

Tonal Stabilization of Video (*Zeev Farbman, Dani Lischinski*)

MVA 2014 - Imagerie numérique

Alexandre THIS

20 janvier 2015

Table des matières

1	Introduction	2
1.1	Problématique	2
1.2	Motivations et travaux connexes	2
2	Méthode proposée	3
2.1	Sélection d'images de références	3
2.2	Lissage avec préservation des bordures	3
2.3	Mise en correspondances de pixels	3
2.4	Carte d'ajustement	4
2.5	Ré-échantillonnage et alignement tonal	5
3	Implémentation et tests	6
3.1	Validation de la méthode de Nystrom pour l'évaluation de W	6
3.2	Tests de stabilisation tonale	7
4	Conclusion	10

1 Introduction

1.1 Problématique

La problématique de la stabilisation tonale est une problématique qui a été peu étudiée. En effet, les appareils personnels de prise de vue se sont démocratisés il y a peu de temps, et l'introduction dans ces appareils de système de balance de couleur et de changement d'exposition automatique fait que ce problème est très récent. La figure 1 montre deux images à moins d'une seconde d'intervalle. On peut remarquer une variation importante de la tonalité de l'image alors que la scène a très peu changé.



FIGURE 1 – Instabilité tonale

Dans les appareils personnels, des ajustements temps réel de l'exposition des vidéos et de balance des blancs sont fait et, d'après les auteurs de l'article, désactiver l'automatisation de ces réglages d'exposition ne peut être une solution à cause de la dynamique des scènes qui peut être très variable notamment.

La récente démocratisation de ces caméras, et des moyens de diffusion de vidéos, apporte un intérêt certain à la problématique de la correction de ces fluctuation tonales. Cependant, chaque appareil peut avoir ses propres algorithmes de corrections, et les paramètres tonaux de ces caméras varient aussi. Ainsi, les auteurs de l'article ne veulent pas introduire dans leur solution un modèle avec un fort a priori.

1.2 Motivations et travaux connexes

Les auteurs souhaitent donc réaliser un algorithme (semi) automatique de correction de ces fluctuations tonales. Les auteurs notent que l'on pourrait utiliser des méthodes permettant de modéliser la réponse tonale des caméras afin de l'inverser, cependant les méthodes existantes nécessitent d'utiliser des images recalées pour lesquelles la seule variation est l'exposition.

Un second type de méthode qui pourrait être utilisé consiste à effectuer un transfert de couleur global entre une image ancre et l'image que l'on souhaite stabiliser. Cependant, les mouvements de la caméra et les objets de la scène perturbent les statistiques de l'image, et rendent donc ces méthodes peu efficaces. Les méthodes locales de transfert de couleur, elles, sont soumise à la problématique de la mise en correspondance de zone dans l'image, problème non trivial.

2 Méthode proposée

La méthode proposée par l'article [1] se décompose en un certain nombre d'étapes qui seront décrites dans ce chapitre.

2.1 Sélection d'images de références

Afin de réaliser la stabilisation tonale de la séquence d'image, l'utilisateur doit choisir une (ou des) image de référence. Ces images sont les images que l'utilisateur choisit comme référence de tonalité. Les auteurs proposent de choisir des images de références par paires, ces paires délimitant des parties de la séquence dans lesquelles une importante variation de la tonalité a lieu.

2.2 Lissage avec préservation des bordures

Dans un second temps, un lissage des images avec un filtre préservant les bordures est appliqué. Cela a pour impact d'atténuer les variations qui ne proviennent pas des paramètres de tonalité de la caméra.

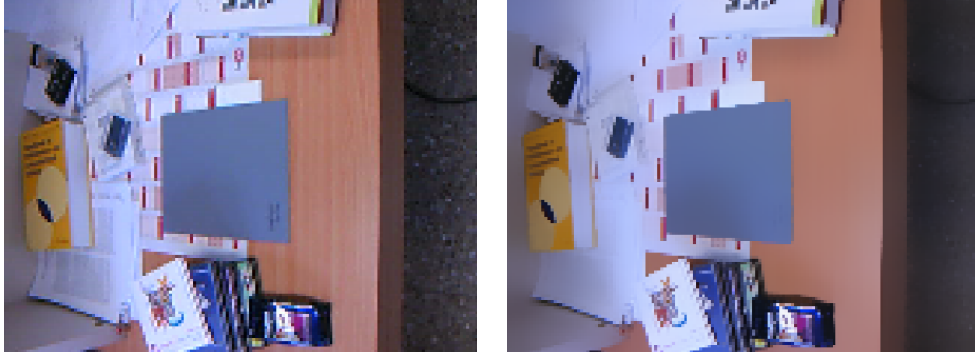


FIGURE 2 – Left : Image d'origine sous échantillonnée, Right : Filtre bilatéral avec les paramètres recommandés

2.3 Mise en correspondances de pixels

Comme, dans une vidéo, les images sont cohérentes spatialement et temporellement, une grande partie des pixels de l'image peuvent être mis en correspondances. Afin de réaliser cela, les auteurs proposent d'utiliser la luminance de deux images consécutives afin de sélectionner ces pixels.

A partir des composantes R,G,B d'une image, la luminance peut s'écrire sous la forme $L = 0,299R + 0,587G + 0,114B$. Les auteurs proposent de conserver dans l'ensemble des correspondances (*robust set*) les pixels dont la luminance a peu changé entre deux images successives. Cette condition est exprimée par l'équation 1

$$R_{i/i+1} = \{x \text{ s.t. } |(L_i(x) - \mu(L_i)) - (L_{i+1}(x) - \mu(L_{i+1}))| < 0.05\} \quad (1)$$

On peut observer sur l'image 3 le *robust set* créé à partir de deux images successives.

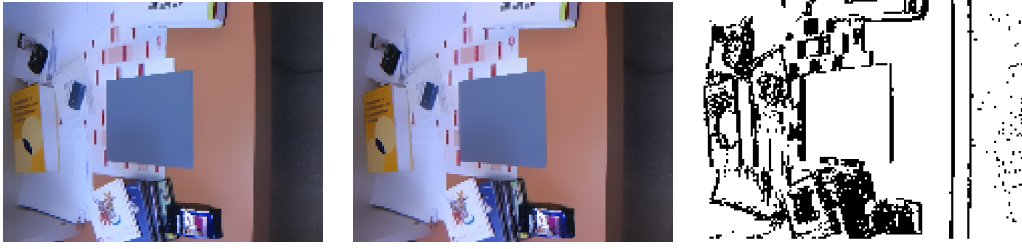


FIGURE 3 – Deux images successives, et le *robust set* qui en découle

2.4 Carte d’ajustement

Initialisation de la carte

La carte d’ajustement est initialisée pour tout les pixels appartenant au *robust set* précédemment calculé. Cette initialisation consiste à ajouter à la carte d’ajustement de l’image précédente la différence de couleur entre les deux images. De cette manière, un décalage graduel de la tonalité est compensé.

$$\hat{A}_{i+1}(x) = \begin{cases} A_i(x) + (f_i(x) - f_{i+1}(x)) & \text{for each } x \in R_{i/i+1} \\ 0 & \text{else} \end{cases} \quad (2)$$

Complétion de la carte

Pour les pixels qui ne font pas partie du *robust set*, les auteurs partent du principe que les fluctuations tonales sont globales, c’est à dire que les modifications sont les même pour toute l’image. Ainsi, les auteurs proposent d’utiliser les ajustements présents dans l’initialisation de la carte pour interpoler les valeurs manquantes. Afin de réaliser cela, les valeurs manquantes sont une moyenne pondérée des ajustements déjà présents.

$$A_{i+1}(x) = \frac{\sum_{r=0}^N w(x, x_r) \hat{A}_{i+1}(x_r)}{\sum_{r=0}^N w(x, x_r) \chi_{\hat{A}_{i+1}(x_r)}} \quad (3)$$

$\chi_{A_{i+1}} \equiv R_{i/i+1}$, et $w(a, b) = \exp(-\frac{\|c(a) - c(b)\|^2}{2\sigma_c^2})$, $c(x)$ étant la couleur du pixel x dans l’espace couleur CIE Lab. Le passage dans l’espace couleur CIE Lab permet de considérer des distance entre couleurs qui ont un sens en regard de la perception des couleurs.

On peut réécrire l’équation 3 sous la forme matricielle 4 :

$$A_{i+1}(x) = \frac{W \hat{A}_{i+1}}{W \hat{\chi}_{A_{i+1}}} \quad (4)$$

On note que la matrice W est de taille $N \times N$. Elle est donc très coûteuse à calculer. De plus, l’évaluation de 4 est coûteuse. Les auteurs proposent d’utiliser la méthode Nyström [2] afin de calculer une approximation de W : $\tilde{W} = U \tilde{D} U^T$ avec \tilde{D} une matrice diagonale contenant les plus grandes valeurs propres de W .

Ainsi on a :

$$A_{i+1}(x) \approx \frac{U\tilde{D}U^T \hat{A}_{i+1}}{U\tilde{D}U^T \hat{\chi}_{A_{i+1}}} \quad (5)$$

2.5 Ré-échantillonnage et alignement tonal

Enfin, l'alignement tonal consiste à appliquer la carte d'ajustement tonal à la séquence d'image. Dans le cas où une paire d'images de référence a été choisie, on applique les cartes d'ajustements de chaque image de référence sur la séquence d'images, et on fusionne les deux séquences d'image résultantes selon la distance à l'image de référence.

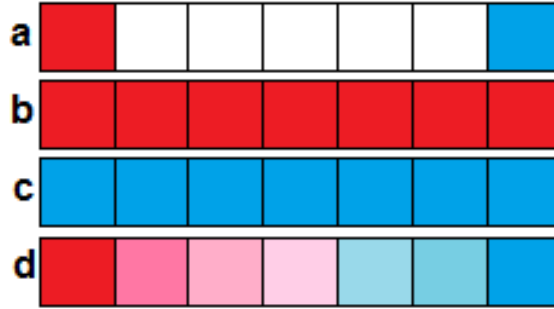


FIGURE 4 – a. Séquence de départ avec ancres, b. Séquence alignée sur l'ancre rouge, c. Séquence alignée sur l'ancre bleue, d. Séquences b et c fusionnées

3 Implémentation et tests

Le but de ce travail était tout d'abord d'implémenter la méthode décrite dans l'article [1], puis dans un second temps d'évaluer la méthode sur un certain nombre de jeux d'exemples.

3.1 Validation de la méthode de Nystrom pour l'évaluation de W

Afin d'implémenter la méthode décrite dans [1], il a tout d'abord fallu implémenter la méthode de Nystrom pour estimer une approximation de la matrice de similarité W .

Afin de valider l'algorithme, une image de petite taille en niveau de gris synthétique a été créée (voir image 5).

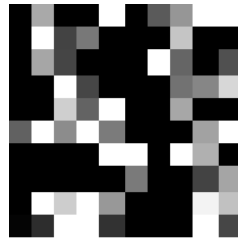


FIGURE 5 – Image synthétique en niveau de gris de taille 10×10

Le but de la validation est de comparer la véritable matrice de similarité W avec l'approximation obtenue avec la méthode de Nystrom. La figure 7 compare la matrice de similarité réelle avec une approximation à 1, 3, 5, 7, et 9 vecteurs propres. On peut observer que l'on arrive à une approximation de W_{reel} sans avoir besoin de beaucoup de valeur et vecteurs propres. La figure 6 représente l'évolution de la RMSE selon le nombre de valeurs propres choisies.

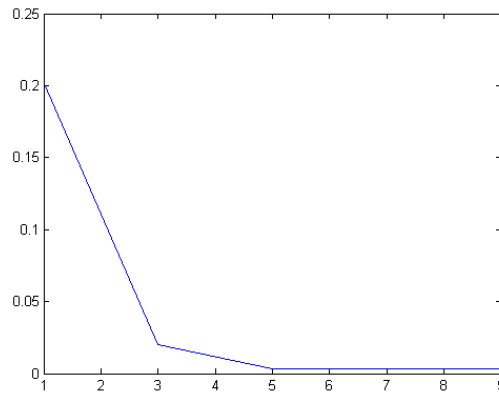


FIGURE 6 – RMSE en fonction du nombre de valeur propres pour l'estimation de W

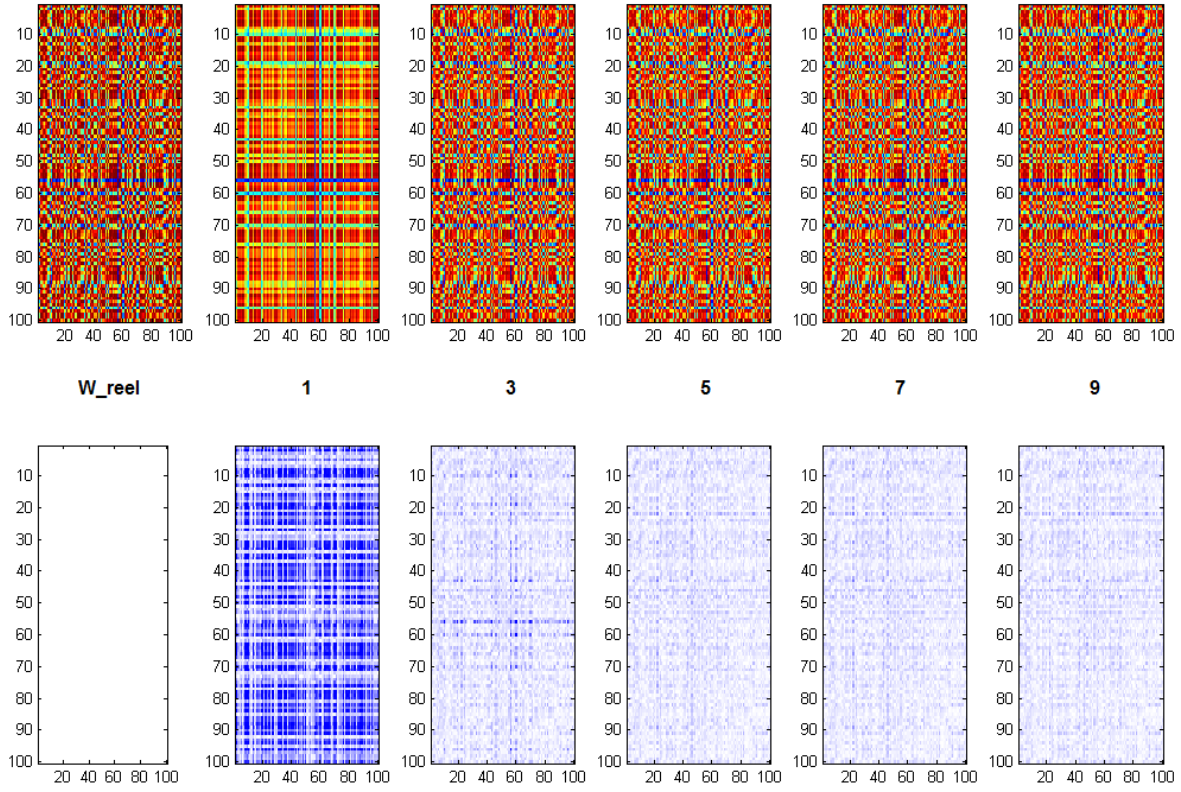


FIGURE 7 – Haut : Carte de similarité W , et \tilde{W}_n (approximation de W à la $n^{\text{ème}}$ valeur propre). Bas : Erreur par rapport à W_{reel}

3.2 Tests de stabilisation tonale

Les séquences vidéos présentées dans l'article [1] sont disponible en téléchargement. Ces données ont donc été utilisées afin de réaliser les tests.

Tout d'abord, la séquence initiale *graycard.avi* a été alignée afin de conserver une tonalité froide (figure 8), puis sur une tonalité chaude.

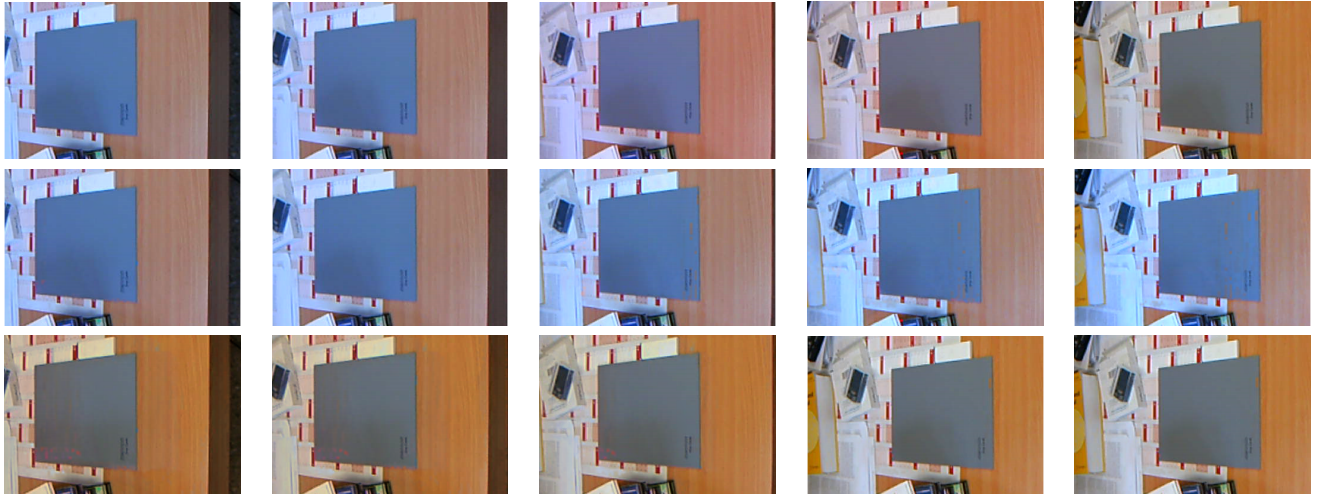


FIGURE 8 – Haut : Séquence initiale, Milieu : Séquence alignée sur la première image, Bas : Séquence alignée sur la dernière image. Les images sont prises à 10 images d'intervalle par rapport à la séquence complète.

On peut voir sur la figure 8 que l'on a effectivement réussi à stabiliser les fluctuations tonales de la vidéo autour. On note tout de même quelques artéfacts de couleur sur l'image.

On peut cependant se poser la question de l'impact de l'algorithme de stabilisation tonale sur une séquence pendant laquelle le changement d'illumination ne serait pas dû à une correction de l'appareil photo mais à une baisse ou une hausse de luminosité naturelle de la scène. L'image 9 montre le résultat d'une séquence dans laquelle la lumière a été baissée progressivement, et la stabilisation de la séquence sur une image sombre.

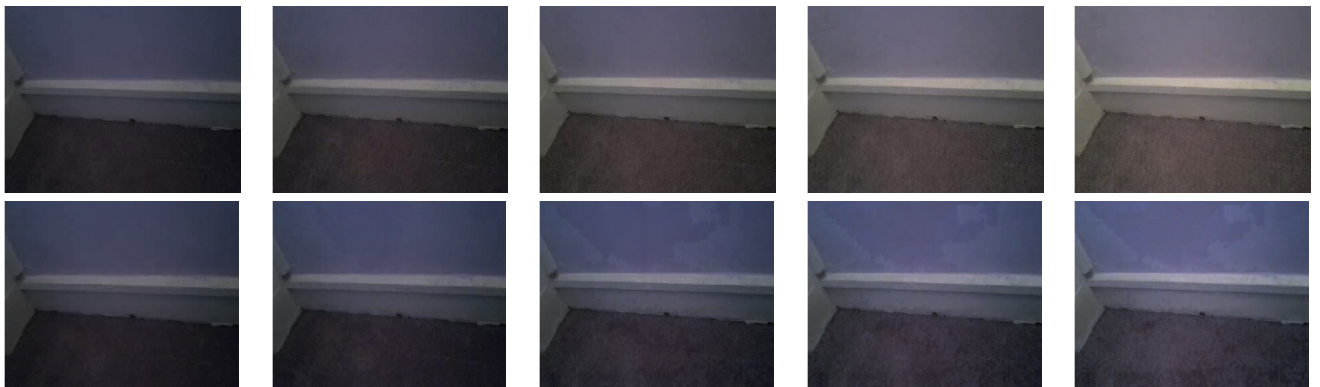


FIGURE 9 – Haut : Séquence initiale, Bas : Séquence alignée sur la première image

On peut tout d'abord noter que la séquence a été stabilisée alors qu'aucune correction n'avait été apporté par l'appareil. D'autre part, on note la présence d'artéfacts sur l'image. Ces artéfacts ont

plusieurs origines. Tout d'abord, les images ont été sous-échantillonnées pour le calcul des cartes d'ajustements et le nombre de pixels est donc très bas. D'autre part, dans des conditions de très basse lumière comme celle-ci, les images ne sont pas nettes.

Enfin, un test a été réalisé afin d'observer le comportement de l'algorithme en regard de la cohérence temporelle de la séquence d'image. Afin de réaliser ce test, deux bouts de séquences ont été concaténées afin de créer une coupure artificielle de la vidéo. On peut remarquer sur l'image 10 que la perte de cohérence temporelle influe énormément sur le résultat de l'algorithme.

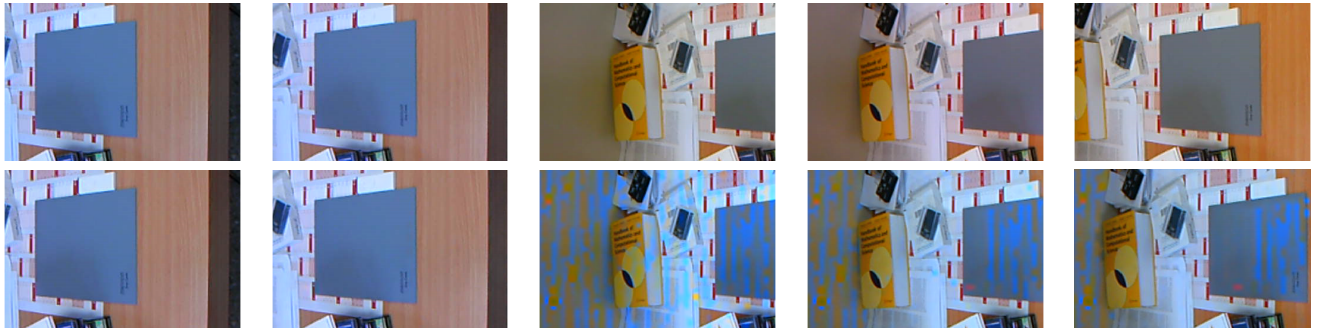


FIGURE 10 – Haut : Séquence initiale, perte de cohérence temporelle, Bas : Séquence alignée sur la première image

4 Conclusion

Dans ce rapport, une présentation succincte de l'algorithme de stabilisation tonale de Farbmán et al. ([1]) a été présentée. Dans un premier temps l'algorithme a été décrit, puis dans un second temps des tests ont été réalisés afin de vérifier la validité de l'implémentation, ainsi que pour mettre en valeur différents exemples d'utilisations.

On a pu remarquer que dans un certain nombre de cas (précisés dans l'article d'origine par l'auteur), l'algorithme ne se comporte pas comme il serait souhaité. En effet, la perte de cohérence temporelle de la vidéo, a donné de très mauvais résultats. D'autre part, les changements naturels d'illumination sur une séquence a, contrairement à ce qui aurait été souhaité, été la source d'une stabilisation tonale.

Malgré cela, cet algorithme présente de très nombreux avantages. Tout d'abord, il permet à l'utilisateur de corriger les séquences avec très peu d'interaction utilisateur. En effet, l'utilisateur a uniquement à choisir les images de références qui seront utilisées par la suite. D'autre part, l'algorithme ne fait aucune hypothèses a priori sur les paramètres de la caméra.

Afin de poursuivre le travail présenté, deux pistes d'améliorations pourraient être suivies. Tout d'abord, on pourrait se pencher sur la question de la sélection automatique des images de références. Cela permettrait à l'algorithme d'être complètement autonome pour réaliser la stabilisation tonale. D'autre part, une détection automatique de perte de cohérence temporelle permettrait d'éviter les résultats observés dans la partie précédente.

Références

- [1] Zeev Farbman and Dani Lischinski. Tonal stabilization of video. In *ACM Transactions on Graphics (TOG)*, volume 30, page 89. ACM, 2011.
- [2] Charless Fowlkes, Serge Belongie, Fan Chung, and Jitendra Malik. Spectral grouping using the nystrom method. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2) :214–225, 2004.