

Compression des séquences vidéo

INF4710 Introduction aux technologies
multimedia

Plan

- ▶ **Compression**
 - Par trame individuelle
 - Par redondance temporelle

Compression vidéo

▶ Séquence vidéo:

- C'est une suite d'images affichées successivement à de courtes intervalles.
- Typiquement d'une taille d'au moins 320x240 avec un taux de trames de 30 trames par seconde.
- Un image 320x240 fait 255 Kilooctets, donc multiplié par 30, on obtient un total de 7.5 Mégaoctets pour une seule seconde.
- La compression est donc encore plus importante pour les séquences vidéo. En effet, 320x240 est une petite taille selon les standards actuels.

Compression vidéo

- ▶ Format progressif vs entrelacé
 - L'œil humain nécessite un taux de trames assez élevé pour percevoir un mouvement comme étant fluide (au moins 30, et 60 à 72 trames/s encore mieux).
 - Taux de trames élevé = grande bande passante requise (ou gros fichier).
 - Truc: utiliser la propriété du système visuel humain de la pertinence rétinienne pour diminuer l'information requise.
 - Lorsqu'on ferme les paupières les images persistent environ 50 millisecondes.

Compression vidéo

- ▶ Format progressif vs entrelacé
 - Dans le format progressif, on affiche toute les lignes et toutes les colonnes des trames selon le taux de trames.
 - Dans le format entrelacé, on alterne l'affichage des lignes paires et impaires selon le taux de trames. Grâce à la pertinence rétinienne, le système visuel perçoit une image complète.
 - Donc, on compresse l'information, car pour un même taux de trames (1080i60 vs 1080p60), on a besoin d'enregistrer ou transmettre seulement des demi-trames (compression de 2:1).

Compression vidéo

- ▶ Format progressif vs entrelacé
 - Le truc de l'entrelacement est en voie de disparition, car:
 - les écrans ne fonctionnent plus de cette façon;
 - utile seulement si la séquence vidéo est captée à 60 trames par seconde, par exemple pour du 320i60. Pour les films, ils sont filmés à 24 trames par seconde...
 - Pour faire la compression d'une vidéo entrelacée, il y a perte d'efficacité (moins de redondance spatiale et temporelle). Donc, les taux de compression sont moins bons, si bien que 1080p50 et 1080i50 requiert le même taux de bits.

Compression vidéo

- ▶ Méthodes de compression par trame individuelle
 - Ce que nous avons vu précédemment indique qu'il faut compresser avec d'autres stratégies que l'entrelacement.
 - Il existe des méthodes de compression par trame individuelle. Exemples:
 - M-JPEG
 - HuffYUV

Méthodes de compression par trame individuelle

▶ M-JPEG (Motion JPEG)

- Chaque trame individuelle est compressée à l'aide de la transformée de cosinus (DCT). On utilise seulement la redondance spatiale.
- Donne des ratios de compression jusqu'à 1:20
- Ne dépend pas du mouvement dans la scène.
- Avantage: Excellent support. Format universel. Nécessite peu de mémoire et de puissance de calcul.
- Inconvénient: On peut obtenir de meilleur ratio de compression en utilisant la redondance temporelle.

Méthodes de compression par trame individuelle

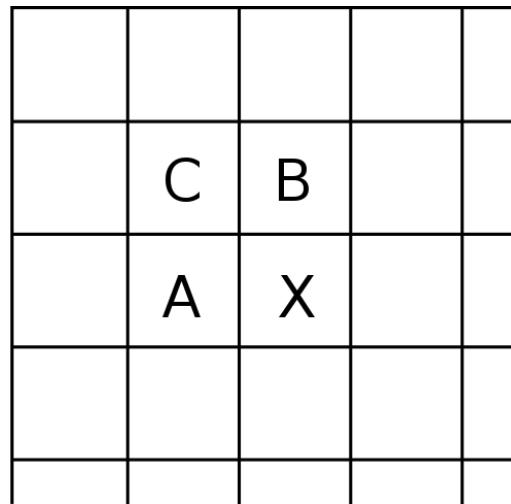
▶ HuffYUV

- Méthode de compression sans perte
- Comprime une séquence vidéo dans le format Y'CbCr. Fonctionne aussi en RGB.
- Utilise le codage prédictif et le codage de Huffman
- Le taux de compression est de 2:1

- Note: Il n'y a aucune perte si la séquence vidéo est originalement dans le même espace couleur que la version compressée.

Méthodes de compression par trame individuel

▶ HuffYUV



- La valeur de X est prédite à partir de A,B,C selon une formule du type $A+B-C$ (Predict gradient)
- Les erreurs de prédiction sont ensuite codées avec Huffman.

Méthodes de compression utilisant les redondances temporelles

▶ Compression vidéo

- Puisqu'elle est constituée d'une suite d'images, on peut aussi utiliser la redondance temporelle.
- En effet, pour deux trames consécutives, habituellement ce n'est pas tous les pixels qui changent.
- C'est ce que font les méthodes telles que CorePNG et MPEG1 / MPEG2.

Méthodes de compression utilisant les redondances temporelles

▶ CorePNG

- Deux types de trames:
 - Trame I: Trame complète codée sous le format PNG
 - Trame P: (facultative) Trame résultant de la différence entre deux trames consécutives. Codée aussi sous le format PNG. Normalement, les différences sont petites, donc moins de bits sont requis pour les coder.
- Donc, certaines trames contiennent les informations complètes sur une image, alors que d'autres trames contiennent seulement les différences par rapport à la trame précédente.
- Maximum de 1000 trames P entre deux trames I.

Méthodes de compression utilisant les redondances temporelles

▶ MPEG1 / MPEG2

- On peut aller encore plus loin dans l'utilisation des redondances temporelles. On peut aussi prédire le mouvement dans la vidéo afin de minimiser l'amplitude des différences, et donc le nombre de bits requis.
- C'est ce que fait, entre autres, MPEG1 et MPEG2, et la plupart des méthodes récentes de compression vidéo.

Méthodes de compression utilisant les redondances temporelles

- ▶ **MPEG1 / MPEG2**
 - Étapes de la compression
 1. Conversion vers Y'CbCr
 2. Division de la trame en macroblocs
 3. Si la trame encode seulement les différences temporelles, on calcule d'abord des vecteurs de mouvement minimisant les différences entre macroblocs. Ensuite, on calcule l'image des différences en tenant compte des vecteurs de mouvement.
 4. Encodage des macroblocs par la DCT.
 5. Quantification
 6. Codage entropique

MPEG1 / MPEG2

- ▶ 1. Conversion vers Y'CbCr
 - La chrominance est sous-échantillonnée selon le format 4:2:0
 - Pour les étapes suivantes, chaque composante (Y', Cb, Cr) est encodée indépendamment (mais de la même façon).

MPEG1 / MPEG2

▶ 2. Division en macroblocs

- Puisque l'image sera codée avec la transformée en cosinus (DCT), l'image doit être divisée en bloc de 8x8
- Toutefois, à cause du sous-échantillonnage de la chrominance, un bloc 8x8 de Y' équivaut à un bloc de 4x4 de Cb et Cr. Pour avoir des blocs 8x8 pour Cb et Cr, il faut donc, un bloc de 16x16 pour Y'. Note: 4x4 ne permet pas une bonne compression, 16x16 donne des effets de bloc visibles.
- Un macrobloc est donc défini comme étant 4 blocs de 8 x 8 de Y', et 1 bloc de 8 x 8 de Cb et 1 bloc de 8x 8 de Cr. La surface de l'image est 16 x 16, mais le codage sur 8x8.

MPEG1 / MPEG2

▶ 2. Division en macroblocs

- Structure d'un macrobloc

+	-	-	+	-	-	+	-	-	+	-	+					
	ADDR		TYPE		QUANT		VECT		CDP		B0		B1		... B5	
+	-	-	+	-	-	+	-	-	+	-	+	-	-	+	-	+

ADDR: Coordonnée du macrobloc dans l'image

TYPE: Type du macrobloc (I, B, ou P)

QUANT: Information de quantification

VECT: Vecteur de mouvement

CDP: 6 bits qui indiquent quels blocs sont codés.

B0..B5: Les blocs: 4 Y', 1 Cb, 1 Cr

MPEG1 / MPEG2

▶ 3. ...vecteurs de mouvement

- Pour l'étape 3, celle-ci est appliquée si le type de trame est approprié.
- Il y a 3 types principaux de trames en MPEG
 - Trame I (I-frame, intra-frame): Encodée et décodée indépendamment des autres trames. L'encodage est comme JPEG. Normalement, on a une trame I à toutes les 14 à 17 trames.
 - Trame P (P-frame, Predicted-frame, Forward-predicted frames): Exploite la redondance temporelle avec la trame (P ou I) qui la précède immédiatement.
 - Trame B (B-frame, bidirectional-frame): Exploite la redondance temporelle avec la trame (P ou I) qui la précède immédiatement, et la trame (P ou I) qui la suit immédiatement.

MPEG1 / MPEG2

▶ 3. ...vecteurs de mouvement

- Une séquence vidéo MPEG est donc définie comme une séquence de groupes de 15 à 18 trames.
- Un groupe de trames (GOP: group of pictures) est composé d'une trame I et plusieurs trames P ou B.
Exemple:
 - IPPPPPBPBPPPBBB

MPEG1 / MPEG2

- ▶ 3. ...vecteurs de mouvement
 - L'étape 3 est appliquée aux trames B et P. Elle vise à exploiter la redondance temporelle, et donc à encoder seulement les différences avec la DCT. Différences petites = moins de bits pour le codage.
 - Voici le principe de l'utilisation du mouvement entre deux images.

Image originale



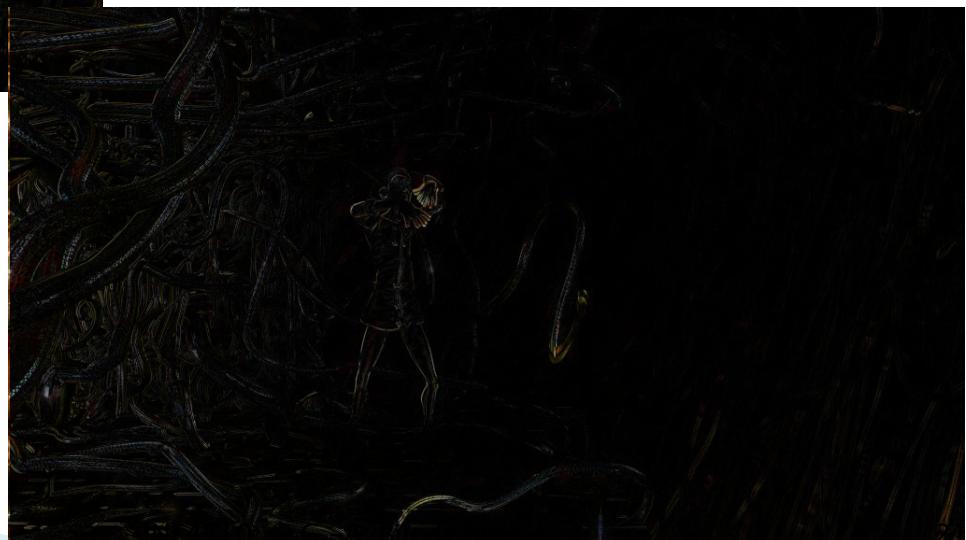
MPEG1 / MPEG2

► 3. ...vecteurs de mouvement



Différences entre 2 trames consécutives

Différences entre 2 trames consécutives après compensation du mouvement



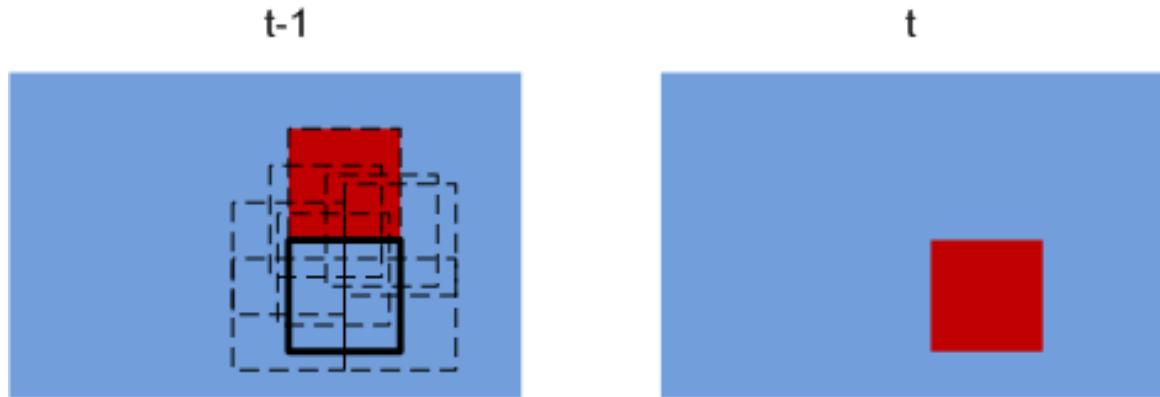
MPEG1 / MPEG2

► 3. ...vecteurs de mouvement

- Les vecteurs de mouvement sont calculés afin de minimiser les différences entre 2 trames consécutives. On encode seulement les macroblocs où il y a une différence significative.
- Pour trouver le vecteur de mouvement du macrobloc dans une trame t , on le déplace dans la trame précédente ($t-1$) dans un rayon de taille fixe autour de sa position initiale dans la trame t . Les déplacements testés sont par incrément de $\frac{1}{2}$ pixel.

MPEG1 / MPEG2

► 3. ...vecteurs de mouvement



- Par exemple le macrobloc rouge à t provient du macrobloc rouge à $t-1$ suite à une translation en y de 16 pixels. Le macrobloc bleu en haut du rouge à t provient de n'importe quel macrobloc bleu voisin à $t-1$.
- Si on applique les vecteurs de mouvement, la trame $t-1$ devient similaire à la trame t , et donc la différence entre la trame t et $t-1$ est zéro (cas très idéal).

MPEG1 / MPEG2

► 3. ...vecteurs de mouvement

- Après estimation du mouvement, les macroblocs contiendront:
 - Un vecteur de mouvement (peut être (0,0))
 - les différences avec un macrobloc de la trame précédente après l'application du vecteur de mouvement. S'il n'y a aucun macrobloc correspondant, le macrobloc est codé comme dans une trame I (les différences == contenu de la trame).
- Les vecteurs de mouvement sont encodés à l'aide d'un codage prédictif, car plusieurs macroblocs ont souvent le même déplacement.
- Les différences par la DCT et une quantification.

MPEG1 / MPEG2

► 3. ...vecteurs de mouvement

- Pour les trames B, on fera aussi l'estimation du mouvement à l'aide d'une trame subséquente (donc, et avec une précédente, et avec une suivante). Parfois, la différence est plus petite avec la trame suivante qu'avec la trame précédente.
- Cela a comme conséquence que lors du codage du flux MPEG, les trames doivent être décodées dans le désordre, car pour reconstruire une trame B, on a besoin de la trame I ou P suivante.
- Note: Un macrobloc d'une trame B peut avoir deux vecteurs de mouvement (donc provenir de la combinaison de deux macroblocs).

MPEG1 / MPEG2

- ▶ 4. Encodage des macroblocs par la DCT.
 - ▶ L'encodage est fait exactement comme pour JPEG.

MPEG1 / MPEG2

▶ 5. Quantification

- La quantification est faite exactement comme pour JPEG.

MPEG1 / MPEG2

▶ 6. Codage entropique

- Le codage entropique est fait exactement comme pour JPEG.
 - RLE
 - Huffman

Méthodes de compression utilisant les redondances temporelles

▶ Différences entre les méthodes standards

- H.120 (1984): Pas d'estimation de mouvement. Pas de DCT. Codage prédictif sur chaque pixel des trames. Quantification. Jamais vraiment utilisé.
- H.261 (1988): Utilise DCT, estimation de mouvement (précis au pixel), zigzag, quantification. Utilisé beaucoup pour la vidéoconférence et la vidéotéléphonie.
- MPEG-1 (1993): Utilisé pour les Vidéo-CDs

Méthodes de compression utilisant les redondances temporelles

- ▶ Différences entre les méthodes standards
 - H. 262, MPEG-2 (1995): Supporte les séquences vidéo entrelacées. Meilleure estimation du mouvement. Plus de contrôle sur la précision de la DCT. Utilisation: DVD Vidéo, Blu-Ray, Télévision numérique, SVCD
 - H.263 (1996): Correction d'erreurs, vecteurs de mouvement sans restriction. Codage arithmétique. Utilisation: Vidéoconférence, Vidéotéléphonie, Vidéo sur téléphone mobile, YouTube.

Méthodes de compression utilisant les redondances temporelles

- ▶ Différences entre les méthodes standards
 - MPEG4 Part 2 (1999): Estimation de mouvement au $\frac{1}{4}$ de pixel. Compensation de mouvement global. Utilisation: Vidéo sur internet (DivX, Xvid, etc.)
 - H.264/MPEG-4 AVC (2003): Modification à la DCT pour supporter du 4x4, codage sans perte de certains macroblocs, Estimation de mouvement à partir d'un max. de 16 trames. Estimation de mouvement pour bloc de taille entre 16x16 et 4x4 pour plus de précision. Estimation de mouvement pondéré (pour effet de gradation), nouvelles matrices de quantification. Permet des mises à l'échelle (la séquence vidéo est encodée sous deux résolutions). Permet d'inclure deux vidéos synchronisées codées à l'aide leur redondance mutuelle (pour la stéréoscopie). Codage arithmétique. Résistance aux pertes d'information (sur réseau) Utilisation: Blu-ray, télévision numérique, vidéo iPod, HD DVD.

Méthodes de compression utilisant les redondances temporelles

- ▶ Différences entre les méthodes standards
 - VC-2 (Dirac) (2008): Utilise la transformée en ondelettes. Utilisation HDTV, Vidéo sur internet.