튜 토 리 얼 III

영상처리 개론

▷ 연 사 : **허 영** 박사 (한국전기연구소)
▷ 사 회 : **황 병 연** 교수 (가톨릭대)

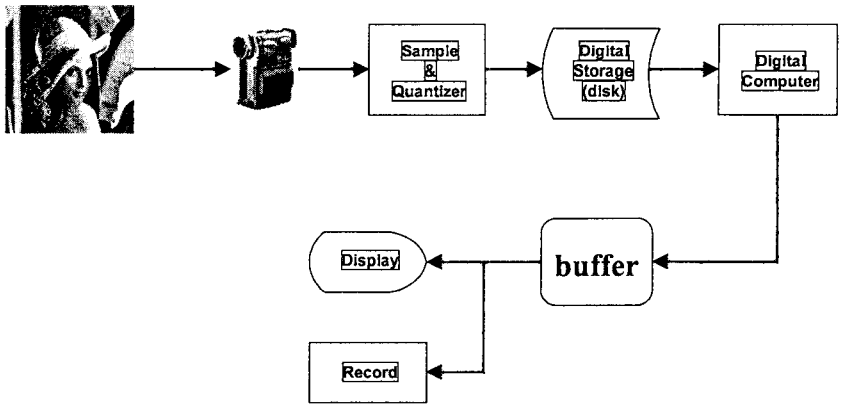
Digital Image Processing for Multimedia

Nov. 18 2000

Young Huh
yhuh@keri.re.kr

Korea Electrotechnology Research Institute

Basic Imaging System Block



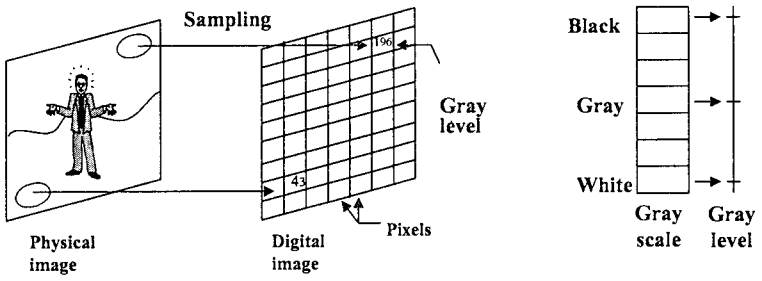
Areas of Image Processing

- Image representation and Modeling
- Image Enhancement
- Image Restoration
- Image Analysis
- Image Reconstruction
- Image Data Compression

Applications of Image Processing

- Remote Sensing via satellites
- Image Transmission and Stroage
- Medical Image Processing
- Robotics
- Automated Inspection

What is the digital images ?



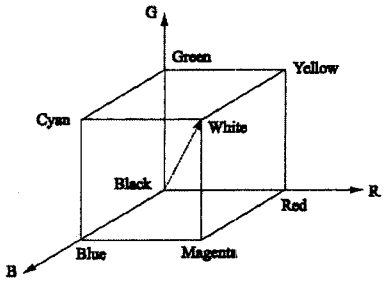
- Each cell => picture element = pixel = pel
- Each pel conveys information of brightness with n bits (256 levels = 8 bit = 1 byte/pel)

Color formats

- RGB : Basic format provided by camera
- YUV : NTSC/PAL/SECAM
- YIQ : NTSC
- YCbCr : Digital Video version
- SMPTE240M : US full-digital HDTV

RGB Color Coordinate

- Tristimulus RGB color cube
- RGB: equal bandwidths(8bits each)
- Human eye is more sensitive to Luminance than Chrominance



Gamma Correction

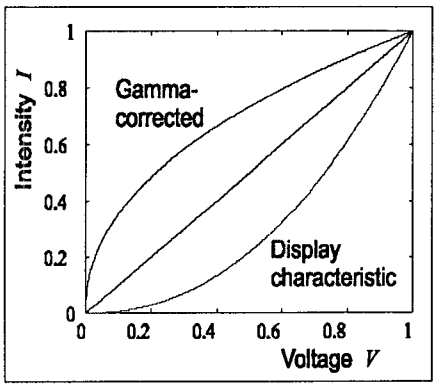
- Nonlinear characteristics of CRT is

$$I_{display} = V_{received}^{\gamma}$$

where gamma is typically
2.2 for NTSC
2.8 for PAL/SECAM

- Correction prior to transmit

$$I_{transmit} = I_{input}^{0.45}$$



$0 \leq V \leq 1$

YCbCr Color Coordinate

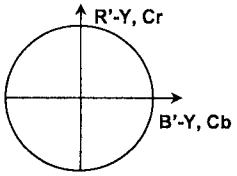
- **Scaled, offset version of YUV**

601 :

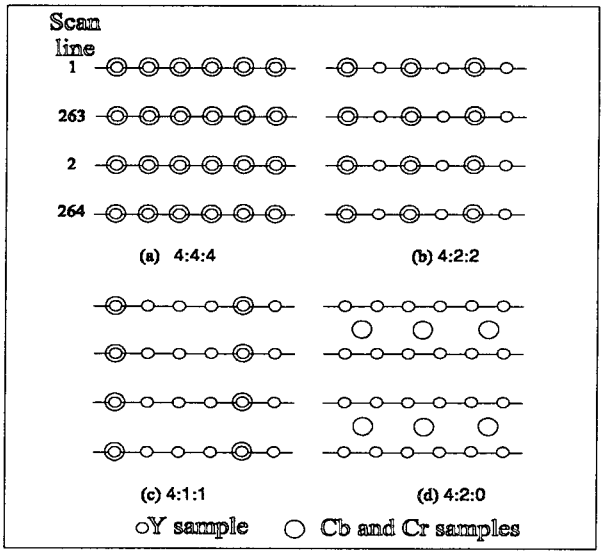
$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{pmatrix} 0.30 & 0.59 & 0.11 \\ 0.50 & -0.42 & -0.08 \\ -0.17 & -0.33 & 0.50 \end{pmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

$$\begin{bmatrix} Cr \\ Cb \end{bmatrix} = \begin{pmatrix} 0.00 & 0.71 \\ 0.56 & 0.00 \end{pmatrix} \begin{bmatrix} B' - Y \\ R' - Y \end{bmatrix}$$

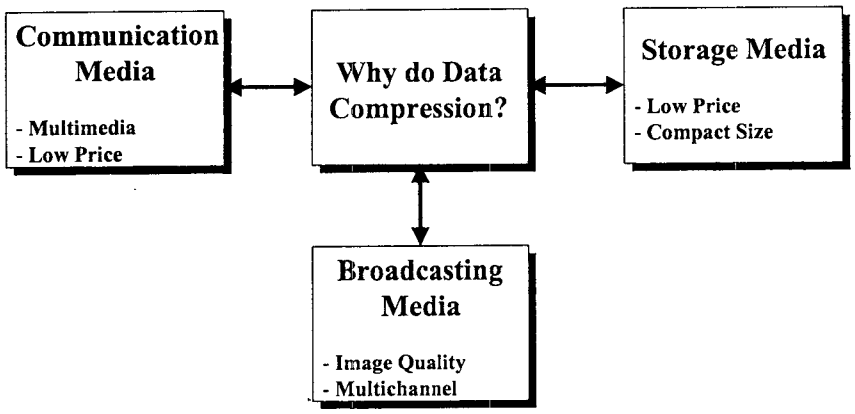
- **In 8-bit implementations,**
  - Y occupies 220 levels : [16, 235]
  - Cr, Cb occupy 225 levels : [16, 240]



Sampling



The Needs for Data Compression



Multimedia Information & Multimedia Compression

- Amount of Information Media

Information Media	Typical Value	Information Amount (bits)	Bit Rate (bits/sec)
Telephone	1 hour	230 M	64 k
Music (CD)	1 hour	5.0 G	1.4 M
VCR (VHS-quality)	1 hour	109 G	30 M
Current TV (Receipt-quality)	1 hour	360 G	100 M
HDTV (Studio-quality)	1 hour	4,320 G	1.2 G

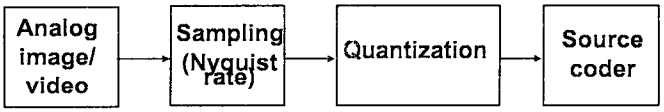
- Bit Rate after Multimedia Compression

Information Media	Original Bit Rate (bits/sec.)	Compression Ratio	Bit Rate after Compression (bits/sec.)
Telephone	64k	4	16k
Music (CD)	1.4 M	5.5	256 k
VCR (VHS quality)	30 M	20 ~ 30	1 ~ 1.5 M
Current TV (Receipt-quality)	100 M	20 ~ 25	4 ~ 6 M
HDTV (Studio-quality)	1.2 G	40 ~ 80	15 ~ 30 M

Video compression basis

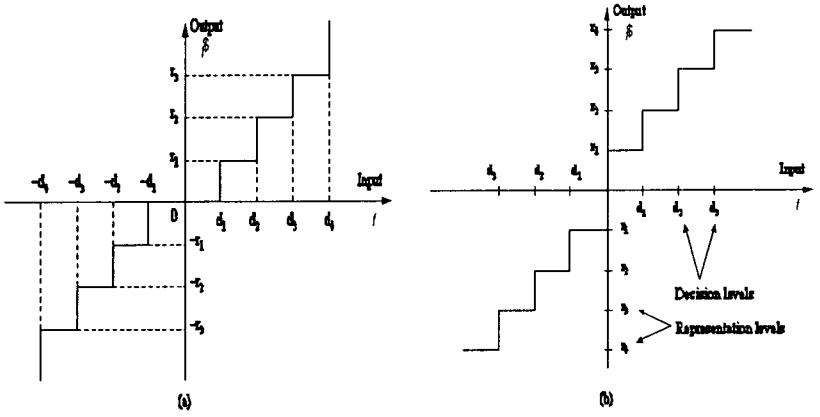
- **Remove spatial and temporal redundancy that exist in natural video imagery**
  - Correlation itself can be removed in a lossless fashion
  - Application of data compression are primarily in transmission and storage of information
- **Exploit limitations in Human Visual System(HVS)**
  - -Limited luminance and very limited color response
  - -Reduced sensitivity to noise in high frequencies (e.g., edges of objects, high detailed areas)
  - -Reduced sensitivity to noise in brighter areas
  - -Goal is to throw away bits in a psychovisually lossless manner
  - 50:1 or more compression efficiency

Quantization (I)



- **Uniform & nonuniform**
- **Midtreader & midriser**
- **Scalar & Vector**
- ***Quantization*** **results in** ***Compression***

Quantization (II)



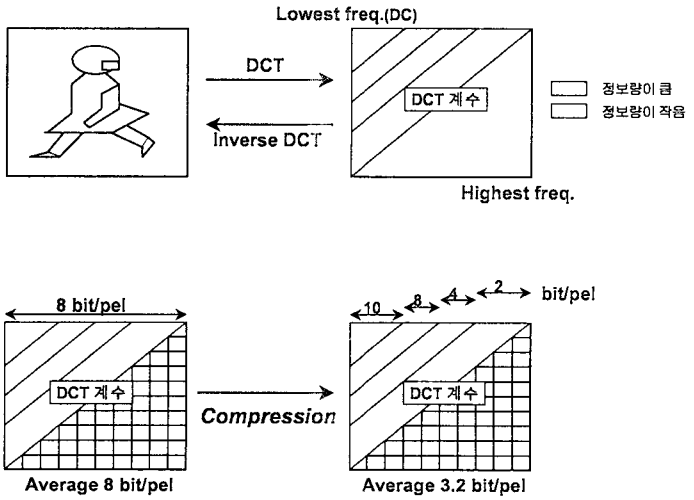
Uniform symmetric midtreader | Uniform symmetric midriser

Transforms

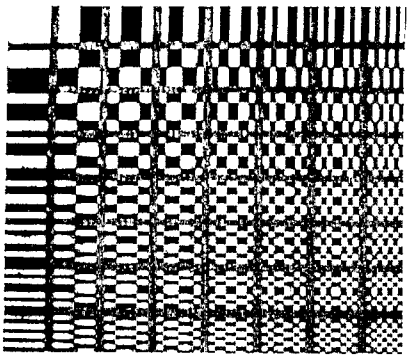
- **KLT (Karhunen-Loeve Transform)**
- **DFT (Discrete Fourier Transform)**
- **DCT (Discrete Cosine Transform)**
- **DST (Discrete Sine Transform)**
- **WHT (Walsh Hadamard Transform)**
- **DHT (Discrete Haar Transform)**
- **LOT (Lapped Orthogonal Transform),**

  ELT (Extended Lapped Transform)

Energy in DCT domain



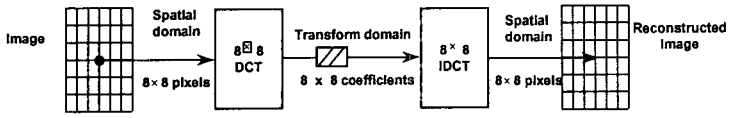
DCT Basis Images



Outer product of 1-D basis vectors

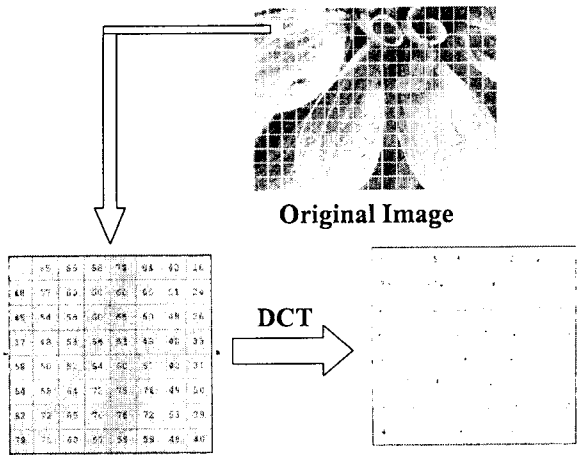
Shows 64 possible images according to data structure

Transform Coding using Discrete Cosine Transform

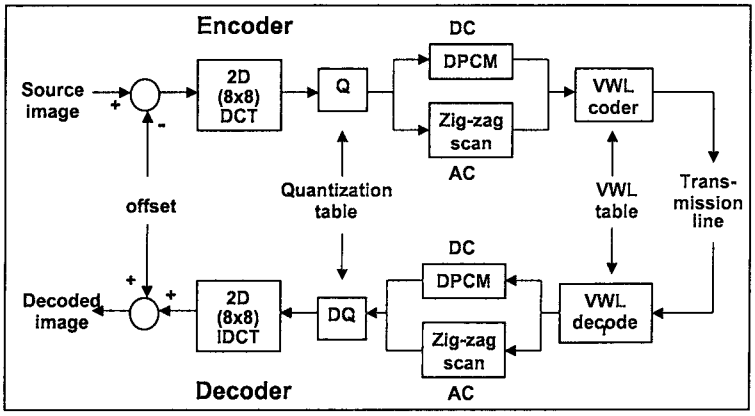


- DCT is an orthogonal transformation
- 2-D DCT is separable in x and y dimensions
- Has good energy compaction properties
- Close to Karhunen-Loeve Transform (KLT), which is optimal but depends on image statistics.
- Efficient hardware realization based on the fast algorithm

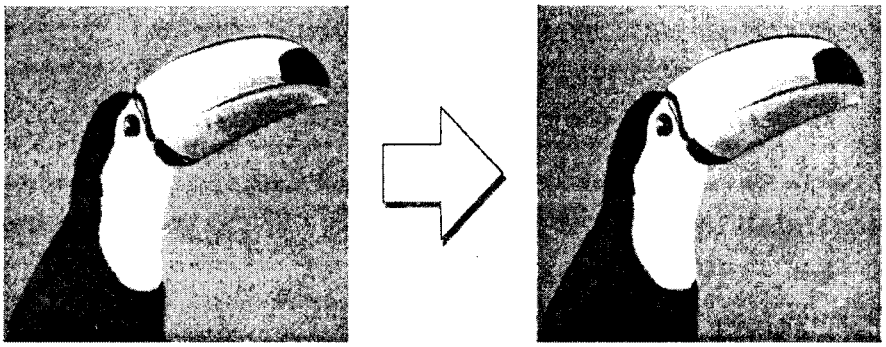
Example of DCT Transform



JPEG baseline system



Example of JPEG Image



Original Image

(2MByte)

JPEG Compressed Image

(40KByte)

JPEG 2000

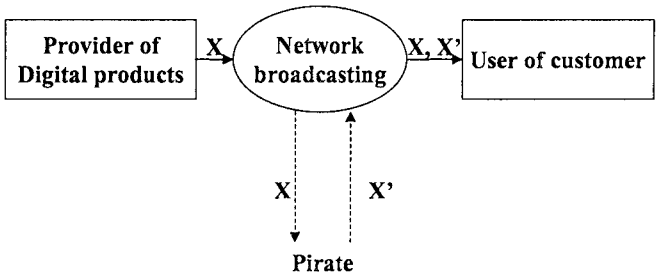
- **JPEG : O.K., but NOT enough!**
- 향상된 압축 효율 및 다양한 기능성을 제공하는 새로운 형태의 정지영상 압축의 표준
- **JPEG**의 압축 성능 한계, 새로운 압축형식의 경쟁, 다양한 응용 분야에 대한 적용의 한계

JPEG 2000 Technical Status

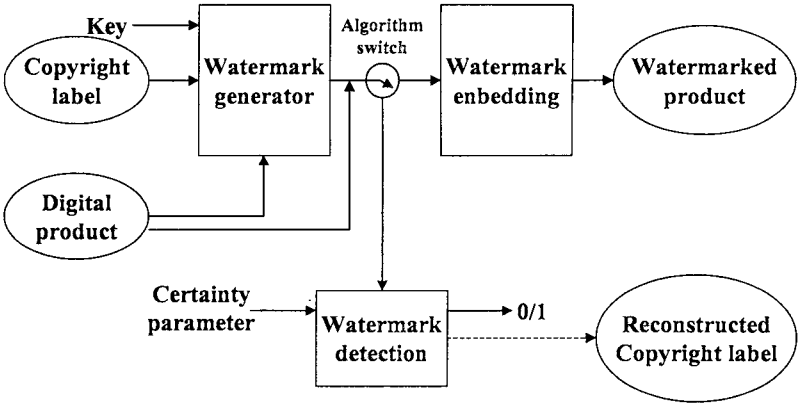
- **Digital Watermarking**
- **Wavelet based**
  (although DCT could be improved, major emphasis and expertise in Wavelet technology)
- **MPEG-4 VTC (Visual Texture Coding) Compatibility**

Digital watermarking

- Copyright and content protection
- Authentication and integrity verification
- Image tagging (tracing original that has been illegally copied)
- Security (e.g., passport photos)
- Metadata tagging (info about content, website address, etc.)
- Secret communication

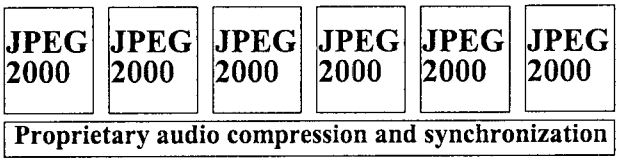


Digital watermarking algorithm



Motion JPEG 2000 (MJ2K)

- Intra-based coding
  - easy editing
  - robustness to error prone environments
- Single decompression architecture
- Scalability



Market and application

1. DSC(Digital Still Camera) with Moving Video
2. Internet Video
3. Mobile Phone/PDA
4. Video Capturing System
5. High Quality Digital Video Recording System
6. Remote Surveillance System
7. Medical Imaging
8. Remote Sensing
9. High Resolution Mapping

DSC with Moving Video

1. Target compression ratio : Programmable from lossless to 100:1
2. H/W requirement : Low memory, low power consumption
3. Wavelet kernel : Integer wavelet is preferred
4. Chroma format : 4:2:0 YUV or 4:2:0 YUV + Alpha
5. Frame rate : From 1 to beyond 30 fps

Bitrate at the compression ratio(1/30, 1/60) for DSC

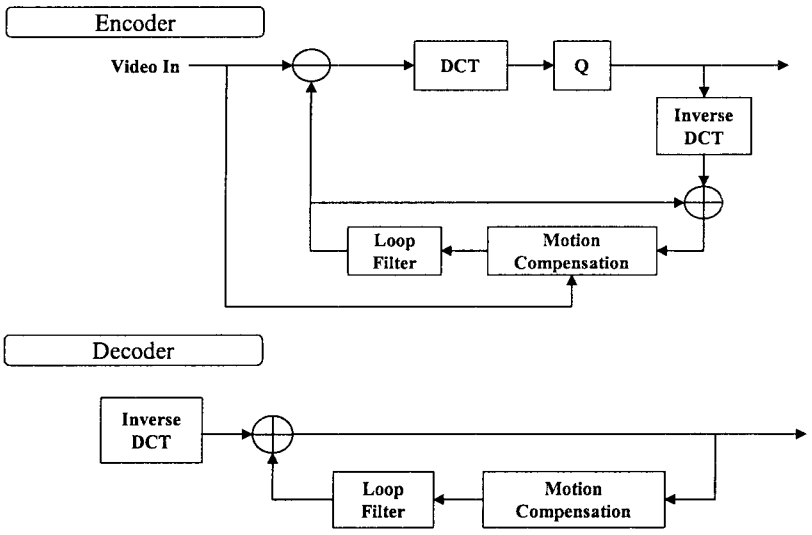
Image Format	VGA	SIF	QSIF
Resolution [Pixels x lines]	640 x 480	320 x 240	160 x 120
Bitrate 1*(4:2:0) 10 fps	1.23 [Mbps]	307.2 [Kbps]	76.8 [Kbps]
Bitrate 2*(4:2:0) 10 fps	614.4 [Kbps]	153.6 [Kbps]	38.4 [Kbps]

1* : Compression ratio is 1 to 30
2* : Compression ratio is 1 to 60

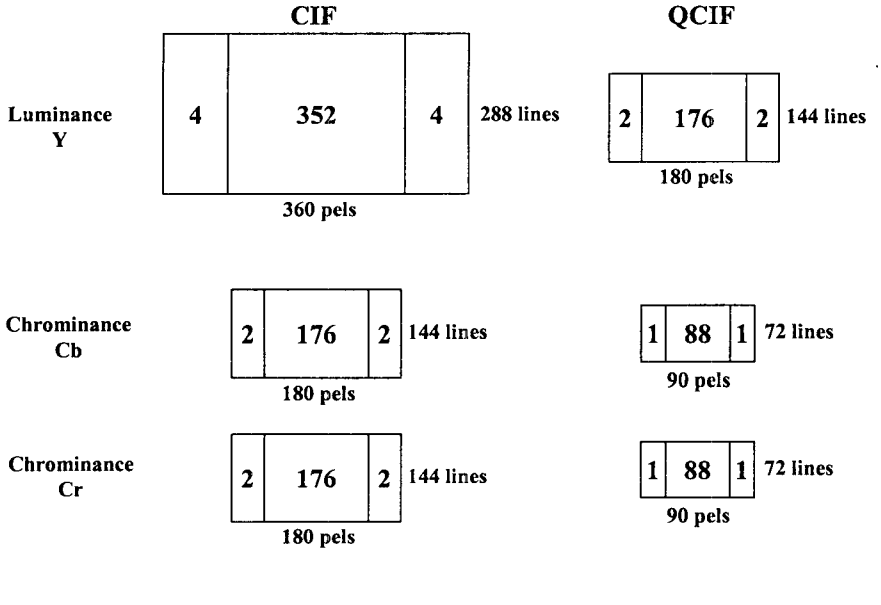
H.261

- Motion-compensation and transform coding
  - Basis of most video coding standards
- ITU-T Study Group 15
  - Videophone and video conferencing
    - Low bit rates, low delay
  - 1984: audiovisual services at m 384 kbit/s (m=...5)
  - 1988-90: p 64 kbit (p=1...30)

Block Diagram of H.261 Codec



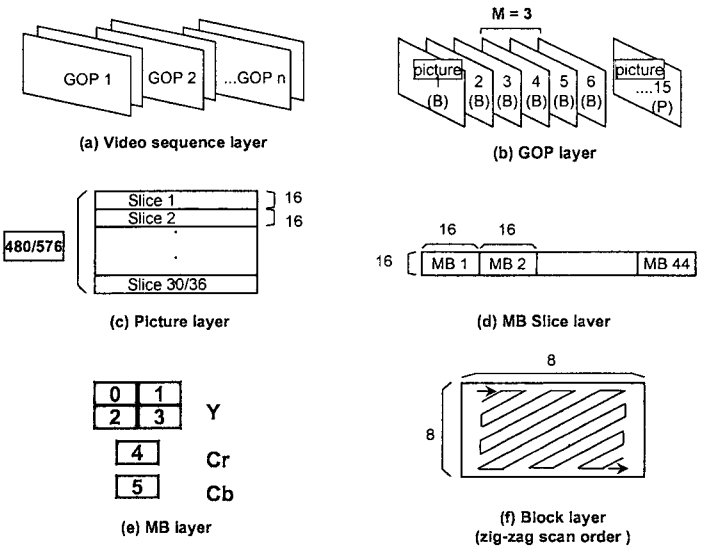
Picture Format and Resolution for CIF/QCIF



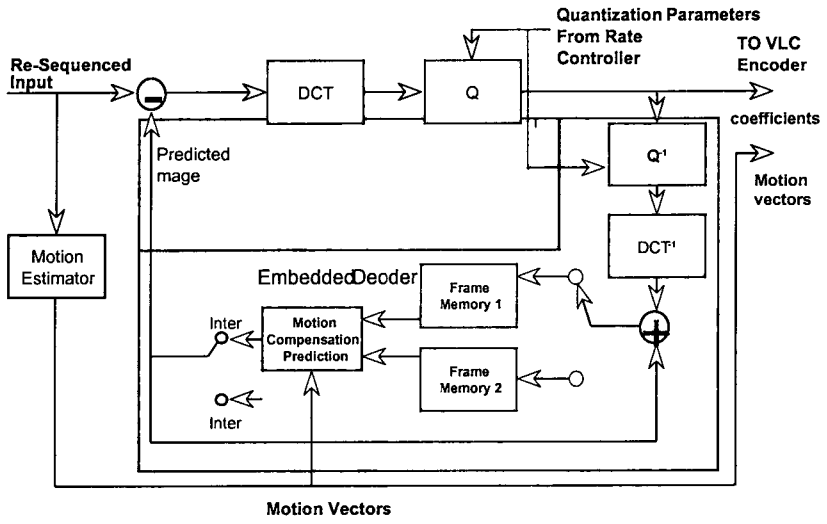
What is the MPEG ?

- MPEG = Moving Picture Experts Group
- Aim was to create the best video compression standards for multimedia and broadcast applications
- MPEG-1 Video aimed at SIF resolution
  - -352 x 480, 30Hz, non-interlaced, 1.5 Mb/s
  - - CD-ROM applications
- MPEG-2 Video aimed at CCIR-601 resolution
  - -720 x 480, 30Hz, interlaced, 4-10 Mb/s
  - - broadcast applications, including HDTV

MPEG video structure



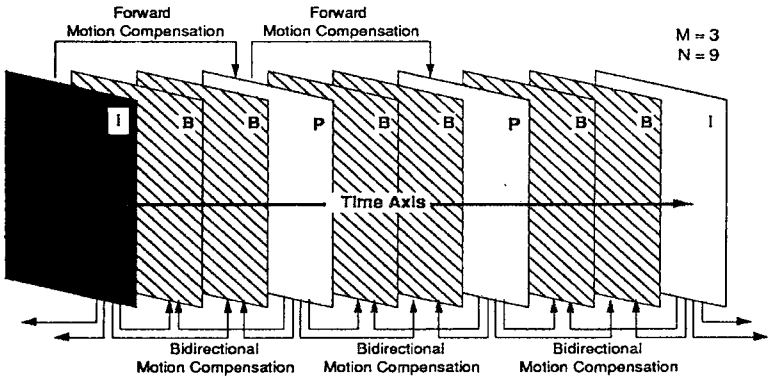
MPEG-1 Encoder



MPEG-2 : A superset of MPEG-1

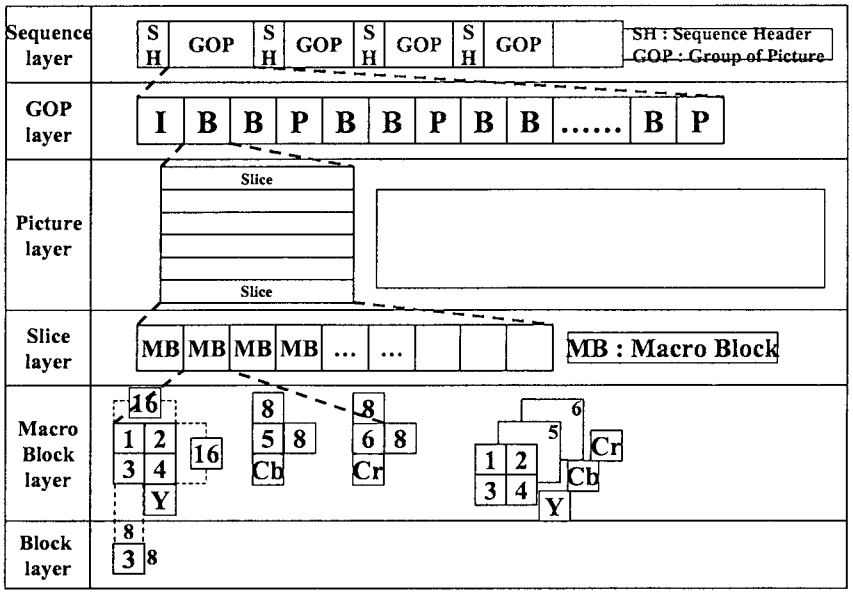
- **MPEG-2 = MPEG-1 Syntax Elements**
  + Interlace Tools
  + New Syntax Structures
  + Scalable Modes
  + Profiles & Levels

Group of Pictures



GOP structure example.: IBBPBBPBB (M=3, N=9)

Layer Structure of MPEG-2 Video



Level and Profile for MPEG-2 Video

Profile	Level	수평 size (pels)	수직 size (pels)	Frame rate (Hz)	Bit rate (Mbps)	VBV size (M bit)	motion vector 범위 (pels)
Simple	Main	720	576	30	15	1.835	-128 ~ 127.5
Main	Low	352	288	30	4	0.489	- 64 ~ 63.5
	Main	720	576	30	15	1.835	-128 ~ 127.5
	High-1440	1440	1152	60	60	7.340	-128 ~ 127.5
	High	1920	1152	60	80	9.787	-128 ~ 127.5
SNR	Low	352	288	30	3 (4)	0.367 (0.489)	- 64 ~ 63.5
	Main	720	576	30	10 (15)	1.223 (1.835)	-128 ~ 127.5
Spatial	High-1440	720 (1440)	576 (1152)	30 (60)	15 (40) (60)	1.835 (4.893) (7.340)	-128 ~ 127.5
High	Main	352 (720)	288 (576)	30 (30)	4 (15) (20)	0.489 (1.835) (2.447)	-128 ~ 127.5
	High-1440	720 (1440)	576 (1152)	30 (60)	20 (60) (80)	2.447 (7.340) (9.789)	-128 ~ 127.5
	High	960 (1920)	576 (1152)	30 (60)	25 (80) (100)	3.036 (9.787) (12.233)	-128 ~ 127.5

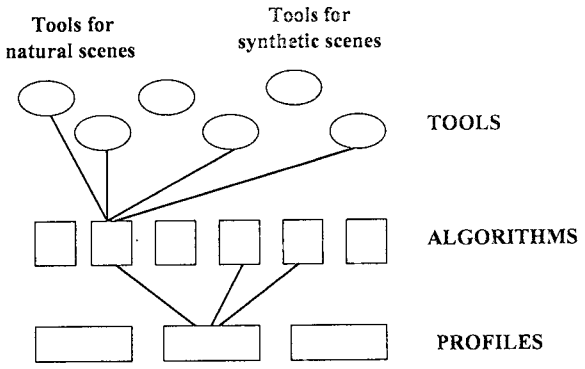
()는 확장 Layer를 참조

MPEG-4

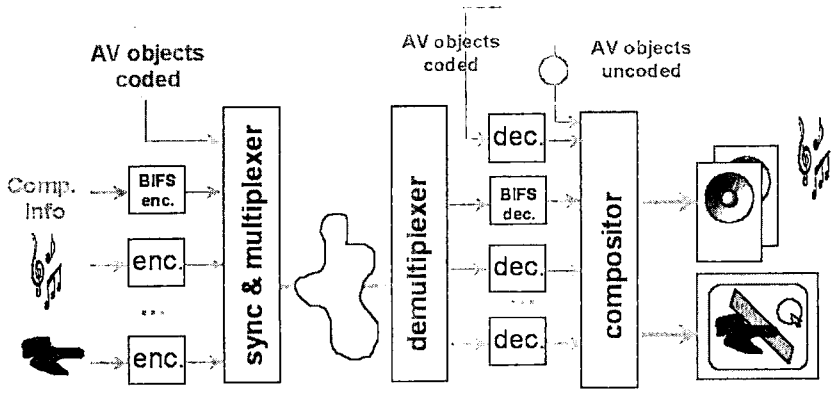
- **A/V Object Compression**
  - **General toolbox for application**
  - **Interactivity based on contents**
- **Characteristics**
  - **Accessibiliy**
  - **Usability**
  - **Interactivity**
  - **Error Resilience**
  - **Scalability**

MPEG-4 Elements

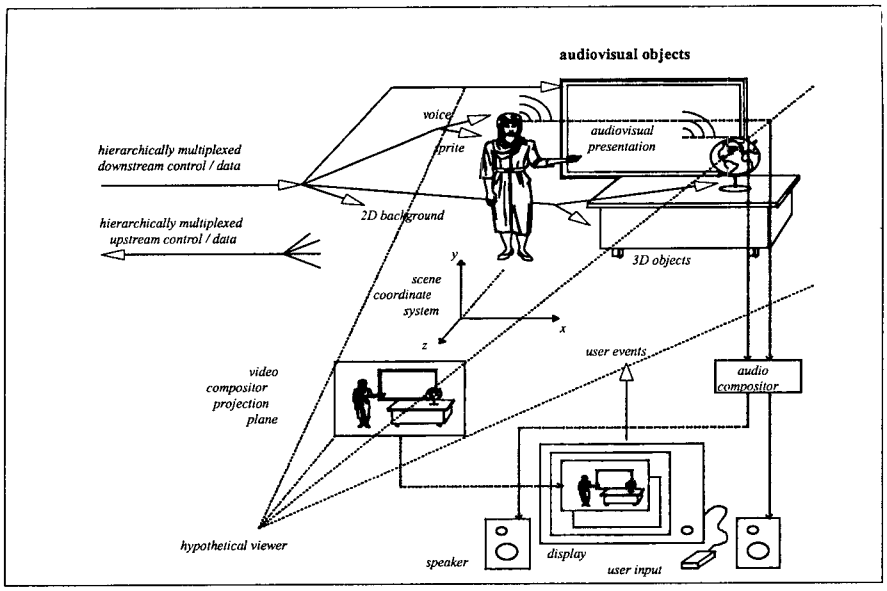
- Communication access and manipulation of digital audio-visual data



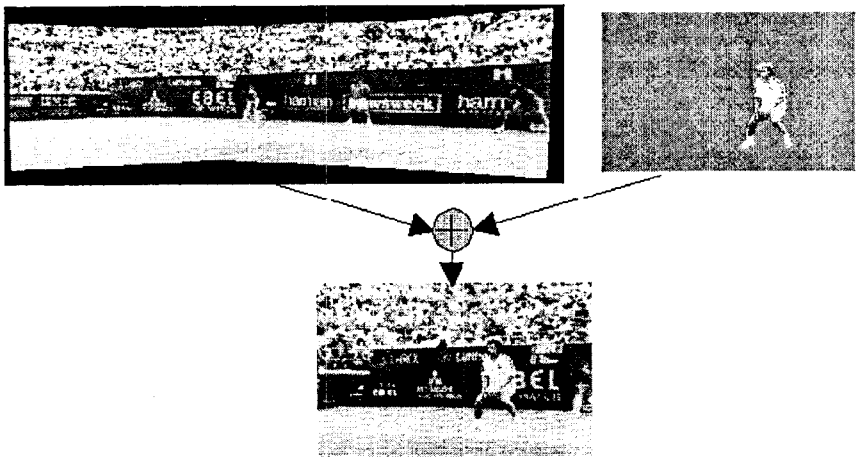
MPEG-4 A/V Systems



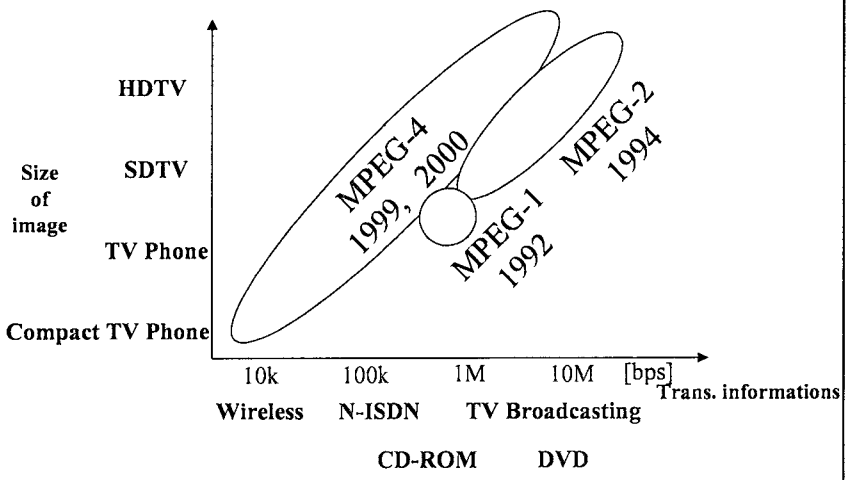
MPEG-4 Scenes



Static Sprite Coding Tools



The Position of the MPEG-4



Relations between Standards

Decoder \ Bitstream	MPEG-1	MPEG-2	H.263	MPEG-4
MPEG-1	○	x	x	x
MPEG-2	○	○	x	x
H.263	x	x	○	x
MPEG-4	x	x	○*	○

○ : decoding

x : Not mandatory

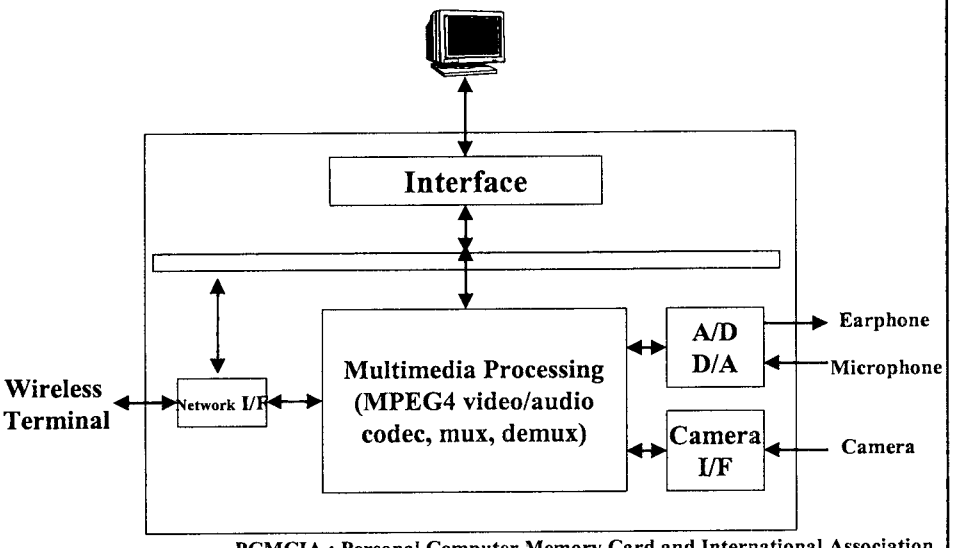
* : Baseline bitstream which is not use the option of H.263

Comparison of MPEG-4, DV and Motion JPEG2000

Format	Compression Performance	Complexity	Easiness of Editing	Scalability
Motion JPEG2000	0	+	++	++
DV	-	+	++	None
MPEG-4	++	0	0	+

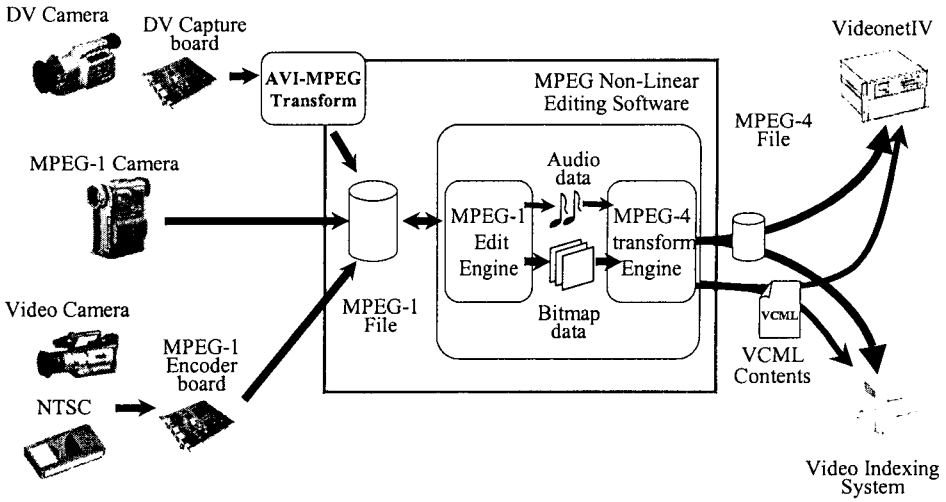
(++ : Very Good, + : Good, 0 : Average, - : Poor)

Configuration of MPEG4 Mobile Information Terminal



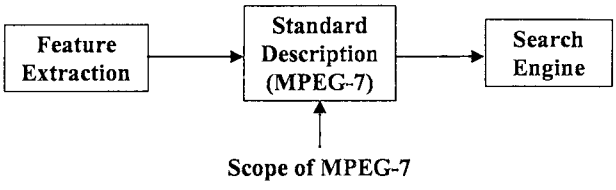
PCMCIA : Personal Computer Memory Card and International Association

Video Editing based on MPEG-4



MPEG-7

- **Multimedia Content Description Interface**
- **Multimedia Data in Various forms of media**
- **Need for efficient search / retrieval, and maintenance of multimedia data**



People wants

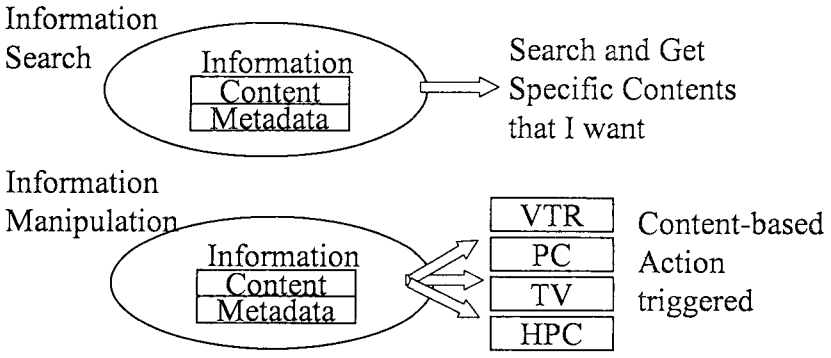
- to find a picture of "the Motorbike from Terminator II"
- to search a sequence where "King Lear congratulates his assistants on the night after the battle"
- to search for "twenty minutes of video according to my preferences of today"

⟹ Need for **CONTENT-based** search technology

Object of MPEG-7

- **Goal :**
  - **allowing efficient search for multimedia content using standardized descriptions**
- **MPEG-1, 2, 4 :**
  - **Representation of contents itself**
- **MPEG-7 :**
  - **Representation of information about the content**
    - **"Metadata"**
    - **"Bits about Bits"**

Scope of MPEG-7



What to Standardize ?

- A set of **Descriptors**
- A set of **Description Schemes**
- **Description Definition Language**
  - A Language to specify DSs (and possibly Ds)
- One or more ways to **Encode** descriptions

Applications of MPEG-7

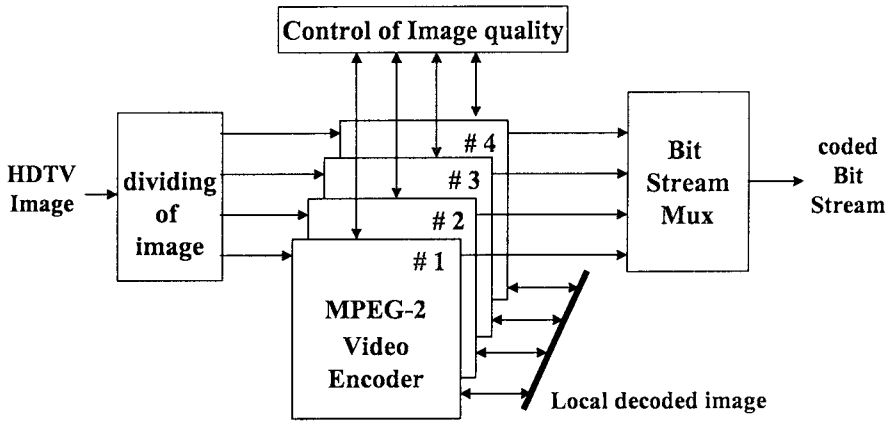
- **Push** applications
  - broadcasting, web-casting
  - indexing and retrieval --> selection and filtering
  - streamed descriptions
  - *personalized TV service*

- **Pull** Applications
  - audio-visual archives, databases, web search
  - initial motive for MPEG-7,
    - AV material "as searchable as text is today"

International Standards

Standards	Bit rate	Source	Application	Year
JPEG (Still Image)	10 – 20:1	Image	Storage, Compressed TX	1992
H.261 (Slow Motion)	P x 64 Kbps	Video	Video phone, Video conference	1990
MPEG -1 (Moving Picture)	1 – 1.5 Mbps	Video/Audio	Storage(DSM)	1992
MPEG -2 (H.262)	≤ 100 Mbps	Video/Audio	Transmission (DTV, HDTV)	1994
MPEG -4 (H.263)	≤4 Mbps (≤ 64 Kbps)	Video/Audio	Video phone, Mobile, Internet	Ver. 1 : 1999.02 Ver. 2 : 2000.02
FCC GA, ATSC (HDTV, DTV)	≤ 100 Mbps	Video/Audio	HDTV, DTV	1993
JPEG 2000	50:1	Image/Video *	Internet/WWW Imagery, DSC *	2000.12
MPEG -7	Multimedia content Description Interface	Video/Audio	Visual Content Search	2001.09

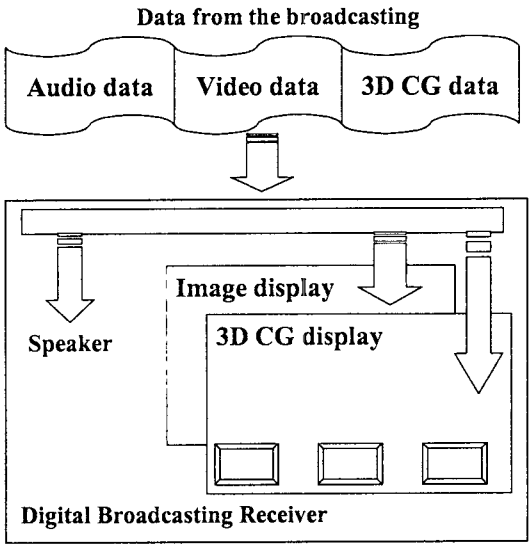
Video Encoder for HDTV



Example of HDTV Codec Spec.

Image	Resolution		1,440 pixel x 1,035 line x 30 frame
	Color format		4 : 2 : 0 / 4 : 2 : 2
	codec	MPEG-2 Simple Profile@High 1440 Level	
		Picture	Frame structure
		Motion compensation	Frame / Field / Dual frame
		DCT	Frame / Field
	BPS		20 ~ 60 Mbit/s
audio	Codec		MPEG-1 Layer 2
	number of channel		4 ch
	BPS/Ch		192 kbit/s /ch
Mux	MPEG-2 transport stream		
	Error correction		Reed-Solomon
Bit rate	BPS		44.736 Mbit/s or 59.648 Mbit/s

3D processing of Digital Broadcasting Receiver



Conclusion

- MJ2K: Easy editing, robust, low cost for the video engine in the internet appliances
- Universal Accessibility
  - Interoperability based on WWW
  - Scalability, Error Prone
- Interactive Service
  - Manipulation and design for the contents
- Multimedia contents description
- Multimedia Framework (MPEG-21)
  - More Information (http://drogo.cselt.it/mpeg/)