

Codage Vidéo Scalable

Marco Cagnazzo

Département Traitement du Signal et des Images
TELECOM ParisTech

27 mai 2013

Plan

- 1 Le problème de la scalabilité
 - MJ2K : codeur vidéo scalable
- 2 Scalabilité dans les codeurs hybrides
 - MPEG-2
 - MPEG-4
 - SVC

Plan

- 1 Le problème de la scalabilité
 - MJP2K : codeur vidéo scalable

- 2 Scalabilité dans les codeurs hybrides
 - MPEG-2
 - MPEG-4
 - SVC

Transmission de la vidéo sur réseau

- Optimisation du codage vidéo
- Hypothèses implicites : le codeur connaît :
 - Bande du canal entre codeur et décodeur
 - Puissance de calcul du décodeur
- Donc le codeur peut optimiser la qualité pour un débit R_{\max} donné

Transmission de la vidéo sur réseau

Réseau hétérogènes

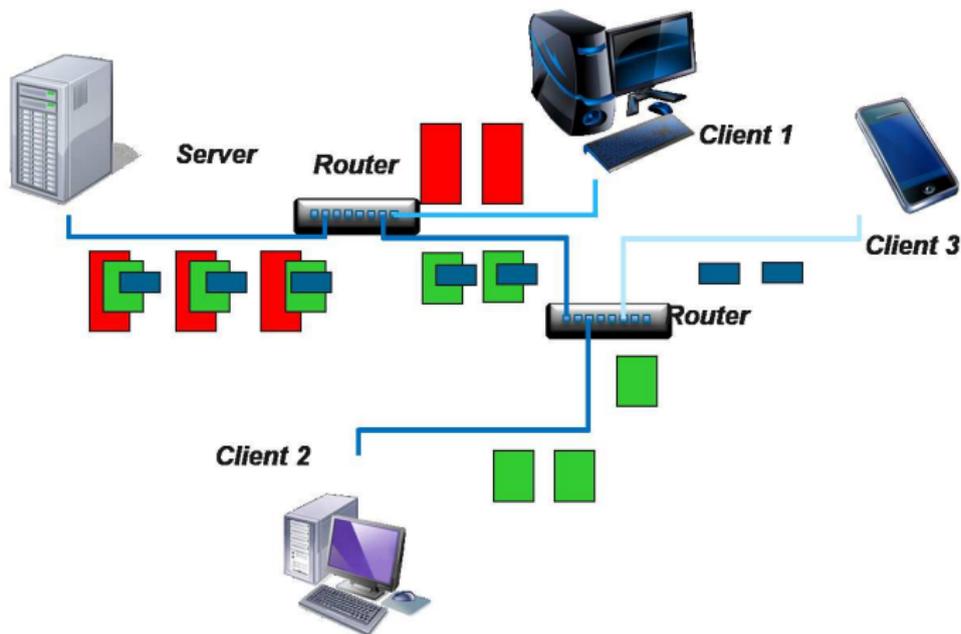
- Utilisateurs
 - Puissance de calcul disponible
 - Résolution demandée
 - Qualité requise
 - Débit d'images par second requis
- Réseau
 - Bande
 - Taux d'erreurs

Transmission de la vidéo sur réseau

Cas *Simulcast*

- Chaque utilisateurs veut accéder à la vidéo avec des caractéristiques différentes
- Solution *Simulcast*
 - Le codeur crée N version différentes de la vidéo
 - On envoie à chaque client la vidéo désirée
 - Solution simple
 - Mauvaise utilisation du réseau
 - Complexité au codeur

Transmission Simulcast : exemple



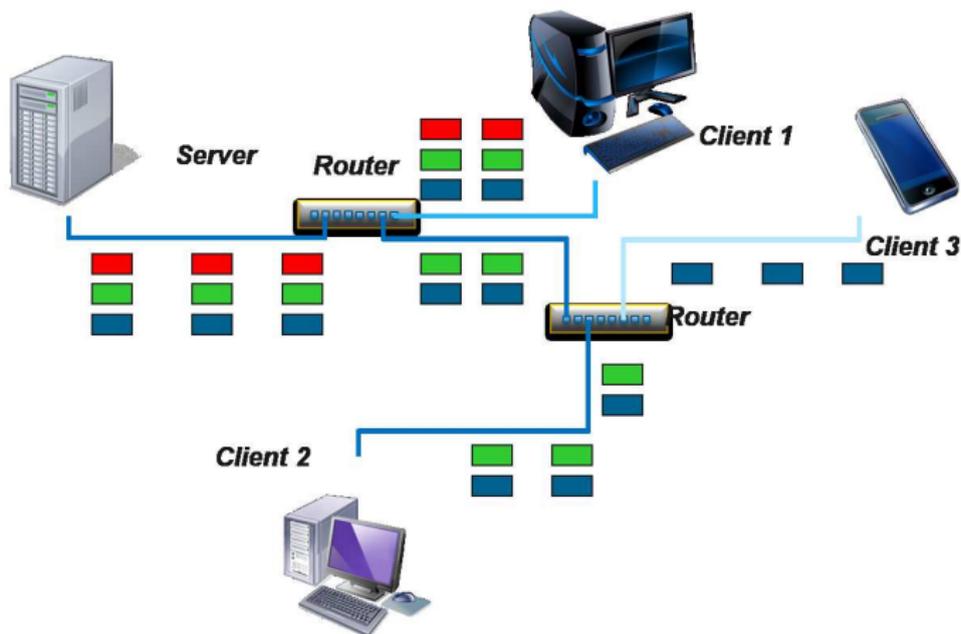
Transmission de la vidéo sur réseau

Cas *scalable*

- Un seul codage par couches (*layers*)
- Couche *base* : qualité (résolution spatiale, temporelle) minimale assurée
- Couches de détails (*enhancement*) : amélioration
- Représentation compacte de la vidéo
 - Un ensemble de vidéo Simulcast n'est pas une vidéo scalable !

Scalabilité : exemple

Cas scalable



Avantages de la scalabilité

- Vidéo codée une seule fois
- Utilisation optimale du réseau
- Chaque utilisateur a l'information voulue
- Robustesse aux erreurs
 - Si les détails sont perdus, pas d'impact sur le niveau base
 - Mais si le niveau base est perdu ?
- Débit de décodage : choisi après codage

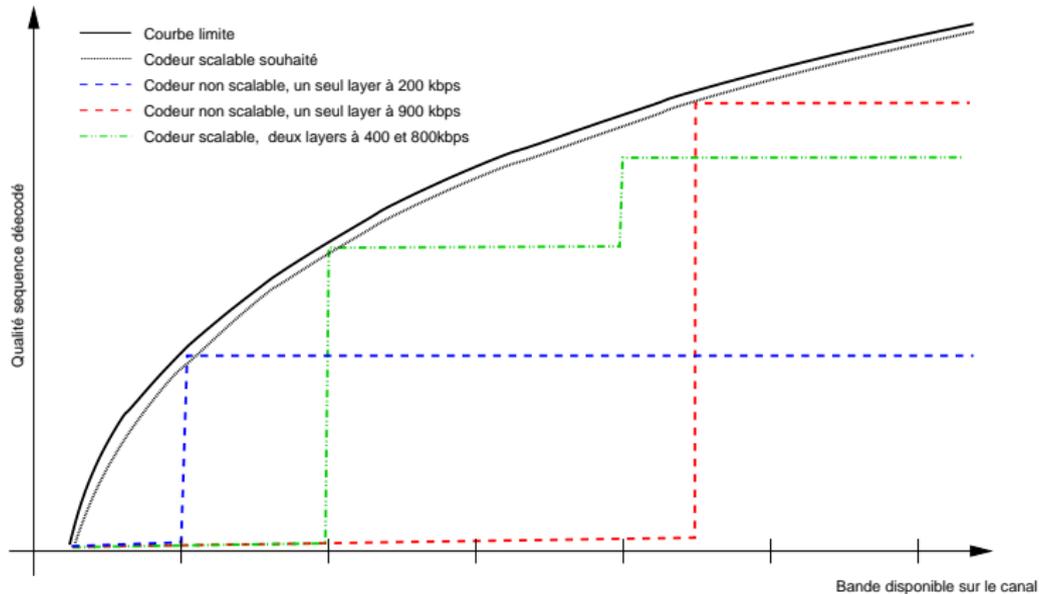
Inconvénients de la scalabilité

- Augmentation de la complexité
- Perte des performance de compression
 - Une vidéo scalable décodé à débit R peut avoir une qualité inférieure au cas non-scalable au même débit
- Augmenter le nombre de niveaux normalement empire les pertes
- Souvent il n'est pas possible de coder la vidéo sans connaître les débits cibles

Types de scalabilité

- Scalabilité en qualité (ou en SNR)
- Scalabilité dans l'espace (ou en résolution)
- Scalabilité dans le temps (ou en frame-rate)
- Scalabilité par objets

Codeurs scalables et non scalables



Scalabilité en résolution et en qualité

Exemple



MJP2K : codeur vidéo scalable

- Pas de ME/MC : faible complexité, faible taux de compression
- Chaque image codée avec JPEG2000
- Partie 3 du standard JPEG2000
- Utilisé pour le cinéma numérique et la vidéo communication en temps réel (video chat)
- Scalabilité
 - En qualité : obtenue grâce aux couches de JP2K
 - En résolution : obtenue grâce au *codeblock* de JP2K
 - En temps : scalabilité totale, car chaque image est codée séparément
- Limite important : taux de compression faible

Plan

- 1 Le problème de la scalabilité
 - MJ2K : codeur vidéo scalable

- 2 Scalabilité dans les codeurs hybrides
 - MPEG-2
 - MPEG-4
 - SVC

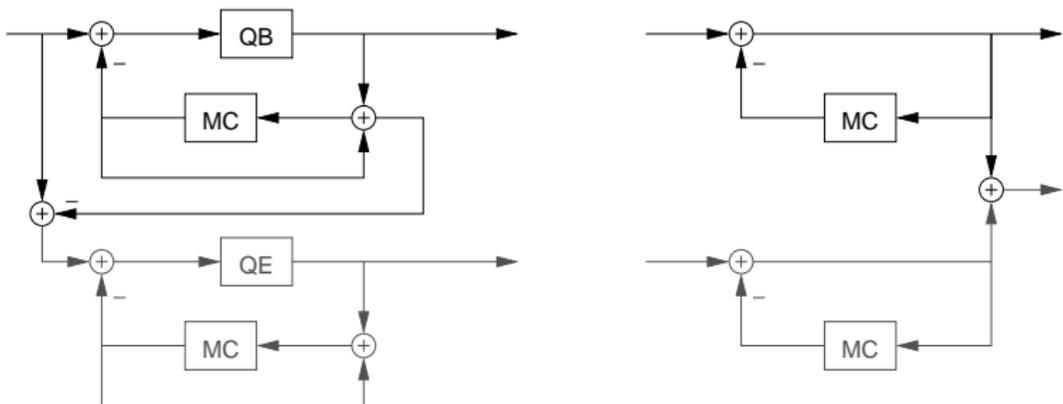
La scalabilité dans les codeurs hybrides

Le dilemme base/détails

- Comment effectuer les prédictions en sorte d'avoir le meilleures performances ?
 - 1) À niveau N , on code la vidéo seulement avec les informations des niveaux $1, \dots, N$
 - 2) À niveau N on code la vidéo avec toute l'information de tout niveau
- Dans le cas 1 le niveau base est codé à la même qualité qu'il aurait dans le cas Simulcast ; le niveau enhanced peut perdre de la qualité car il est prédit par la base, quand il pourrait être prédit par toute l'info
- Dans le cas 2 le niveau enhanced a la meilleure qualité ; le niveau base peut se desynchroniser (drift)

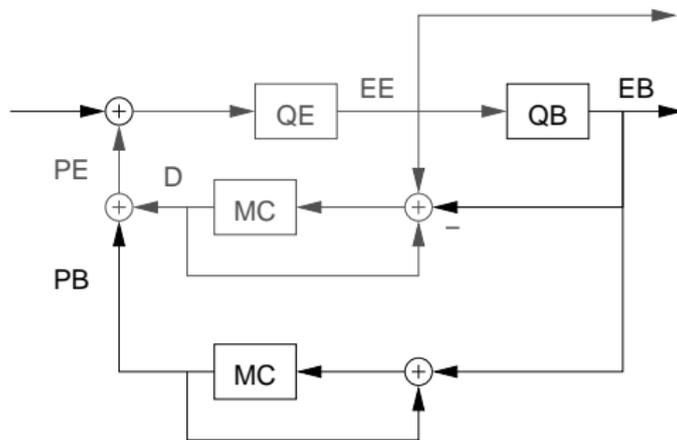
La scalabilité dans les codeurs hybrides

Schéma avec 2 anneaux de prédiction : pas de drift



La scalabilité dans les codeurs hybrides

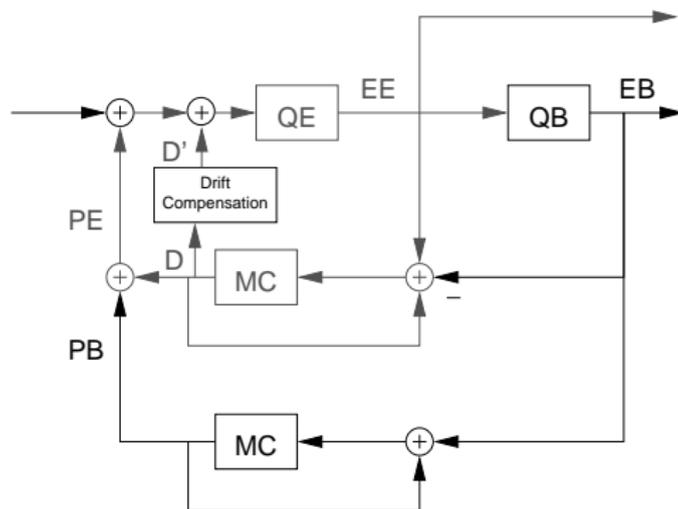
Schéma avec drift



- D n'est pas disponible pour un décodeur base
- Cela entraîne un *drift* entre codeur et décodeur
- Le niveau enhanced est codé avec toute l'info disponible

La scalabilité dans les codeurs hybrides

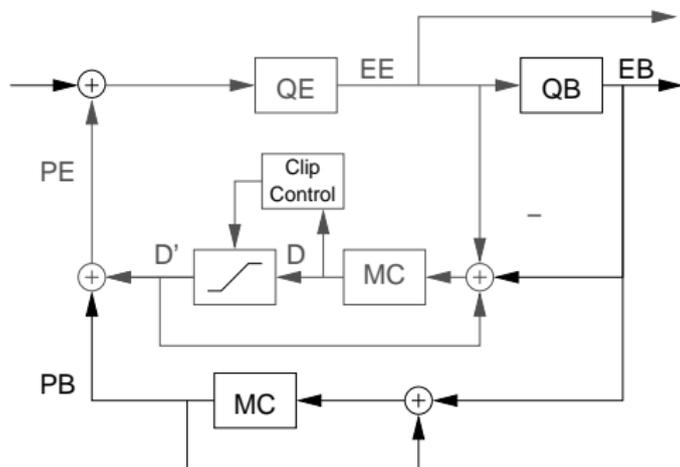
Schéma avec contrôle du drift



- Si $D' = -D$: pas de drift, cas 1)
- Si $D' = 0$: free drift, cas 2)
- Le décodeur n'est pas obligé à connaître D'
- Choix de D' : compromis entre qualité base et enhanced

La scalabilité dans les codeurs hybrides

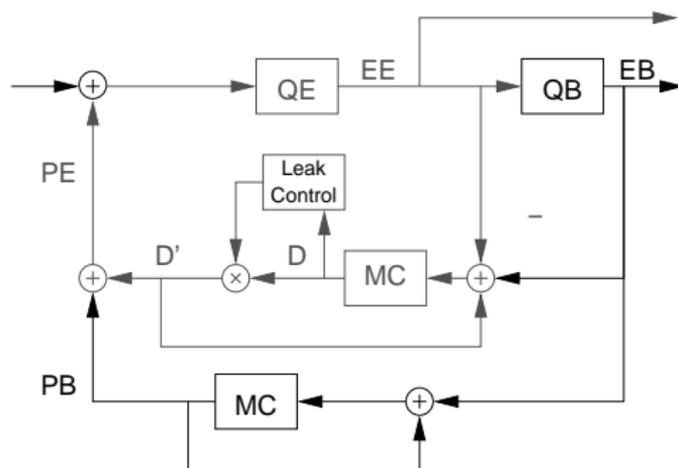
Stratégies de contrôle du drift : Clip



- Le drift est limité à une valeur D_{\max}
- Le décodeur doit effectuer le même seuillage
- Performance meilleure du codeur "multiple-loop"
- Si $D_{\max} = 0$, pas de drift
- Si $D_{\max} = \infty$, free drift

La scalabilité dans les codeurs hybrides

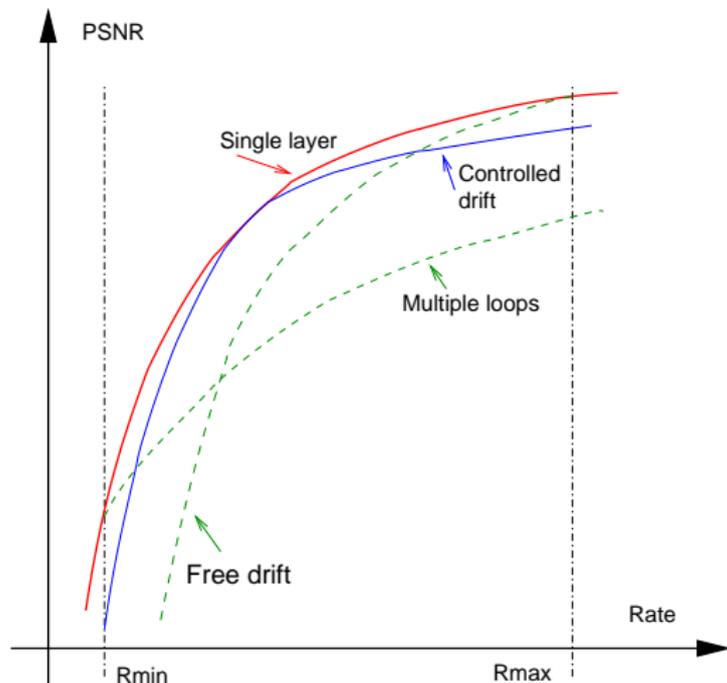
Stratégies de contrôle du drift : Leak



- Le drift est multiplié par un facteur $\alpha \leq 1$
- Le décodeur doit effectuer la même mise à l'échelle
- Performance meilleure du codeur "multiple-loop"
- Si $\alpha = 0$, pas de drift
- Si $\alpha = 1$, free drift

La scalabilité dans les codeurs hybrides

Performance du contrôle du drift



La scalabilité dans les codeurs hybrides

Résumé

- La scalabilité en SNR a toujours un impact sur les performances de compression dans les codeurs hybrides
- Cela est dû à la structure récursive et au drift
- La perte en performance dépend de la vidéo
- Le contrôle du drift permet d'optimiser le point de travail du codeur

La scalabilité en MPEG-2

- MPEG-2 est le premier standard de codage vidéo qui supporte la scalabilité
- Scalabilité en SNR (ou en qualité)
- Scalabilité spatiale (ou en résolution)
- Scalabilité temporelle (ou en frame-rate)

Standard MPEG-2

Profilscalables

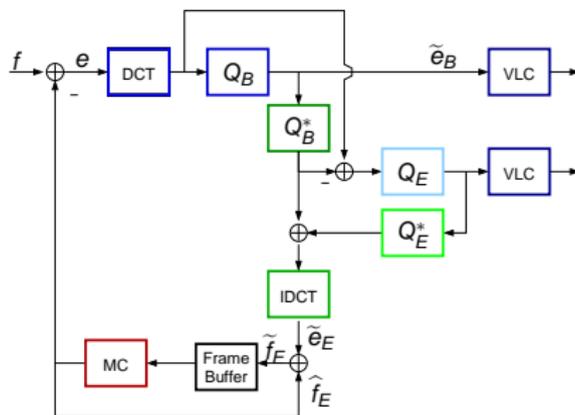
Le scalabilité est disponible seulement dans certains profils

Niveau	width [pixel]	height [pixel]	frame rate [frame/s]	bit rate [Mbps]
Low	352	288	30	4
Main	720	576	30	15
High-1440	1440	1152	60	60
High	1920	1152	60	80

Profil	fonctionnalités
Simple	Pas de scalabilité; vidéo entrelacée; pas de trames B
Main	Simple + Trames B
SNR scalable	Main + Deux/trois niveaux de scalabilité en qualité
Spatial scalable	SNR + Deux/trois niveaux de scalabilité en résolution
High	Space + Chroma sur-échantillonnée

Standard MPEG-2

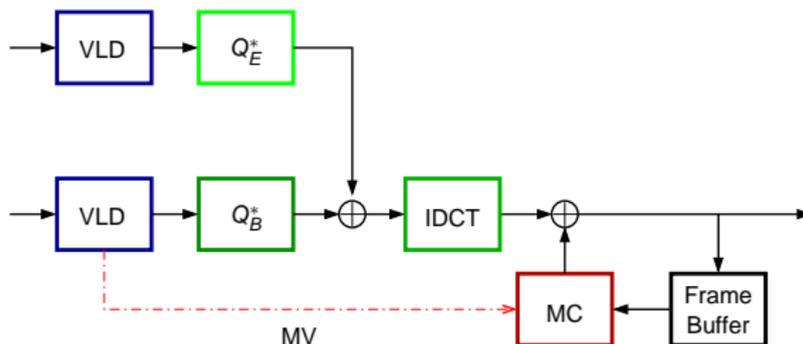
Scalabilité en SNR : codeur



- Raffinement des coefficients DCT
- Pas de contrôle du drift
- Bonne qualité du niveau enhanced
- Qualité du niveau base pas toujours satisfaisante

Standard MPEG-2

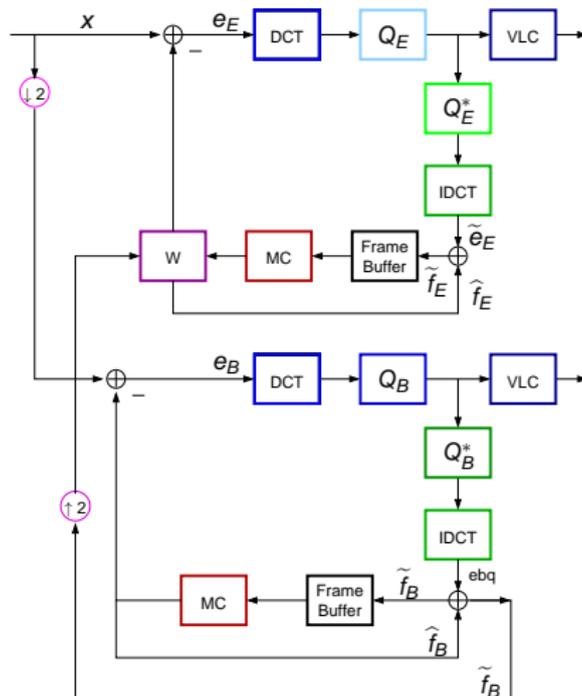
Scalabilité en SNR : décodeur



- Pas de contrôle du drift
- Un seul champ de vecteurs de mouvement

Standard MPEG-2

Scalabilité en résolution : codeur



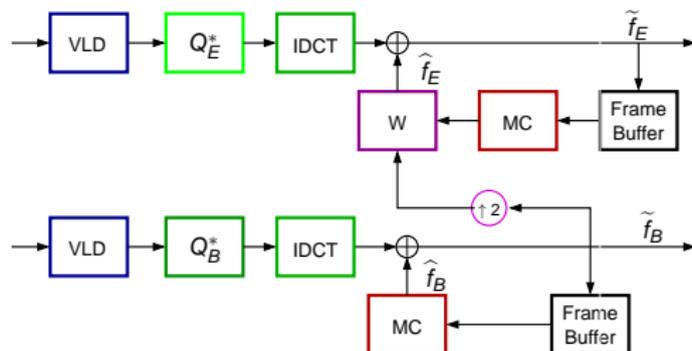
Standard MPEG-2

Scalabilité en résolution : codeur

- Double boucle : pas de drift
- La vidéo d'entrée est filtrée et sous-échantillonnée
- La prédiction de niveau enhanced est la somme pondérée de :
 - L'image de niveau base interpolée
 - La prédiction ME/MC
- Le choix du poids est faite MB par MB et cette info est codé dans le train binaire

Standard MPEG-2

Scalabilité en résolution : décodeur



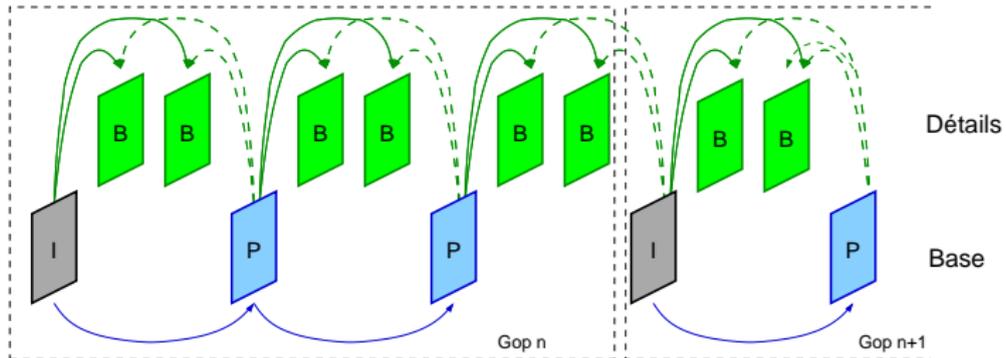
Standard MPEG-2

Scalabilité en résolution : décodeur

- L'image à basse résolution est interpolée et pondérée
- Interpolation linéaire
- On obtient les informations nécessaires pour le décodage du niveau enhanced

Standard MPEG-2

Scalabilité en temps



Standard MPEG-2

Scalabilité hybride, 1/2

- Les formes de scalabilité disponibles peuvent être combinées entre elles, jusqu'à 3 niveaux en total
- Trois type de scalabilité hybride sont possibles
- SNR + spatiale
 - 1 SDTV/CIF, basse qualité
 - 2 HDTV/SDTV, basse qualité
 - 3 HDTV/SDTV, haute qualité

Standard MPEG-2

Scalabilité hybride, 2/2

- spatiale + temporelle
 - 1 SDTV entrelacée
 - 2 HDTV entrelacée
 - 3 HDTV progressive
- SNR + temporelle
 - 1 HDTV entrelacée, basse qualité
 - 2 HDTV entrelacée, haute qualité
 - 3 HDTV progressive, haute qualité

Standard MPEG-2

Data partitioning

- Outil pour la robustesse aux erreurs
- Les blocs de 64 coefficients DCT sont divisés en sous-ensembles (DC+basses fréquences, hautes fréquences)
- Le sous-ensemble plus critique peut être envoyé sur un canal plus robuste (Unequal error protection)
- Il est nécessaire que le décodeur supporte le Data partitioning (pas de *backward compatibility*)

Le standard MPEG-4

- Développé en 1993-1998
- Parties du standard
 - 5 parties principales (comme MPEG-1 et 2)
 - 18 parties supplémentaires, couvrant une variété de problèmes
 - P.e. MPEG4/part 10 coïncide avec H.264/AVC

Standard MPEG-4

Scalabilité

- Frame-rate et résolution : comme en MPEG-2
- Qualité : *fine grain scalability* grâce au codage par plans de bits
- Objet : la scène peut être composée en utilisant un sous-ensemble des AVOs disponibles

Standard MPEG-4

Scalabilité spatiale

- Deux anneaux de prédiction
- Décodage base indépendant du niveau supérieur
- Permet d'utiliser comme référence pour la prédiction
 - L'image à basse résolution
 - L'image précédente à haute résolution
- Le choix est effectué comme dans le cas de MC bidirectionnelle

Standard MPEG-4

Fine grain scalability

- Scalabilité en qualité de MPEG-2 : perte de performance a basse qualité ou à haute qualité
- Problème lié aux boucles de ME/MC
- Avec la FGS on a toujours 2 layers
- Mais le deuxième peut être coupé n'importe où avec un amélioration des performances partiale

Standard MPEG-4

Fine grain scalability

- Les plans de bits des coefficients DCT sont codés séparément
- Le décodage de chaque plan de bit correspond à une quantification plus fine des coefficients
- La représentation des coefficients est donc plus précise quand on dispose de un nombre supérieur de plans de bits
- Les plans de bit peuvent être codé par un codeur entropique qui en exploite la corrélation
- Le codage par plan de bits est plus efficace que le codage par run-length et valeur
 - Jusqu'au 20% de débit en moins pour QP élevé
 - Les plans de bit ont des statistiques indépendantes de QP par contre dans le codage classique il est impossible de trouver des tableau VLC adaptés à toute situation

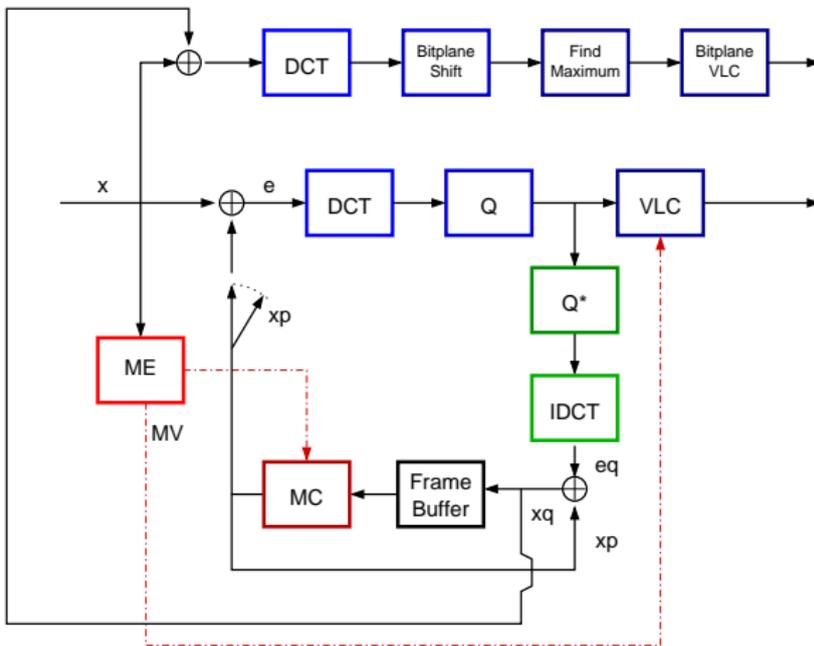
Standard MPEG-4

Fine grain scalability

- Base layer et enhancement layer (EL)
- Base layer codé avec le codeur non-scalable
- L'EL contient la difference entre vidéo originale et base layer
- L'EL est codé par plans de bits
- L'EL peut être coupé à n'importe quelle position
- Le décodeur doit être capable de décoder un EL coupé à une position arbitraire
- La qualité de l'EL augmente pour chaque nouveau bit décodé

Standard MPEG-4

Fine grain scalability : codeur



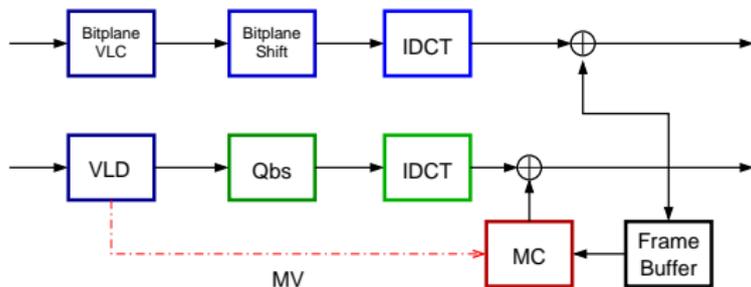
Standard MPEG-4

Fine grain scalability : codeur

- Nombre différent de plan de bit pour chaque composante couleur : trouver le maximum pour Y, U et V
- Codage par plan de bit : quatre tableau VLC pour le MSB0, MSB1, MSB2 et les autres
- Les statistiques des planes de bits ont été étudiées avec grand soin pour optimiser les performances des VLC

Standard MPEG-4

Fine grain scalability : décodeur



Standard MPEG-4

Fine grain scalability : performance

- FGS vs Multilayer SNR scalability
 - Scalabilité en résolution avec upsampling égal à 1
 - Open-loop scalability
 - FGS a un gain de 2dB en PSNR (en moyenne)
- FGS vs non scalable : le coût de la scalabilité peut arriver à 2 dB
- FGS vs Simulcast
 - Simulcast : on génère N versions codée de la vidéo
 - La courbe débit/PSNR du simulcast monte plus lentement et par “marches”

Standard MPEG-4

Fine grain scalability : advanced features

- Frequency weighting (qualité visuelle)
- Selective enhancement (ROI)
- Robustesse aux erreurs : marqueurs à chaque bit-plane
- Scalabilité temporelle et FGS

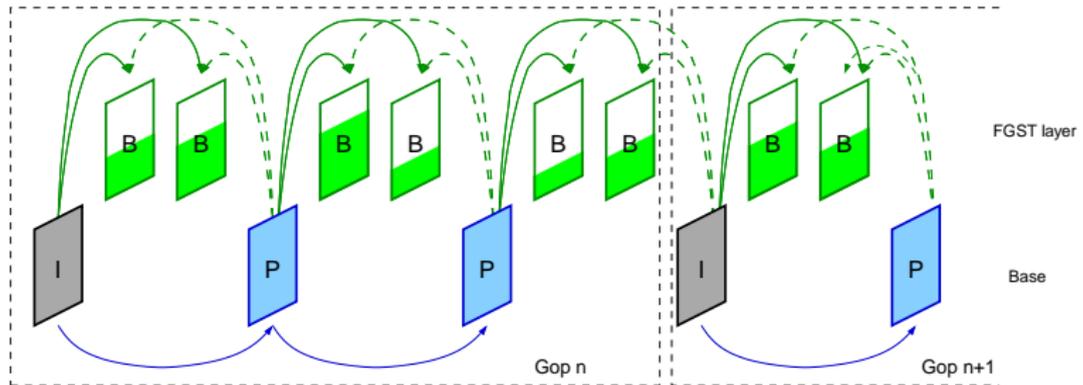
Standard MPEG-4

Fine grain scalability / Temporal

- Scalabilité hybride avec l'axe temporel
- FGST : les images du layer temporel supérieur sont codées directement en FGS
- En conséquence, n'importe quelle partie du layer supérieur peut être utilisé
- Réduction de la complexité

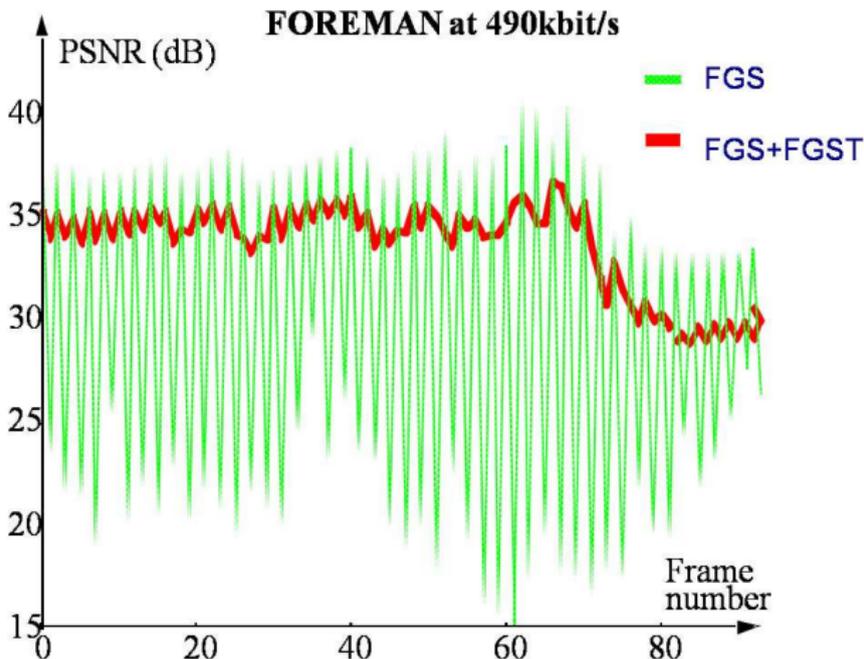
Standard MPEG-4

Fine grain scalability / Temporal



Standard MPEG-4

FGST rate control



Le standard H.264/SVC

Scalable Video Coding

- Extension scalable de H.264/AVC
- Développé en 2003-2007
- Joint Video Team (ITU+ISO/IEC)
- Codeur hybride
- Extension scalable de H.264 : niveau base compatible avec H.264/AVC

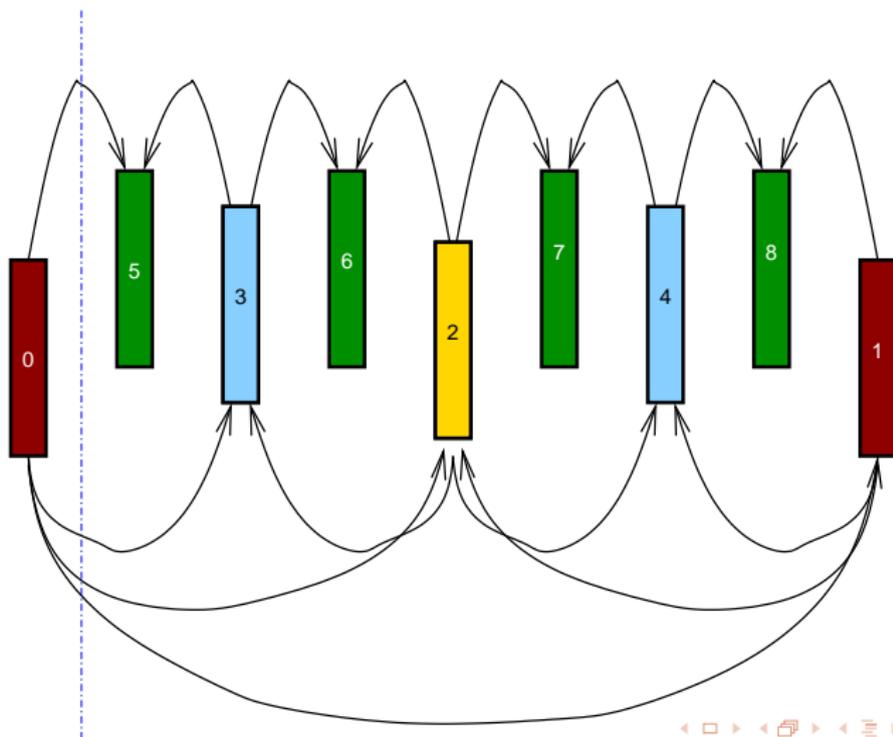
Le standard H.264/SVC

Scalabilité temporelle

- Il suffit d'organiser les trames B de H.264 dans une structure hiérarchique
- La prédiction est toujours faite à partir d'une image du même niveau ou de niveau inférieur
- Les niveaux supérieurs ne sont pas nécessaires pour le décodage

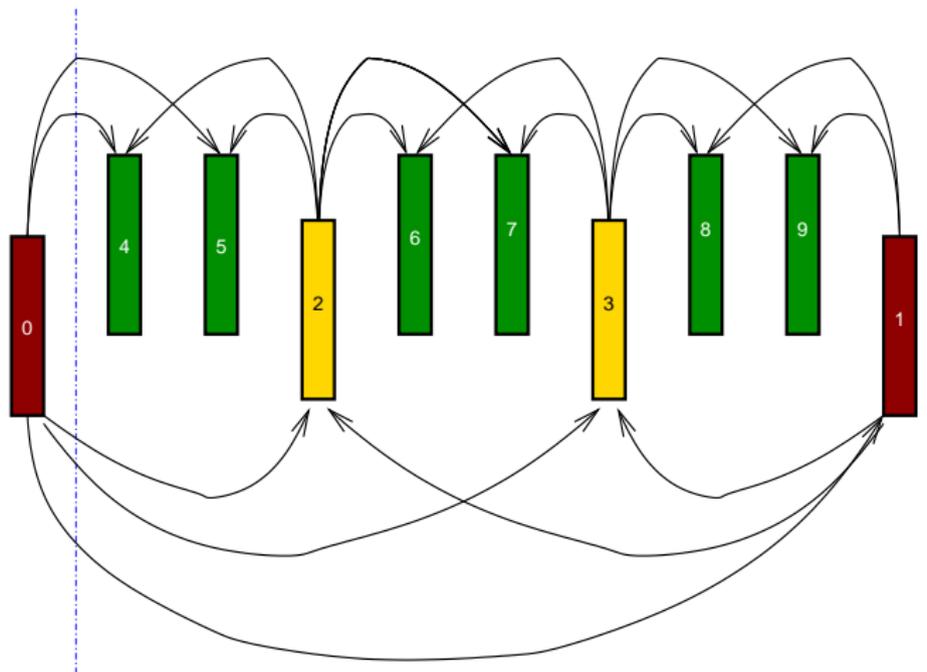
Le standard H.264/SVC

Scalabilité temporelle : structure hiérarchique



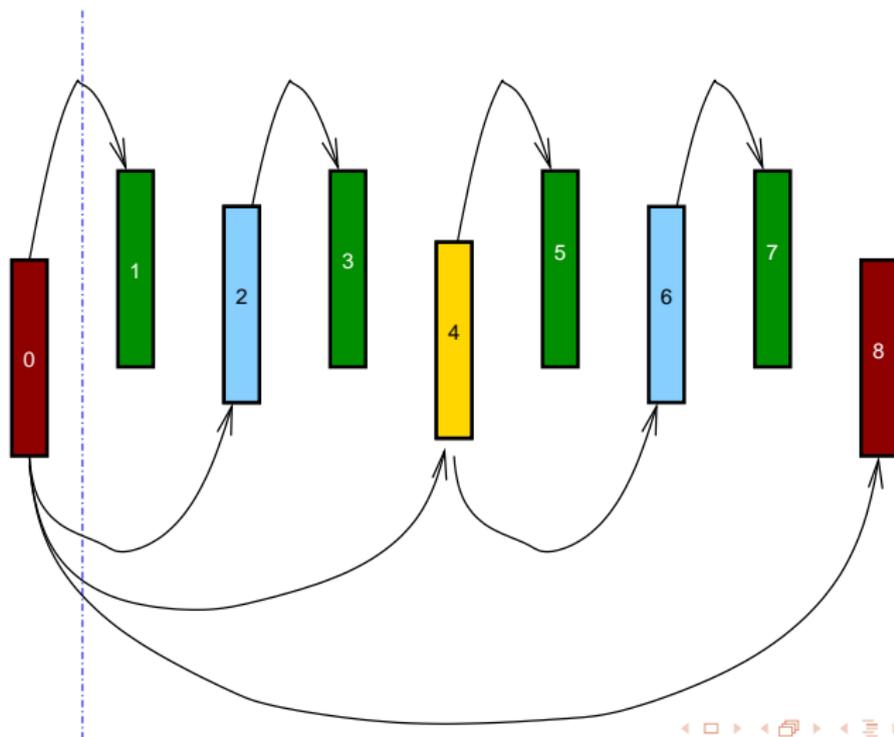
Le standard H.264/SVC

Scalabilité temporelle : structure non-dyadique



Le standard H.264/SVC

Scalabilité temporelle : structure sans retard



Le standard H.264/SVC

Scalabilité temporelle : performance

- Performance de compression : choix des pas de quantification
- Niveaux inférieurs finement quantifiés car ils servent de référence aux niveaux supérieurs
- Les valeurs optimales des pas de quantification dépendent du signal
- Formule simple pour le pas de quantification à niveau T :

$$QP_T = QP_0 + 3 + T$$

- Meilleurs résultats si le retard est admis

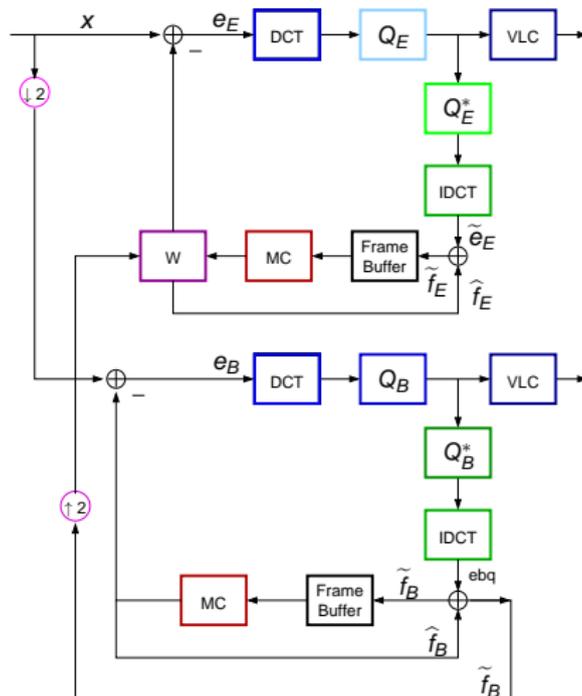
Le standard H.264/SVC

Scalabilité spatiale

- Schéma pyramidale comme dans MPEG-2
- Nouveauté technique : la prédiction *inter-layer*
 - Les paramètres de codage d'un layer peuvent être prédits de un layer inférieur
 - Inter-layer motion prediction
 - Inter-layer intra et residual prediction
- Scalabilité spatiale généralisée : le niveau supérieur peut avoir n'importe quelle résolution
- Certaines contraintes sont imposées sur la prédiction inter-layer pour réduire la complexité

Le standard H.264/SVC

Scalabilité spatiale : codeur



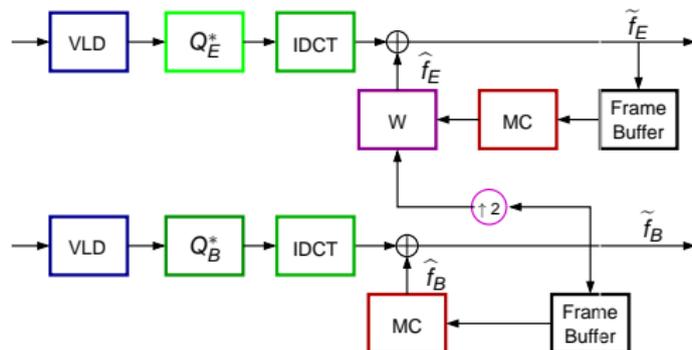
Le standard H.264/SVC

Scalabilité spatiale

- Double boucle : pas de drift
- La vidéo d'entrée est filtrée et sous-échantillonnée
- La prédiction de niveau enhanced est la somme pondérée de :
 - L'image de niveau base interpolée
 - La prédiction ME/MC
- Le choix du poids est faite MB par MB et cette info est codé dans le train binaire

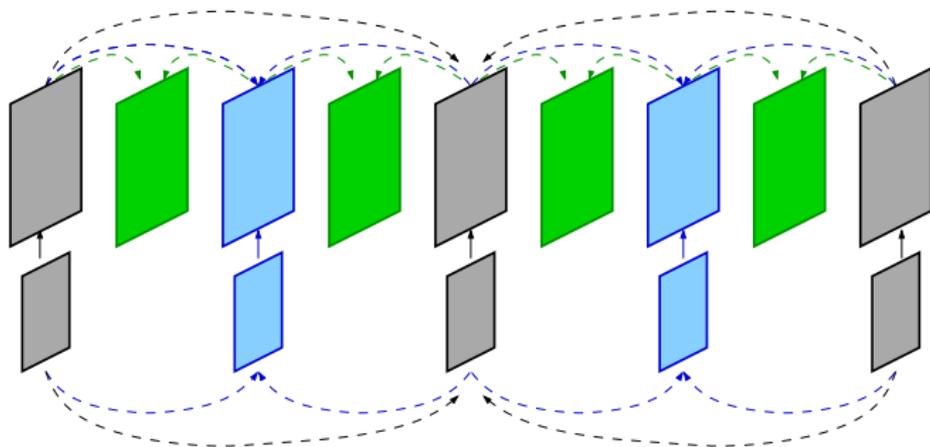
Le standard H.264/SVC

Scalabilité spatiale : décodeur



Le standard H.264/SVC

Scalabilité spatiale et temporelle



Le standard H.264/SVC

Scalabilité en qualité

- Coarse grain scalability
- Medium grain scalability

Le standard H.264/SVC

Scalabilité en qualité : CGS

- Implémentée comme la scalabilité résolution, mais sans changer d'échelle entre niveaux
- Inter-layer Intra Prediction et Inter-layer Residual Prediction dans le domaine transformé
- Pas de drift
- Petit nombre de niveaux (courbe à marche)

Le standard H.264/SVC

Scalabilité en qualité : MGS

- On peut définir des point d'accès aux niveaux de scalabilité
- On introduit du drift pour gérer le compromis entre performances des différents niveaux
- On peut choisir entre les extrêmes : pas de drift (comme en MPEG-4/SNR) et drift libre (MPEG-2/SNR)
- Key picture : pas de drift au niveau plus bas, drift admis au niveaux supérieurs

Conclusion

- Scalabilité : propriété importante pour la diffusion de la vidéo sur réseau
- Scalabilité en SNR, résolution, frame-rate
- Problèmes de drift avec les codeurs hybrides
- Pas de drift avec la TO, mais pas de compatibilité non plus
- SVC utilise un codeur hybride, avec compatibilité H.264

Conclusion

Références

- Weiping Li, “Overview of Fine Granularity Scalability in MPEG-4 Video Standard”. IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 3, March 2001.
- Jens-Rainer Ohm, “Advances in Scalable Video Coding”. Proceedings of the IEEE, vol. 93, no. 1, January 2005.
- Heiko Schwarz, Detlev Marpe, and Thomas Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard”. IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, September 2007
- Al Bovik Ed., “Handbook of Image and Video Processing” Texas Academic Press , Austin