Chapter 3 Digital Video Fundamentals

This portion of lecture notes are courtesy of Prof. A. Murat Tekalp of Koc University, Istanbul, Turkey, which should give us a brief review of digital video concepts, systems, and the standards. Some of these topics will be studied in detail during the semester.

ANALOG VIDEO

One or more analog signals that contain time-varying 2-D intensity (monochrome or color) and timing information to align the pictures.

- **Component Analog Video (CAV)**
  - RGB
  - YCrCb (YUV or YIQ)
- **Composite Video**
  - NTSC (National Television Standards Committee)
  - PAL (Phase Alternating Line)
  - SECAM (SEquential Color And Memory)
- **S-Video (Y/C video)**
  - NTSC / PAL / SECAM

Color Spaces for Analog Video

- Color is specified in terms of three primaries, commonly R, G, B.
- Not all R, G, B spaces are the same. They may have different primaries.
- R, G, B primaries themselves are specified in terms of X, Y, Z values.

SMPTE C (NTSC)

Colorimetry	Red	Green	Blue	White D65
x	0.630	0.310	0.155	0.3127
y	0.340	0.595	0.070	0.3290

EBU (PAL, SECAM)

Colorimetry	Red	Green	Blue	White D65
x	0.640	0.290	0.150	0.3127
y	0.330	0.600	0.060	0.3290

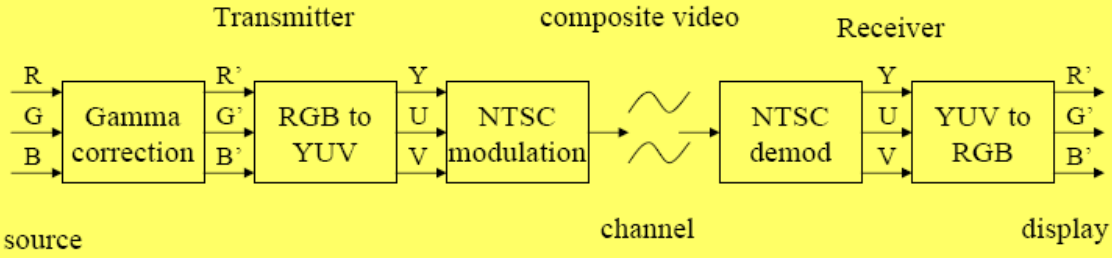
- R, G, B values are converted to Y, U, V before transmission/storage using the following transformation.

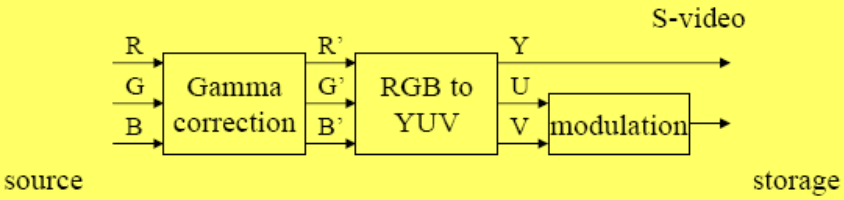
• PAL	• NTSC
Y = 0.299R + 0.587G + 0.114B	Y = 0.299R + 0.587G + 0.114B
U = -0.147R - 0.289G + 0.436B	I = 0.596R - 0.274G - 0.322B
V = 0.615R - 0.515G - 0.100B	Q = 0.211R - 0.523G + 0.311B

CRT Display: Gamma Correction

- CIE luminance, chrominance vs. NTSC luma, chroma
- Physics of CRT: The intensity on the screen is proportional to the voltage input (video signal) raised to the power 2.5
- To maintain the correct tone scale, this nonlinearity of CRT must be compensated at acquisition by a 0.45 power law. R'=R0.45 G'=G0.45 B'=B0.45
- Perceptual uniformity
  - CRT voltage-intensity function is nearly the inverse of luminance-perceived luminance function.
- Noise sensitivity, SNR
  - Noise makes a larger contribution to small signal values

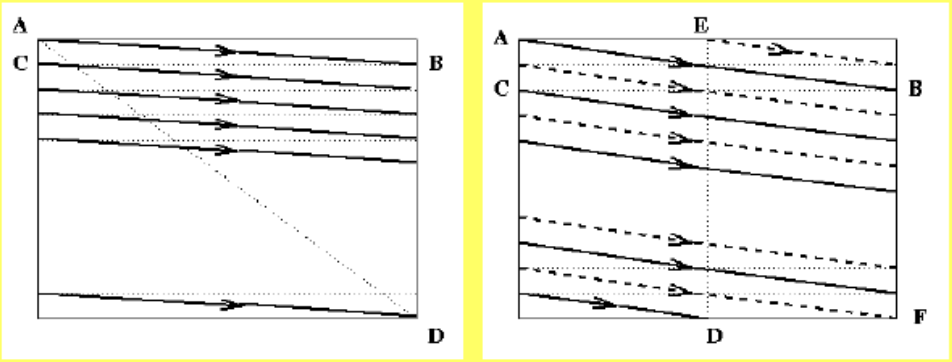
Generation of Composite and S- Video





Scanning

Scanning determines a mapping between spatial position and time.



- ***Progressive scan****:* Each frame is made up of lines.
- ***Interlaced scan***: Each frame is split into two *fields*. This provides a tradeoff between temporal and vertical resolution.

Temporal Rate and Flicker

- **Frame/field rate and flicker***:* Minimum *refresh rate* for flicker-free viewing is 50 Hz (temporal sampling).
- **Field vs. frame rate**:
  - Motion picture Progressive, 24x2=48 Hz
  - TV (NTSC) Interlaced, 60 Hz
  - TV (PAL/SECAM) Interlaced, 50 Hz
  - Computer monitor Progressive, > 72 Hz
- **Viewing conditions**
  - Dim vs. bright environment

Viewing distance = (3400/Lines) x Picture Height one pixel subtends to one minute arc (1/60 of a degree)

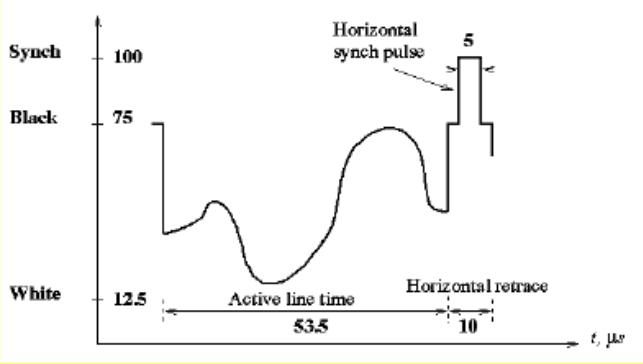
International Analog TV Standards

	Aspect Ratio	Interlace	Frames/s	Total/Active Lines	BW (MHz)
NTSC (USA,Japan, Canada,Mexico)	4:3	2:1	29.97	525/480	4.2
PAL (Great Britain)	4:3	2:1	25	625/580	5.5
PAL (Germany, Austria, Italy)	4:3	2:1	25	625/580	5.0
PAL (China)	4:3	2:1	25	625/580	6.0
SECAM (France,Russia)	4:3	2:1	25	625/580	6.0

Image aspect ratio = image width /image height

Synchronization

Scanning at the display device must be synchronized with that at the source.



NTSC video signal for one full line.

- *Blanking pulses* are inserted during the retrace intervals to blank out retrace lines on the receiving CRT.
- *Sync pulses* are added on top of the blanking pulses to synchronize the receiver's horizontal and vertical sweep circuits.

Resolution and Bandwidth

$$BW = \frac{1}{2} \frac{FR \times NL \times HR}{\rho}$$

FR = Frame Rate

NL = Number of Lines/Frame

HR = Horizontal Resolution

ρ = fraction of time allocated to active video signal per line

- *Example*: Video Bandwidth of the NTSC signal

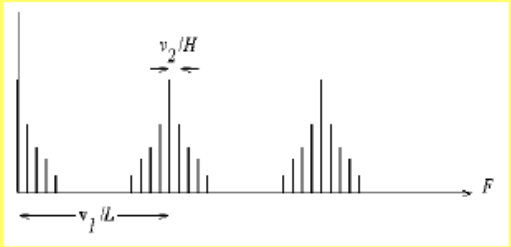
  ρ = 53.5 / 63.5 = 0.84

  BW = 4.2 MHz

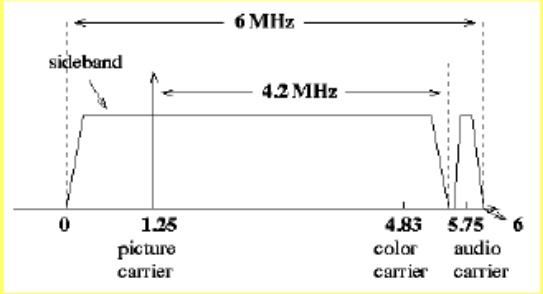
  Line Rate = FR x NL = 29.97 x 525 = 15,734

  $$HR = \frac{2 \times 4.2 \times 10^6 \times 0.84}{15{,}734} = 448 \text{ pixels}$$

Spectral Content



- Spectrum of the scanned video signal for still images.



- Spectrum of the NTSC video signal.

Analog Video Recording

Video Format	*Tape Format*	*Number of Lines*	*Luma Bandwidth*
Composite	VHS, 8mm	240	3.0 MHz
	U-matic	330	4.0 MHz
S-Video	S-VHS, Hi8	400	5.0 MHz
Component	Betacam SP	480	4.5 MHz

DIGITAL VIDEO

- Digital data communications (e.g., ftp, e-mail)

and

- Digital audio (e.g., CD players, digital telephony)

- *What is next?*

  Digital video - as a form of computer data

  Products such as: digital TV/HDTV, video/cell phone, multimedia PCs, are alpready in the marketplace.
- ``Digital video,'' **IEEE Spectrum Magazine**, pp. 24-30, Mar. 1992 to see what was predicted in 1992.

Why Digital Video?

- Digital representation is robust: Error correction minimizes the effect of transmission/storage media distortion, noise and other degradations.
- Digital video requires lower bandwidth than analog video of equivalent subjective quality by using compression.
- Digital video enables integration of interactive networked multimedia, broadcast TV, and real-time communications in a unified system architecture.
- Digital video provides flexibility for signal processing for enhancement, standards conversion, composition, special effects, nonlinear editing, etc.

What is the Bottleneck?

Let's look at the raw data rates for digital audio and video:

- CD quality digital audio (mono)
- 44 kHz sampling rate x 16 bits/sample 700 kbps
- High definition video (from the GA-HDTV proposal)
- 1280 pels x 720 lines luma; 640 pels x 360 lines chroma x 60 frames/s x 8 bits/pel/channel 663.5 Mbps
- A picture is worth 1000 words!!
- A good source: Inglis and Luther, Video Engineering, McGraw Hill, pp. 160-178, 1996.

Color Spaces for Digital Video

- ITU-R BT.601

  Y-Cb-Cr are shifted and scaled versions of the analog Y-U-V components.

  $$Y = 0.257R + 0.504G + 0.098B + 16$$
  $$Cb = -0.148R - 0.291G + 0.439B + 128$$
  $$Cr = 0.439R - 0.368G - 0.071B + 128$$

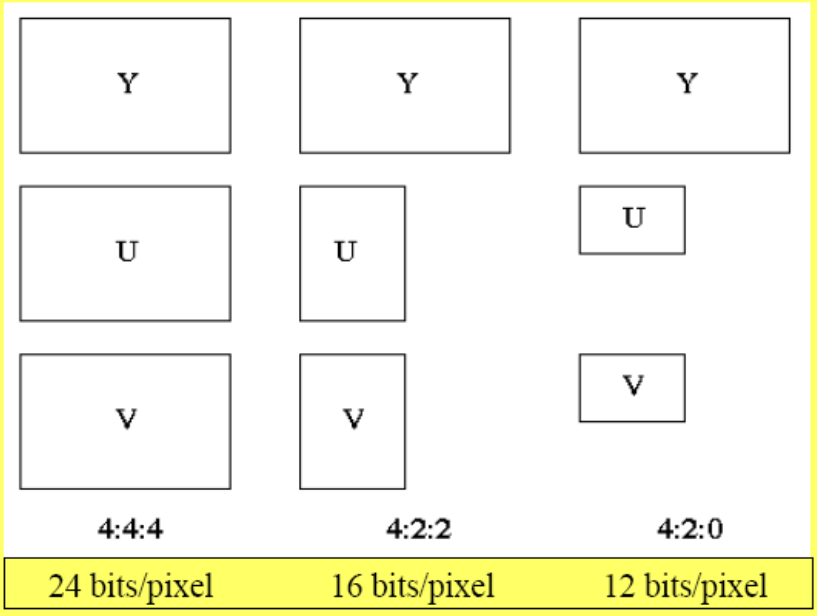
  where R, G, B are in the range (0-255).

- ITU-R BT.709 (HDTV monitors)

Colorimetry	Red	Green	Blue	White D65
x	0.640	0.300	0.150	0.3127
y	0.330	0.600	0.060	0.3290

$$Y = 0.2215R + 0.7154G + 0.0721B$$
$$Cb = -0.1145R - 0.3855G + 0.5000B$$
$$Cr = 0.5016R - 0.4556G - 0.0459B$$

Chrominance Formats for Digital Video



Digital Video Standards

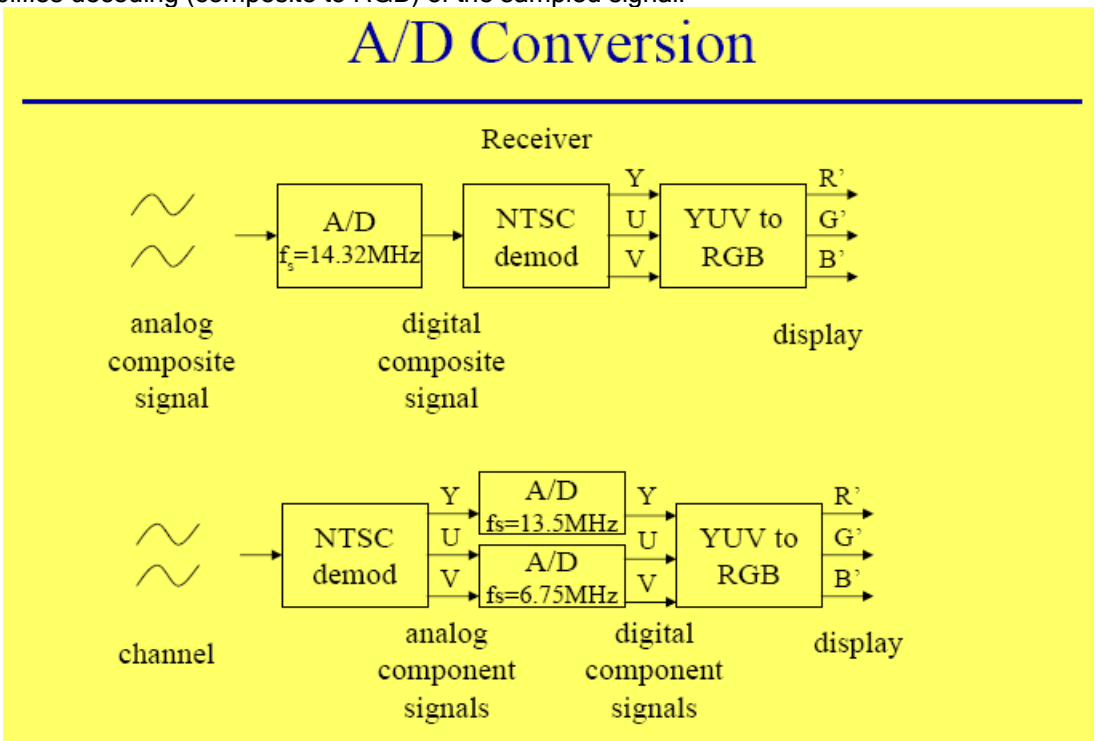
	ITU-R 709	ITU-R 601 625/50 (525/60)	CIF	QCIF
Number of active pels/line				
Lum (Y)	1920	720	360	180
Chroma (U,V)	960	360	180	90
Number of active lines				
Lum (Y)	1080	576 (480)	288	144
Chroma (U,V)	1080	576 (480)	144	72
Interlace		2:1	1:1	1:1
Pictures/sec	25 (30)	25 (30)	30	5-15
Aspect ratio	16:9	4:3	4:3	4:3
Raw data (Mbps)	884.7	165.9	37.3	

Analog-to-Digital Conversion:

- The minimum sampling frequency is 4.2 x 2 = 8.4 MHz (Nyquist rate)
- Sampling rate should be an integral multiple of the line rate, so that samples in successive lines are aligned.
- To sample *component signals*, there should be a single rate for 525/30 and 625/50 systems; i.e., the sampling rate should be an integral multiple of both 29.97 x 525 = 15,734 and 25 x 625 =15,625.
- To sample *the composite signal*, the sampling frequency must be an integral multiple of the subcarrier frequency.

This simplifies decoding (composite to RGB) of the sampled signal.

A/D Conversion



Sampling Component Signals

		525/59.94 SMPTE125M	*625/50*
Luminance	Sampling frequency	13.5 MHz	13.5 MHz
	Total/Active Samples line	858/720	864/720
	Bitrate	108 Mbps	108 Mbps
Chrominance 4:2:2	Sampling frequency	6.75 MHz	6.75 MHz
	Total/Active Samples line	429/355	432/358
	Bitrate	54 Mbps	54 Mbps

Total bitrate= 216 Mbps; Active picture area = 165.9 Mbps

Sampling the Composite Signal

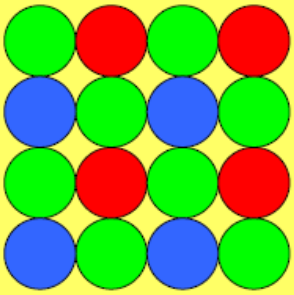
	NTSC 3 fsc	*NTSC SMPTE 244M*	*PAL 4 fsc*
Bandwidth (MHz)	4.2	4.2	5.5
Subcarrier/ Sampling frequency	3.58/10.74 MHz	3.58/14.32 MHz	4.43/17.72 MHz
Total/Active Samples lines	682/576	910/768	1134/939
Bitrate (Mbps)	85.9	114.5	141.8

Aspect Ratio:

- Image aspect ratio (IAR) = image width / image height
- Pixel aspect ratio (PAR) = (IAR × No. of lines)/(No. of pels/line)
- Computer monitors (square pixels, PAR=1)
  IAR = (No. of pixels/line) / No. of lines
- SDTV (rectangular pixels)
  NTSC: IAR = 4:3; PAR= (4/3 × 483) / 720 = 0.89
  PAL: IAR = 4:3; PAR= (4/3 × 576) / 720 = 1.07
- HDTV, IAR = 16:9 = 1.777
- Motion pictures, IAR = 1:85:1 or 2.35:1

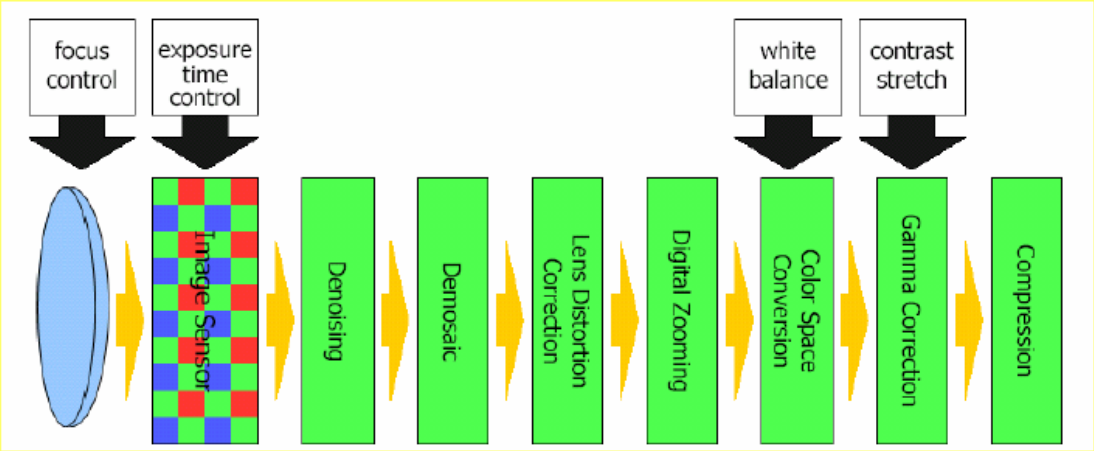
CCD Video Cameras

- Three-sensor array cameras
- Single-sensor array cameras



Bayer pattern

Digital Camera



Image/Video Compression Standards

ISO JPEG2000	Still frame gray scale and color images
ITU-T H.261	Video for ISDN applications (px64 kbps)
ITU-T H.263	Video for PSTN applications (less than 64 kbps)
ISO MPEG-1	Video for optical storage media (1.5 Mbps)
ISO MPEG-2	High quality generic video (4-20 Mbps)
ISO MPEG-4	Object-based video (10 kbps – 2 Mbps)
ISO/ITU-T JVT	

Proprietary File Formats

Quicktime	Apple Computer
AVI (Audio/Video Interleaved) **Video for Windows** **ASF (Advanced Streaming Format)**	Microsoft
Real Video	Real Networks
PhotoCD	Eastman Kodak Co.

SVGA Computer Monitors
1024 pixels × 720 lines @72fps

APPLICATIONS OF DIGITAL VIDEO:

- **Consumer Electronics: VideoCD, DVD**
  @ 1.5 Mbits/s CD-ROM or harddisk storage
- **Digital TV (Broadcast, Cable, or Satellite)**
  HDTV @ 20 Mbits/s over 6 MHz channels
  SDTV @ 4-6 Mbits/s
- **Video telephony**
  ISDN @ 384 kbits/s using p x 64 kbits/s
  PSTN @ up to 56 kbits/s using the copper network
  Wireless @ 10 kbits/s using GSM
- **Video over IP: Internet and wireless**
  bitrate depends on the connection speed
- Other
  Surveillance Imaging (military/ law enforcement) , Intelligent V. Highway Systems
  Telemedicine, Telepresence, Distance Learning

Digital Strorage / Consumer Electronics:

- VideoCD (MPEG-1) vs. DVD (MPEG-2)
- CD-ROM holds 650 Mbytes

- DVD

  Single-sided, single-layer ~4.7 Gbytes
  Single-sided, dual-layer ~8.5 Gbytes (133 minutes of MPEG-2 compressed movie at 8 Mbps)
  Double-sided, single-layer ~9.4 Gbytes
  Double-sided, dual-layer ~17 Gbytes

Digital TV: ATSC Standard:

- 18 video formats - includes SDTV and HDTV
- Colorimetry: 4:2:0 chrominance sampling; 16:9 aspect ratio
- MPEG-2 video compression and transport (188 byte packets)
- 8-VSB for terrestrial and 16-VSB for high data rate cable transmission
- Digital HDTV requires about 884:20 = 44:1 compression

ATSC Video Formats

No. of Lines	*No. of Pixels*	*Aspect Ratio*	*Picture Rate*
1080	1920	16:9	60I 30P 24P
720	1280	16:9	60P 30P 24P
480	704	16:9 4:3	60I 60P 30P 24P
480	640	4:3	60I 60P 30P 24P

DVB:

- **Europe** developed a variation of the ATSC standard known as Digital Video Broadcast (DVB) based on MPEG-2 video compression. DVB does not support HDTV mode.

Real-Time Communications

- Videoconferencing/phone over ISDN: up to 2 Mbps using H.261 or H.263
- Videophone over PSTN: 8 - 32 kbps using H.263 or H.263+
- Videoconferencing/phone over LAN/Internet: Heterogeneous networks

Compression Requirements to Reach 10 kbps

Frames/s	*ITU-R 601 720 x 576*	*CIF 352 x 288*	*QCIF 176 x 144*
7.5	4979:1	915:1	229:1
10	6637:1	1216:1	304:1
15	9952:1	1824:1	456:1
30	19904:1	3648:1	912:1

Transmission / Access Networks

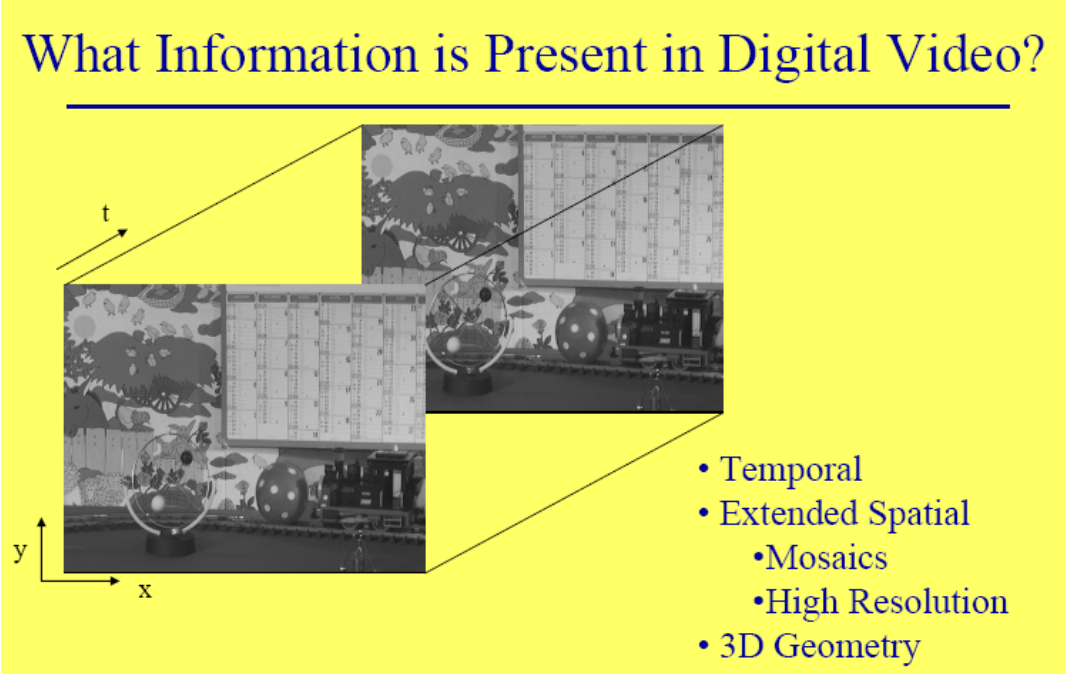
Analog Telephone Modem	28.8 – 56 kbps
ISDN	128 kbps – 2 Mbps (p x 64 kbps)
T-1	1.5 Mbps
DSL	1.5 – 6 Mbps (downstream)
Cable Modem	30 Mbps (downstream) shared by multiple users
Ethernet (LAN)	10-100 Mbps shared by multiple users
Fiber Backbone (B-ISDN/ATM)	55 Mbps – 1 Gbps
GSM	10 kbps
GPRS	30-40 kbps
CDMAOne	80 kbps
UMTS	384 kbps

Convergence of Video/Voice/Data Applications:

- Network infrastructures
  - Telephone networks (PSTN, ISDN, ADSL)
  - CableTV networks (Cable Modem)
  - Internet (network of networks) (TCP/IP)
  - Digital wireless (Mobile)
- Services:
  - Real-time two-way communications, video phone and conferencing
  - Digital TV, interactive TV
  - Interactive multimedia over the internet, Web-based services

Video over IP: Internet and Wireless:

- Standards-Based vs. Proprietary
- Compression Issues
  - Error resilience
  - Scalability
  - Buffer Control
- Protocol Issues
  - TCP/IP
  - RTP
- Packetization Issues



Challenges in Digital Video Processing:

Exploit temporal redundancy by motion analysis

- Motion Analysis
  - 2-D motion/optical-flow estimation and segmentation
  - 3-D motion, structure estimation and segmentation
  - Object tracking, occlusion, deformations
- Filtering and Standards Conversion
  - Deblurring, noise filtering, edge sharpening
  - Frame rate conversion and deinterlacing, resolution enhancement
- Compression
  - JPEG, H.261/H.263, MPEG 1-2-4
  - Subband/wavelet and content-based coding

Convergence of Digital Video, Computer Vision and Graphics:

- DSP
  - Transforms - DFT, DCT, JPEG/MPEG compression
  - Linear and nonlinear filtering, restoration
- Computer Vision
  - Video object segmentation and tracking
  - 3D motion modeling - Structure from motion
- Computer Graphics
  - Animation, texture mapping.