



Institut
Mines-Télécom



Journée Futur et Ruptures

Communications

immersives :

Enjeux et perspectives

Béatrice Pesquet-Popescu

Télécom ParisTech, Département TSI

5 mars 2015



Tendances actuelles

■ Plus, plus, plus ...

- Résolutions d'écran (Digital Cinema: 4K x 2K)
- Ultra High Definition (UHD : 8K x 4K) en plein développement
- Résolutions HD pour les mobiles (720p, mais aussi 1080p)
- High dynamic range (HDR)
- Echantillonnage couleur 4:4:4
- High Frame Rate (HFR) : 120-600fps
- 3D , multi-vues, free viewpoint video, ...



**Cameras et écrans prêts pour ce “saut”
L’infrastructure de transmission peine à assurer
les débits associés !**

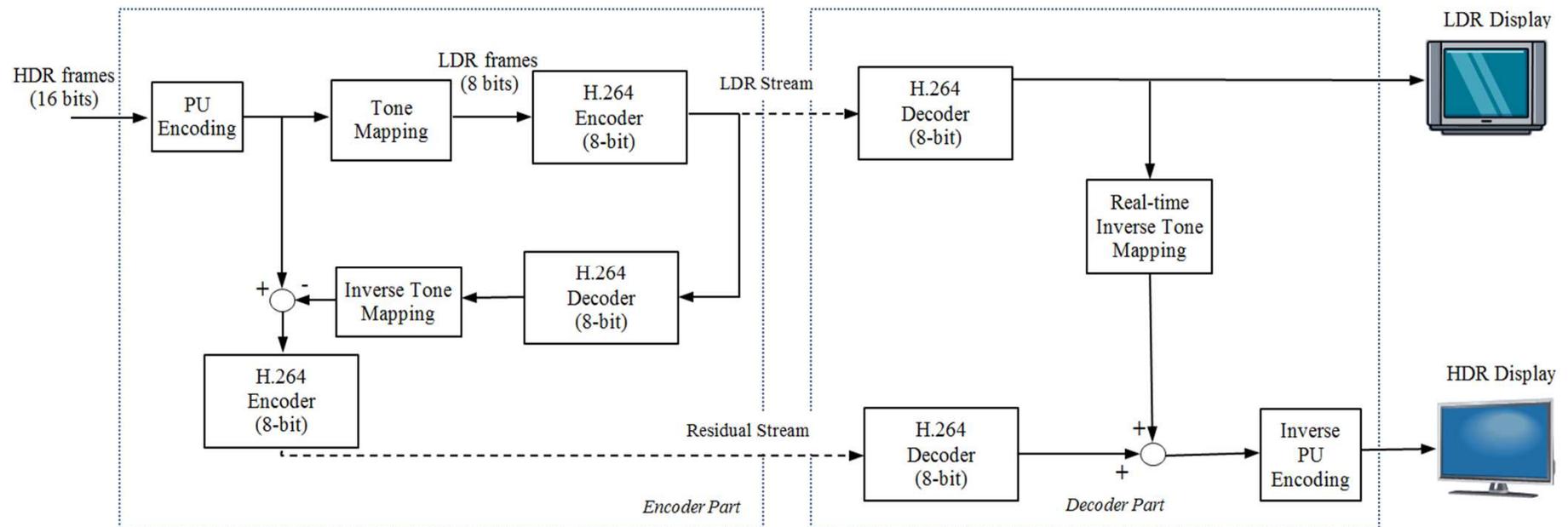


Expérience Multimédia Immersive

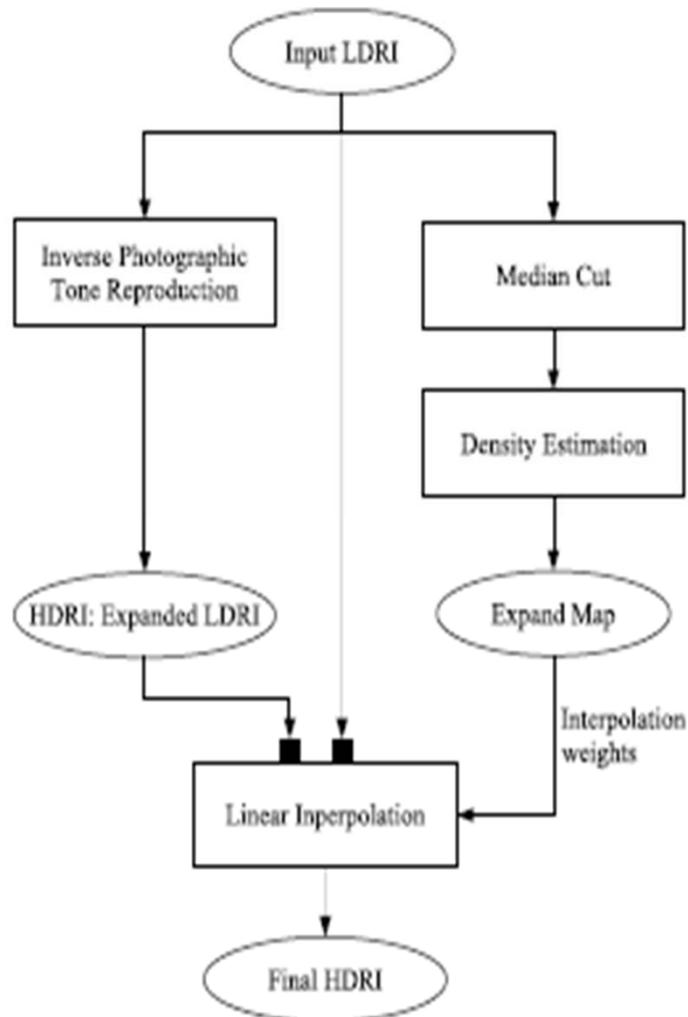
- Les avantages de chaque amélioration dépendent du type de contenu, d'écran, d'environnement de visualisation
- Quels développements futurs de ces formats vidéo conduiraient à une qualité d'expérience maximale pour l'utilisateur ?
- Comment la chaîne de la production à la consommation peut être modifiée pour supporter ces nouveaux formats ?
- Comment ces formats doivent être utilisés (convertis) pour correspondre aux capacités des différents réseaux et dispositifs de visualisation ?

High dynamic range video

- Luminosité des scènes réelles varie de 10^{-6} cd/m² (noir) à 10^8 cd/m² (soleil brillant)
- Représenter les valeurs de la luminosité (cd/m²) pour générer une image indépendante de l'écran utilisé pour la visualisation
- Représentation 16bits ou plus : nécessite une compression efficace
- Problème de retrocompatibilité/adaptation à la dynamique de l'écran



High dynamic range: Inverse Tone Mapping

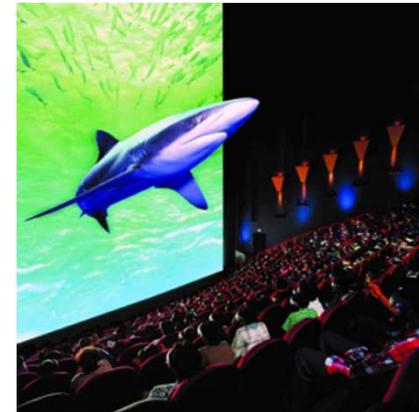


■ Cohérence temporelle des traitements

- Estimation de mouvement
- Critère global de distorsion sur GOP, régularisation non locale
- Approche d'optimisation convexe

Vidéo 3D

- Intérêt constant depuis ...100 ans !
 - Amélioration de l'expérience utilisateur
 - Perçue comme un avancement technologique majeur
 - Encore des barrières technologiques et d'usages ...
- Applications et enjeux :
 - Télévision 3D, théâtres 3D (cinéma)
 - Téléprésence immersive
 - Vision par ordinateur (jeux, réalité virtuelle et augmentée, ...)
 - Médecine, ingénierie, ...

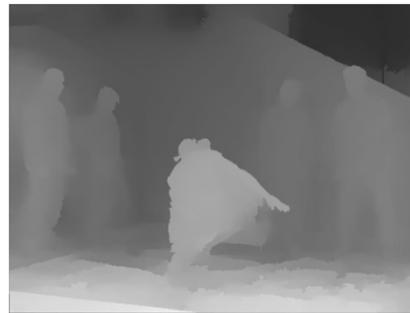


Acquisition vidéo 3D : le cas stéréo

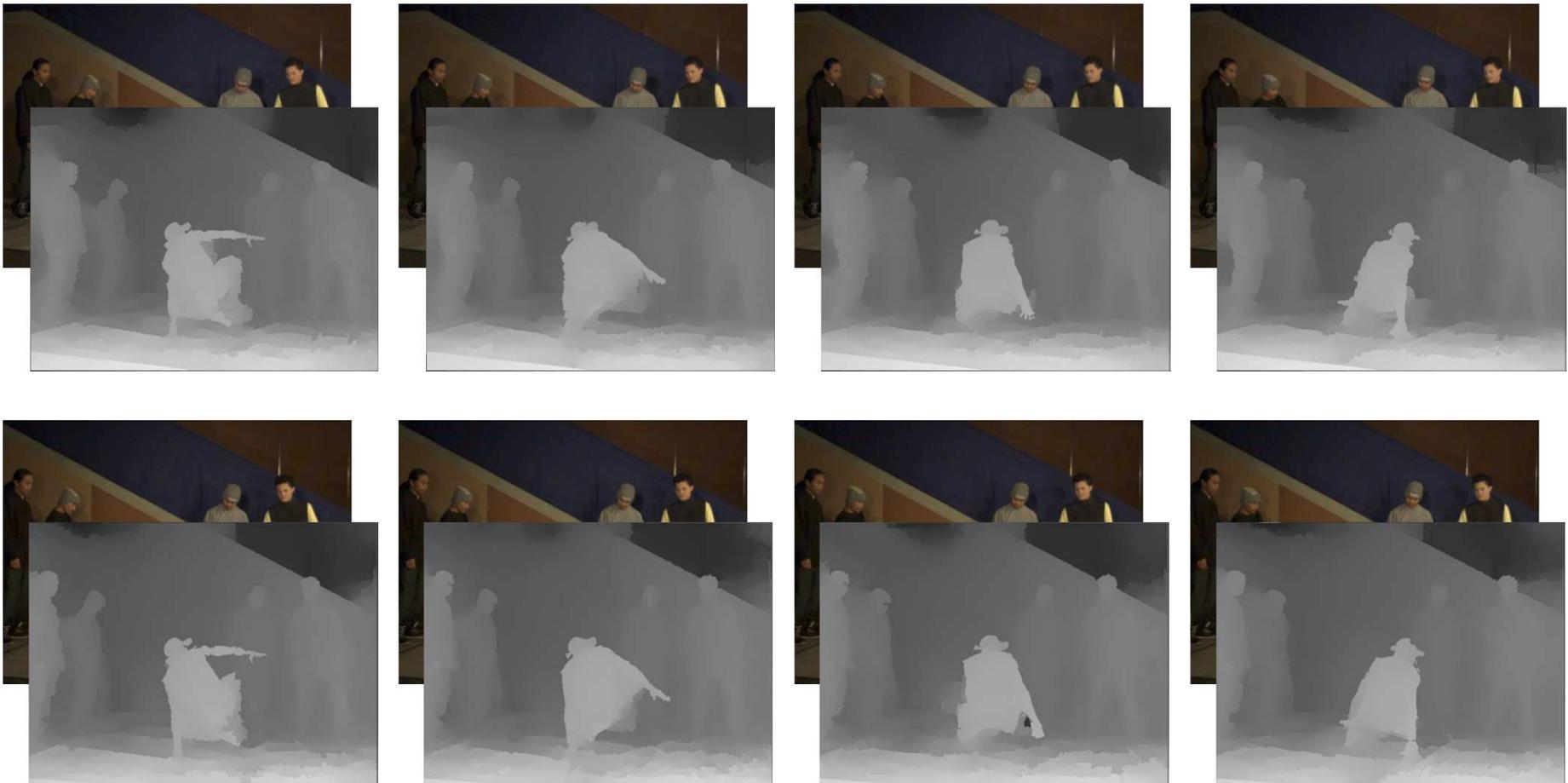




Acquisition vidéo 3D : texture 2D + profondeur



Acquisition Vidéo 3D : multi-vues + multi-profondeurs (MVD)



Représentations vidéo 3D

stéréo



gauche



droite

vidéo + profondeur



texture



profondeur

multivues



vue 1

...



vue N

multivues + profondeur



vue 1

...



vue N

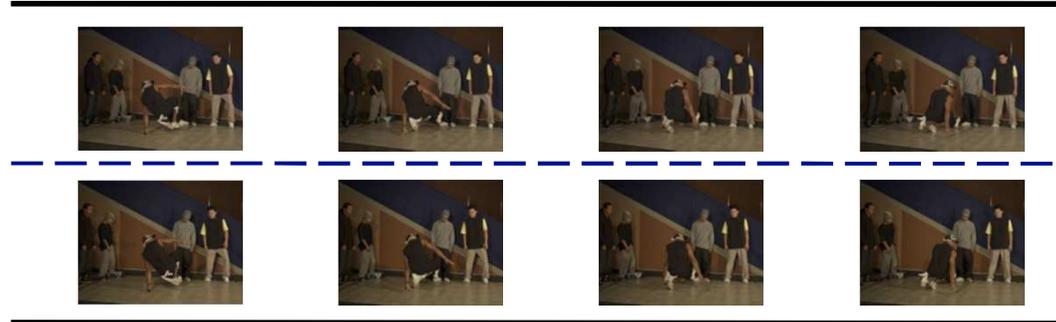
4D visual feature processing

- Vers une génération et consommation massive du contenu vidéo 4D (3D+temps)
 - Comment rechercher et analyser ce type de données ?
- Descripteurs visuels
 - Maturité des descripteurs 2D (standardisation CDVS, CDVA)
 - Plusieurs représentations possibles pour 3D + temps (texture+profondeur, multivues, *layered depth video*, maillages variant dans le temps, ...)
 - **Attributs indépendants de la représentation ?**
- **Impact: recherche par le contenu visuel, détection d'objets et d'événements, classification de scène, systèmes de recommandation ...**
- **Définition, extraction et codage efficace des descripteurs 4D**

Codage de la vidéo 3D : choix de la représentation ?



- Encodage séparé de chaque vue (simulcast)



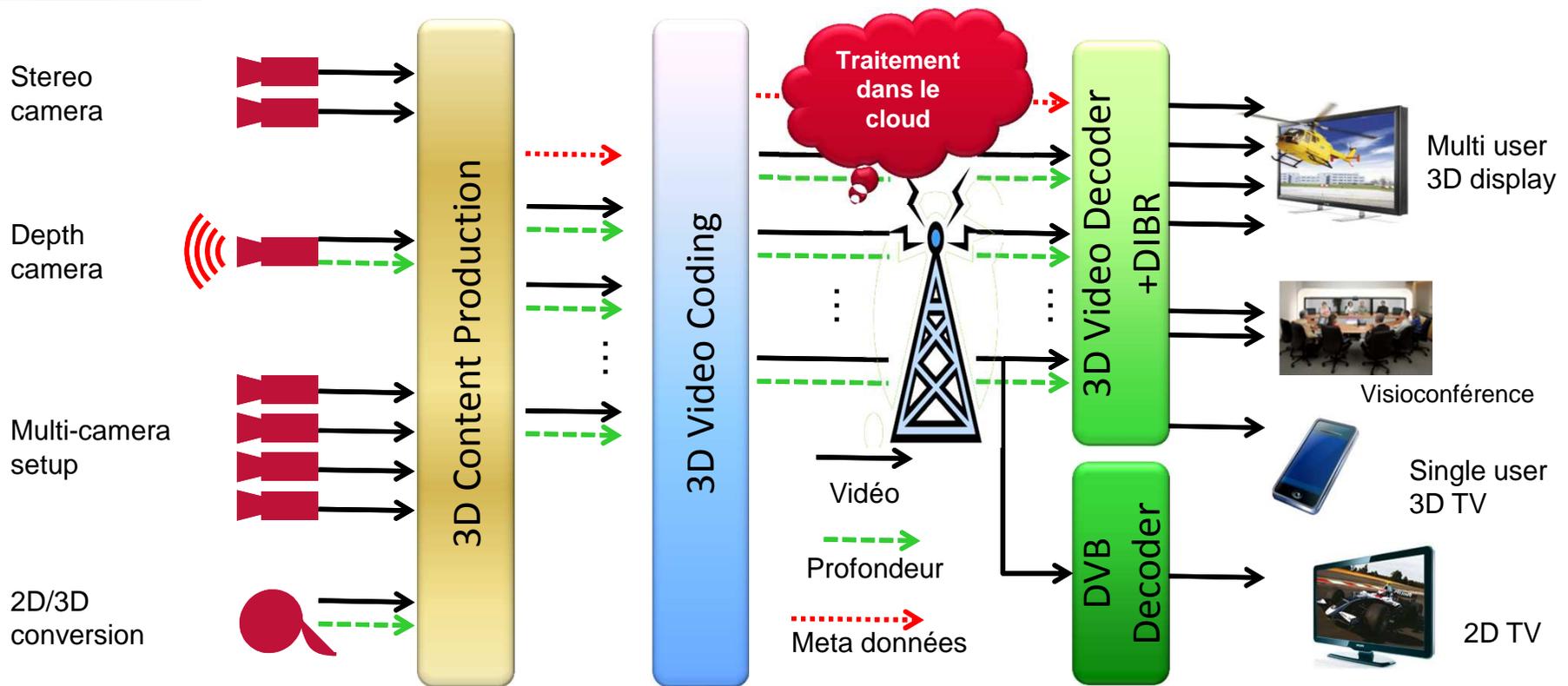
- Encodage conjoint de plusieurs vues
 - Prédiction de la vue courante



- Encodage d'une vue et de la profondeur associée
 - Prise en compte des dépendances texture-profondeur (estimation et codage)



Eco-système de la vidéo 3D



Enjeux/Verrous :

- Nouveaux systèmes d'acquisition/transformation des contenus existants
- **Bande passante**
- Compatibilité arrière pour l'affichage
- **Formats de transmission**
- **Adaptation au décodeur et/ou dans le cloud**

Estimation de la disparité



Image gauche

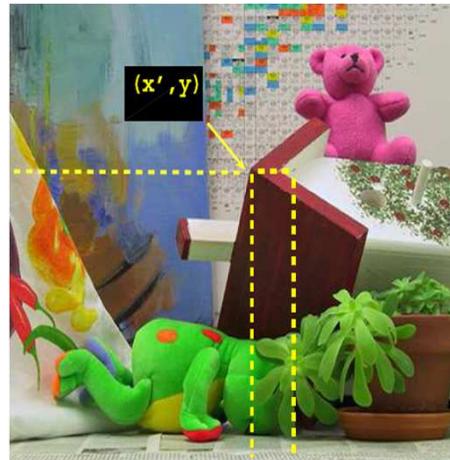
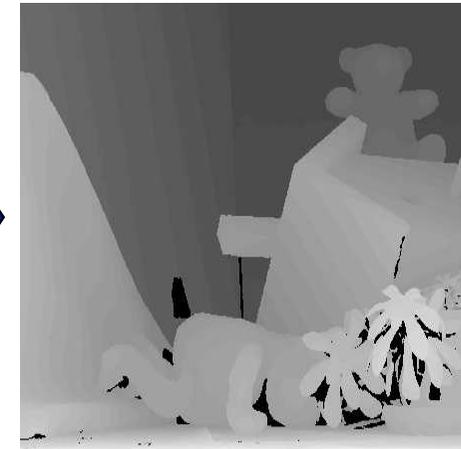
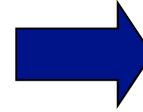


Image droite



Carte de disparité

- **Synthèse de vues virtuelles - fluidité de la visualisation**
- **Mise en correspondance : Estimation d'une carte de disparité dense** (directement liée à la profondeur des objets)
- **Estimation conjointe mouvement-disparité**
- **Estimation simultanée de plusieurs disparités dans un système multi-caméras**
- **Méthodes d'optimisation convexe**



Streaming Vidéo Multivues Interactif

■ Motivation :

- Vidéo multi-vues disponible sur un serveur
- Requêtes dynamiques d'un client pour changer de point de vue

■ Verrous :

- La suite de changements n'est pas connue au moment de l'encodage / n'est pas la même pour tous les utilisateurs !
- Comment exploiter la redondance temporelle ?

■ Interactive multiview video streaming

- Bande passante vs. optimisation de l'espace de stockage
- Efficacité de compression vs. accès aléatoire au contenu
- Streaming robuste sur des réseaux hétérogènes

Codage Vidéo Distribué (DVC)

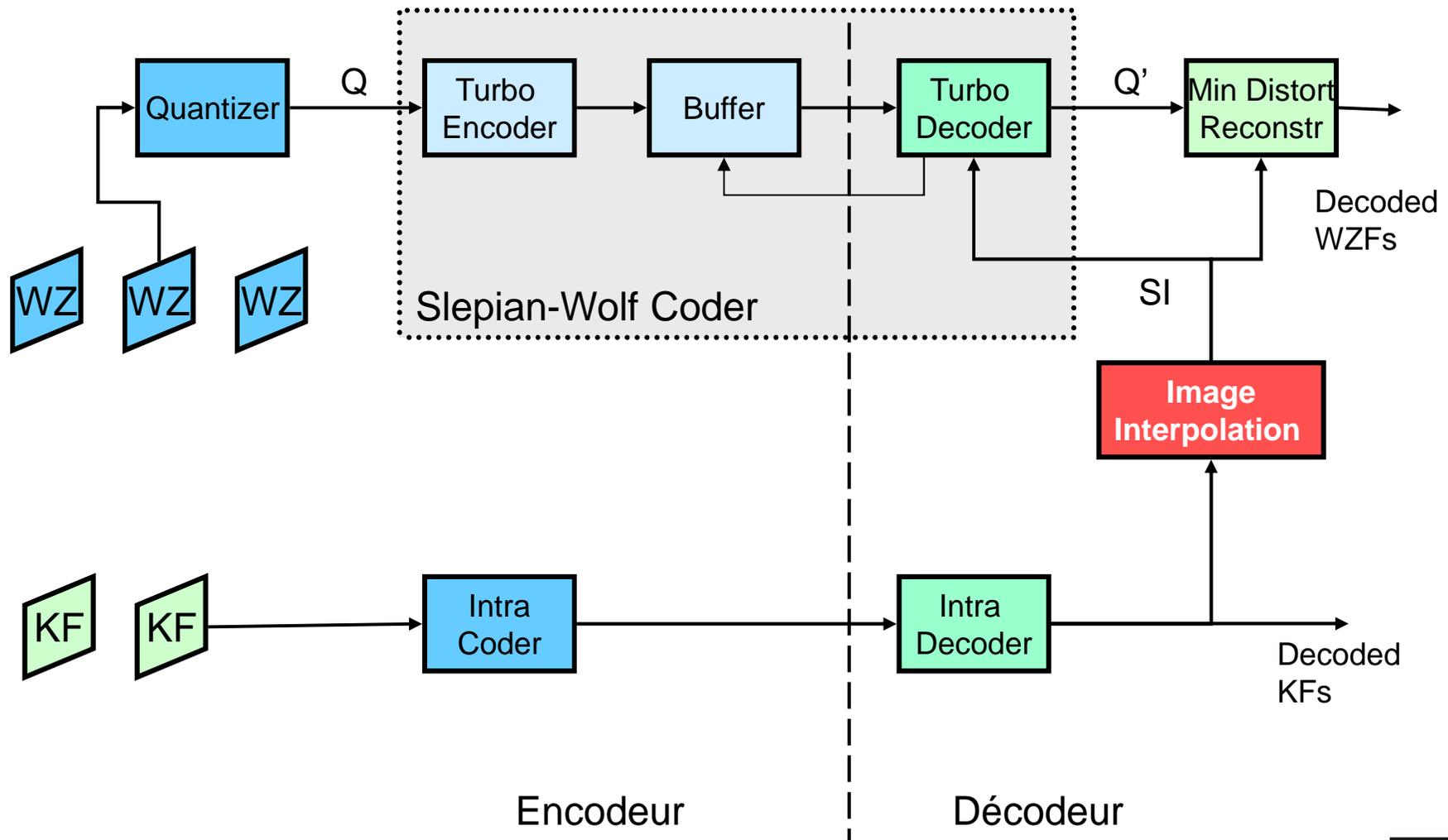
- Nouveau paradigme : Codage indépendant de plusieurs sources corrélées

Exemples d'applications:

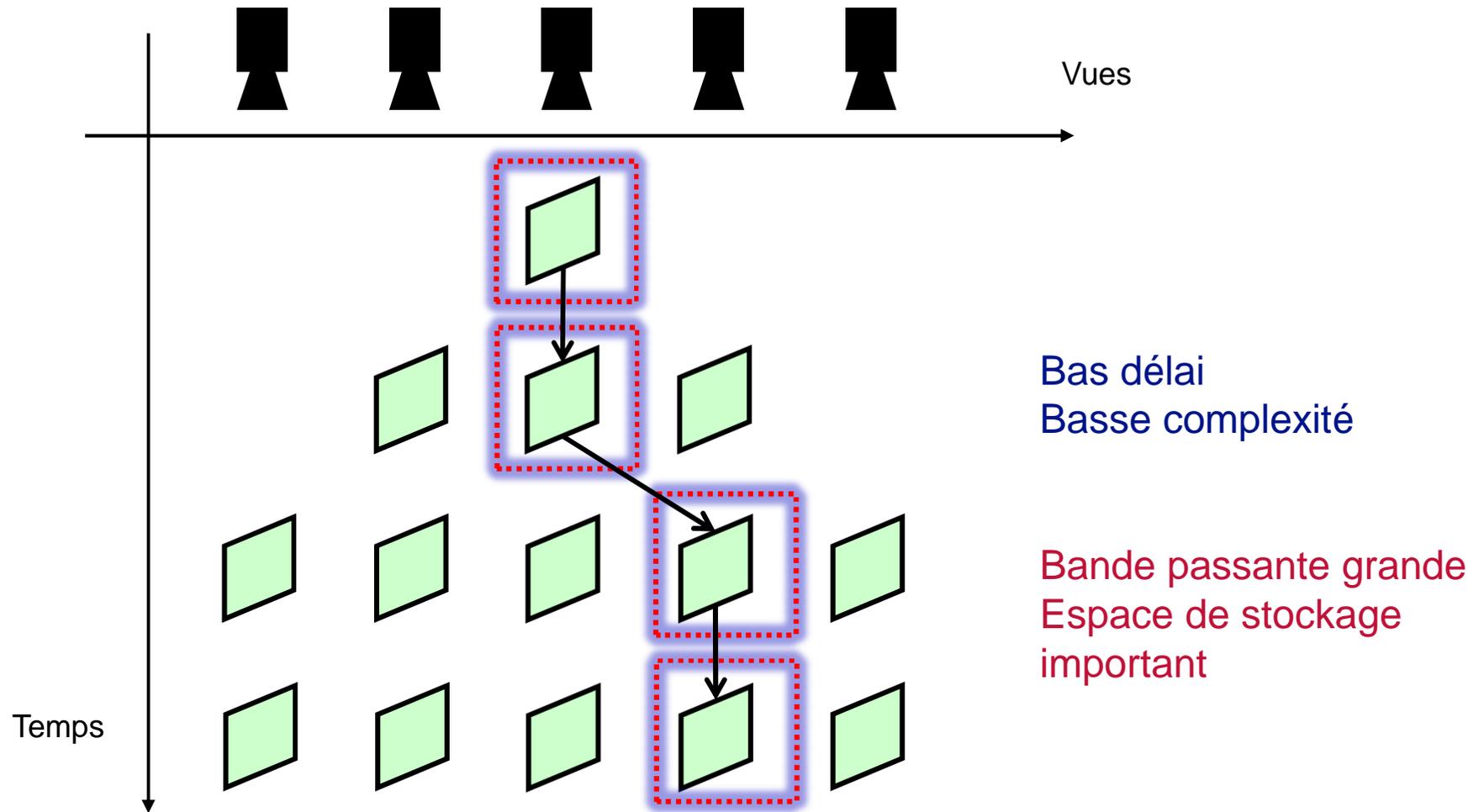
- Compression vidéo basse complexité (pas d'estimation de mouvement à l'encodage)
 - Réseaux de capteurs, y compris micro-satellites et applications médicales
 - Vidéo multi-vues (pas de communication entre les caméras)
- Théorie de l'information : mêmes performances RD que le codage conventionnel (prédictif)
 - En pratique, on se rapproche ...



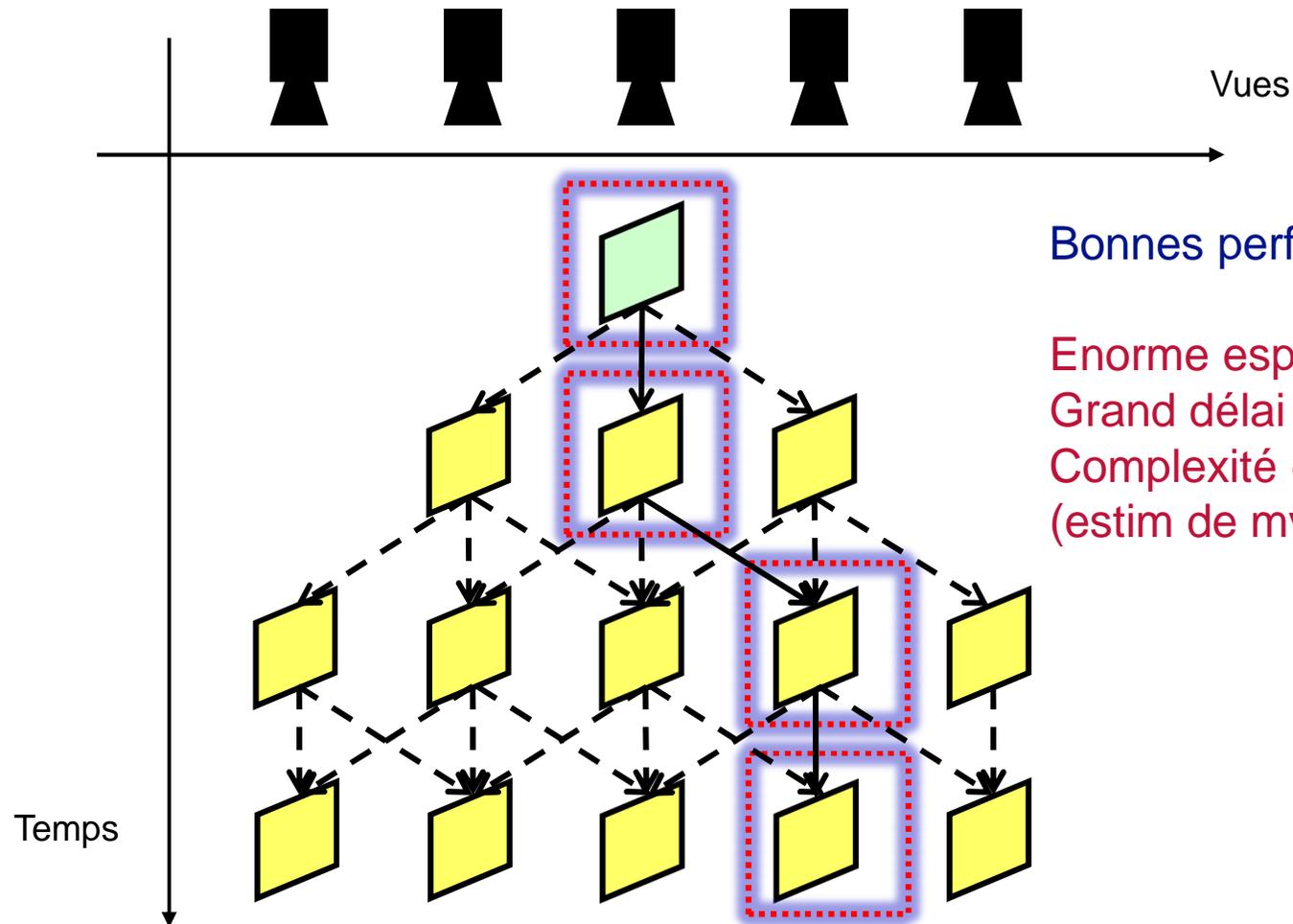
Architecture d'un codec vidéo distribué



Interactive Multiview Video Streaming : Codage Intra



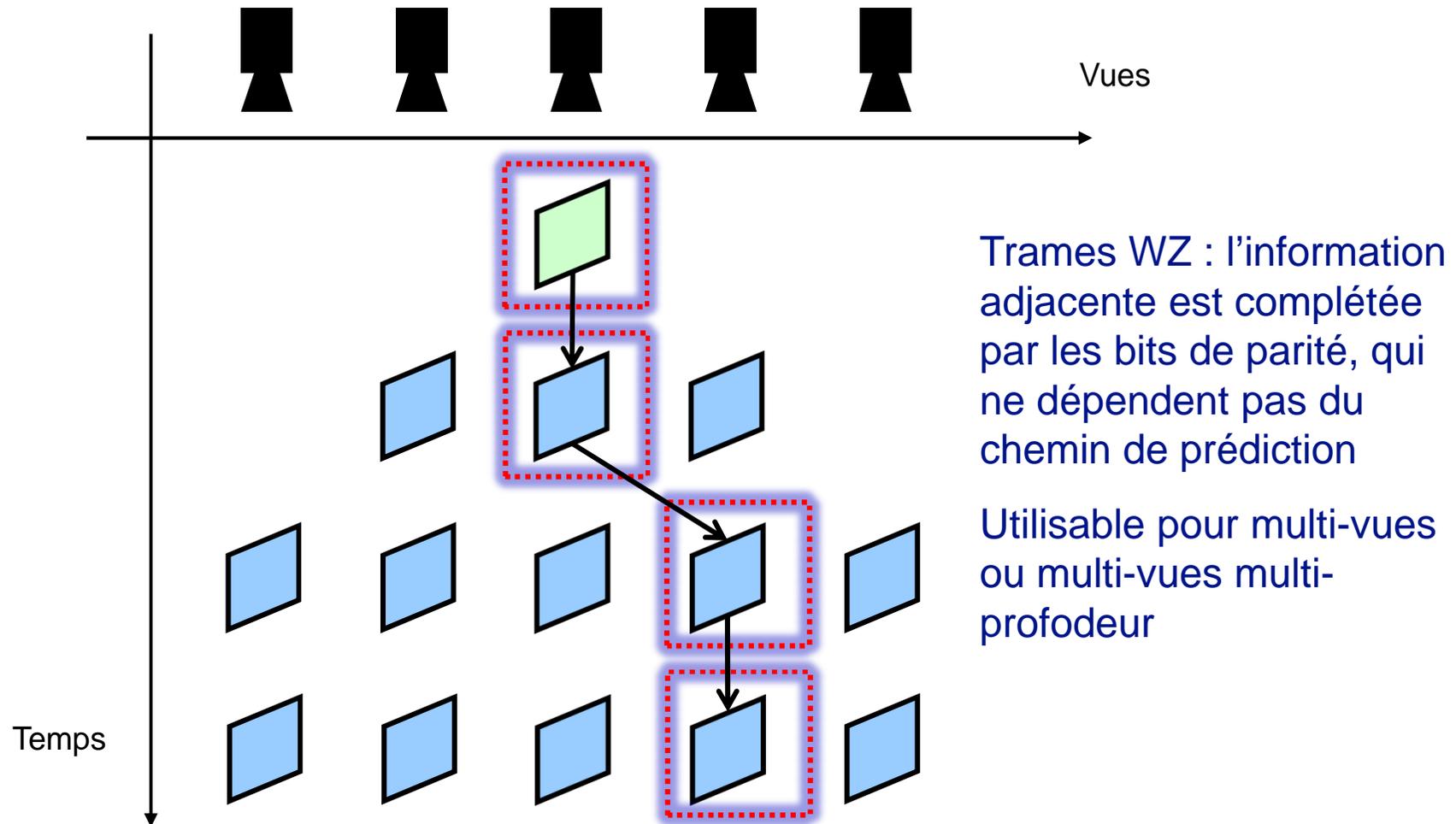
Interactive Multiview Video Streaming : encodage de tous les chemins



Bonnes performances RD

Enorme espace stockage
Grand délai d'encodage
Complexité de l'encodeur
(estim de mvnt et de disparité)

Interactive Multiview Video Streaming : Codage Vidéo Distribué





Traitements vidéo dans le *cloud*

- « Nuage informationnel » vs CDNs traditionnelles : capacité à réaliser les traitements online sur les contenus est le caractère différentiateur
- *Adaptive delivery of non-linear content scenario (3D VoD for heterogeneous users)* : conversion des vidéos stéréo en contenus 3D multi-vue, free point of view ou SMV
- Super-resolution à partir de plusieurs versions de vidéo basse résolution, compressées basse qualité – techniques d’optimisation convexe
- Conception & développement d’algorithmes optimisés : estimation de disparité et synthèse de vues sur plateformes de calcul parallèle et hétérogènes (multi-core, multi-GPU)
- Intégration et tests sur le cluster Nova de Bull
- Perspectives : parallélisations avec Hadoop



Vidéo 3D : futur proche

■ Emergence de nouveaux formats :

- Super Multi-vue (SMV)
- Imagerie Intégrale (Holoscopie)

Nouveau format 3D: Super Multi-Vue

- 10+ or 100+ points de vue
- Représentation “champ de lumière”

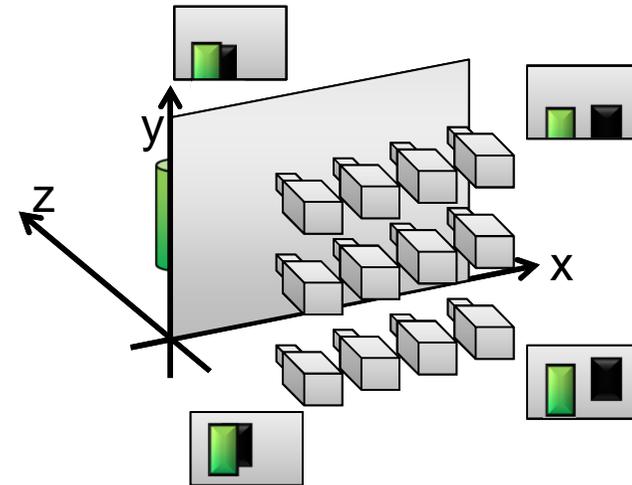
→ Sans lunettes

→ Parallaxe de mouvement amélioré

- Horizontalement
- Complet (horizontal + vertical)

→ Explosion de la quantité d’information à traiter !

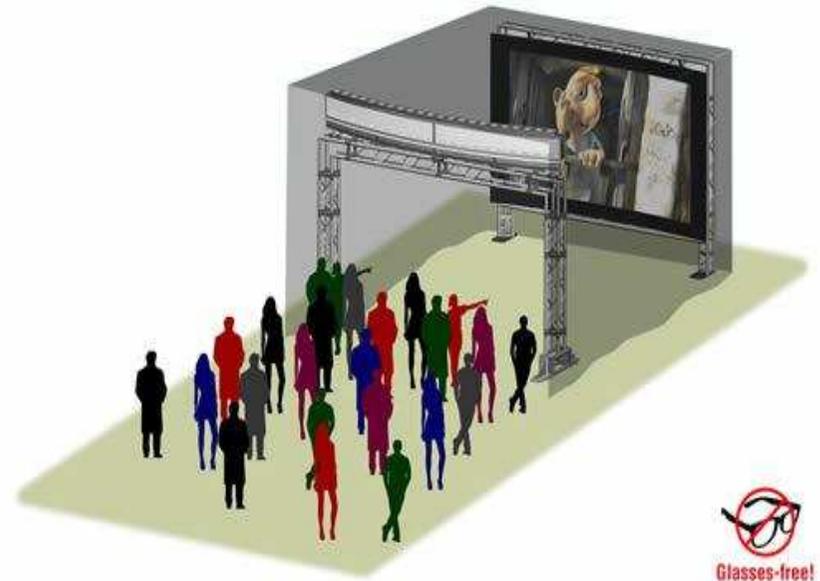
→ **Besoin de technologies de compression efficaces**



23

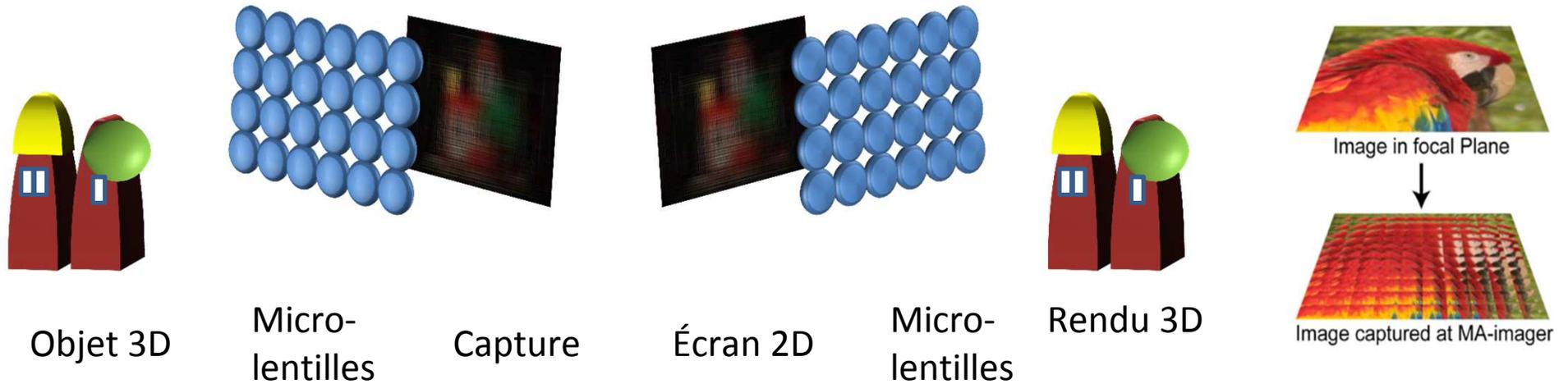
Intérêt pour le Super Multi-Vues

- A l'étude dans MPEG FTV depuis Oct 2013
- Systèmes d'affichage lightfield
 - Technologies variées (projection, écrans TV, *tabletop*, holoscopique, ...)
 - Prototypes ou à l'état expérimental (ex: NICT, NHK)
 - Produits commercialisés (ex: système **Holografika Hologvizio**)
- Exemples d'applications
 - Live broadcast d'événements sportifs
 - Téléprésence immersive, free navigation etc.



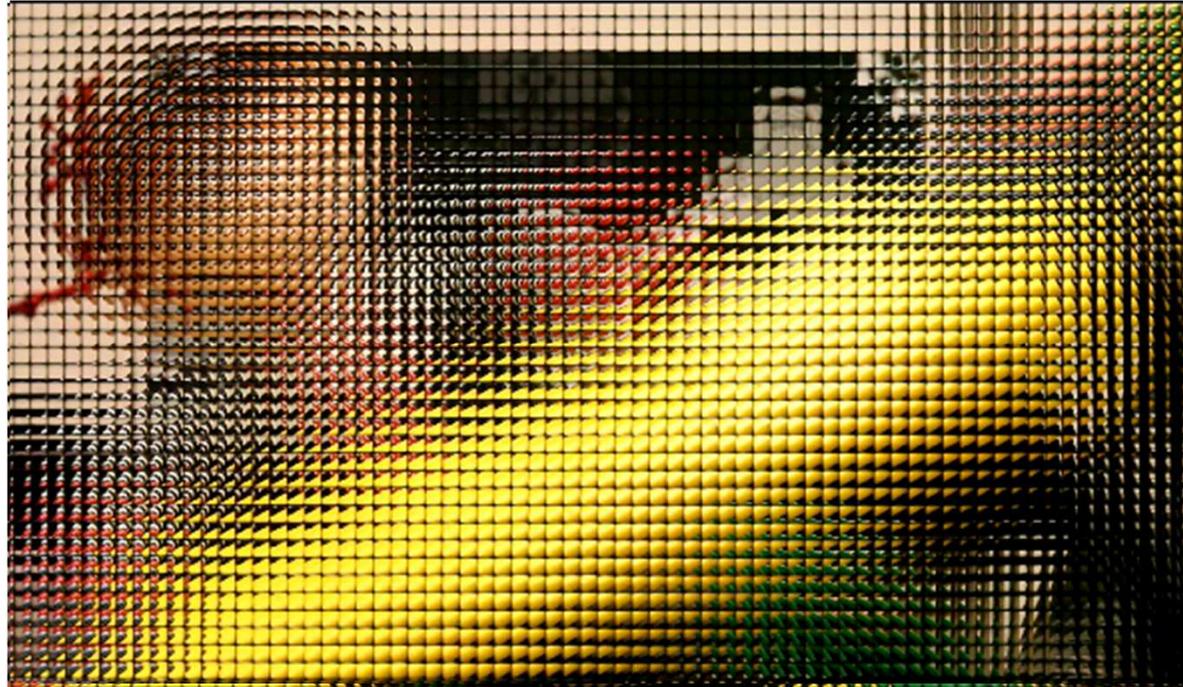
→ Phase de normalisation - technologies de codage SMV

Imagerie intégrale (holoscopie)



- Image en « œil de mouche »
- Sans conflit accommodation/convergence
- Rendu 3D réaliste sans lunettes

Exemple vidéo holoscopique



Verrous :

- Compression ? (prise en compte des vues extraites ?)
- Qualité ? (forme et densité du réseau lenticulaire, moins de points exploitables)
- Artefacts ?



Perspectives

■ Foisonnement de nouveaux formats vidéo

- Ubiquité, immersivité
- Interactivité

■ Outils théoriques puissants pour le traitement et l'optimisation distribués

- Enjeux : *Big Data*, Traitement dans le *cloud*, Modélisation

■ Interactions multimodales



Merci de votre attention !