



Titre :

Algorithme d'appariement de blocs basé sur l'optimisation de la compensation de disparité pour le codage d'image stéréoscopique

Auteurs :

Imen KADRI^{+,-}, Gabriel DAUPHIN⁺, Anissa MOKRAOUI⁺, Zied LACHIRI⁻
⁺L2TI, Institut Galilée, Université Paris 13 Sorbonne Paris Cité
 99, avenue Jean-Baptiste Clément 93430 Villetaneuse, France
⁻LSITI, Ecole Nationale d'Ingénieurs de Tunis, Université de Tunis El Manar
 BP 37, Campus Universitaire, Le Belvédère, 1002, Tunis, Tunisie

Résumé :

Cet article présente un algorithme d'appariement de blocs basé sur la compensation optimisée de disparité pour le codage d'image stéréoscopique. Notre algorithme propose d'estimer une première carte de disparité basé sur l'algorithme classique de mise en correspondance des blocs similaires dans les deux vues en minimisant l'erreur quadratique moyenne entre la vue originale et sa version reconstruite sans compensation de disparité. Cette carte est ensuite modifiée de façon à ce que la vue prédite puis compensée donne de meilleures performances en termes de débit-distorsion comparées au schéma classique de codage. Pour un bloc donné, l'algorithme propose de tester l'ensemble des disparités de la fenêtre d'appariement choisie en tenant compte de l'impact de l'erreur résiduelle décodée sur la vue reconstruite. Parmi les disparités testées, l'algorithme sélectionne alors celle qui améliore la vue reconstruite après compensation pour un débit cible. Basé sur ce principe, les disparités de la carte initiale sont toutes mises à jour. Les simulations montrent que notre algorithme non seulement réduit la redondance inter-vue mais également améliore la qualité de la vue reconstruite et compensée par rapport à la méthode classique de codage avec compensation de disparité.

Mots-clefs :

Image stéréoscopique, mise en correspondance de blocs, disparité, compensation, codage

1. Contexte et état de l'art :

Une image stéréoscopique est composée de deux vues acquises selon deux points de vue légèrement différents de la même scène. De ce fait, les images stéréoscopiques nécessitent deux fois plus d'information comparée à une image 2D. Il s'avère alors nécessaire de se préoccuper de la question du codage de ces images stéréoscopiques pour les stocker ou les transmettre. Dans cet objectif, de nombreux travaux se sont intéressés à la réduction de la redondance inter-vue [1]. L'idée sous-jacente étant de reconstruire l'une des deux vues en utilisant la deuxième vue (vue de référence) et la carte de disparité associée à la vue à reconstruire. En effet la carte de disparité est moins coûteuse en termes binaire que le codage d'une des deux vues. Cette dernière, associée à l'une des deux vues, est généralement estimée par l'algorithme classique d'appariement de blocs similaires dans les deux vues. Notons que la qualité de la vue reconstruite dépend fortement de la carte de disparité estimée. De ce fait, pour tenir compte des erreurs de reconstructions trop importantes dues par exemple aux changements de luminosité, aux zones texturées ou aux occultations, la majorité des travaux proposent des solutions de codage basées sur le principe de la compensation de disparité. L'objectif de cet article est d'améliorer les performances en termes de débit-distorsion de ce schéma classique de codage.

2. Algorithme d'appariement de blocs basé sur l'optimisation de la compensation de disparité :

Dans cet objectif, nous proposons d'estimer la carte de disparité par blocs (notée $\{d(k, l, Q)\}$) qui minimise l'erreur quadratique moyenne entre l'image originale de la vue droite $I_d(i, j)$ et sa version prédite compensée par l'image résiduelle décodée notée $\hat{I}_d(i, j) = I_c(i, j, s, Q)$. L'image résiduelle, donnée par $(I_d(i, j) - \hat{I}_g(i, j + d(k, l, Q)))$ avec \hat{I}_g l'image gauche codée, est codée par JPEG selon le facteur de qualité Q choisi. L'estimation de la carte de disparité est alors formalisée par l'équation suivante :

$$d(k, l, Q) = \underset{s \in S}{\operatorname{argmin}} \left(\sum_{(i,j) \in B_{k,l}} (I_c(i, j, s, Q) - I_d(i, j))^2 \right),$$

où S représente la fenêtre d'appariement choisie, $B_{k,l}$ est le bloc à appairer de coordonnées (k, l) .

Pour résoudre cette équation, nous proposons la démarche brièvement résumée ci-dessous qui consiste à estimer une première carte de disparité. Cette dernière cherche à minimiser l'erreur quadratique moyenne entre l'image originale associée à la vue droite et sa version uniquement prédite $\hat{I}_d(i, j) = I_g(i, j + s)$:

$$d(k, l) = \underset{s \in S}{\operatorname{argmin}} \left(\sum_{(i,j) \in B_{k,l}} (I_g(i, j + s) - I_d(i, j))^2 \right).$$

L'ensemble des disparités de cette carte initiale sont mises à jour, une à une, en testant toutes les disparités possibles de la fenêtre d'appariement et plusieurs facteurs de qualité Q pour comprimer l'image résiduelle. Pour une disparité choisie dans S , l'algorithme opère comme suit : (i) l'image de la vue droite est reconstruite ; (ii) l'image résiduelle est à son tour codée selon JPEG en utilisant plusieurs facteurs de qualité Q ; (iii) pour chaque facteur de qualité, l'image résiduelle est décodée pour compenser l'image prédite. Les performances sont mesurées en termes de débit-distorsion. Le débit binaire tient compte du coût de codage entropique de la carte de disparité ainsi que le coût de codage du résiduel associé à l'image droite I_d . La distorsion s'appuie sur le PSNR où l'erreur quadratique moyenne est calculée entre l'image de la vue droite originale I_d et sa version prédite compensée par le résiduel décodée. Cette procédure est répétée pour toutes les valeurs possibles de la fenêtre d'appariement. On retiendra la disparité qui donnera le meilleur PSNR pour le débit cible. Ce processus est itéré tant que les disparités de la carte initiale n'ont pas toutes été mises à jour.

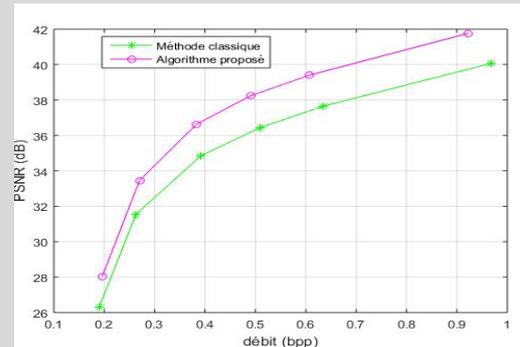
3. Résultats de simulation :

Dans les simulations réalisées, le décodeur dispose de l'image gauche originale. Les tests sont réalisés sur les images extraites de la base Middlebury [2] et Demos. Le débit binaire tient compte du coût de codage de la carte de disparité (codage arithmétique) et de l'image résiduelle par JPEG. La taille des blocs est fixée à 8×8 . Les performances de notre algorithme sont comparées à celles d'un schéma de codage classique. Le tableau ci-dessous présente les performances obtenues sur l'image droite. La métrique Bjontegaard montre que la valeur moyenne du gain en PSNR à débit donné de l'algorithme proposé par rapport au schéma classique est en faveur de notre algorithme. Nous observons que notre algorithme permet une très nette amélioration pour certaines images. Les courbes débit-distorsion de la Figure 1 montrent un gain en termes de PSNR relativement constant sur un large intervalle de valeurs de débit (entre 0.4 bpp et 1 bpp).

Tableau 1 : Métrique Bjontegaard

Image	Δ PSNR (dB)	bpp (%)
Bowling	+1,85	-20,8 %
Books	+1,65	-27,5 %
Teddy	+1.1	-21 %
Barn	+0.44	-13 %
Voiture	+0,1	-4,4 %

Figure 2 : Comparaison des performances en termes de débit-distorsion.



4. Conclusion et perspectives :

Les bonnes performances de l'algorithme proposé dans cet article montrent bien l'intérêt d'estimer les disparités en tenant compte de la prédiction après compensation effective. La grande complexité numérique de cet algorithme nous amène à réfléchir sur l'utilisation de techniques d'apprentissage pour estimer plus simplement les disparités tout en maintenant les bonnes performances avec un temps de calcul relativement raisonnable.

Référence :

[1] F. Dufaux, B. Pesquet-Popescu, M. Cagnazzo: Emerging Technologies for 3D Video: creation, coding, transmission and rendering, 2013, John Wiley & Sons.
 [2] <http://vision.middlebury.edu/stereo/data/>