



Back to the future of video compression

Touradj Ebrahimi

Touradj.Ebrahimi@epfl.ch



It's all about representation

- From a keynote by M. Vetterli at ICIP2004:
 - 256x256 pixels, 24 bits/pixel= $2^{(256 \times 245 \times 24)}$
(D. Field)
 - All images on the internet
 - 5-50 billion or 33 to 36 bits
 - All pictures taken in the history of mankind (M. Vetterli)
 - 100 years: 10^{10} circa 44 bits
 - Learn from nature
 - 20 bits/sec processed by the human brain (D. Gabor)
 - Current state of the art
 - JPEG: 200 Kbytes
 - JPEG 2000: 30% better in average
 - There is still a lot of room for improvements



Time-Space-Frequency representations

- 1D: Time-Frequency representation:
 - DCT, sub-band, wavelets, ...
 - MP3, AAC
- 2D: Space-Frequency representation:
 - 2D separable DCT (MPEG, JPEG, ITU-T)
 - 2D separable wavelets (MPEG, JPEG, ITU-T)
- 3D: Time-Space-Frequency representation
 - Straightforward 3D separable extension (Karlsson et al.)
 - Mainly hybrid schemes adopted in state of the art
 - Motion compensated 3D extension (Jens et al.)
 - Basis of Scalable Video Coding



New representation methods

□ Non-linear approximation

- Sub-space depends on the signal to be decomposed
- Take the M largest coefficients versus the first M coefficients
- $A(x+y) \neq A(x) + A(y)$

□ Non-separable decomposition

- Geometric transforms



Extensions to wavelets

□ Signal adapted schemes

- Bandelets [LePennec & Mallat]: wavelet expansions centered at discontinuity as well as along smooth edges
- Non-linear tilings [Cohen, Mattei]: adaptive segmentation
- Tree structured approaches [Shukla et al, Baraniuk et al]



Extensions to wavelets

□ Bases and frames theory

- Wedgelets [Donoho]: Basic element is a wedge
- Ridgelets [Candes, Donoho]: Basic element is a ridge
- Curvelets [Candes, Donoho] Scaling law: width $\sim \text{length}^2$
- Contourlets and Multidirectional pyramids [Do et al] Discrete-space
- Directional wavelet transforms and frames [Velisavljevic et al]



Extensions to wavelets

- Coping with discontinuities:
 - Fourier is not good for singularities
 - Wavelets are good for point singularities
 - X-lets are good for curve and line singularities
- Rate of convergence
 - Fourier: $O(M^{-1/2})$
 - Wavelet: $O(M^{-1})$
 - Contourlet: $O(M^{-2})$



Open questions

- How much does it cost to completely code an x-let representation
 - Lessons learned from region/object based representation
- How to apply x-lets on 3D (video=space+time) and how much gain
 - Lessons from 3D subband decomposition of video
- How much gain comes from the choice of x-lets as transform block versus the rest of compression scheme (quantization, entropy coding)?
 - JPEG proposed AMD



JPEG proposed AMD

- There is an attempt to replace wavelet transform by DCT in JPEG 2000
- Rationale:
 - The 30% average gain in compression comes from entropy coding
 - Lossless transcoding from and to JPEG
 - Existing products based on a similar approach (Stfflt)



Is there a Moore's law of compression?

- The compression efficiency doubles every five years
 - H.263 twice better than H.261
 - MPEG-4 AVC part 10 (H.264) twice better than MPEG-4 version 1
- Developments on future directions in video compression will show if such a law exists