

Détection et reconstruction des éléments hautes fréquences appliquées au désentrelacement

Jérôme ROUSSEL^{1,2}, Pascal BERTOLINO¹, Marina NICOLAS²

¹GIPSA-Lab, INPG-CNRS
ENSIEG, Domaine universitaire, Grenoble, France.

²STMICROELECTRONICS Grenoble.
12 rue Jules Horowitz, Grenoble, France.

jerome.rousseau@lis.inpg.fr, pascal.bertolino@inpg.fr, marina.nicolas@st.com

Résumé – Cet article présente un nouvel algorithme de désentrelacement spatial qui pourra être facilement intégré dans un système plus complet de désentrelacement comme le système spatio-temporel adaptatif. L'interpolation spatiale ne conserve pas la continuité des lignes horizontales ce qui provoque souvent des artefacts visuels. Notre système préserve cette continuité en prenant en compte que les problèmes se situent aux extrema locaux de l'image. Notre algorithme repose sur l'utilisation de listes chaînées et sur la construction de graphes. La nouvelle méthode est restreinte à ces seules structures car pour le reste de l'image un système de désentrelacement classique donne de bons résultats. Bien que le nombre de pixels affectés par notre interpolation soit faible, l'amélioration de la qualité subjective est importante. De plus, notre solution permet une amélioration proche des méthodes de compensation de mouvements, sans leur grande complexité.

Abstract – This article presents a new algorithm for spatial deinterlacing that could easily be integrated in a more complete deinterlacing system, typically a spatio-temporal motion adaptive one. The spatial interpolation part often fails to reconstruct close to horizontal lines with a proper continuity, leading to highly visible artifacts. Our system preserves the structure continuity taking into account that the misinterpolated points usually correspond to local value extrema. The processing is based on chained lists and connected graph construction. The new interpolation method is restricted to such structures, for the rest of the image, a proper traditional directional spatial interpolation gives satisfactory results already. Although the number of pixels affected by the extrema interpolation is relatively small, the overall image quality is subjectively well improved. Moreover, our solution allows to gain back one of the major advantages of motion compensation methods, without having to afford their complexity cost.

1 Introduction

Le signal vidéo est diffusé dans le monde entier en trames entrelacées : l'image est séparée en deux trames qui représentent les champs pairs et impairs. Cependant, aujourd'hui, les écrans plats (plasma, L.C.D) ont un affichage progressif qui nécessite l'affichage d'une image entière. Les méthodes permettant de passer de l'entrelacé au progressif sont appelées désentrelacement. Elles peuvent être classées en deux grandes familles, les méthodes sans compensation et avec compensation de mouvement [1]. Nous nous intéresserons ici aux méthodes sans compensation de mouvement. Celles-ci peuvent à leur tour être décomposées en méthodes temporelles, spatiales et adaptatives [2]-[3]. La solution spatiale est la seule possible sur les zones en mouvement de l'image mais elle possède différentes limitations. Les méthodes proposées pour faire de l'interpolation spatiale sont nombreuses [4]. Une des solutions les plus répandues est d'interpoler les pixels manquants en considérant la direction des contours. Cette méthode appelée E.L.A. (Edge Line Average)[5] détecte dans une fenêtre centrée sur le pixel à interpoler la meilleure direction possible, puis effectue l'interpolation selon cette direction. Malgré une assez bonne interpolation des contours, elle présente de nombreux défauts. En effet, la corrélation se fait au ni-

veau local et reste très sensible aux bruits. La direction des contours est donc parfois erronée ce qui donne des artefacts gênants. De nombreuses variantes de cette méthode [6] permettent de résoudre ce problème sur une majorité des pixels de l'image : la corrélation est effectuée sur des groupes de pixels et non plus pixel à pixel (*figure 1*). Toutefois, ces méthodes sont dépendantes de leur taille de fenêtre qui limite l'angle de reconstruction des contours. En outre, plus la fenêtre est grande, plus le risque de mauvaise interpolation est élevé [7]. Différentes métriques existent pour essayer d'agrandir la fenêtre et rajoutent des poids pour réduire le nombre de fausses directions [8]-[9]. Mais la complexité augmente fortement pour calculer ces poids de manière efficace.

La méthode proposée s'affranchit des méthodes actuelles dans la mesure où celle-ci n'est plus basée sur une recherche dans une fenêtre. Elle vient en complément d'une méthode classique de désentrelacement spatial qui pourra elle-même être intégrée à une solution adaptative. Son terrain d'action est limité aux zones de l'image qui possèdent certaines caractéristiques posant des problèmes aux méthodes existantes.

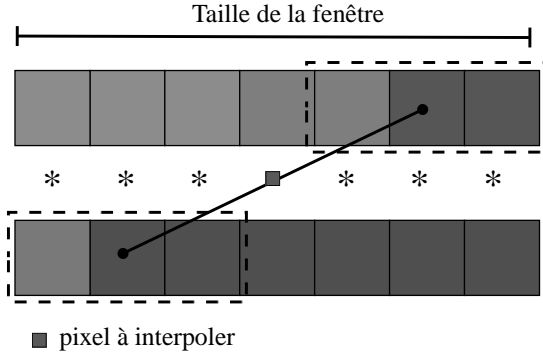


FIG. 1 – Détection sur une fenêtre donnée de la direction du contour pour une bonne interpolation



FIG. 2 – Exemple d’une partie d’image où les lignes noires sont à interpoler

2 Détection

La difficulté consiste à localiser puis reconstituer la continuité des structures hautes fréquences (*figure 2*). Elles peuvent être assimilées à des droites épaisses partiellement détruites et systématiquement déconnectées par le sous-échantillonnage horizontal. Leur détection est réalisée sur les lignes connues de l’image en comparant la valeur de chaque pixel $f(i, j)$ avec les valeurs des lignes inférieures et supérieures voisines $f(i-2, j)$ et $f(i+2, j)$. (i, j) sont les coordonnées spatiales d’un pixel, où i représente les lignes et j les colonnes. On peut ainsi facilement obtenir deux ensembles distincts de pixels qui sont \mathcal{H} l’ensemble des pixels de type maxima locaux et \mathcal{L} l’ensemble des pixels de type minima locaux :

$$\mathcal{H} = \{f(i, j) > \max(f(i-2, j), f(i+2, j)) + T\} \quad (1)$$

$$\mathcal{L} = \{f(i, j) < \min(f(i-2, j), f(i+2, j)) - T\} \quad (2)$$

T est une valeur de contraste minimum autorisé ($T = 16$ dans nos expérimentations).

Sur une même ligne, les extrema locaux d’un même type peuvent former des composantes connexes (ou segments) au sens de la 2-connexité horizontale (*figure 3*). Comme la suite de la méthode n’est pas fondée sur le parcours et le traitement de pixels mais sur le parcours et le traitement de segments, la structure bi-dimensionnelle classique de l’image n’est plus appropriée et est abandonnée au profit d’une structure de plus haut niveau : le segment. Chaque segment est une entité caractérisée par ses coordonnées (ligne, colonne de départ), sa longueur et son type (minimum ou maximum). La structure de données choisie est un compromis entre la taille mémoire requise et la complexité pour parcourir \mathcal{H} et \mathcal{L} .



FIG. 3 – En noir, les segments de minima extraits à partir de la figure 2. Les lignes grises sont connues, les lignes blanches sont à interpoler

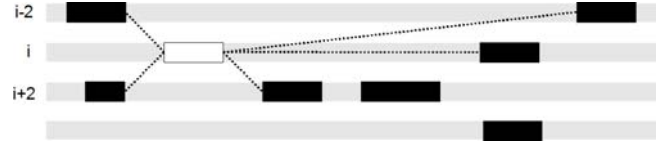


FIG. 4 – Le segment marqué en blanc possède 5 voisins directs

3 Liaison

A cette étape nous disposons d’un tableau de lignes (une entrée par ligne de trame), chaque ligne étant une liste chaînée des segments extraits sur la ligne correspondante de la trame. Notre but est de réussir à inter-connecter entre eux les segments de \mathcal{H} (resp. de \mathcal{L}) afin de constituer un ou plusieurs graphes connexes de maxima (resp. un ou plusieurs graphes connexes de minima) représentant des droites épaisses déconnectées de l’image.

Nous effectuons pour ceci une recherche des plus proches voisins comme décrit ci-dessous : soit un segment S sur une ligne i , il possède au plus 6 voisins directs de même type : 3 du côté ouest et 3 du côté est sur chaque ligne $i-2$, i et $i+2$ (*figure 4*). La distance entre deux segments voisins de même type S_1 et S_2 est la distance euclidienne $d(S_1, S_2)$, calculée entre les extrémités les plus proches de S_1 et S_2 .

Pour un côté donné (ouest ou est), S est connecté à son voisin qui est à la plus faible distance. Si deux voisins sont les plus proches à la même distance, S est connecté aux deux (*figure 5*). Un seuil adapté aux longueurs des deux segments candidats à la connexion est utilisé afin d’éviter des connexions peu fiables car trop longues. Ce seuil adaptatif est défini comme suit : L_1 et L_2 étant les longueurs respectives de S_1 et S_2 , la distance $d(S_1, S_2)$ doit vérifier la condition suivante pour que S_1 et S_2 soient connectés :

$$d(S_1, S_2) < \min(L_1, L_2) + \delta \quad (3)$$

Dans nos expérimentations δ est fixé à 2. Les connexions d’un segment avec ses voisins sont stockées dans le segment lui-même (sous forme de pointeurs sur les segments voisins). Chaque composante connexe de segments est un graphe (*figure 5*).

4 Simplification des graphes

Certains graphes extraits peuvent contenir des connexions indésirables qui doivent être supprimées avant l’étape de reconstruction (*figure 5*), ce qui provoque leur simplification et leur division en plusieurs sous-graphes. La suppression des connexions doit d’une part privilégier des



FIG. 5 – Graphe connexe correspondant aux segments de la figure 3



FIG. 6 – Les connexions restantes après la simplification du graphe de la figure 5

sous-graphes ayant chacun une direction prédominante et d'autre part supprimer les faux-positifs (connexions effectuées à tort).

Lors du parcours en profondeur du graphe, on appelle direction d'entrée celle par laquelle le segment est accédé. Cette direction est inexistante pour le segment de départ du parcours. Les directions de sortie correspondent à toutes les connexions du segment sauf la direction d'entrée. La simplification respecte les règles suivantes :

1. S'il existe 2 connexions de sortie du même côté, elles sont supprimées. Cette règle permet de ne pas relier à tort des structures potentiellement différentes.
2. Si une connexion de sortie est du même côté que la connexion d'entrée, elle est supprimée. Cette règle permet de ne conserver que des structures étirées et non en zig-zag.

Les sous-graphes résultant du parcours et des règles de simplification sont des arbres à une seule branche (figure 6). De longues structures rectilignes ou courbes peuvent être ainsi reconstruites.

5 Reconstruction

La reconstruction des informations manquantes est la dernière étape (figure 7). Elle est réalisée lors d'un parcours en avant (de l'ouest vers l'est) de chacune des branches. Chaque segment connecté à un autre (noir) permet de définir un nouveau segment (gris) les reliant. Les valeurs des pixels dans ce nouveau segment seront obtenues par une interpolation linéaire des valeurs appartenant aux deux segments voisins, comme explicité ci-dessous.

Soient S_1 et S_2 deux segments connectés de longueur L_1 et L_2 , ayant comme coordonnées de départ (Y_{S_1}, X_{d_1}) et (Y_{S_2}, X_{d_2}) . Les pixels à interpoler avec notre méthode correspondent au segment S_I dont les abscisses extrémités X_{d_I} et X_{f_I} , sont interpolées linéairement des abscisses extrémités des segments S_1 et S_2 .



FIG. 7 – Les segments en gris foncé représentent les pixels de maxima interpolés grâce aux segments connexes noirs

$$X_{d_I} = \left\lceil \frac{X_{d_2} + X_{d_1}}{2} \right\rceil \quad (4)$$

$$X_{f_I} = \left\lceil \frac{X_{f_2} + X_{f_1}}{2} \right\rceil \quad (5)$$

$$\begin{aligned} \tilde{f}(i, j) = & \frac{1}{2}f\left(Y_{S_1}, X_{d_1} + \left\lceil \frac{(j - X_{d_I}) \times (L_1 - 1)}{(L_I - 1)} \right\rceil\right) \\ & + \frac{1}{2}f\left(Y_{S_2}, X_{d_2} + \left\lceil \frac{(j - X_{d_I}) \times (L_2 - 1)}{(L_I - 1)} \right\rceil\right) \end{aligned} \quad (6)$$

$$j \in [X_{d_I}, X_{f_I}]$$

L_I est la taille du segment à interpoler et vaut $X_{d_I} - X_{f_I} + 1$.

6 Résultats

Pour les tests, une interpolation des extrema est réalisée avec notre méthode, la méthode E.D.I.[8] étant utilisée pour le reste de l'image. Les premiers résultats ont été obtenus sur des images fixes avec de nombreux détails horizontaux, telle que l'image du phare où la méthode reconstruit bien la continuité des structures (figure 8). Sur la séquence "tennis", l'amélioration est également très visible, et lorsque la vidéo est jouée en temps réel les lignes horizontales en mouvement deviennent plus stables. D'une manière générale, les structures quasi-horizontales qui ne sont pas reconstruites par les méthodes classiques sont ici presque identiques à l'original. Les résultats obtenus présentent une amélioration importante proche des méthodes de compensation de mouvements, en conservant la complexité des méthodes spatiales classiques.

En effet, le temps de traitement diffère très peu de celui de la méthode ELA. En moyenne le pourcentage de pixels interpolés par notre méthode est de l'ordre de 2% de l'image entière. Ce faible pourcentage explique les mesures de PSNR (table 1). En effet, la différence entre les méthodes est plus significative lorsque le PSNR est calculé seulement sur les pixels interpolés par notre méthode. Enfin, l'amélioration subjective est visible d'autant plus que l'œil est très sensible à la continuité des structures rectilignes et à leur scintillement. Néanmoins, l'interpolation ici est uniquement spatiale et ne peut pas reconstruire des contours horizontaux qui ont été supprimés par l'entrelacement comme peut le faire un système avec compensation de mouvement beaucoup plus coûteux.

7 Conclusion et perspectives

Notre méthode ne s'appuie pas sur le principe de méthodes existantes. Elle pallie ainsi les problèmes entraînés par ces dernières. La méthode s'attache à corriger les artefacts les plus désagréables pour l'œil en détectant leur

TAB. 1 – Comparaison (1) PSNR sur l'image entière (2) PSNR sur les pixels interpolés par notre méthode

PSNR		Linéaire	E.D.I.	Notre Méthode
Phare	(1)	30.28	30.23	31.48
	(2)	18.17	18.09	23.35
Tennis de Table	(1)	24.23	24.62	25.34
	(2)	17.09	17.96	21.16
Starwars	(1)	38.04	38.16	38.39
	(2)	24.24	24.39	26.34
Course Auto	(1)	30.29	30.54	31.14
	(2)	20.31	20.70	22.66

origine. Elle est fondée sur la continuité des objets afin de les reconstruire. De ce fait, les structures à fort contraste sont plus stables. Enfin, notre méthode peut se greffer à toutes les méthodes classiques pour améliorer leurs défauts sans un surcoût élevé.

Le réglage du seuil pour la détection des extrema dépend de la dynamique locale ou globale de l'image. Une solution pour s'affranchir de ce seuil pourra être de pré-filtrer l'image. Ainsi le seuil ne dépendrait plus des basses fréquences de l'image.



FIG. 8 – Comparaison des résultats sur l'image Phare

Références

- [1] K. Sugiyama, H. Nakamura. "A method of de-interlacing with motion compensated interpolation", IEEE Trans. on Consumer Electronics, vol 45 p 611-616 (1999).
- [2] T. Koivunen. "Motion detection of an interlaced video signal", IEEE Trans. on Consumer Electronics, vol 40 p 753-760 (1994).

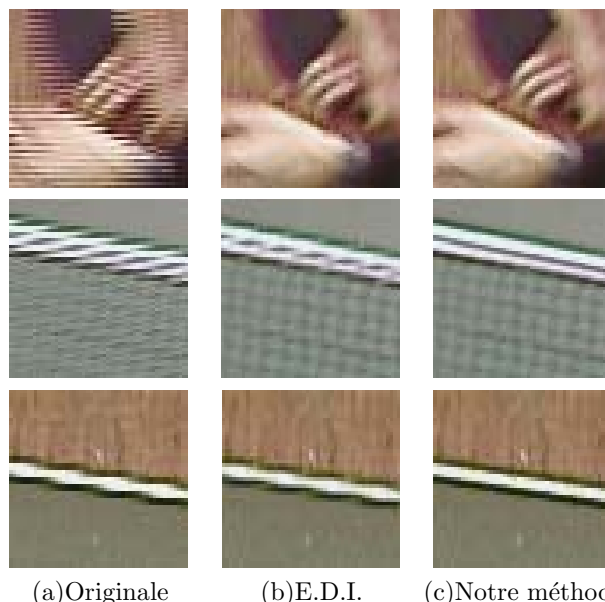


FIG. 9 – Comparaison des résultats sur la séquence Tennis de table

- [3] G. de Haan, E.B. Bellers. "De-Interlacing of video data", IEEE Trans. on Consumer Electronics, vol 43 p 819-825 (1997).
- [4] G. de Haan, E.B. Bellers. "Deinterlacing - An overview", Proceedings of the IEEE, vol 86 p 1839-1857 (1998).
- [5] T. Doyle. "Interlaced to sequential conversion for EDTV applications", in Proc. 2nd International Workshop Signal Processing of HDTV, p 412-430 (1988).
- [6] T. Chen, H. Ren Wu, Z. Hua Yu. "Efficient deinterlacing algorithm using edge-based line average interpolation", Optical Engineering, vol 39 p 2101-2105 (2000).
- [7] H. Yoo, J. Jeong. "Direction-oriented interpolation and its application to de-interlacing", IEEE Trans. on Consumer Electronics, vol 48 p 954-962 (2002).
- [8] M.K. Park, M.G. Kang, K. Nam, S.G. Oh. "New edge dependent deinterlacing algorithm based on horizontal edge pattern", IEEE Trans. on Consumer Electronics, vol 49 p 1058-1512 (2003).
- [9] M. Byun, M.K. Park, M.G. Kang. "EDI-Based Deinterlacing Using Edge Patterns", IEEE International Conference on Image Processing, vol 2 p 1018-1021 (2005).