



Epreuve de Moyenne Durée

Exercice 1 : (6 points)

Répondre aux questions suivantes :

- Pourquoi la compression vidéo se fait sur le format YUV et quelle est la signification de 4:1:1 pour le format YUV. (1,5pts)
YUV est choisi pour traiter le Y à part entière et réduire les dimensions de U et V car la composante chromatique qui est moins sensible à la perception que la composante luminance (Y).
4 : 1 : 1 cela signifie un down sampling des composantes chromatiques U et V, qui seront réduites à 1/4.
- Quel est l'intérêt d'introduire les frames de type B dans le GOP (Group Of Pictures) dans le standard mp4. (1,5pts)
Réduire le résidu et donc d'augmenter le taux de compression.
- Comment se fait le calcul des éléments de la DCT pour optimiser le temps de calcul. (1,5pts)
On calcule les produits des cosinus en offline et on les sauvegarde dans un fichier.
- Quels sont les principes de base utilisés dans la compression mp3 ? (1,5pts)
Masquage fréquentiel et temporel (prendre le texte du cours)

Exercice 2 : (8 points)

Le standard de compression vidéo utilise des blocs de taille 16x16.

- (a) Justifiez ce choix de taille. (02 pts)

**Augmenter cette taille, revient à réduire le temps de recherche de blocs similaires, augmente le résidu et donc décroît le taux de compression.
Et Inversement si on réduit la taille du bloc : plus de vecteurs, meilleur qualité, plus de temps de calcul. La taille 16 est un compromis.**

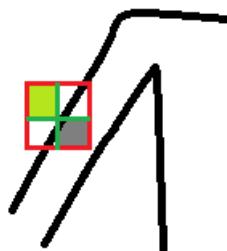
- (b) Dans quelles circonstances, on réduirait cette taille ? Pourquoi ? Donnez un exemple d'applications. (02 pts)

**La taille est réduite lorsque peu de changements sont présents dans une frame.
On réduit vers des tailles inférieures pour diminuer le résidu. Dans certaines circonstances, le bloc 16x16 contiendra uniquement quelques parties qui changent, le reste est inchangé.**

Autre possibilité, c'est lorsque la taille de l'objet occupe moins de 16x16 pixels (cas de petits objets, ou objets lointains).

Exemple d'application :

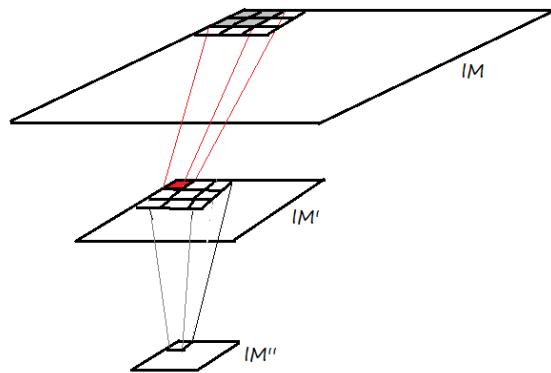
- (1) Surveillance de scène, cas de régions qui bougent occupent moins de pixels (8x8) ou (4x4) en raison de leurs dimensions ou de leurs positions relativement à la caméra.
 - (2) Visio-Conférence : Le premier plan (Foreground) est statique en globalité à l'exception de quelques parties qui changent (face).
- (c) On veut concevoir un codec pour lequel la taille est choisie de façon dynamique avec une taille fixe pour les blocs d'une frame. Quel type d'algorithme choisir pour implémenter cette solution ? (02 pts)
Il faudra détecter les régions en mouvement en utilisant un algorithme de traitement d'images (Background subtraction pour localiser la zone de l'objet).
- (d) Proposez une exemple d'application pour cet algorithme. (01 pt)
Un exemple d'application, vidéo surveillance de scène et suivi des objets en mouvement. Un algorithme de traitement d'image détecte les objets intrus. En fonction de leurs tailles, la compression utilisera la taille du bloc standard ou 8x8 ou 4x4.
- (e) La division du bloc en quatre pour le calcul des vecteurs de mouvement est-il bénéfique ? si oui dans quels intérêts ? (01 pt)
Oui il est bénéfique pour réduire le résidu et donc augmenter le taux de compression (cas du codec DIVX). Le cas d'utilisation est en visio-conférence, lorsque le bloc traité englobe la frontière de la personne, sa subdivision permet de trouver des sous blocs similaires avec petites valeurs des résidus pour les sous blocs internes ou externes (voir exemple des sous blocs coloriés illustrés par la figure ci-dessous.).



Exercice 3 : (6 points)

La recherche de blocs similaires et donc du vecteur de mouvement pour la compression vidéo peut être réalisée en optant pour la recherche hiérarchique.

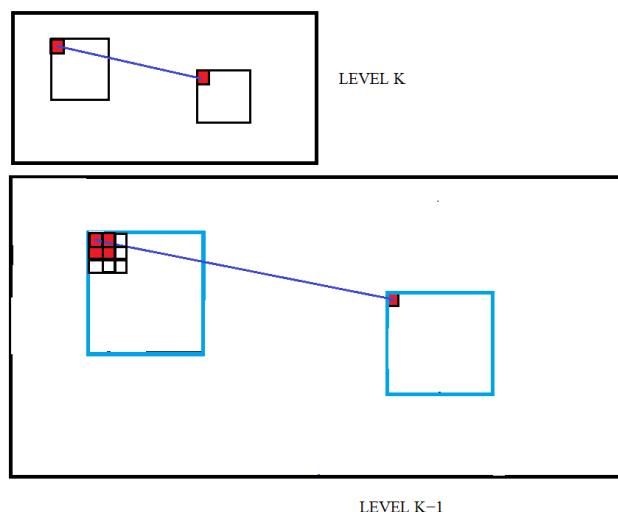
- (a) Expliquer comment se fait le raffinement du vecteur de mouvement. (4 pts, distribués comme suivant :)

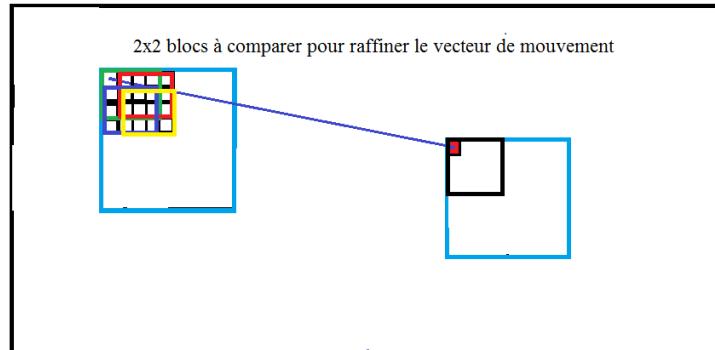


Chaque frame de dimensions (n,m) sera convertie vers une frame $(n/2, m/2)$ où chaque quadruplet de pixels sont mappés vers un seul pixel. On peut aussi considérer 9 pixels aussi au lieu de 4 pixels. On produit ainsi (récursivement) 3 ou 4 frames (niveaux 1, 2, 3 et 4).

La zone de recherche est aussi réduite de moitié à chaque niveau. (1 point)

Supposons que le vecteur de mouvement a été calculé au niveau k comme indiqué par la figure ci-dessous. Au niveau $k-1$, chacun des pixels du niveau k a été calculé moyennant les 4 pixels du niveau $k-1$.





LEVEL K-1

Le bloc similaire est calculé à nouveau dans cette zone en prenant comme début du bloc un quatre pixels. Ceci est dû au fait que le pixel résultat (début du bloc apparié) est en réalité la moyenne des 04 pixels du niveau k-1.

On refait cette opération au niveau k-2 pour le pixel sélectionné, en utilisant les blocs de taille 8x8 (4 comparaisons). De même pour le niveau initial.

Calcul de la complexité : (2 points)

Pour le niveau initial, et pour un espace de recherche de taille k, il faut $(2k+1)^*(2k+1)$ comparaisons de blocs (MSE), taille du bloc est égale à 16.

Pour high resolution video, k=32, donc 4225 MSE.

Pour le niveau 4, taille du bloc est 2, et k=4, il faut $(2^*4+1)^*(2^*4+1)=81$ MSE (avec taille du bloc égale à 2), plus 3*3 MSE pour le raffinement (3 avec la taille du bloc égal à 4, 3 avec taille 8 et 3 avec taille 16).