Theme: Special Topics
Lecture **2**

Scalable Image and Video Coding

José M. Dana

CERN

Inverted CERN School of Computing, 3-5 March 2008

Outline

- Introduction
- Image coding (JPEG)
- Video coding (MPEG-4)
- Scalable image coding (JPEG2000)
- Scalable video coding (FSVC)
- Video samples

Introduction

José M. Dana

CERN

Inverted CERN School of Computing, 3-5 March 2008

Rule #1

Data ≠ Information

Why do we need compression?

Multimedia Data	Size/Duration	Bits/Pixel or Bits/Sample	Uncompressed size	Transmission bandwidth
A page of text	11“ x 8.5“	Varying resolution	4-8 KB	32-64 Kb/page
Telephone quality speech	10 sec	8 bps	80 KB	64 Kb/sec
Grayscale image	512 x 512	8 bpp	262 KB	2.1 Mb/image
Color image	512 x 512	24 bpp	786 KB	6.29 Mb/image
Medical image	2048 x 1680	12 bpp	5.16 MB	41.3 Mb/image
SHD image	2048 x 2048	24 bpp	12.58 MB	100 Mb/image
Full-motion video	640 x 480 1 min (30 fps)	24 bpp	1.66 GB	221 Mb/sec

Data coding systems

- Lossless
  - LZW, LZ77, LZ78 (universal data compression algorithms)
  - FLAC (audio coding system)
  - LS-JPEG, PNG (image coding systems)
  - Special algorithms for medical, astronomical, etc. purposes
- Lossy
  - MP3, AAC, Ogg Vorbis (audio coding systems)
  - JPEG, JPEG2000 (image coding systems)
  - MPEG-2, MPEG-4, H.264 (video coding systems)

The human visual system

- Our eyes are our "data acquisition system"
- We’re limited by resolution and bandwidth, therefore some data can be ignored

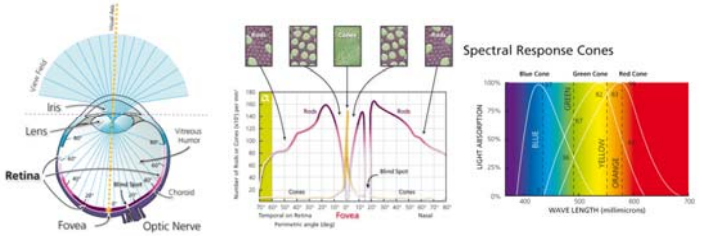


Image coding

José M. Dana

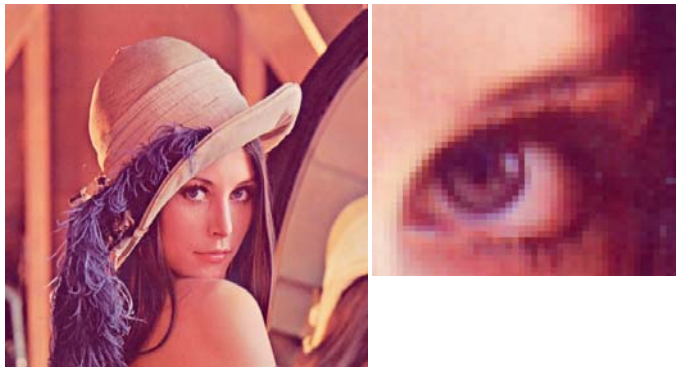
CERN

Inverted CERN School of Computing, 3-5 March 2008

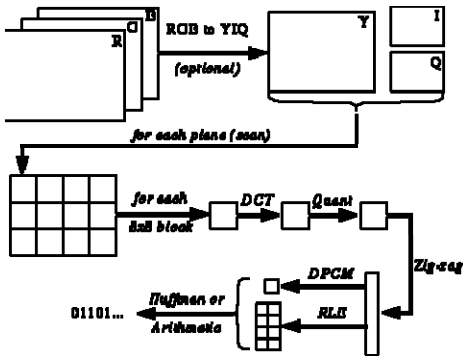
Some definitions

- **Digital Image:** *An image stored in binary form and divided into a matrix of pixels, each consists of one or more bits of information that represent either the brightness, or brightness and color, of the image at that point.*
- **Pixel:** *A contraction of the words picture element. The smallest unit of information in an image or raster map. Referred to as a cell in an image or grid.*

Spatial redundancy



JPEG coding

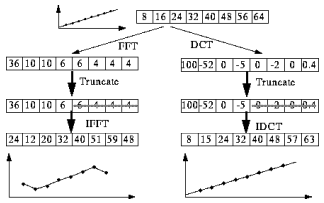


JPEG - Colorspaces

- Valid color spaces
  - Grayscale
  - YIQ (Y=Luminance; I,Q=Chrominance)
  - YCbCr (Y=Luminance, Cb=Blue/Yellow axis, Cr=Red/Green axis)
  - CMYK (Cyan-Magenta-Yellow-Key)
- Not valid color spaces (transformation needed)
  - RGB (Red-Green-Blue)
  - RGBA (Red-Green-Blue-Alpha)
  - YUV (Y=Luminance; U,V=Chrominance)
  - Etc.

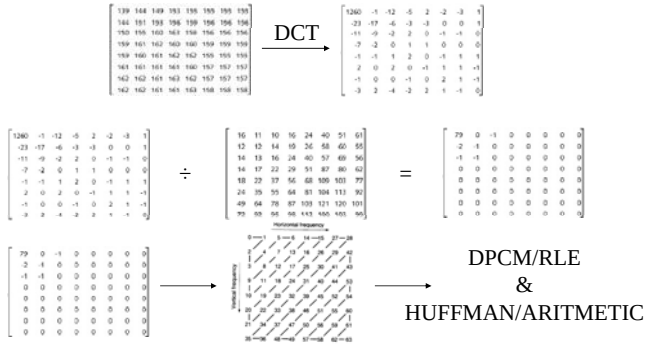
Energy compaction capability

- When applying the DCT to a signal, a higher ratio of the energy is concentrated in a small number of coefficients relative to the FFT or other similar transforms



FFT vs. DCT

JPEG - Quantization and "zig-zag"



Block artifacts

Original image

DC component

Video coding

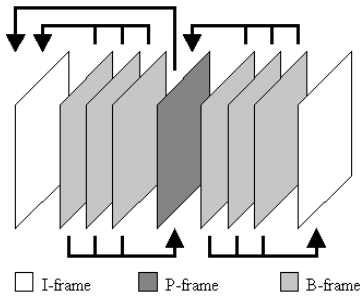
José M. Dana

CERN

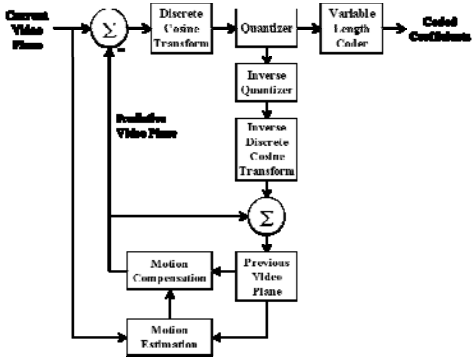
Inverted CERN School of Computing, 3-5 March 2008

Inter-frame dependencies

- *GOP*: Group of pictures
- *I-frame*: Intracoded frame
- *P-frame*: Forward predicted frame
- *B-frame*: Bi-directional predicted frame
- **Problem: Error propagation!**

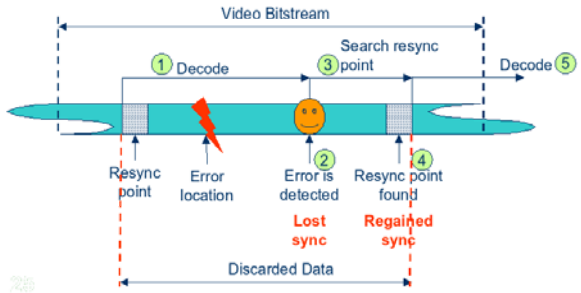


Generalized video codec (MPEG-4)

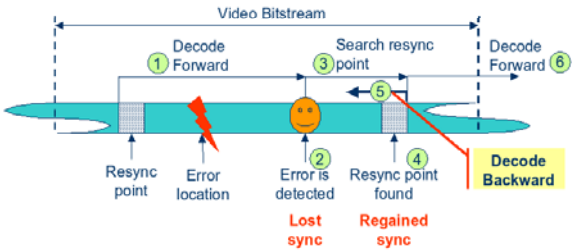


Video Packet Mode (MPEG-4)

Resync points (MPEG-4)



Reversible variable length codes (MPEG-4)



Scalable image coding

José M. Dana

CERN

Inverted CERN School of Computing, 3-5 March 2008

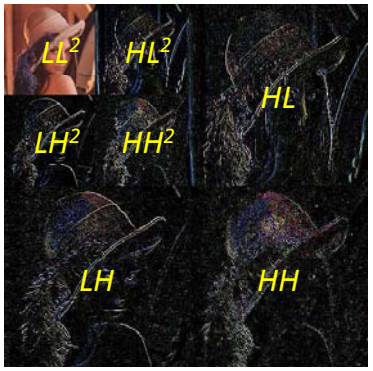
What does scalability mean?

- *Encode Once, Display/Stream Anywhere*
- Current digital video applications require at least three types of scalability features:
  - Quality scalability
  - Spatial resolution scalability
  - Temporal (frame rate) scalability

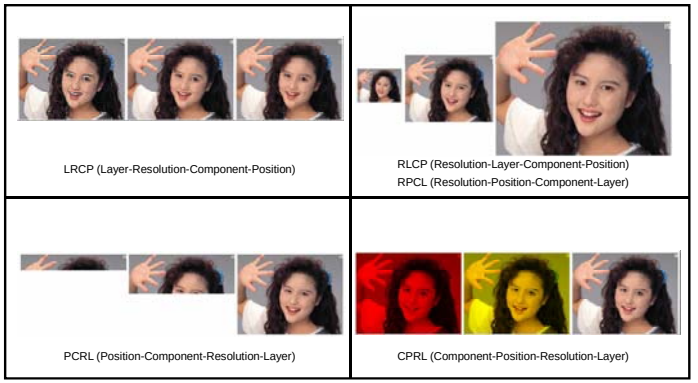
The Discrete Wavelet Transform

- The DCT previously carries out a division into squared blocks (8x8 pixels in JPEG) while the 2D-DWT works in its totality
- The decomposition into subbands gives a higher flexibility in terms of scalability in resolution and distortion
- The DWT returns a multiresolution representation in a joint spatial-spectral domain
- Better error resilience

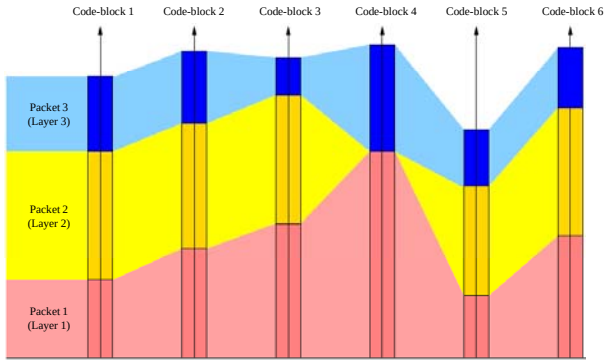
DWT – Dyadic decomposition



JPEG2000 - Progressions



JPEG2000 - Quality Layers



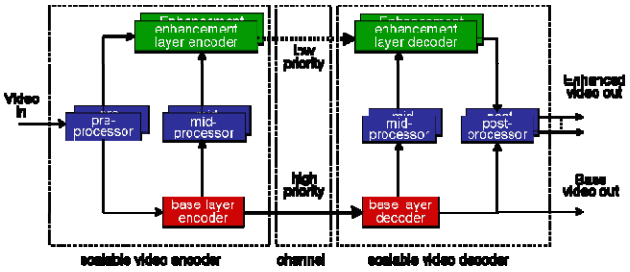
Scalable video coding

José M. Dana

CERN

Inverted CERN School of Computing, 3-5 March 2008

Generalized scalable video codec

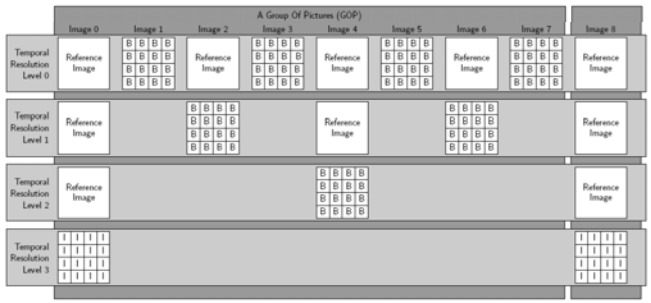


- **Problem #1: The base layer**
- **Problem #2: Drift**

Problem #2: Drift

- In video coding, **drift error** refers to the **continuous decrease** (in picture quality) when a group of motion-compensated interframe pictures have been decoded using frames of reference that are different from the ones used during the encoding step (motion vector field mismatch)
- Scalable video coders have traditionally avoided using enhancement layer information to predict the base layer, so as to avoid so-called “drift”
- As a result, they are less efficient than a one-layer coder

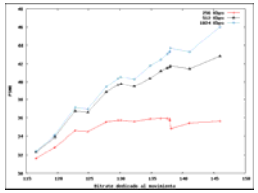
Temporal scalability (FSVC)



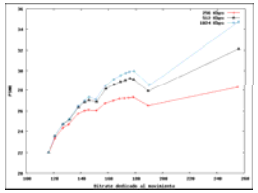
Motion information scalability

- We have been speaking about scalable spatial information but… what about scalable motion information?
- Redundancy can be also found in motion vector fields but it is much more "special" (zoom, accelerated motion, etc.)
- **Problem #1 (again!): The base layer**

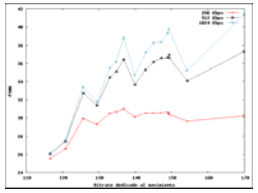
The importance of motion information



Akiyo



Foreman



Container

That's all folks!