

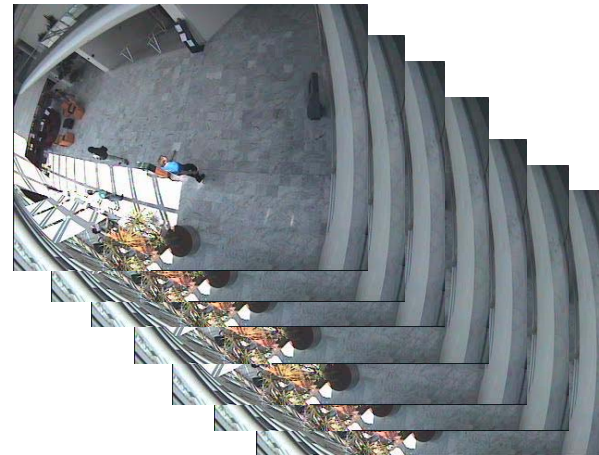
La vidéo, format et compression

La vidéo numérique

- Une vidéo est une succession d'images à une certaine cadence.
- Avec la cadence (20 images par seconde), il est possible de tromper l'œil et de lui faire croire à une image animée.
- On caractérise la fluidité (vitesse) d'une vidéo par le nombre d'images par secondes (en anglais frame rate), exprimé en FPS (Frames per second, en français trames par seconde).

La vidéo

- Une vidéo est une succession d'images à une certaine cadence. Avec la cadence de 20 images par seconde, il est possible de tromper l'œil et de lui faire croire à une image animée.



551 images
25 fps

La vidéo

- On caractérise la fluidité d'une vidéo par le nombre d'images par secondes (en anglais *frame rate*), exprimé en *FPS* (*Frames per second*, en français *trames par seconde*).



110 images
Enregistrement à : 5 fps
Affichage à : 25 fps

La vidéo

- Vitesse

Enregistrement à: 25 fps



Affichage à : 10 fps



Affichage à : 70 fps

La vidéo numérique

Une séquence vidéo brute est une suite d'images fixes, qui peut être caractérisée par trois principaux paramètres :

- **Résolution en luminance,**
- **Résolution spatiale,**
- **Résolution temporelle.**

Compression de la vidéo

Méthodes simples de réduction :

- Réduction de la taille de l'image
- Réduction du nombre d'images par seconde

Compression de la vidéo

Méthodes avancées :

- Les images contiennent beaucoup d'informations redondantes.
Enlever les redondances
- Réduire l'information de chrominance (couleur) car l'œil humain est plus sensible à la luminance

Compression de la vidéo

Types de compression:

- **Compression intra-trame (compression spatiale)** : Prise en compte de la similarité dans les zones de la même image (redondances spatiales). Cette compression ne prend pas en considération la corrélation entre les différentes trames.
- **Compression inter-trames (compression temporelle)**: Prise en compte de la similitude d'une image et celles qui l'entourent dans la séquence (précédente, suivante), on ne tiendra en compte que la différence en elles (redondances temporelle).

Sous échantillonnage de la chrominance

Principe:

L'œil humain est plus sensible aux variations de la luminance qu'à la chrominance. On conserve alors plus d'informations de luminance que de chrominance.

Espaces colorimétriques:

- L^*a^*b
- RVB
- CMJN

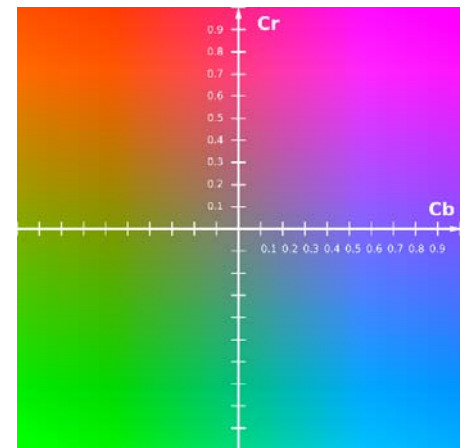
Sous échantillonnage de la chrominance

Les couleurs sont transportés dans un autre mode de représentation que le RGB. Le système YCbCr où Y représente la luminance et Cb, Cr la chrominance.

La luminance est obtenue en combinant les trois couleurs:

$$Y = 0,299 R + 0,587 G + 0,114 B$$

Les composantes Cb et Cr sont des combinaisons du Y avec le Bleu et le Rouge, respectivement.



Sous échantillonnage de la chrominance

On utilise la notation 4:1:1, 4:2:0 ou 4:2:2 pour indiquer les proportions respectives de Y, Cb et Cr.

Y Cb Cr	Y	Y	Y	4:1:1
Y Cb Cr	Y	Y	Y	

Y Cb Cr	Y	Y Cb Cr	Y	4:2:2
Y Cb Cr	Y	Y Cb Cr	Y	

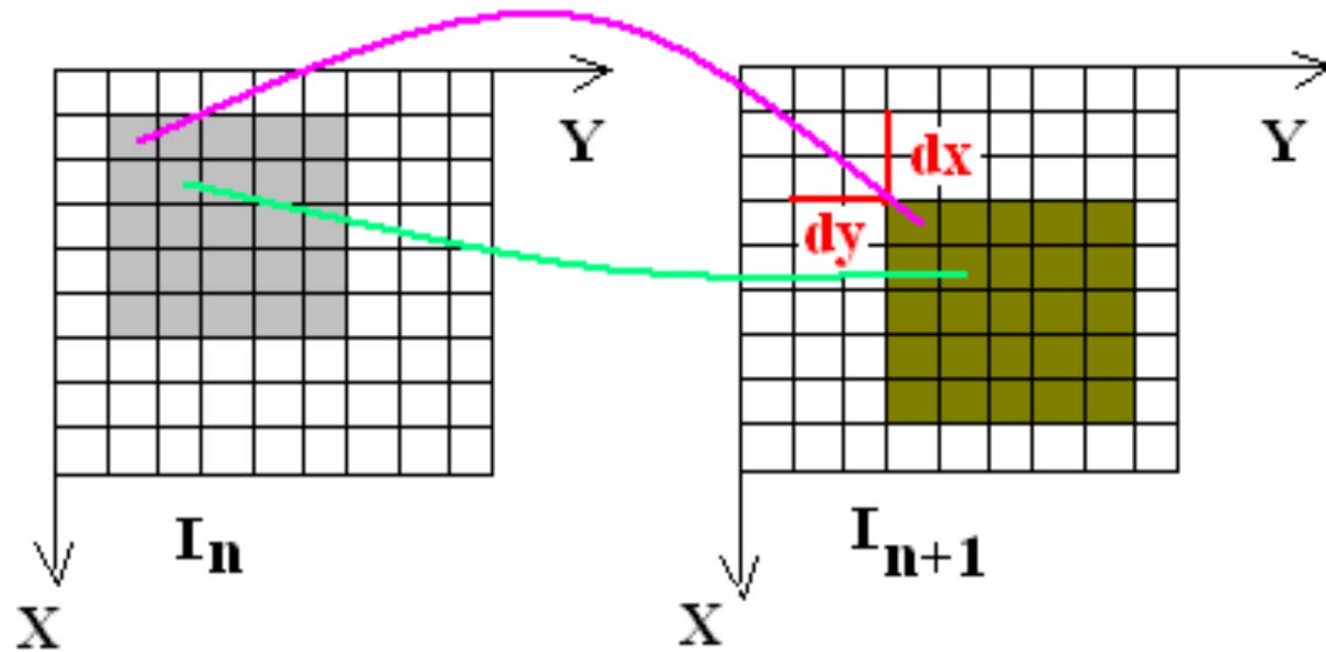
Y Cb Cr	Y	Y Cb Cr	Y	4:2:0
Y	Y	Y	Y	

La compensation de mouvement

Principe:

Le principe de la compression par compensation de mouvement est de coder une trame en JPEG (trame de référence), trouver la différence entre chaque trame et la trame de référence puis coder cette différence.

Block matching



Block matching

Pour trouver la différence entre deux images, on recherche un bloc donné de la première dans une zone de recherche de la deuxième. On calcule l'erreur moyenne quadratique (Mean Squared Error) entre les deux blocs.

$$\text{MSE} = \frac{1}{N * M} \sum_{x,y} (I_{n+1}(x + dx, y + dy) - I_n(x, y))^2$$

Pour chaque bloc B_i de l'image I_n chercher le bloc B_j correspondant dans l'image I_{n+1} permettant d'avoir le MSE minimum.

Vecteur de mouvement et résidu

- Ce vecteur est utilisé par le décodeur pour retrouver l'emplacement du bloc sur la trame n à partir du bloc sur la trame précédente ($n-1$).
- La différence entre le bloc codé et le meilleur bloc correspondant dans la trame précédente est connu par le terme résidu.
- Ce résidu doit être codé et transmis avec le vecteur mouvement pour pouvoir reconstruire le bloc.

Prédiction d'images en MPEG

1-Les trames de type I (Intracodée)

Pour ce type de frame, seule la redondance spatiale est utilisée pour la compression (jpeg). Dans la norme MPEG, une trame I est insérée chaque 12 trame. Cela permet d'éviter la propagation d'erreur.

2-Les trames de type P (Prédictive)

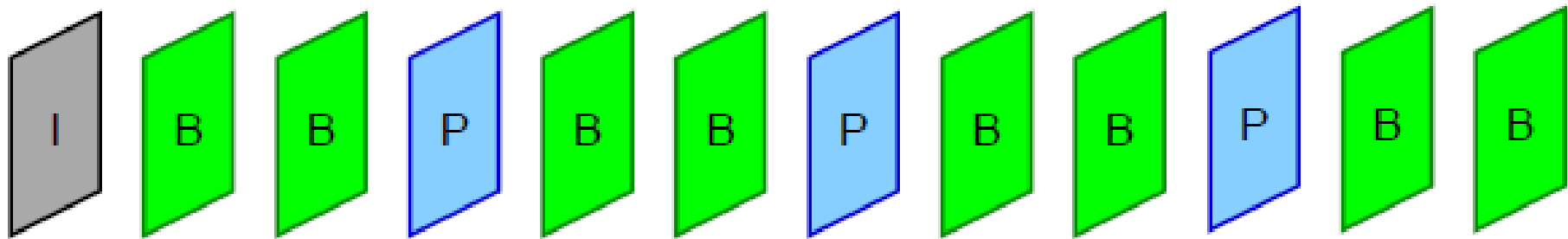
Elles sont codées en utilisant la redondance temporelle en la comparant avec la trame précédente I ou P.

3- Les trames de type B (Bidirectionnelle)

Elles sont codées en utilisant la redondance temporelle en les comparant avec la trame précédente et suivante (I,P) ou (P,P).

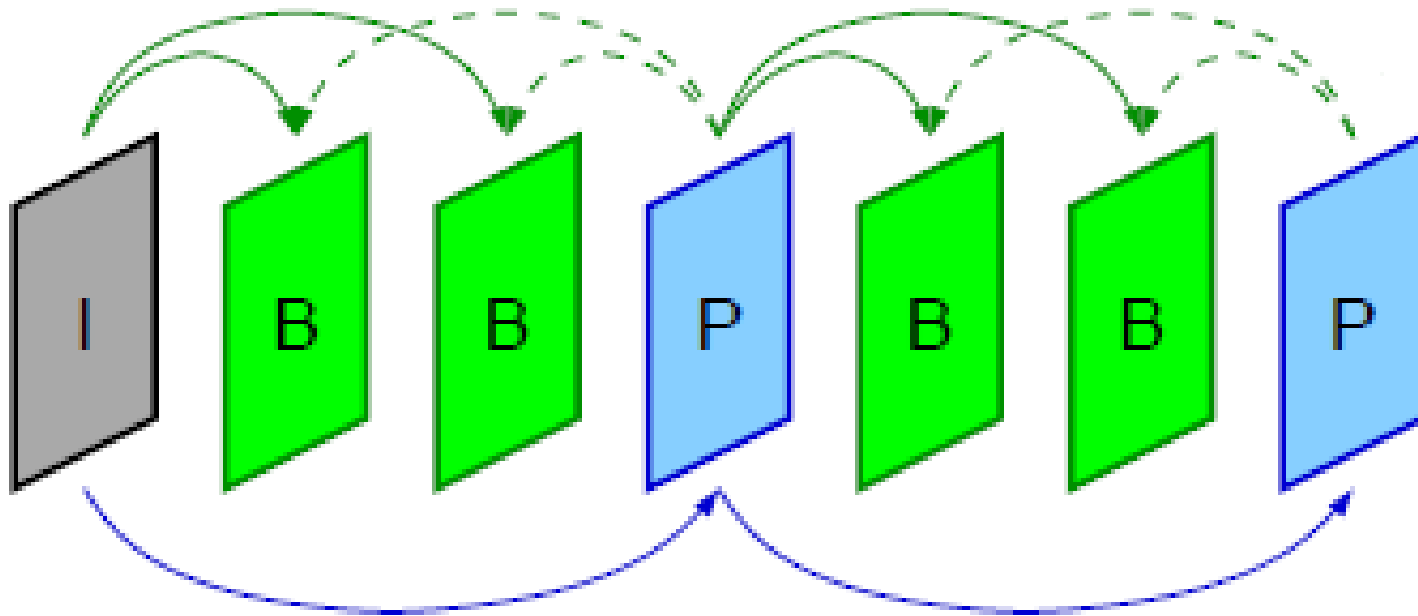
Prédiction d'images en MPEG

Séquence d'images type (GOP)

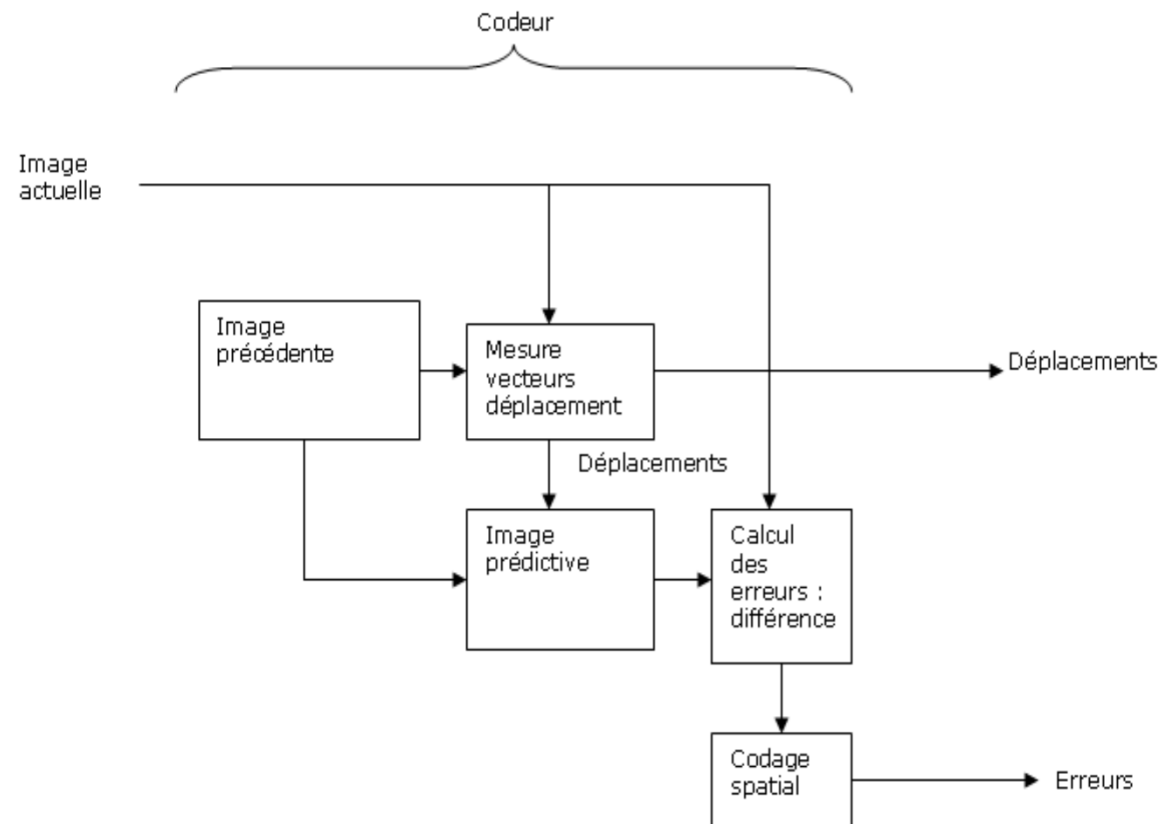


Prédiction d'images en MPEG

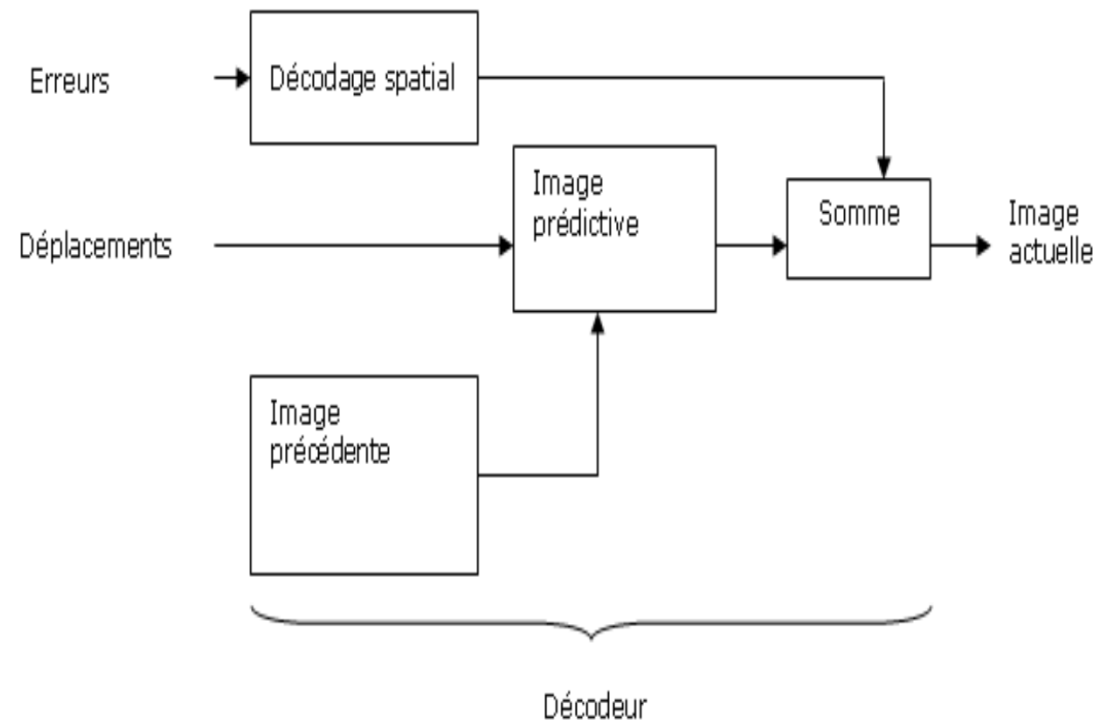
Prédiction



La compensation de mouvement



La compensation de mouvement



Codecs

C'est un dispositif capable de compresser et décompresser un signal numérique (code-decode en anglais). Parmi les codecs existant :

- **M-Jpeg (Moving Jpeg) :**

La compression M-Jpeg applique le format jpeg à chaque image de la vidéo.

Ce format est utilisé pour le montage vidéo car on accède à chaque image séparément mais est inadapté au téléchargement ou au streaming.

Codecs

- **MPEG (Moving Pictures Experts Group):**

Format de fichier audio et vidéo utilisant la compression avec perte.
Utilisé pour la diffusion en continu sur Internet (le streaming).

Plusieurs normes ont été définies : MPEG-1, MPEG-2, MPEG-4, MPEG-21.

Codecs

- **DivX**

Basé sur la compression MPEG-4. Sa compression est plus performante. Il est compatible avec la plupart des logiciels.

Les conteneurs

Les formats conteneur englobent tous les éléments de la vidéo : audio, vidéo et autre information (chapitres, sous titres, les méta données, ...).

Parmi les formats existant, on peut citer : AVI (Audio Video Interleave), MPEG, QuickTime, ASF, MP4, 3GP.

Les conteneurs

- **AVI (Audio Video Interleave):**

Il a été créé par Microsoft pour Windows. C'est le conteneur le plus répandu, il fonctionne sur tous les lecteurs vidéo. Ce format est utilisé par la majorité des appareils photo numériques.

Un fichier vidéo en AVI peut être encodé avec un codec quelconque: DivX est souvent utilisé, MPEG2/4, etc. et pour l'audio: mp3, PCM, etc .

Il permet de réunir en un seul fichier une piste vidéo et 99 pistes audio au maximum, ce qui permet de bénéficier, par exemple, de plusieurs langues pour un même film.

Les conteneurs

- **QuickTime :**

C'est un format souple, crée par Apple puis diffusé gratuitement sur les autres plateformes. Il peut contenir des pistes audio, vidéo, et des textes (pour les sous titres). Une piste peut être également un stream (diffusion en temps réel par internet).

Les conteneurs

- **ASF (Advanced Streaming Format) :**

Lié à Windows Media, que l'on retrouve plutôt sous forme de fichiers WMV. ASF ne spécifie pas les formats de compression des flux audio et vidéo qu'il contient, mais uniquement la structure de ces flux.

Ce qui implique que les fichiers au format ASF peuvent être encodés avec un très grand nombre de codecs.

Les conteneurs

- **MP4 :**

Quicktime amélioré, souvent utilisé pour les DivX. Très souple, peut contenir des images.

- **3GP :**

Dérivé du MP4 pour les téléphones mobiles.