La correction temporelle du bruit de moustique

Claire Mantel, Patricia Ladret, Thomas Kunlin

▶ **To cite this version:**

Claire Mantel, Patricia Ladret, Thomas Kunlin. La correction temporelle du bruit de moustique. MajecSTIC 2009 - MAnifestation des JEunes Chercheurs en Sciences et Technologies de l'Information et de la Communication, Nov 2009, Avignon, France. hal-00449365

HAL Id: hal-00449365
https://hal.science/hal-00449365

Submitted on 21 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MajecSTIC 2009
Avignon, France, du 16 au 18 novembre 2009

La correction temporelle du bruit de moustique

Claire Mantel[1,2], Patricia Ladret[1] et Thomas Kunlin[2]

1 : GIPSA-Lab - Department of Signal and Images, Grenoble Institute of Technology - UMR CNRS 5216, 961 rue de la Houille Blanche, 38402 Grenoble - France.
2 : ST Microelectronics S.A., 12 Rue Jules Horowitz B.P. 217, Grenoble - France.

Contact : claire.mantel@gipsa-lab.inpg.fr

Résumé
Actuellement, l'utilisation de traitements temporels pour améliorer la qualité est rare, y compris pour les vidéos. Ceci est en grande partie lié à l'espace mémoire et au temps de calcul qu'ils nécessitent. Dans cet article, nous traitons d'un bruit de compression dont l'aspect temporel est primordial : le bruit de moustique. Sa nature à la fois spatiale et temporelle est détaillée dans une première partie, puis les différents traitements disponibles dans la littérature sont exposés. Un correcteur spatio-temporel et indépendant de la compression est ensuite proposé. Enfin, l'intérêt d'un traitement spatio-temporel par rapport à un spatial est montré à travers différents tests d'évaluation.

Abstract
Due to the required memory space and computing time, temporal processing is rarely used even in video processing. However, when the temporal aspect of a compression noise is major, it cannot be treated spatially only. The following article presents an example of such a noise : the mosquito noise. The equally spatial and temporal nature of the mosquito artifact is first explained through its definition. It is followed by the different ways already used to treat it. A spatio-temporal and compression-independent mosquito noise corrector is then introduced. Finally, the interest of a temporal corrector over purely spatial ones is shown through tests results.

Mots-clés : Correction, Bruit de moustique, Compression vidéo, Filtrage temporel

Keywords: Mosquito Noise, Video, Correction, Temporal Filtering, Temporal Artifact

---

1. Introduction

A cause de la complexité que cela impliquerait, les vidéos décompressées sont rarement débruitées par des procédés temporels. Pourtant Anderson et al. ont montré dans [2] que le Système Visuel Humain (SVH) ne voit pas les images d'une vidéo pour ce qu'elles sont : une succession discontinue d'images statiques, mais que les informations temporelles qu'elles contiennent sont perçues comme des informations réelles (c'est-à-dire continues). En d'autres termes, les mouvements présents dans une vidéo sont traitées de la même façon qu'un mouvement naturel. Il apparaît donc clairement qu'une vidéo est plus qu'une simple suite d'images statiques et aborder cet aspect temporel revient donc à émuler le SVH.
Les variations temporelles d'une séquence sont prises en compte depuis les tous débuts de la compression vidéo : dès 1984 le codec H.120, premier standard de compression, utilise la redondance temporelle d'une vidéo pour gagner de la bande-passante. En effet il détecte les zones sans mouvement et ne les encode que dans la première image où elles apparaissent. Depuis 1988 et le codec MPEG1, tous les standards compensent en mouvement les zones de l'image courante pour diminuer au maximum la quantité d'information à coder.
La compensation de mouvement et l'utilisation d'une transformation basée sur la TCD sont communs aux normes de compression MPEG2, MPEG4 et H.264. Dans la mesure où ces principes de base de la compression vidéo n'ont pas changé depuis un moment (de 1995 à nos jours), les types de dégradations résultant de la compression n'ont pas non plus évolué. Toutefois l'essor des écrans LCD ainsi que des

sources et des moniteurs Haute Définition modifie à la fois notre perception et notre atttente vis-à-vis de la qualité vidéo. Plus précisément, c'est le degré de gêne respectif des artéfacts les uns par rapport aux autres qui a complètement changé. Par exemple alors que la réponse passe-bas des écrans cathodiques atténuait plutôt les bruits hautes fréquences, l'augmentation de la taille des écrans et l'acuité des grilles de pixels des écrans LCD les accentuent au contraire.
Cet article traite d'un artéfact de compression de ce type : le bruit de moustique (BM). Alors que d'autres défauts de nature purement spatiale (comme l'effet de block ou le ringing) ont fait l'objet de nombreuses études approfondies, ce dernier est peu adressé dans la littérature. Ce bruit illustre parfaitement l'intérêt du filtrage temporel car c'est sa variation au cours du temps qui est surtout visible. En effet, il n'est pas si gênant sur une image statique. Notre objectif était donc de réaliser un correcteur qui tienne compte de l'aspect temporel de ce défaut : il comporte donc une partie spatiale et une partie temporelle. Pour avoir le plus large champ d'action possible, nous avons choisi une approche totalement indépendante du type et des paramètres de compression.
La première partie de cet article définit la nature à la fois spatiale et temporelle du bruit de moustique, expliquant ainsi l'intérêt de le traiter temporellement. Dans un second temps, les différentes approches documentées dans la littérature pour corriger cet artéfact seront présentées. Ensuite nous décrirons notre Filtre Inverse aux Variations Temporel avant de finir par une comparaison avec d'autres algorithmes.

2. Le bruit de moustique : une définition complexe

La première difficulté dans l'étude du bruit de moustique est de cerner le sujet car il existe autant de définitions de ce défaut que d'articles sur le sujet. Dans la recommandation de l'ITU qui pose les principes d'un système d'évaluation de la qualité vidéo [8], le BM est défini de la façon suivante : *"Forme de distortion présentant une activité des contours parfois associée au mouvement, caractérisée par des artéfacts en mouvements et/ou un schéma de bruit tacheté se superposant aux objets (ressemblant à un moustique volant autour de la tête et des épaules d'une personne)."* et l' "activité des contours" comme une *"Distortion concentrée près des bords des objets, plus précisément caractérisée par ses caractéristiques spatiales et temporelles."*. Alors que ces définitions devraient faire référence, elles sont malheureusement trop vagues pour être suffisantes.

2.1. Les similitudes avec d'autres artéfacts

La définition plutôt générale de l'ITU réunit de nombreuses caractéristiques qui s'appliquent également à d'autres sortes d'artéfacts. Cela fait ressortir une des difficultés de la spécification du bruit de moustique : non seulement il est délicat à définir mais en plus il est facile de le confondre avec d'autres sortes de distorsion. Il peut par exemple être assimilé à du ringing car tous deux sont présents près des bords d'objets. Cela pousse certains auteurs comme Segall et al. dans [9] à intégrer les variations temporelles du ringing au bruit de moustique, alors que d'autres comme Coudoux et al. dans [4] ne font clairement pas de différences entre les deux défauts. Il n'existe en réalité pas de définition du ringing dans la documentation de l'ITU [8]. Cet artéfact est lui purement spatial et possède effectivement des ressemblances avec le BM : ils apparaissent tous deux dans des zones uniformes proches des frontières d'objets.
Un autre obstacle pour étudier le BM est que son origine n'est pas clairement déterminée. La quantification étant la seule source de perte d'information lors de la compression, elle est forcément liée au BM ; mais ceci est également vrai pour tous les autres bruits de compression.

2.2. Les aspects spatiaux

La quasi-totalité des articles sur le BM le présente comme proche des contours d'objets. La majorité des auteurs ajoutent que les zones contenant du BM doivent être voisines d'une région d'activité importante. Certains, comme Coudoux et al. dans [4], précisent encore que le BM apparaît en général aux frontières de zones uniformes ou quasi-uniformes (dites "'plates"). D'autres spécifient également que les blocs présentant du BM doivent englober à la fois une partie plate et une texturée (c.f. Westen et al. dans [10]). Enfin DelCorso et al. affirment dans [6] que le BM est situé à proximité des objets en mouvement.
Le second trait majeur du BM est sa très faible amplitude. L'ensemble des articles concourt sur ce point : les erreurs de niveau de gris dues au BM oscillent faiblement autour de la valeur exacte (c.f. [4, 6, 1]). Ceci explique notamment pourquoi ce bruit n'est pas visible dans les régions riches en détails ou texturées.
En résumé, malgré les divergences qui apparaissent dans la littérature quant à la localisation du BM, une propriété globalement reconnue est la faiblesse du bruit *par rapport* à la variation de niveau de gris

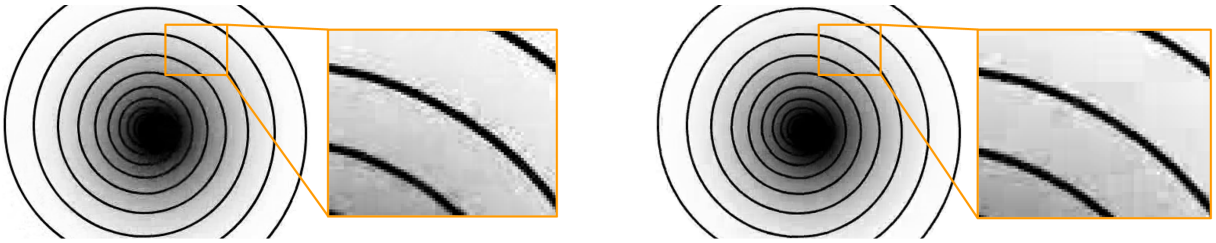

FIGURE 1 – Le bruit de moustique sur une vidéo de spirale en MPEG2 (gauche) et H-264 (droite)

dans l'image (contour ou texture) entourant le BM.

2.3. Les aspects temporels

Il existe deux façons d'aborder les caractéristiques temporelles du BM. La première, l'approche perceptuelle, part du constat que le bruit indispose à cause de sa variation au cours du temps, comme il est dit dans [9]. En effet, ses oscillations ne participent pas d'un mouvement consistant et leur fréquence est identique à celle de la vidéo donc le SVH ne peut pas se stabiliser sur une zone de BM. C'est réellement ce que Coudoux appelle la "discontinuité le long de l'axe temporel" dans [4] qui est perçue comme une gêne. Comme la vision périphérique est extrêmement sensible au mouvement, ce bruit est remarqué n'importe où dans l'image et pas seulement dans les zones saillantes.
Le second axe d'étude est son lien avec l'aspect temporel de la compression. Dans la plupart des normes de compression, des blocs sont codés en mode inter (ou prédictif). Cela signifie qu'ils sont compensés en mouvement par rapport à une image de référence. Seules les différences entre le bloc courant et sa version compensée sont ensuite codées. C'est ce processus que l'on désigne par "aspect temporel". L'ensemble des articles sur le sujet s'accordent pour dire que le codage des résidus après la compensation de mouvement amplifie le BM. Toutefois, les avis sont divisés sur le lien avec le mouvement présent dans une vidéo : alors que Fenimore affirme dans [7] qu'il peut y avoir du BM dans une séquence immobile, DelCorso et al. insistent sur la nécessité d'un objet mouvant sur un fond fixe. Nous pouvons néanmoins en déduire une corrélation entre la quantité de mouvements dans une séquence et la présence de BM.

2.4. Vidéos de spirales

Dans [7] Fenimore et al. choisissent une approche originale pour présenter leur métrique de BM. Pour résoudre la difficulté de séparer le BM des autres artéfacts afin de le mesurer seul, ils adaptent leurs vidéos et non leur métrique. Ils utilisent des vidéos qui ne font apparaître que du BM à la compression. Ces séquences sont constituées de spirales noires se déplaçant sur un fond blanc.
La réalisation de telles vidéos nous a permis de nous assurer que l'artéfact le plus gênant lors de leur compression en MPEG2 est bien le bruit de moustique (figure 1). Comme les résultats étaient moins évidents pour le H.264, nous avons créé de nouvelles séquences où les spirales effectuent des rotations et des translations. Le compensateur de mouvement d'un encodeur ne gère pas les rotations, il ne peut donc pas estimer correctement ces mouvements. Il apparaît alors clairement qu'en le mettant en difficulté, la quantité de BM augmente fortement (figure 1). Sachant que la modification du mouvement présent dans une vidéo ne peut affecter que les blocs codés en mode inter, ceci prouve le lien existant entre le type d'encodage d'un bloc et le BM.
La meilleure explication de l'origine du BM est celle donnée par Abbas et al. dans [1] : il n'est pas dû à une cause unique mais à la combinaison de plusieurs facteurs. Cette approche explique d'une part la diversité d'interprêtation et clarifie surtout l'ambivalence spatialo-temporelle : l'origine et la perception du BM sont à la fois spatiales et temporelles.

3. État de l'art du débruitage

Cet article traitant de la correction du BM, la revue de littérature ne portera pas sur les pré-traitements (comme [10]) mais sur les méthodes de post-traitements indépendantes de la compression.
La proximité du BM avec les contours signifie concrètement que les algorithmes doivent être suffisem-

ment précis pour corriger sans les dégrader. Il existe pour cela plusieurs stratégies. Dans [4], Coudoux et al. marquent comme pixel appartenant à un contour ceux ayant un gradient supérieur à 80 puis utilisent ces marques pour classer chaque bloc 8x8 en bloc de contour ou non. Ils filtrent ensuite dans les "blocs de contours" les pixels qui ne sont pas marqués comme pixels de bords avec un filtre de Wiener en 2-D. Nous partageons l'avis d'Abbas et Karam dans [1] que cela pose deux problèmes : l'hypothèse de non-corrélation du filtre de Wiener n'est pas respectée car le bruit dépend de l'image (c.f. partie 2.2) et un pixel peut être filtré avec des pixels appartenant à d'autres objets, floutant ainsi les frontières.
Dans [1] Abbas et al. utilisent un détecteur de Canny pour localiser précisément les contours puis, filtrent l'ensemble de l'image sauf les pixels de contour. Le choix d'un filtre-ϵ est en phase avec les caractéristiques spatiales du BM : il s'agit en effet d'un simple filtre moyenneur dont le support est restreint aux pixels distincts d'au plus ϵ du pixel courant. Leur REF (Recursive Epsilon Filter) utilise une version récursive du filtre-ϵ appliqué pour corriger le BM en utilisant l'écart-type d'un bloc comme approximation de celui du bruit. De même que Coudoux et al., les pixels de contours ne sont pas filtrés mais peuvent servir à filtrer des pixels appartenant à un objet différent car l'écart-type peut dans certains cas être suffisemment élevé pour inclure de tels pixels dans le support du filtre. Par exemple lorsque la limite entre deux objets coupe un bloc en deux ou lorsque les contours ne sont pas nets. Il en resulte du flou.
Comme il a été dit section 2.3, DelCorso et al. focalisent dans [6] leur correcteur sur les contours mobiles : leur traitement est appliqué sur l'ensemble de l'image et détecte les mouvements mais pas les contours. Ils font cependant un choix très surprenant : afin de ne pas altérer les zones en mouvement, ils ne les filtrent pas. Ils motivent cette décision par la visibilité plus grande du BM sur le fond à côté des objets mouvants que sur les objets eux-même. Leur partitionnement de l'image en blocs 8x8 leur fait dans tous les cas manquer les pixels potentiellement bruités contenus dans des blocs mobiles. Ils se fondent sur la faible amplitude du BM pour affirmer que corriger la valeur moyenne de chaque bloc 8x8 est suffisant. Ils appliquent tout d'abord un filtre médian sur les composantes continues des blocs courant, supérieur et inférieur. Le résultat est filtré encore une fois par un médian avec les coefficients DC du bloc équivalent dans les images précédente et suivante.
Au final, les caractéristiques temporelles du BM sont soit complètement ignorées, comme dans [1] et [4], ou prises en compte uniquement pour les zones immobiles dans [6].

4. Le Filtre Inverse aux Variations Temporel

Le correcteur présenté ici, le Filtre Inverse aux Variations Temporel (T-VIF), adresse les trois aspects de la définition du BM. Sa structure est celle d'un filtre spatial adapté aux bruits de faible amplitude, le Filtre Inverse aux Variations (FIV), décrit section 4.1. Le support de ce filtre est ensuite sélectionné selon les autres caractéristiques du BM : sa proximité avec les bords d'objets (c.f. Section 4.2) et sa variation temporelle (c.f. Section 4.3). La figure 2 schématise la structure globale du T-VIF. Enfin, une version temporelle rapide du T-VIF, particulièrement adaptée au BM, est expliquée paragraphe 4.4.

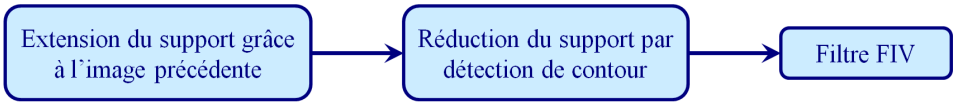

FIGURE 2 – La structure du filtre T-VIF.

4.1. Faibles variations de niveau de gris

Le Filtre Inverse aux Variations est un filtre 1-D créé par Frédérique Crété-Roffet dans [5] pour corriger le ringing. Dans ce cadre il est prévu pour filtrer verticalement puis horizontalement. Nous allons décrire ici le procédé vertical de filtrage. Le concept d'origine répond à l'un des problèmes majeurs de la correction de ringing : faire efficacement la distinction entre le bruit et un contour. L'amplitude du bruit dépendant directement de la force du contour, la dynamique d'un bruit de ringing dû à un contour fort peut être plus élevée que celle d'un contour faible présent dans la même image. La première étape du

FIV est donc d'adapter la force du filtre à celle du bord. Ce filtre étant dédié à des vidéos encodées en MPEG2, la normalisation $Norm$ se fait selon les formules suivantes, pour (i,j) variant de 1 à 8, c'est-à-dire parcourant l'intégralité d'un bloc 8x8 dans l'image courante Im :

$$D_{i,j} = |Im_{i,j+1} - Im_{i,j}|\ ,\ \ Max = \max(D_{i,j})\ \ et\ \ Min = \min(D_{i,j})$$

$$Si\ Max \neq Min,\ \ Norm_{i,j} = Seuil_{max} \cdot \frac{D_{i,j} - Min}{Max - Min}$$

Si Max et Min sont égaux, les variations sont constantes et le bruit ne peut être distingué des bordures donc nous ne filtrerons pas. $Seuil_{max}$ représente le nombre de niveaux possibles.
Le VIF effectue ensuite un filtrage adaptatif. L'idée de base est que plus un pixel est distinct du pixel courant, plus il est probable qu'ils appartiennent à des objets différents. Le poids donné à chaque pixel dépend donc de l'inverse de sa différence avec le pixel courant. Les coefficients $W_{i,j\pm1}$ des pixels (i,j-1) et (i,j+1), qui sont utilisés pour filtrer verticalement le pixel (i,j), sont calculés selon les équations subséquentes, où $Filter_Norm$ est la norme $\mathcal{L}_1$ du filtre :

$$W_{i,j\pm1} = \frac{1}{Filter_Norm} \cdot \frac{1}{Norm_{i,j\pm1} + 1}$$

Ainsi qu'il a été dit section 2.2, l'amplitude du BM est faible comparée à l'amplitude du contour voisin. Le FIV ayant été conçu pour corriger un bruit de faible dynamique par rapport à la variation due à un contour, il est donc également adapté aux caractéristiques spatiales du BM. Le VIF devait néanmoins être adapté à 3 dimensions (2-D et le temporel) de même qu'à d'autres partitions de l'image qu'aux blocs 8x8 du MPEG2. Les gradients ayant déjà été calculé pour la détection de contours, nous avons avantageusement remplacé Max et Min respectivement par le maximum et le minimum des gradients dans un voisinage 15x15 du pixel courant. Le résultat du filtrage $F_{i,j}$ est calculé de la façon suivante :

$$F_{i,j} = \frac{1}{Filter_Norm} \cdot \left(1 + \sum_{(k,l)\in\{-1,1\}^2} \frac{1}{Norm_{i+k,j+1} + 1}\right)$$

4.2. A proximité des contours

Nous avons vu section 2.2 que le BM est présent près des bords d'objet. Le T-VIF comporte donc un détecteur de contour pour deux raisons : premièrement, déterminer pour chaque pixel s'il risque de présenter du BM. C'est le cas s'il est dans le même bloc de quantification qu'un pixel de contour pendant l'encodage, en pratique s'il se trouve dans le voisinage 8x8 d'un pixel de contour ; autrement ce n'est pas nécessaire de le filtrer. Filtrer uniquement les pixels potentiellement bruités permet de gagner du temps et de ne pas dégrader de zones saines n'ayant pas besoin d'être corrigées.
La détection de contour permet ensuite d'adapter le filtre au contenu de l'image. Pour éviter de détériorer ou flouter les contours, nous limitons le support du filtre aux pixels auxquels le détecteur de contour a assigné la même étiquette (contour/pas contours). Le but n'est pas ici de filtrer les bords avec les bords et les objets avec les objets. Cela n'aurait assurément aucun sens car, excepté pour les vidéos synthétiques, il n'existe pas de bord épais d'un pixel autour d'un objet : le "bord" fait partie de l'objet.
Le support du filtre est ainsi réduit pour être certain de filtrer avec des pixels appartenant tous au même objet. Nous avons choisi le détecteur de Canny car il fournit des bords larges d'un pixel. Le filtre étant appliqué sur un voisinage 3x3, deux pixels peuvent faire partie d'objets différents seulement s'ils sont séparés par un pixel de contour. Comme les contours obtenus par le détecteur de Canny sont connexes, cela signifie alors que le pixel central est un pixel de bord. Dans notre algorithme, il sera donc filtré uniquement avec d'autres pixels de bord.
De plus, le FIV garantit aussi une sécurité en cas d'erreur du détecteur de contours : l'influence d'un pixel extérieur inclus dans le support serait minimisée de part sa différence avec le pixel courant.

4.3. Les variations temporelles

La particularité première du T-VIF est son aspect temporel car, comme il a été dit section 3, il existe actuellement un seul correcteur temporel qui il ignore les contours et se restreint aux zones immobiles. Le BM varie d'une image à l'autre autour de la valeur juste. Nous avons donc inclus l'image précédente

dans le support du filtre pour converger vers la valeur exacte. En outre, l'utilisation des pixels précédents dans le temps aide à lisser les discontinuités temporelles. Pour améliorer la fluidité temporelle de la séquence filtrée, le FIV a donc été modifié en filtre spatio-temporel par compensation de mouvement. Pour ce faire, l'estimation de flot optique présentée par Black et Anandan dans [3] a été utilisée. Elle fournit pour chaque pixel un vecteur compensant le mouvement par rapport à une image de référence et permet de créér une image compensée en mouvement. Le voisinage 3x3 de l'image compensée en mouvement du pixel courant est adjoint au support du filtre. Les pixels sont aussi triés selon leur appartenance à un contour ou non avant d'être ajoutés.

4.4. Filtrage temporel rapide

La compensation de mouvement est une opération coûteuse à la fois en terme de ressources et de temps de calcul. Nous avons donc établit une version du T-VIF qui s'adapte au mouvement, nommée Fast-T-VIF. Le fait que le BM soit de faible amplitude par rapport au contour et qu'il varie d'image en image a permis une approche simple mais efficace. L'idée est de trier les pixels de l'image de référence d'après leur appartenance au même objet et non selon leur mobilité. Pour s'en assurer le maximum des gradients dans un voisinage 15x15, assimilé à un gradient de contour, sert de nouveau d'étalon de mesure pour évaluer les autres pixels.
Une image de différence entre l'image courante et l'image de référence est tout d'abord calculée (figure 3). Les pixels avoisinnants de l'image de référence sont adjoints au support du filtre si cette différence est plus petite que le gradient maximal. De même que pour les méthodes précédentes, les pixels sont triés selon leur appartenance ou non à un contour avant d'être ajoutés.

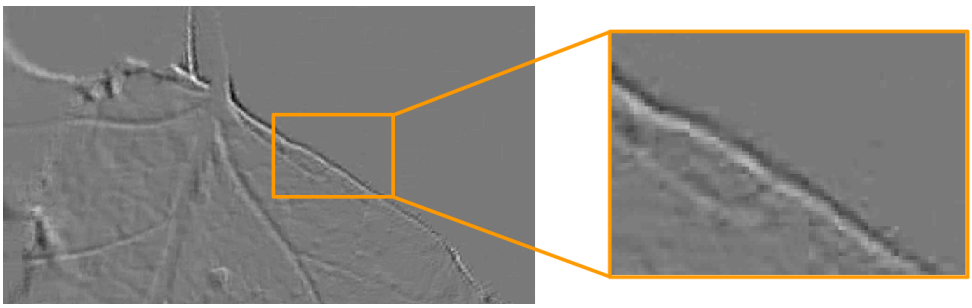

FIGURE 3 – Différences entre les images 16 et 17 de la séquence Aspen

5. Tests

5.1. Vidéos utilisées

Nous avons comparé les différentes versions du T-VIF à l'algorithme REF. Les séquences utilisées pour les tests sont Aspen, InToTree (du site du VQEG www.its.bldrdoc.gov/vqeg/) et Life (tirée du site www.hdgreetings.com). Les deux premières sont des vidéos naturelles avec respectivement beaucoup et peu de mouvement, la dernière est en image de synthèse avec une quantité de mouvement moyenne. Les vidéos de référence sont de qualité "parfaite", elles ont été réduites à une résolution de 720p puis encodée avec l'encodeur X264 (encodeur H264 en libre accès) qui permet de désactiver le filtre de deblocking. Les débits étaient respectivement de 1Mb/s pour Aspen et 600Kb/s pour les autres. Les séquences ont ensuite été réduites à une résolution de 500x500 pour l'analyse.

5.2. Analyse spatiale des résultats

L'analyse spatiale est faite en calculant la SSIM et le PSNR entre les images originales et les corrigées. La courbe de SSIM pour la vidéo Life, figure 4, est clairement en faveur du T-VIF et du Fast-T-VIF.
Les résultats pour Aspen sont légèrement meilleurs pour nos algorithmes que pour le REF. Pour la vidéo InToTree, les valeurs obtenues pour les correcteurs sont toutes du même ordre que celle de la vidéo non

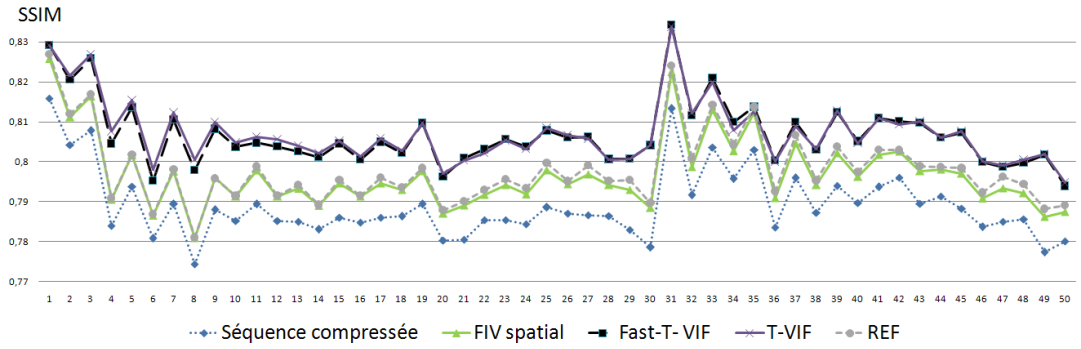

FIGURE 4 – Courbe SSIM pour la séquence Life.

corrigée bien qu'elles semblent de meilleur qualité. Les corrections sont évidentes sur l'images de la figure 5. On peut également y voir que les versions temporelles du FIV suppriment autant de bruit que le REF tout en conservant une meilleur dynamique des niveaux de gris.

(a) Vidéo compressée

(b) REF

(c) T-VIF

(d) Fast-T-VIF

FIGURE 5 – Image 19 de la séquence Aspen.

5.3. Analyse temporelle des résultats

L'ajout majeur du T-VIF est sa partie temporelle dont le but est d'améliorer la fluidité de la vidéo. Pour étudier ceci, nous avons appliqué le PSNR et la SSIM aux différences entre deux images successives. L'idée est ici d'étudier l'effet des filtres sur les variations temporelles.
Pour les vidéos Aspen et Life, le T-VIF et le F-T-VIF améliorent nettement la continuité temporelle. Pour InToTree l'amélioration est trop légère pour être significative. La courbe de PSNR de la vidéo Aspen est montrée figure 6.

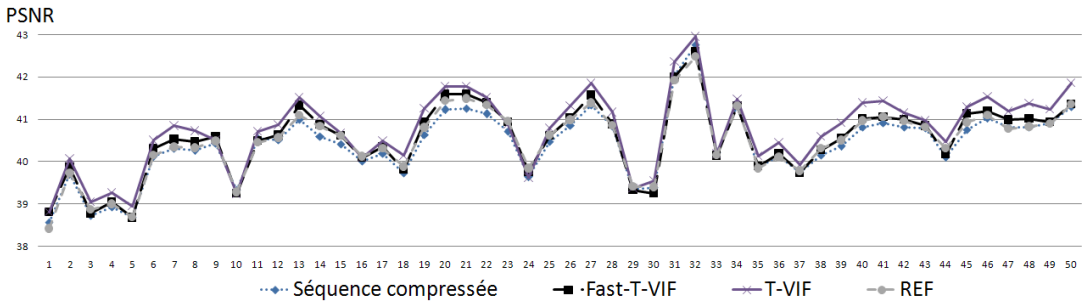

FIGURE 6 – Courbe de PSNR sur les différences entre images successives pour Aspen.

6. Conclusion

Sur le plan spatial, d'après les métriques de fidélité, notre méthode fournit une amélioration significative sur l'une des vidéos testées, légère pour une autre et aucune pour la troisième. Sur le plan temporel, elle améliore légèrement la continuité. L'analyse devra être étendue à une base de donnée plus large pour dégager des tendances plus marquées.
Mais la question principale pour estimer l'efficacité d'un correcteur demeure la méthode à utiliser. En effet, la localisation précise et la faible amplitude de ce bruit tendent à neutraliser les métriques globales et à compliquer l'évaluation temporelle. Afin de valider le T-VIF malgré ces difficultés, une expérience subjective est prévue pour déterminer la qualité visuelle des vidéos corrigées. Pour passer outre ces problèmes, la prochaine étape est d'élaborer une métrique dédiée au bruit de moustique. Nous avons commencé à travailler sur une telle métrique prenant en compte les aspects perceptifs et temporels du bruit de moustique.

Bibliographie

1. H. Abbas et L.J. Karam. Suppression of mosquito noise by recursive epsilon-filters. *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 1 :I–773–I–776, 15-20 April 2007.
2. J. Anderson et B. Anderson. The myth of persistence of vision revisited. *Journal of Film and Video*, 45(1) :3–12, 1993.
3. M. J. Black et P. Anandan. The robust estimation of multiple motions : Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1) :75–104, January 1996.
4. F.X. Coudoux, M.G. Gazalet, et P. Corlay. A post-processor for reducing temporal busyness in low-bit-rate video applications. *Signal Processing : Image Communication*, 18 :455–463, July 2003.
5. F. Crété-Roffet. *Estimer, mesurer et corriger les artefacts de compression pour la télévision numérique*. Thèse de doctorat, GIPSA-Lab - Université Joseph Fourier, 2007.
6. S. Del Corso, C. Miro, et J. Jung. Mnr : a novel approach to correct mpeg temporal distortions. *Consumer Electronics, IEEE Transactions on*, 49(1) :229–236, Feb. 2003.
7. C. P. Fenimore, J. M. Libert, et P. Roitman. Mosquito noise in mpeg-compressed video : test patterns and metrics. *Human Vision and Electronic Imaging V*, 3959(1) :604–612, 2000.
8. Recommendation P.930 ITU-T. Principles of a reference impairment system for video. Technical report, August 1996.
9. C.A. Segall et A.K. Katsaggelos. Pre- and post-processing algorithms for compressed video enhancement. *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, 2 :1369–1373, 2000.
10. S.J.P. Westen, R.L. Lagendijk, et J. Biemond. Adaptive spatial noise shaping for dct based image compression. *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, 4 :2124–2127, 7-10 May 1996.