

Agrandissement adapté des GOP (Group Of Pictures) en vidéo

S. FASSINO A. MONTANVERT

Laboratoire LIS – Fédération ELESA
961, rue de la Houille Blanche – BP 46
38402 Saint Martin d'Hères cedex

sylvain.fassino@lis.inpg.fr

annick.montanvert@lis.inpg.fr

Résumé - Cet article traite de l'agrandissement de séquences vidéo dans le but de diminuer les besoins en capacités de stockage et en débits de transmission. Nous définissons un procédé permettant d'agrandir, en résolution spatiale, les images d'une séquence qui peut ensuite être stockée sous forme de "petit format".

Dans un premier temps, nous traitons l'agrandissement d'images par induction, une méthode de régularisation développée pour les images fixes que nous étendons à un facteur d'agrandissement quelconque.

Ensuite, nous expliquons comment les vecteurs de mouvement et les images d'erreurs obtenus par un algorithme de Block-matching sont utilisés avec l'induction pour réaliser un agrandissement de séquences vidéo de qualité.

Abstract - This article deals with the video enlargement to decrease storage capacity and transmission rate. We define a process to enlarge images in size. Thus, the sequence can be stored in a "small format".

Firstly, we deal with enlargement of images by induction, a regularization method developed on still images, extended to any factor.

Next, we explain the idea to enlarge video with movement vectors and errors images given by a block-matching algorithm with induction, to obtain a good quality video sequence.

1 Introduction

Avec l'essor du multimédia, l'utilisation de séquences vidéo est devenue courante. Malheureusement, des contraintes de capacité de stockage ou de débit de transmission apparaissent. Aussi, les normes actuelles combinent de nombreux procédés pour compresser ces vidéos tout en gardant la meilleure qualité possible [6], [8]. Deux cadres d'applications sont possibles : la compression par réduction/agrandissement où la vidéo initiale est réduite avant d'être stockée ou transmise ; elle sera agrandie au moment de la visualisation. Le deuxième cadre d'application est de réaliser un agrandissement de qualité sans séquence initiale de grande taille.

Nous nous plaçons dans le contexte du codage MPEG-2, et nous avons donc travaillé d'une part à l'agrandissement d'images fixes, et d'autre part à la prise en compte de la prédiction/compensation de mouvement. Parmi les nombreuses méthodes qui peuvent être utilisées pour l'agrandissement d'images fixes [3], nous avons spécifié l'agrandissement par induction [2] pour un facteur quelconque (cf. §.2). Ensuite (cf. §.3), nous montrons comment nous utilisons la prédiction/compensation de mouvement pour réaliser l'agrandissement de séquences vidéo [7], [11], [12].

2 Agrandissement d'images fixes

2.1 Principe

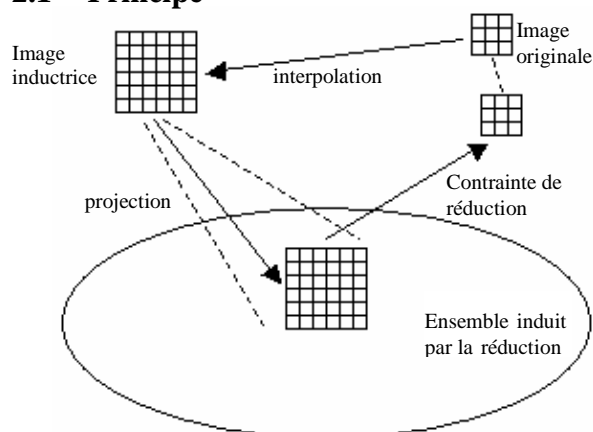


Figure1 – Principe de l'agrandissement par induction.

L'agrandissement par induction [2], est traité comme un problème inverse de la réduction : il s'agit de produire une image agrandie dont la réduction redonne l'image originale (contrainte de réduction). Le schéma précédent (figure 1) illustre ce principe : l'image originale est tout d'abord agrandie par interpolation linéaire pour donner une image appelée « image inductrice ». Cette image est ensuite projetée sur l'ensemble des solutions visuellement admissibles [5], c'est-à-dire l'ensemble des

solutions qui satisfont la contrainte de réduction et qui donnent une image proche de l'image originale. Nous appliquons l'induction (qui est un processus itératif) pour faire converger l'image inductrice, vers l'image qui satisfait au mieux la contrainte de réduction.

De ce fait, cette méthode restaure le contenu fréquentiel de l'image initiale dans la résolution supérieure en cohérence avec son contenu spatial (contours, zones homogènes, etc.).

Initialement, l'agrandissement par induction a été développé pour un facteur de 2 et il utilise un filtre de Burt [1] pour réaliser la contrainte de réduction ; mais ce filtre ne permet pas le passage à une autre valeur du facteur d'agrandissement. Des travaux ont été faits dans ce domaine pour autoriser les facteurs quelconques [10]. Malheureusement, les filtres calculés demandent plusieurs noyaux générateurs et possèdent des degrés de liberté, ce qui rend l'implémentation temps réel difficile. Aussi, nous avons étendu l'agrandissement à un facteur quelconque en utilisant un nouveau filtre obtenu par la discrétisation d'une fonction gaussienne et en prenant en compte des contraintes spatiales.

2.2 La contrainte de réduction (extension à un facteur quelconque)

Nous avons utilisé une gaussienne pour réaliser le filtre de réduction. La gaussienne est la fonction qui assure le meilleur compromis entre localisations spatiale et fréquentielle d'après le principe d'incertitude d'Heisenberg.

Nous avons utilisé la fonction suivante pour définir notre gaussienne :

$$f(x) = A.e^{-\frac{x^2}{2\sigma^2}}$$

L'induction n'améliore pas la qualité de l'agrandissement si nous utilisons un filtre passe-bas classique satisfaisant uniquement des contraintes fréquentielles. Nous proposons d'introduire des contraintes spatiales [1] pour la construction du nouveau filtre. Dans l'approche de Burt, les coefficients du filtre sont calculés uniquement à partir de critères spatiaux (positions spatiales des coefficients du filtre). En plus de la contrainte de séparabilité, Burt a proposé quatre nouvelles contraintes qu'un filtre doit satisfaire : normalisation, symétrie, équicontribution et unimodalité, exprimées ci-dessous pour un facteur deux :

- La contrainte de normalisation impose que toutes les valeurs calculées varient dans le même intervalle que celui des valeurs de l'image originale (conservation de la dynamique de l'image) :

$$\sum_{i=0}^{K-1} \omega(i) = 1$$

- La contrainte de symétrie assure que les éléments situés à une même distance du centre du filtre $h(x,y)$ interviennent avec un poids identique dans le calcul. Ainsi, il n'y a pas de direction privilégiée :

$$\omega(x) = \omega(K-1-x) \quad 0 \leq x < K$$

- La contrainte d'équicontribution, intervient pour que chaque pixel du niveau courant contribue avec le même poids dans la construction du niveau supérieur (Burt avait créé son filtre pour la synthèse de pyramide multi-résolution) :

$$\sum_{i=0}^{K-1} \omega(m+2i) = \frac{1}{2} \quad 0 \leq m \leq 1$$

- Et enfin, la contrainte d'unimodalité impose que le poids d'un élément soit d'autant plus fort qu'il est proche du centre de $h(x,y)$. Si de plus, tous les poids sont positifs, alors les variations de la fonction de niveaux de gris de l'image initiale sont approximativement conservées. Il s'agit d'une condition pour éviter de créer de nouvelles structures (faux contours) lors de la réduction [4], [9]. Cette contrainte s'énonce de la manière suivante :

$$0 \leq \omega(x) \leq \omega(y) \quad x \leq y < \frac{K}{2}$$

La résolution du système formalisé par ces différentes contraintes, avec les poids contrôlés également par la fonction gaussienne, nous a permis d'obtenir les valeurs :

$$A = 0.375 \text{ et } \sigma = 1.1105$$

qui approchent au mieux les contraintes de Burt (en effet, le système posé ne contient pas de solutions exactes).

Les images de la figure 2 montrent les résultats obtenus avec le filtre précédemment calculé. Les résultats sont remarquables au niveau des zones texturées, où l'on peut voir une nette amélioration des détails ainsi que du contraste pour l'image induite



Figure 2 – Image Lena agrandie par interpolation linéaire d'un facteur 1.2 (a) et 1.4 (c), agrandie par induction avec le filtre gaussien d'un facteur 1.2 (b) et 1.4 (d).

3 Agrandissement vidéo

3.1 Principe

Pour fournir une vidéo de plus grande taille, d'une part on ne veut pas avoir à reconstruire la séquence de petite taille, image par image, d'autre part, on veut réduire les temps de calcul de manière à produire un agrandissement en temps réel. Aussi, pour réaliser l'agrandissement d'une séquence vidéo, nous nous sommes inspirés des caractéristiques d'un flux MPEG (prédiction/compensation de mouvement) [6], [8]. Pour chaque GOP (Group Of Picture), nous avons réalisé les opérations suivantes, répétées jusqu'à la fin de la séquence :

- La première image d'un groupe d'images (codage Intra) est agrandie dans sa globalité par induction. Nous avons effectué une mise en œuvre optimisée et simplifiée pour réduire le temps de calcul, tout en obtenant une image de bonne qualité visuelle.
- Les images suivantes du GOP (reconstruite par compensation de mouvement) sont reconstruites bloc à bloc à partir :
 - des vecteurs de mouvement de taille agrandie que nous appliquons à des blocs dont la taille est, elle aussi agrandie.
 - de l'image d'erreurs agrandie par interpolation linéaire pour avoir un faible temps de calcul.

L'apport de cette méthode comparée aux agrandissements proposés par les lecteurs MPEG est indéniable. En effet, les lecteurs actuels intègrent au mieux un agrandissement par interpolation.

3.2 Résultats avec un facteur égale à 2 (ou facteur entier)

L'implémentation pour un facteur entier est assez facile. Les vecteurs de mouvement ainsi que les blocs sont agrandis, et il n'y a aucun problème pour la juxtaposition des blocs (les blocs agrandis ont une taille entière). Les images d'erreurs sont agrandies par une interpolation globale et elles sont ajoutées pour corriger la compensation de mouvement.

Nous présentons en figure 3 une image P reconstruite par notre méthode (facteur d'agrandissement par 2). Pour cette séquence, nous avons choisi des paramètres standardisés pour du PAL [8]. La structure des GOP est la suivante : une image I (codage *Intra*), trois *frames* P (codées par simple prédiction de mouvement), et deux *frames* B (codées par bi-prédiction de mouvement) entre chaque *frame* P.



Figure 3 – Image P reconstruite par compensation de mouvement.

3.3 Résultats pour un facteur quelconque

Le cas de l'agrandissement initial par induction de l'image I ne pose aucune difficulté. Mais, des problèmes interviennent pour l'agrandissement du reste de la séquence avec un facteur quelconque.

Comme une image est un ensemble discret de pixels, une taille non-entière pour un bloc n'est pas concevable.

Nous avons élaboré deux techniques pour essayer de palier ce problème, mais elle trouve rapidement leurs limites. La première consiste à approcher la taille des blocs agrandis au plus proche entier. Aussi, pour correctement couvrir l'ensemble de l'image, la taille des blocs est irrégulière et nous devons être particulièrement attentif à leurs jonctions. Dans la seconde technique, la taille des blocs agrandis est approximée à la valeur supérieure et nous les superposons légèrement les uns aux autres.

Les résultats n'étant pas suffisamment bons, pour améliorer l'efficacité des images d'erreurs, nous avons essayé d'agrandir ces images par bloc. En effet, la prédiction/compensation de mouvement se faisant par

bloc, les images d'erreur présentent elles aussi cette structure. Par conséquent, si nous agrandissons l'image d'erreur par bloc et non dans son intégralité, les discontinuités sont mieux préservées.

Les images suivantes montrent les résultats de cette méthode d'agrandissement de l'image d'erreur. L'effet de bloc qui était visible sur l'image de gauche sur le tronc de l'arbre est réduit sur l'image de droite.

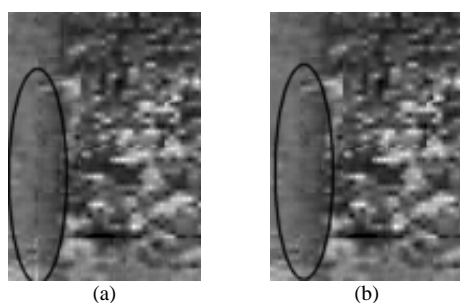


Figure 4 – Reconstruction avec une image d'erreur agrandie dans sa globalité (a) et une image d'erreur agrandie par bloc (b).

La dégradation rapide des images reconstruites par compensation de mouvement au cours de la séquence, va entraîner un effet stroboscopique perceptible. D'où, les perspectives envisagées dans la conclusion.

4 Conclusion et perspectives

L'utilisation conjointe de l'agrandissement par induction, et des vecteurs de mouvement, offre un agrandissement de qualité supérieure à celui fourni par une duplication de pixels ou une simple interpolation linéaire (ceci dans le cadre d'un agrandissement par deux). Les résultats obtenus pour un facteur quelconque sont décevants : en effet, il semble difficile de complètement supprimer un effet de flou sur les contours, en particulier lorsque nous avançons dans la succession des images P. Il semblerait que nous soyons arrivés à la limite de la méthode en utilisant cette technique.

Il est possible de se focaliser sur des points importants de l'image : textures, contours, zones particulières... et de réaliser un agrandissement adapté à chacune de ces zones.

Mais l'étape suivante sur laquelle nous travaillons plus particulièrement maintenant est l'exploitation de la notion d'objet. En effet, en utilisant les objets (formes et textures) et leurs déplacements suivant les spécifications de MPEG-4[6], nous comptons améliorer grandement la qualité des agrandissements de vidéo.

Références

[1] P. J. Burt. "Fast filter transforms for image processing". Computer Graphics and Image Processing, 16:20-51, 1981.

[2] D.Calle and A. Montanvert. "Super resolution inducing of an image". Proceedings of IEEE International Conference on Image Processing ICIP-98, 3, October 1998.

[3] S. Chaudhuri. "Image Super-resolution imaging". Kluwer Academic Publishers, January 2001.

[4] A. Chéhikian. "Algorithmes optimaux pour la génération de pyramides d'images passe-bas et laplaciennes". Traitement du signal, 9(4) :297-307, 1992.

[5] P. L. Combettes, "Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections". IEEE Transactions on image processing, vol. 6, no. 4, pp. 493-506, April 1997.

[6] T. Ebrahimi and C.Horne. "MPEG-4 natural video coding – An overview". http://www.csel.it/leonardo/icjfiles/mpeg-4_si/7-natural_video_paper/7-natural_video_paper.htm.

[7] S. Fassino et A. Montanvert. "Agrandissement video par exploitation de la prediction de mouvement". CORESA 2001, Novembre 2001.

[8] "Generic coding of moving pictures and associated audio, Recommendation H.262". ISO/IEC 13818-2, Draft International Standard, March 1994.

[9] T. Lindeberg. "Scale space for discrete signals". IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(3):234-254, March 1990.

[10] V.Lozano, "Contribution de l'analyse d'image couleur au traitement des images textile". PhD thesis, univ. Jean Monnet, Saint-Etienne, Janvier 1998.

[11] N.R. Shah and A. Zakhor. "Resolution enhancement of color video sequences". IEEE Transactions on image processing, vol. 8, no. 6, pp. 879-885, June 1999.

[12] B.C. Tom. "Resolution enhancement of monochrome and color video using motion compensation". IEEE Transactions on image processing, vol. 10, pp. 278-287, February 2001.