par translation. De plus, lors des tremblements

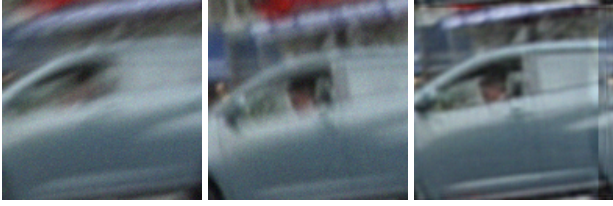


FIGURE 1 – Deux tuiles d’une séquence simulée de 7 images et le résultat de la méthode proposée.

induits par le porteur de la caméra (véhicule ou humain), la rotation le long de l’axe optique est bien plus rare que les décalages par translations.

Soit  $T_1$  et  $T_2$  deux tuiles dans le domaine de Fourier extraites à la même position dans deux images successives. Il est possible de calculer la translation  $(x, y)$  permettant de recalculer  $T_1$  sur  $T_2$  par la méthode de *phase correlation* :

$$PC(\omega) = \frac{T_1(\omega)T_2^*(\omega)}{|T_1(\omega)T_2^*(\omega)|} \approx e^{i(x\omega_x + y\omega_y)}$$

La DFT inverse de  $PC$  produit un dirac dont les coordonnées correspondent à  $(x, y)$ .

Une des limitations du recalage par *phase correlation* est sa sensibilité au flou différent entre les deux images. Villet et al. [3] proposent de remplacer  $PC$  par une version invariante aux noyaux de flou possédant une symétrie centrale, ce qui inclue les noyaux Gaussien et les noyaux composés d’un segment, de type flou de mouvement uniforme. Cependant, ces noyaux ne sont pas représentatifs de ceux présents sur des vidéos réelles et nous observons de meilleurs résultats en utilisant une simple corrélation croisée. Des expérimentations devront donc être effectuées afin de vérifier la pertinence des diverses méthodes dans le cadre d’images floues.

Une fois la translation entre deux tuiles estimée, le recalage peut s’effectuer dans Fourier par décalage de phase. Ce recalage produit une translation circulaire, comme visible à droite sur la figure 1. Cette translation n’a pas d’impact sur la fusion FBA, mais impose qu’une bordure de la tuile résultat doit être supprimée. Pour cela, après la fusion des tuiles, nous extrayons uniquement la partie centrale de la tuile, ce qui permet également d’éviter les artéfacts de périodicité présents dans la méthode FBA originale.

L’information de déplacement entre  $T_i$  et  $T_{i+1}$  peut permettre d’extraire  $T_{i+2}$  à une position dans l’image plus adaptée, en supposant une continuité relative du mouvement de la caméra. De plus, si le recalage est suffisamment précis, il est possible de composer les estimations de déplacements de  $T_i$  à  $T_{i+1}$  et de  $T_{i+1}$  à  $T_{i+2}$  pour éviter d’estimer le déplacement  $T_i$  à  $T_{i+2}$ , ce qui constitue un gain de temps considérable.

## 2.2 Gestion des objets mobiles

Dans le cas où un objet en mouvement est présent dans les tuiles, une agrégation directe produit des artéfacts de type

objets fantômes. Pour palier à cela, une première solution consiste à ne pas considérer les tuiles dont l’estimation de translation a indiqué une faible corrélation. En effet, deux cas peuvent se produire : soit l’estimation de mouvement échoue et la corrélation sera très faible, soit la scène est correctement recalée mais les objets mobiles ne le sont pas, auquel cas le coefficient de corrélation sera également affaibli par rapport à une tuile correctement recalée et sans objets mobiles.

Une seconde solution consiste, lorsque le recalage échoue, à utiliser un algorithme de détection de mouvement de type flot optique dans le domaine spatial et à remplacer les pixels identifiés comme appartenant à un objet mobile par ceux de la tuile de référence. Cette solution correspond en partie à la méthode mise en place dans FBA.

Enfin, Hasinoff et al. [2] rendent robuste leur méthode de fusion multi-images face aux objets mobiles par un filtrage fréquentiel combinable avec FBA sans augmenter significativement le temps de calcul.

## 3 Conclusion

Nous avons proposé des améliorations à la méthode Local Fourier Burst Accumulation. Premièrement, nous remplaçons le recalage par flot optique par un recalage de chaque tuile par translation estimée et effectuée dans le domaine de Fourier. En plus d’être très efficace, effectuer le recalage dans ce domaine permet de directement calculer FBA sur les tuiles sans repasser dans le domaine spatial. Ce recalage se prête également bien à la parallélisation puisqu’il considère l’image comme un ensemble de blocs indépendants. Cependant, le recalage doit être adapté afin de le rendre robuste au flou qui est différent au cours de la vidéo.

Nous proposons également une gestion des objets mobiles simple en utilisant la confiance de l’estimation de mouvement. À travers ces différents changements, nous espérons atteindre une cadence temps réel avec une implémentation GPU.

## Références

- [1] Mauricio Delbracio and Guillermo Sapiro. Hand-held Video Deblurring via Efficient Fourier Aggregation. *IEEE Transactions on Computational Imaging*, 1(4) :270—283, 2015.
- [2] Samuel W. Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T. Barron, Florian Kainz, Jiawen Chen, and Marc Levoy. Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics*, 35(6) :1–12, 2016.
- [3] Ojansivu Ville and Janne Heikkila. Image registration using blur-invariant phase correlation. *Signal Processing Letters, IEEE*, 14(7) :449–452, 2007.