CMPT 365 Multimedia Systems

Introduction

Xiaochuan Chen
Spring 2017

Edited from slides by Dr. Jiangchuan Liu

Outline

- Course information
- What is multimedia? A brief introduction
- Popular multimedia tools
- Summary

Course Information

- **Instructor:**
  - Xiaochuan CHEN

    School of Computing Science
    Office: TASC I 8002
  - E-mail: xca64@sfu.ca
  - Office Hours: Wed 10:45-11:45am
  - E-mail is the best way to communicate with me
    - or send me email for special appointment
- TA
  - Saeedeh Afshari (safshari@sfu.ca)
  - Office Hours: TBA

Course Information

- **Time & Venue**
  - M/W/F 2:30PM - 3:20PM AQ 3149

Why this course?

- Multimedia is cool
  - Media -> Multimedia
  - Everywhere
  - Requires broad knowledge in mathematics, signal processing, communications, networking, software, hardware, ...
- Job opportunities
  - Multimedia is a booming industry
    - in the metro Vancouver area
  - Tons of opportunities created by next-generation standards and emerging applications:
    - JPEG/JPEG 2000
    - MPEG-1/2/4 H.264/265/HEVC 4K UHD 3D/freeview
    - 3G/4G/5G mobile communications
    - Multimedia-enabled smartphone, tablets
    - Social media, Cloud media, Crowd media
    - Online gaming

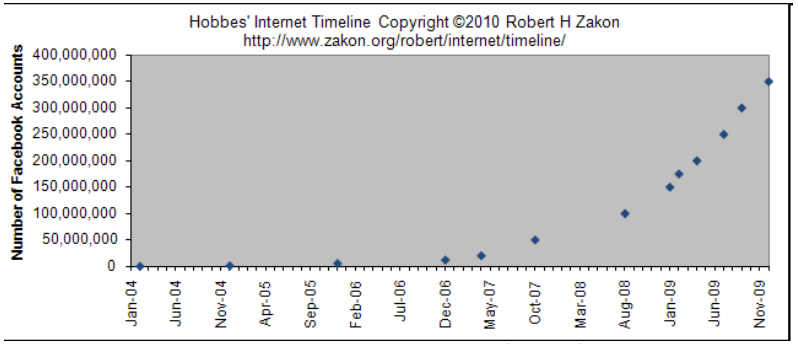
Example

- Old: NTT DoCoMo 3G Mobile Phone:
  - launch in 2001
  - 99% coverage in Japan as of March 2004
  - Up to 384 kbps video downloading
  - 40 times faster than 2G network (comparable to ADSL)
- New: 4G LTE Mobile Phone:
  - 100 Mbps for high mobility communication
  - 1 Gbps for low mobility communication
  - allow 3D virtual reality and interactive video / hologram images
  - Commercial service since in 2010
  - 97% of the population in Canada now

Killer Internet Applications

- Web2.0/Media streaming (Internet TV)
  - YouTube, Netflix
  - HD/UHD video ?
  - 3D video ?
- E-commerce
  - Ebay, Amazon, Craigslist, Groupon
- Online game
  - PS3, XBOX 360, Wii
- …

- Social networking (2004-)
  - Facebook, Twitter, WhatsApp …
- Mobile Internet
  - iPads, tablets …
  - End of PC ? …

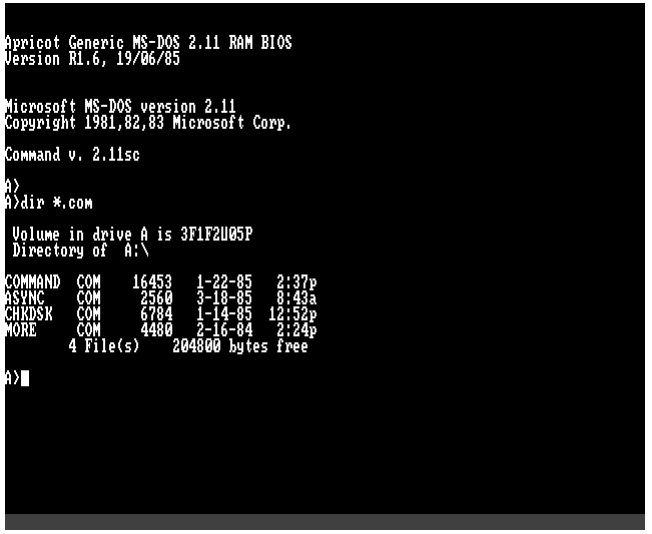


Multimedia Companies

- Microsoft
- Intel
- AMD
- Adobe
- RealNetworks
- Apple
- Google
- Facebook
- Twitter
- Nokia
- NEC
- Sony
- Sharp
- Philips
- Panasonic
- YouTube
- Netflix
- …

What are the objectives of this course?

- Understand what's behind the interface
  - Behind VCD, DVD, BluRay, HDTV, mp3, flac, raw, jpeg ? …
  - 3D, 4K TV ?
- Process multimedia data by yourself (programming projects)
- Have fun!
  - What a life without multimedia ?!
    - A PC with black-white monitor only …



More details

- To understand the methods for multimedia representation and compression

  - *Representation (audio/video)*
    - Digitization
    - Quantization

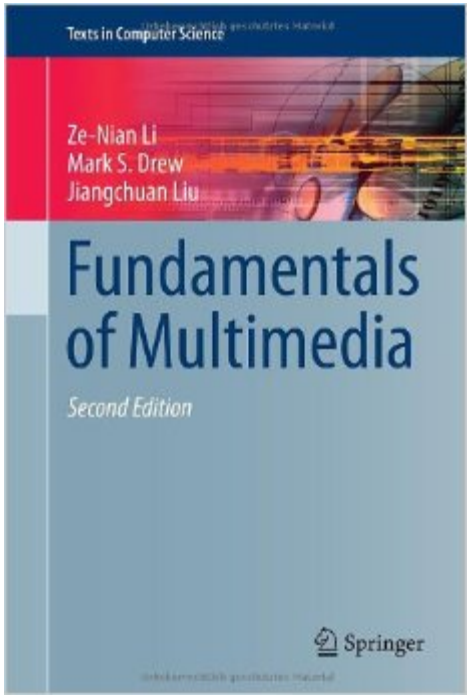
  - *Compression (audio/video)*
    - Transform
    - Entropy Coding
    - Coding Standards

  - *Communication**

- To help you survive a job interview in multimedia
  - Programming assignments
  - C, C++, Java, Python, Matlab could be involved

Books and References

- Textbook
  - *Fundamentals of Multimedia, 2nd Edition, by Z.-N. Li, M.S. Drew, and J. Liu, Springer 2014.*
- Others
  - A reference book on C/C++/Java
- Resource
  - Home page
    - www.sfu.ca/~xca64/cmpt365
  - Pls check your email



What Do You Need To Do?

- Your prerequisites
  - Data structure, algorithms
  - Math (calculus, linear algebra, probability)
  - programming: C/C++, Java
  - basic concepts of operating systems/GUI
- Your workload
  - Homework assignments
    - 2 assignments [written and coding]
    - 1 final programming project
  - One in-class midterm exams, and one final exam

Grading (tentative)

Assignment x 2	20%
Programming work	28%
In-class midterm	20%
Final exam	32%

- Class participation
- More important is what you learn than the grades

What Do You Need To Do?

- Your prerequisites
  - Data structure, algorithms
  - Math (calculus, linear algebra, probability…)
  - Programming: C/C++, Java
  - Basic concepts of operating systems/GUI

- Remember: It's a computer science course

Hard math example (1)

- Suppose:
  - a data source generates output sequence from a set $\{A_1, A_2, \ldots, A_N\}$
  - $P(A_i)$: (Independent) probability of A_i
- First-Order Entropy:
  - the average self-information of the data set

$$H = \sum_i -P(A_i)\log_2 P(A_i)$$

- The first-order entropy represents the minimal number of bits needed to losslessly represent one output of the source.

Hard math example (2)

- Quantization error: $e(x) = x - \hat{x}$
- Mean Squared Error (MSE) for Quantization
  - Average quantization error of all input values
  - Need to know the probability distribution of the input

- Number of bins: M
- Decision boundaries: b_i, i = 0, …, M
- Reconstruction Levels: y_i, i = 1, …, M
- Reconstruction:

$$\hat{x} = y_i \quad \text{iff } b_{i-1} < x \le b_i$$

- MSE:

$$MSE_q = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x) dx = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f(x) dx$$

  - Same as the variance of e(x) if μ = E{e(x)} = 0 (zero mean).
  - Definition of Variance:

$$\sigma_e^2 = \int_{-\infty}^{\infty} (e - \mu_e)^2 f(e) de$$

Hard math example (3)

- MSE

$$MSE_q = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x)dx = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f(x)dx$$

$$= M \frac{1}{2X_{\max}} \int_0^{\Delta} \left(x - \frac{\Delta}{2} \right)^2 dx = \frac{M}{2X_{\max}} \frac{1}{12} \Delta^3 = \frac{1}{12} \Delta^2$$

  - M increases, Δ decreases, MSE decreases

- Variance of a random variable uniformly distributed in [- Δ/2, Δ/2]:

$$\sigma^2{}_q = \int_{-\Delta/2}^{\Delta/2} (x - 0)^2 \frac{1}{\Delta} dx = \frac{1}{12} \Delta^2$$

- Optimization: Find M such that MSE ≤ D

$$\frac{1}{12} \Delta^2 \le D \Rightarrow \frac{1}{12} \left(\frac{2X_{\max}}{M} \right)^2 \le D \Rightarrow M \ge X_{\max} \sqrt{\frac{1}{3D}}$$

Hard math example (4)

$$\mathbf{C}_{i,j} = a\cos\left(\frac{(2j+1)\,i\,\pi}{2N}\right), \quad \mathrm{i, j} = 0, \ldots, \mathrm{N}-1.$$

- Definition:

$$a = \sqrt{1/N} \quad \text{for i} = 0,$$

$$a = \sqrt{2/N} \quad \text{for i} = 1, \ldots, \mathrm{N}-1.$$

- N = 2 (Haar Transform):

$$\mathbf{C}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \mathbf{C}_2\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_0 + x_1 \\ x_1 - x_1 \end{bmatrix}$$

- y_0 captures the mean of x_0 and x_1 (low-pass)
  - $x_0 = x_1 = 1$ ➔ y_0 = sqrt(2) (DC), $y_1 = 0$
- y1 captures the difference of x0 and x1 (high-pass)
  - $x_0 = 1$, $x_1 = -1$ ➔ $y_0 = 0$ (DC), y_1 = sqrt(2).

Hard math example (5)

- Forward transform y = Tx (x is N x 1 vector)
  - Let t_i be the i-th row of T
  - ➔ $y_i = t_i x = \langle t_i^T, x \rangle$ (Inner product)
  - y_i measures the similarity between x and t_i
  - Higher similarity ➔ larger transform coefficient

- Inverse transform:

$$\mathbf{x} = \mathbf{T}^T \mathbf{y} = \begin{bmatrix} \mathbf{t}_0^T & \mathbf{t}_1^T & \dots & \mathbf{t}_{N-1}^T \end{bmatrix} \mathbf{y} = \sum_{i=0}^{N-1} \mathbf{t}_i^T y_i$$

- x is the weighted combination of t_i.
  - Rows of T are called basis vectors.

What Do You Need To Do?

- Your prerequisites
  - Math (calculus, linear algebra, probability…)

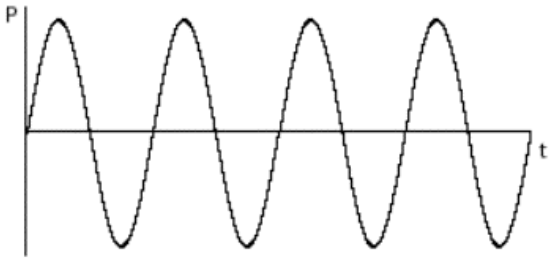
- Never heard of them? -- you'd better drop the course
- Forget ? -- you'd better drop unless you're confident you can pick them up

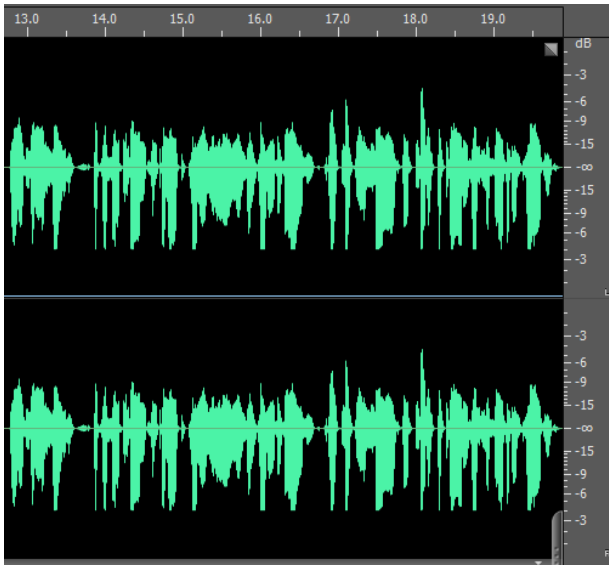
Hard programming example

Interfaces

- Read data byte-by-byte from an input file
  - E.g., read *en.wikipedia.org/wiki/BMP_file_format*
  - Write a program to read a BMP file

- Output an image to screen pixel by pixel





Algorithms

- Do this as fast as possible !

What Do You Need To Do?

- Your prerequisites
  - Data structure, algorithms
  - Programming: C/C++, Java
  - Basic concepts of operating systems/GUI

- Can't do them by yourself?
  - Better drop this course
  - project is NOT group-based!

(The course is NOT about *using* YouTube, Photoshop ….;
rather, it's about *write your own* YouTube, Photoshop …)

Questions?

Outline

- Course information
- What is multimedia? A brief introduction
  - Concepts
  - Representation
  - Compression
  - Communication
- Popular multimedia tools
- Summary

What is "media" ?

Information represented in different formats/media

- text
- graphics
- images

} Discrete media: time independent

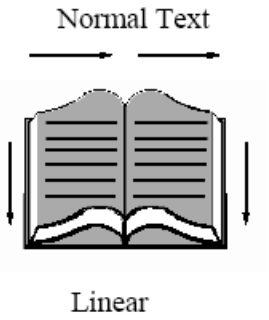
- animation
- audio
- video

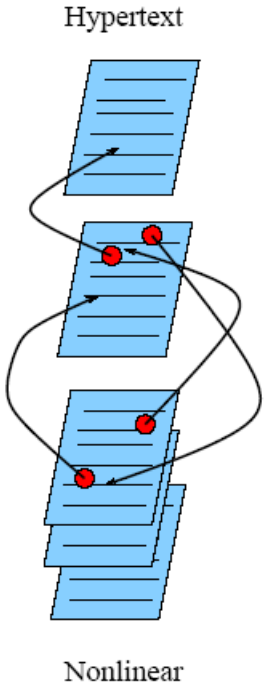
} Continuous media: time dependent

- Analog vs Digital
  - analog format: the time-varying feature (variable) of the signal is a continuous representation of the input, i.e., analogous to the input audio, image, or video signal
  - **Physical world is analog !**

Hyper Text, Hypermedia

- A **hypertext** system: meant to be read nonlinearly, by following links that point to other parts of the document, or to other documents
- HTLM/XML

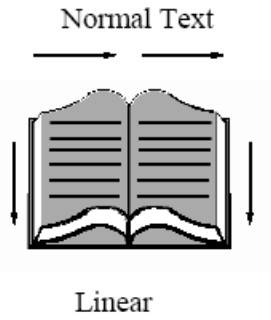


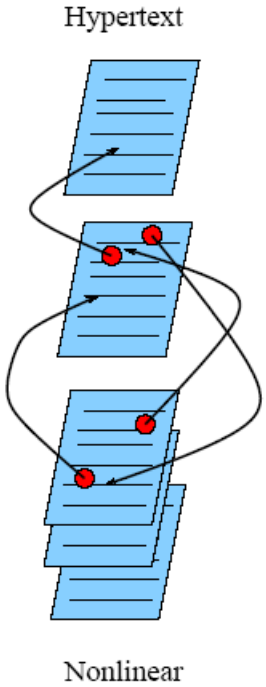



```
<HTML> <HEAD>
          <TITLE>
          A sample web page.
          </TITLE>
          <META NAME = "Author" CONTENT =                    "Cranky
    Professor">
    </HEAD> <BODY>
          <P>
          We can put any text we like here,
          since this is a paragraph element.
          </P>
    </BODY> </HTML>
```

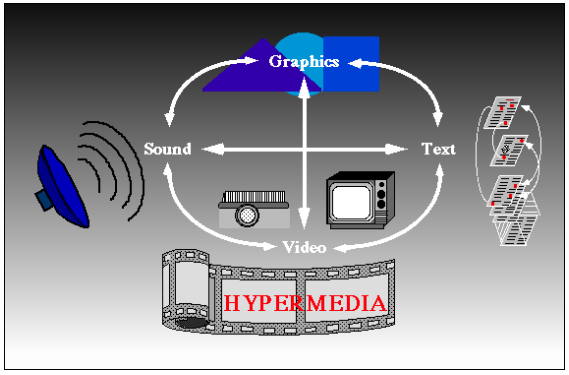
Hyper Text, Hypermedia

- A **hypertext** system: meant to be read nonlinearly, by following links that point to other parts of the document, or to other documents





- **HyperMedia**: not constrained to be text-based, can include other media, e.g., graphics, images, and especially the con tinuous media ¦ sound and video.
  - World Wide Web (WWW) --- the best example



Multimedia System

- **Multimedia:** information represented through audio, graphics, images, video, and animation in an integrated and interactive manner (as contrast to traditional single-modality media, i.e., text and graphics drawing).

- **Multimedia system:** the generation, manipulation, storage, presentation, and communication of multimedia information

Digital Media

- Multimedia **digitized**
  - Captured, stored, transmitted, processing in digital (discrete) domain
  - By general purpose computers or dedicated embedded computers
    - Today's digital cameras' have a number of CPUs inside, many of which are more powerful than a PC of 1990's or even 2000's.

- What do you mean by **digitized** ?
- Why **digitized** ?

(Digital/Computer) Multimedia Systems

- Using computers to present and process multimedia information, in an integrated and interactive manner

- Examples of Multimedia Systems:
  - Digital camera/camcord
  - World Wide Web
  - Video conferencing
  - Video-on-demand
  - Interactive TV
  - Online games
  - Virtual reality
  - Digital video editing and production systems
  - Multimedia Database systems
  - Social media

Different Views

- Different views from different people
  - A PC vendor: a PC that has sound capability, a DVD/BluRay drive, and perhaps the superiority of multimedia-enabled CPU/GPU (Graphical Processing Unit) that understand additional multimedia instructions.
  - A consumer entertainment vendor: interactive cable TV with hundreds of digital channels available, or a cable TV-like service delivered over a high-speed Internet/wireless connection.
  - A Computer Science (CS) student: applications that use multiple modalities, including text, images, drawings (graphics), animation, video, sound including speech; **integration** and **interactivity**.
- Multimedia and Computer Science:
  - Data representation compression
  - Graphics, visualization, computer vision
  - Networking, database systems

Multimedia Research Topics and Projects

- To the computer science researcher, multimedia consists of a wide variety of topics:

1. **Multimedia processing and coding**: multimedia content analysis, content-based multimedia retrieval, multimedia security, audio/image/video processing, compression, etc.
2. **Multimedia system support and networking**: network protocols, Internet, operating systems, servers and clients, quality of service (QoS), and databases.
3. **Multimedia tools, end-systems and applications**: hypermedia systems, user interfaces, authoring systems.
4. **Multi-modal interaction and integration**: web-everywhere devices, multimedia education including Computer Supported Collaborative Learning, and design and applications of virtual environments.
5. …

History of Multimedia

1. **Newspaper**: perhaps the *first* mass communication medium, uses text, graphics, and images.
2. **Motion pictures**: conceived of in the 1830's in order to observe motion too rapid for perception by the human eye.
3. **Wireless radio transmission**: Gugliemo Marconi, at Pontecchio, Italy, in 1895.
4. **Television**: the new medium for the 20th century, established video as a commonly available medium and has since changed the world of mass communications.
5. The **connection** between **computers** and ideas about **multimedia** covers what is actually only a short period: 1945 -- Vannevar Bush wrote a landmark article describing what amounts to a hypermedia system called **Memex**.

History of Multimedia cont'd

1960 Ted Nelson coined the term **hypertext**.

1967 Nicholas Negroponte formed the **Architecture Machine Group**.

1968 Douglas Engelbart demonstrated the **On-Line System** (**NLS**), another very early hypertext program.

1969 Nelson and van Dam at Brown University created an early hypertext editor called **FRESS**.

1976 The MIT Architecture Machine Group proposed a project entitled **Multiple Media** | resulted in the *Aspen Movie Map*, the first hypermedia videodisk, in 1978.

1985 Negroponte and Wiesner co-founded the **MIT Media Lab**.

1989 Tim Berners-Lee proposed the **World Wide Web**

1990 Kristina Hooper Woolsey headed the **Apple Multimedia Lab**.

1991 **MPEG-1** was approved as an international standard for digital video | led to the newer standards, **MPEG-2**, **MPEG-4**, and further **MPEGs** in the 1990s.

1991 The introduction of **PDAs** in 1991 began a new period in the use of computers in multimedia.

1992 **JPEG** was accepted as the international standard for digital image compression | led to the new JPEG2000 standard.

History of Multimedia cont'd

1992 The first **MBone** audio multicast on the Net was made.

1993 The University of Illinois National Center for Supercomputing Applications produced **NCSA Mosaic** -the first full fledged browser.

1994 Jim Clark and Marc Andreessen created the **Netscape**

1995 The **JAVA** language was created for platform-independent application development.

1996 **DVD video** was introduced; high quality full-length movies were distributed on a single disk.

1998 **XML** 1.0 was announced as a W3C Recommendation.

1998 **Hand-held MP3 devices** first made inroads into consumerist tastes in the fall of 1998, with the introduction of devices holding 32MB of flash memory.

2000 WWW size was estimated at over **1 billion pages**.

In the New Millennium

- Year 2000-, your time …
- Image/Audio
  - Huge/cheap flash memory
  - No worry anymore ?
    - 4K UHD 48 Gbps uncompressed

In the New Millennium

- **2001** The first peer-to-peer file sharing system, Napster, was shut down by court order. First commercial 3G wireless network.
- **2003** Skype: free peer-to-peer voice over the Internet.
- **2004** Web 2.0 promotes user collaboration and interaction. Examples include social networking, blogs, wikis.

  Facebook founded.

  Flickr founded .
- **2005** YouTube created.

  Google launched online maps
- **2006** Twitter created: 500 million users in 2012, 340 million tweets/day.

  Amazon launched its cloud computing platform.

  Nintendo introduced the Wii home video game console -- can detect movement in three dimensions.
- **2007** Apple launched iPhone, running the iOS mobile operating system. .

  Goolge launched Android mobile operating system.

In the New Millennium

- **2009** The first LTE (Long Term Evolution) network was set, an important step toward 4G wireless networking.

  James Cameron's film, Avatar, a surge on the interest in 3D video.
- **2010** Netflix migrated its infrastructure to the Amazon's cloud computing platform.

  Microsoft introduced Kinect, a horizontal bar with full-body 3D motion capture, facial recognition and voice recognition capabilities, for its game console Xbox 360.
- **2012** HTML5 subsumes the previous version, HTML4. Able to run on low powered devices such as smartphones and tablets.
- **2013** Twitter offered Vine, a mobile app that enables its users to create and post short video clips.

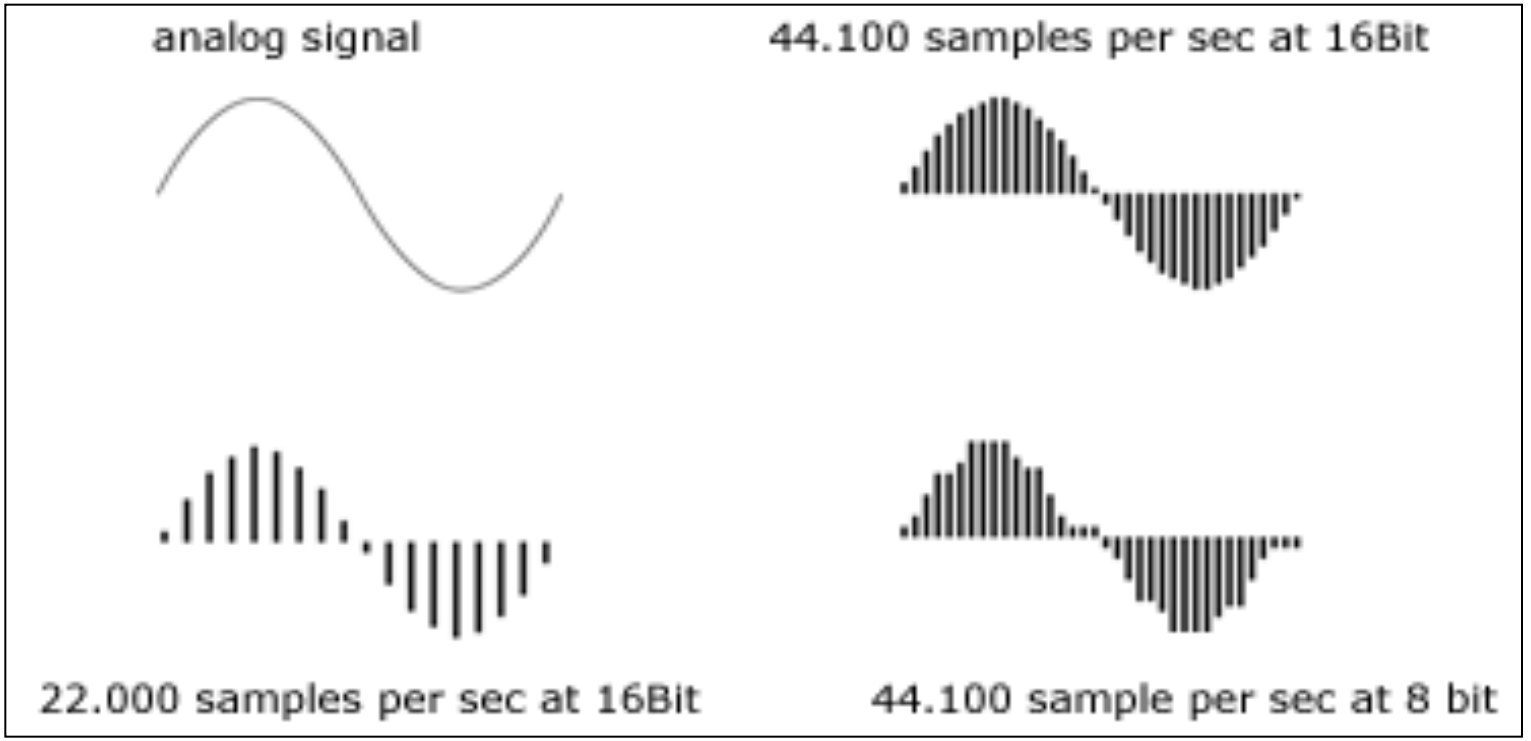
  Sony released its PlayStation 4 a video game console, which is to be integrated with Gaikai, a cloud-based gaming service that offers streaming video game content.

  4K resolution TV started to be available in the consumer market.

Outline

- Course information
- What is multimedia? A brief introduction
  - Concepts
  - Representation
  - Compression
  - Communication
- Popular multimedia tools
- Summary

Audio Digitization (PCM)



Representation ? ➔ Digitization for computers

Digital Media

- What do you mean by **digitized** ?
  - Audio/visual signals from the natural world is Analog
    - Continuous in time and space
    - Conventional storage/playback: LP (audio record), tape, CRT TV (old TV), film
    - Can't be handled by computer
  - A/D conversion
    - to 1/0 discrete signals

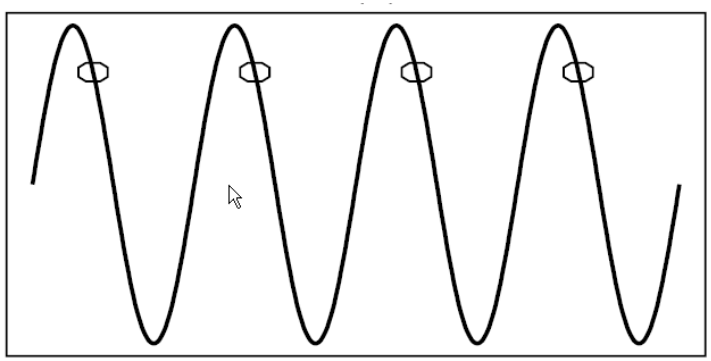
- Why **digitized** ?
  - Bulky storage (space, cost, lifetime)
  - Poor quality
  - Poor/no compression
  - Poor portability/mobility/editibility

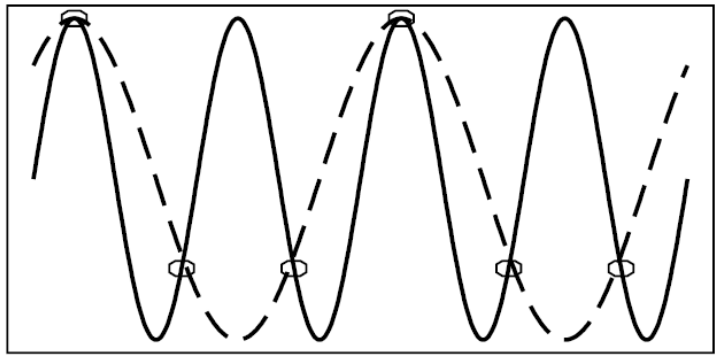
    MP3 player, iPod, YouTube ? No way

    Film -> Polaroid -> Digital camera

Sampling Rate

- Sampling theory – Nyquist theorem

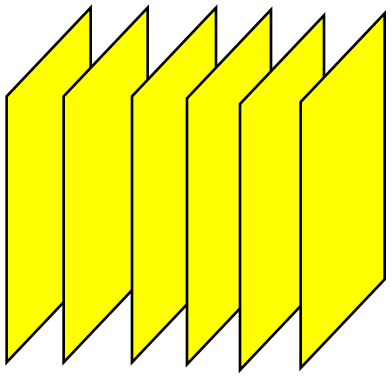




Image/Video Digitization

- Digital image is a 2-D array of pixels
- Each pixel represented by bits
  - R:G:B

  - Y:U:V
    - Y = 0.299R + 0.587G + 0.114B (Luminance or Brightness)
      U = B - Y (Chrominance 1, color difference)
      V = R - Y (Chrominance 2, color difference)
- Video is sequence of images (frames) displayed at constant frame rate
  - e.g. 24 images/sec



Outline

- Course information
- What is multimedia? A brief introduction
  - Concepts
  - Representation
  - Compression
  - Communication
- Popular multimedia tools
- Summary

Why Compression ?

- Multimedia data are too big
  - "A picture is worth a thousand words ! "

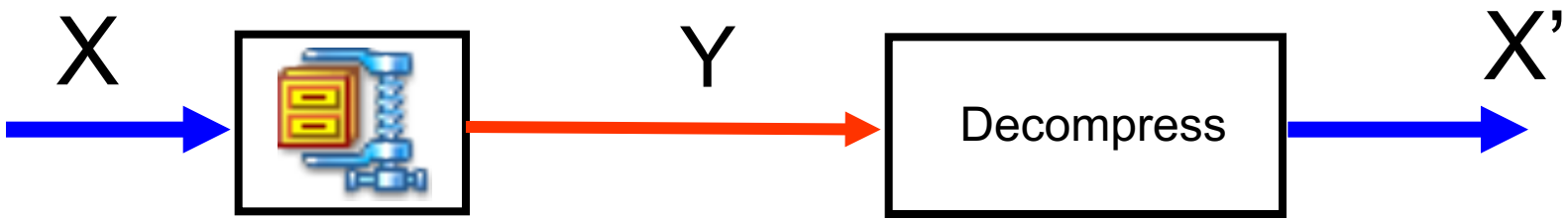
File Sizes for a One-minute Audio CD Clip

Sampling Rate	Resolution	Channels	Bit-rate (bps)	File Size (Bytes)
44,100Hz	16 bits	2	1,411,200	**10,584,000**

File Sizes for a One-minute QCIF Video Clip

Frame Rate	Frame Size	Bits / pixel	Bit-rate (bps)	File Size (Bytes)
30 frames/sec	176 x 144 pixels	12	9,123,840	**68,428,800**

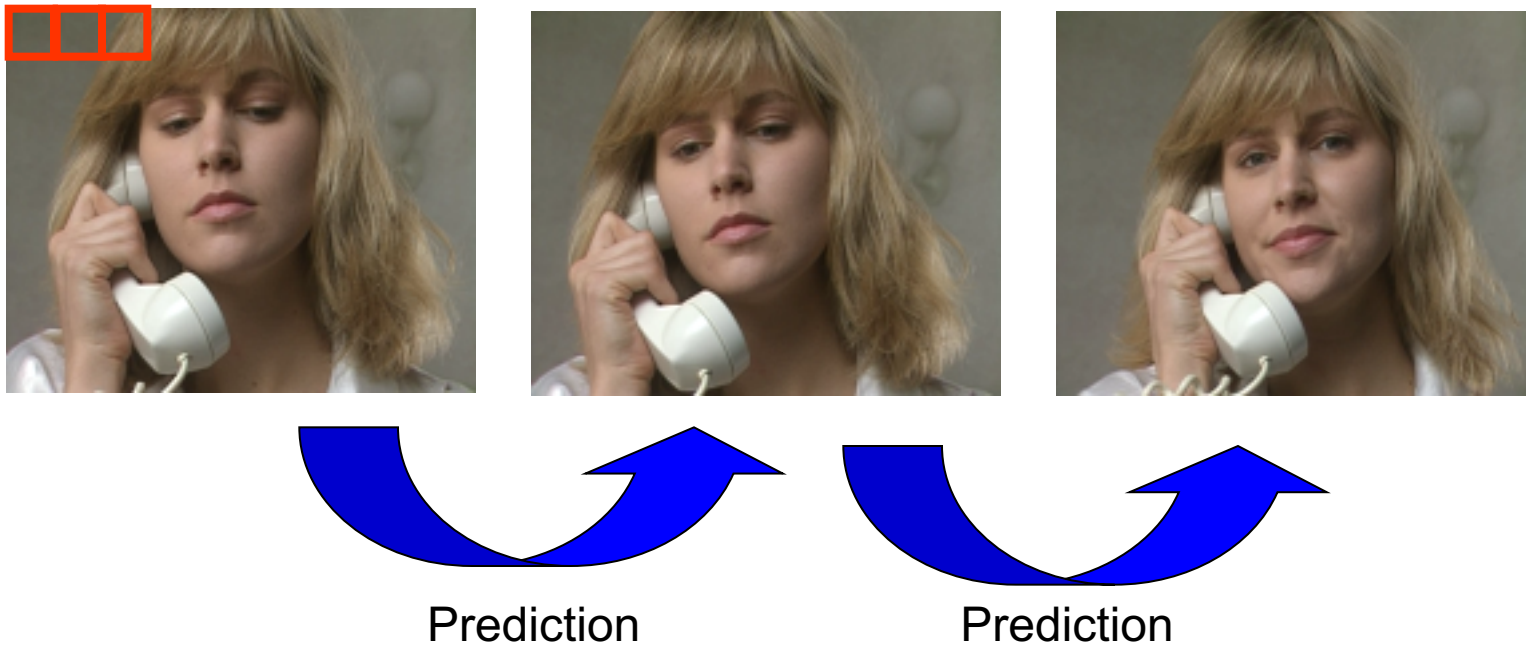
Data Compression



- Lossless Compression: X'=X
  - Example: Computer file compression
  - Low compression ratio
- Lossy Compression: X' ≠ X
  - Many applications do not require lossless compression
  - Our eyes and ears cannot identify some details
  - High compression ratio

Essential of Compression

- Remove redundant information:
  - Spatial redundancy:
    - Neighboring samples have similar values
  - Temporal redundancy:
    - Neighboring frames in a video sequence are similar



A Typical Image Compression System

Transform

Quantization

Entropy coding

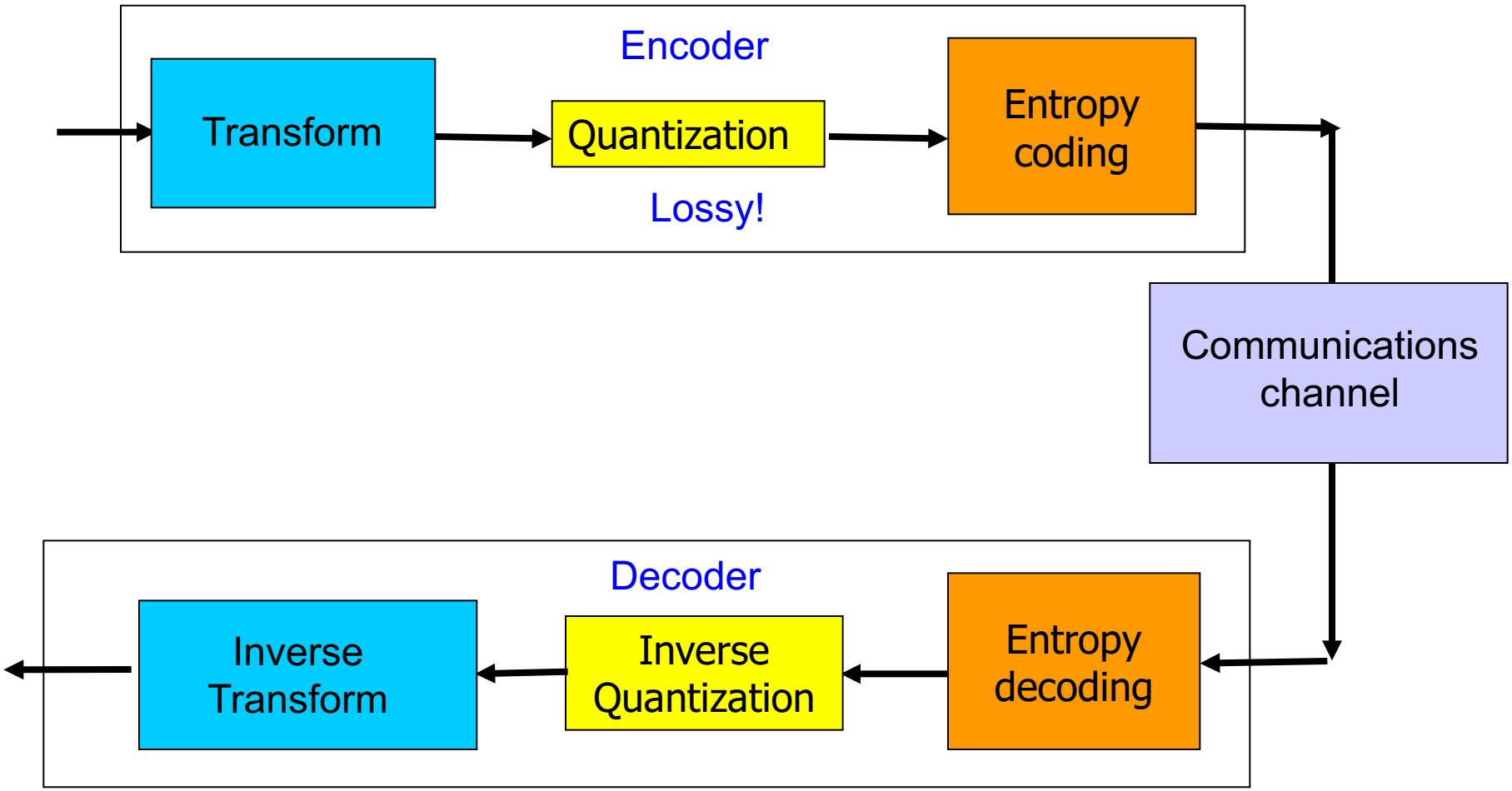
Compressed bitstream

Exploit the spatial redundancy: DCT, Wavelet, lapped transform

Eliminate smaller coefficients that cannot be perceived

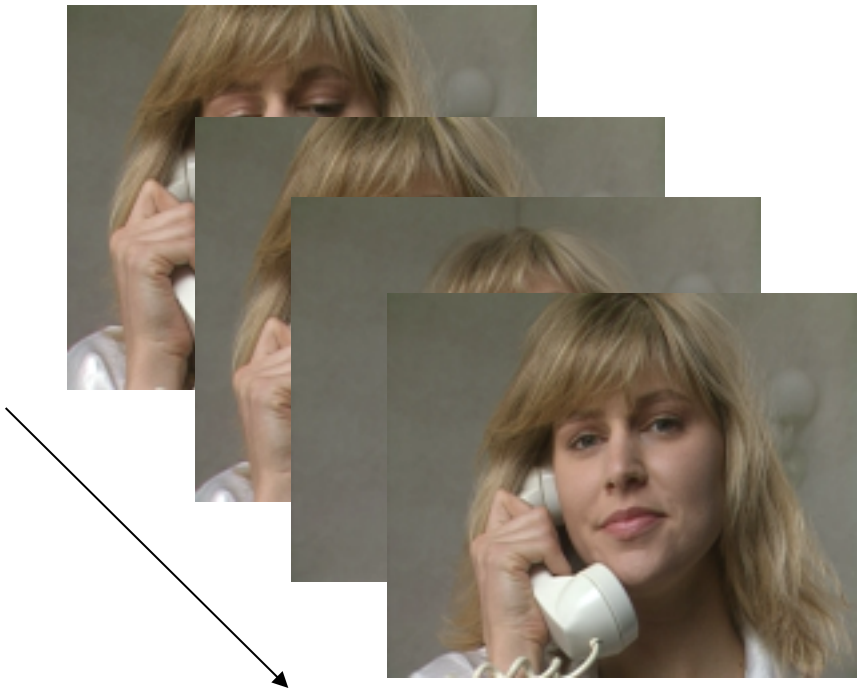
Huffman or arithmetic coding: Assign shorter codes to more probable symbols.

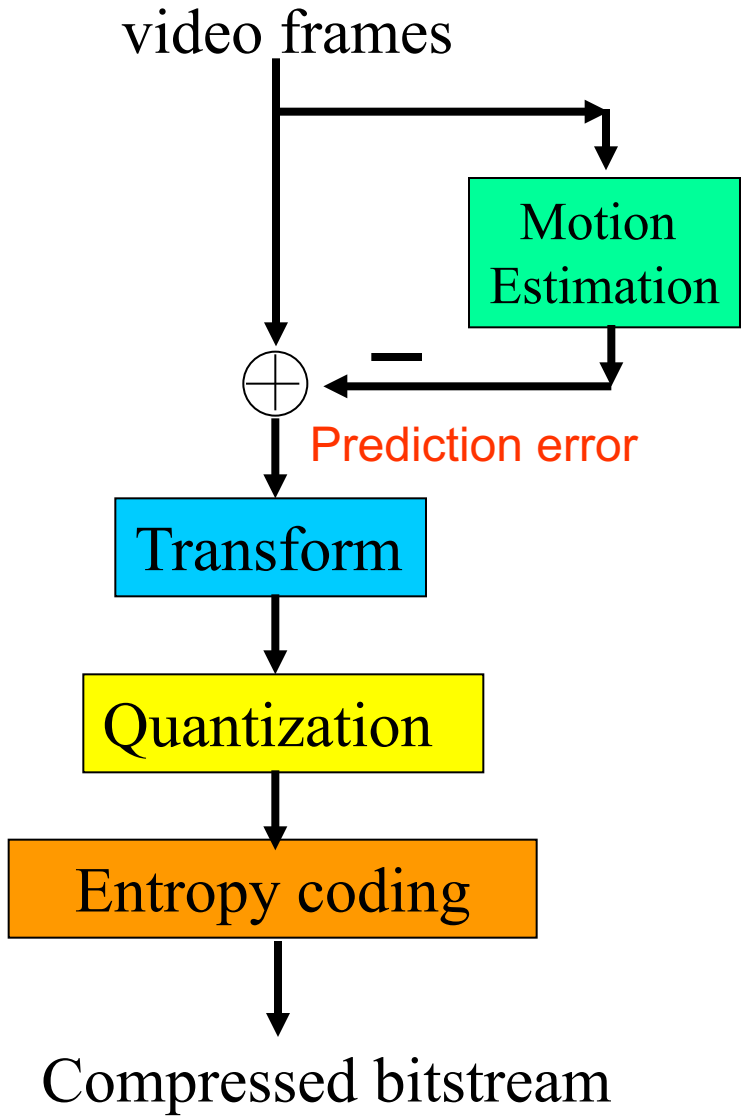
Decoder


Encoder
Transform
Quantization
Lossy!
Entropy coding
Communications channel
Decoder
Inverse Transform
Inverse Quantization
Entropy decoding

A Typical Video Compression System

Remove temporal redundancy





Compression Standards

- Why standards
  - A standard allows products from multiple vendors to communicate
    - Yet, users have flexibility in selecting equipment or software
  - Assures a large market for a particular piece of equipment or software
    - encourages mass production, VLSI technologies etc
    - lower costs.
  - Patent war !

- Standard does not prevent innovation (?)
  - Only decoder is specified by the standard.
  - Encoder can still be improved.
  - MPEG-2:

    Bit rate has been reduced from 8Mbps in 1994 to 2Mbps now, offering the same quality.

Standardization Bodies

- ITU: International Telecommunications Union
  - ITU-T: ITU Telecommunication Standardization Sector (CCITT)
- ISO: International Standards Organization
- IEC: International Electro-technical Commission
- SMPTE: Society of Motion Picture and Television Engineers

- JPEG (ISO/IEC Joint Photographic Experts Group)
- JBIG (ISO Joint Bi-level Image Experts Group)
- MPEG (ISO Motion Picture Experts Group)
- VCEG (ITU-T Video Coding Experts Group)

Image Coding Standards

- JPEG:1993 (JPG file format)
  - DCT-based block transform
- JPEG2000: Dec. 2000
  - Wavelet-based
  - Much more complicated than JPEG
- JBIG: Joint Bi-level Image Experts Group (1993)
  - for lossless bi-level image compression (fax)
  - can also be used for grayscale images
- JBIG2: 1999
  - Supports both lossless and lossy compression

Video Coding Standards

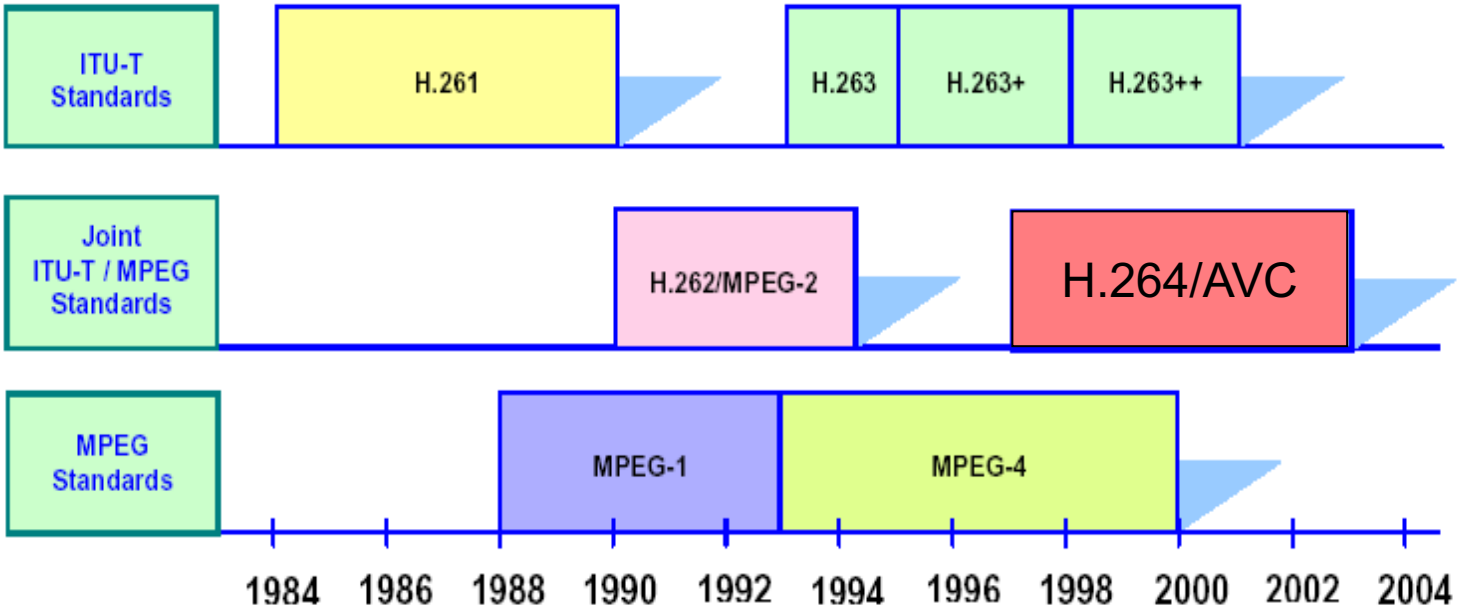
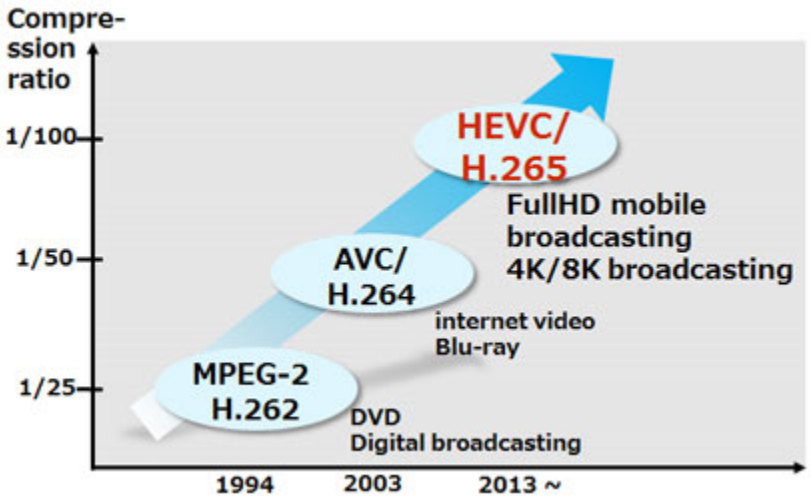


Figure 1. Progression of the ITU-T Recommendations and MPEG standards.

H.264/AVC: ITU-T H.264 / MPEG-4 (Part 10) Advanced Video Coding (AVC)
- Finalized in May 2003 (for general purpose)
- Fidelity Range Extensions (FREXt): 2003-2004 (for professional)

Video Coding Standards

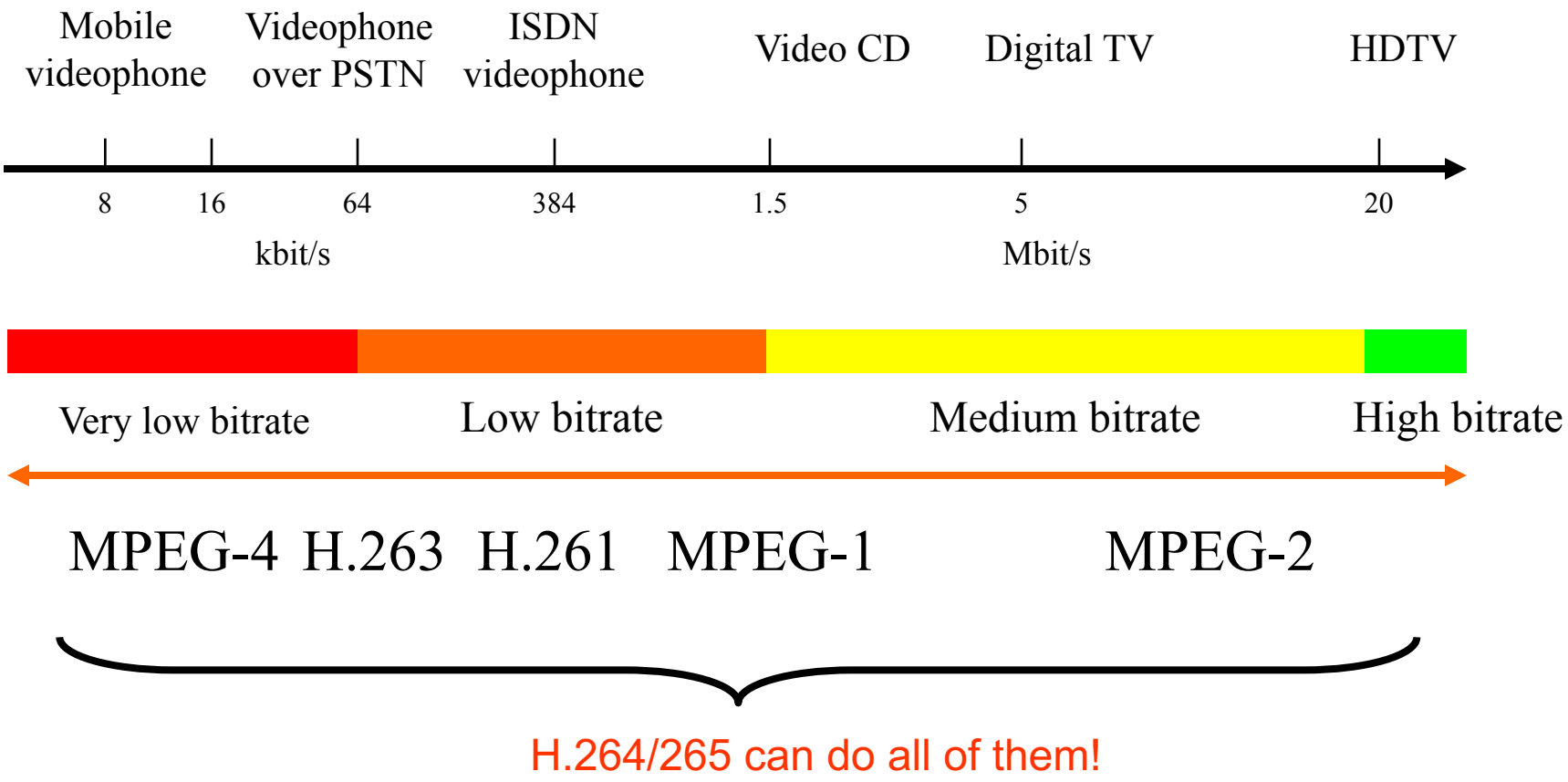


H.265/HEVC (High Efficiency)

- 50% goal (bitrate reduction)
- Start from 2010

February 2012: Committee Draft (complete draft of standard)
July 2012: Draft International Standard
January 2013: Final Draft International Standard (ready to be ratified as a Standard)
April 2013: Standard released

Coding Rate and Standards



Audio coding standards

Range of human' hearing: 20Hz - 20kHz
➔ Minimal sampling rate: 40 kHz (Nyquist frequency)

Format	**Bit Depth**	**Sampling Rate**	**Bit Rate** (2 channels)
CD Audio	16 bits	44.1 kHz	1,411,200 bps
DVD Audio	24 bits	96 kHz	4,608,000 bps

- MPEG-1 audio layer 3 (MP3)
  - CD quality at 10 : 1 compression ratio.
- MPEG-2 AAC (advanced audio coding):
  - used by XM Radio (satellite radio in US)
- MPEG-4 AAC :
  - Up to 48 channels, 96KHz
- ATSC AC-3: 1994
  - Dolby Digital (5.1 channel)
  - ATSC: Advanced Television Systems Committee
  - For DTV, DVD
- iTunes
  - AAC
  - AIFF (Audio Interchange File Format

Outline

- Course information
- What is multimedia? A brief introduction
  - Concepts
  - Representation
  - Compression
  - Communication
- Popular multimedia tools
- Summary

Multimedia communications

- Examples of Multimedia Communication Systems:
  - World Wide Web
  - Video conferencing
  - Video-on-demand
  - Interactive TV
  - Online games

Fundamental Characteristics

- Typically **delay sensitive**
- But can **tolerate occasional loss:**
  - infrequent losses cause minor glitches
- Cf. data transmission: (e.g. FTP)
  - loss intolerant but delay tolerant

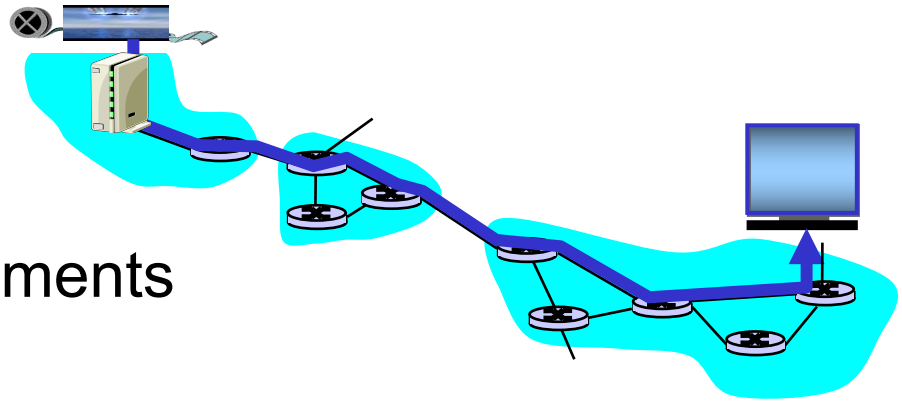
Challenges in Multimedia Communications

- Transmission of Compressed Multimedia:
  - Real-time communications
    - Delay < 0.4 sec in video conference
  - Sequencing within the media
  - Synchronization (e.g., between video & audio)
  - Robustness to transmission error

- We will learn how to
  - Transmit multimedia over Internet and wireless network

Recall: Challenges in Multimedia Communications

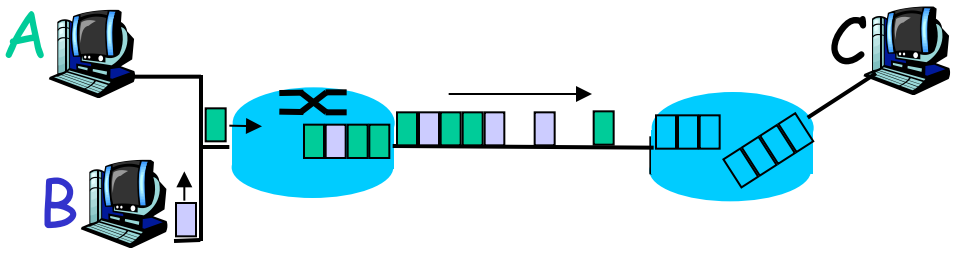
- Real-time communications
  - Delay < 0.4 sec in video conference
- Sequencing within the media
- Synchronization (e.g., between video & audio)
- Robustness to transmission error



Can we achieve these requirements through the Internet?

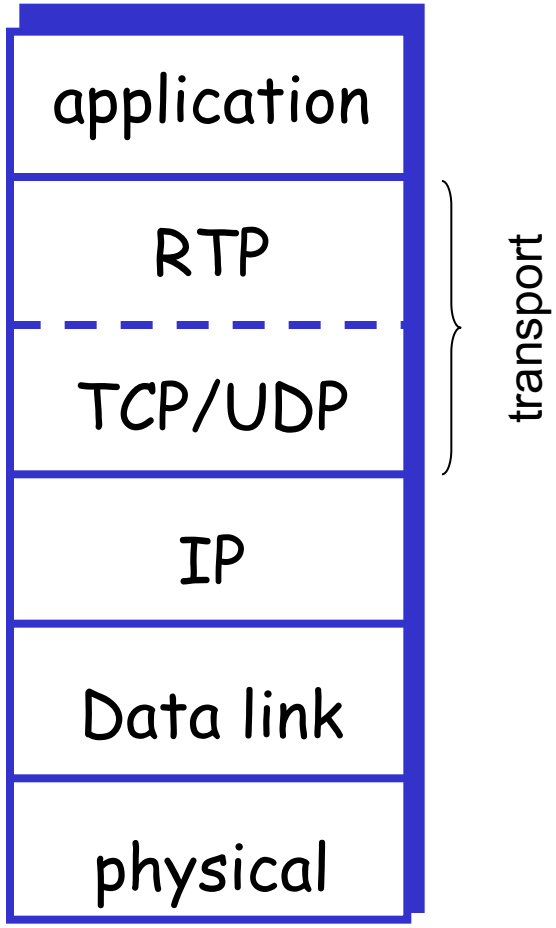
Internet

- Packet-switched network
- Network resources are shared
- Each packet is handled by a series of routers before being received
- Packets can be discarded if the buffer of a router is full
- All packets are treated the same way in congestion



Internet Protocol Stack

- IP: Internet Protocol
  - Best effort (unreliable)!
- TCP: Transmission Control Protocol
  - Provides reliable (but slow) service
- UDP: User Datagram Protocol
  - Provides unreliable (but fast) service
  - Suitable for real-time application
- RTP: Real-time Transport Protocol
  - packet format for multimedia streams
- RTCP: RTP control protocol
  - Monitor/report service quality
- RTSP: Real-time streaming protocol
  - "Internet VCR remote control"



Quality of Service (QoS) Parameters

- End-to-end Delay
  - time required for the end-to-end transmission of a single data element
- Jitter
  - variation in delay
- Packet loss rate
  - the proportion of data elements that are dropped
- Bandwidth: bits / second (bps)
  - rate of flow of multimedia data

QoS Control

- Algorithms to improve the QoS of Multimedia applications
- Policing
  - Control the input rate to network (leak bucket model)
- Scheduling
  - Divide buffers into logic queue
  - Decide which queue to service next

Error Resilience

Improve the decoded quality in the presence of lost data
- often occurs in wireless networks (and also Internet)

- Add redundancy at encoder:
  - Error correction code
  - Layered coding
  - Multiple description coding
- Post-processing at decoder to hide the error
  - Error concealment

Outline

- Course information
- What is multimedia? A brief introduction
  - Concepts
  - Representation
  - Compression
  - Communication
- Popular multimedia tools
- Summary

Popular Multimedia Software Tools

- The categories of software tools briefly examined here are:
    1. **Music Sequencing and Notation**
    2. **Digital Audio**
    3. **Graphics and Image Editing**
    4. **Video Editing**
    5. **Animation**
    6. **Multimedia Authoring**

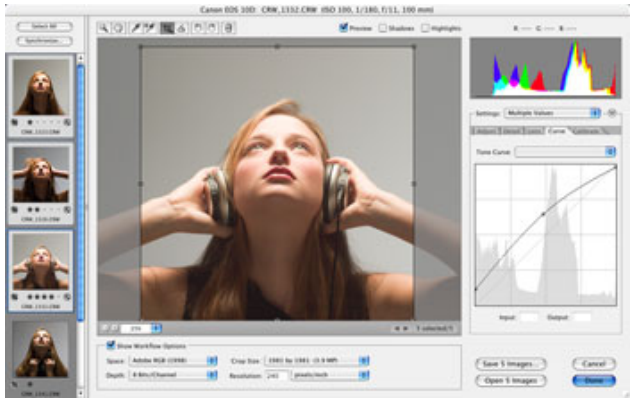
Digital Audio

- **Digital Audio** tools deal with accessing and editing the actual sampled sounds that make up audio:
- **Cakewalk Pro Audio/Adobe Audition (formerly Cool Edit Pro)**
  - Powerful and popular digital audio toolkits; emulates a professional audio studio --- multitrack productions and sound editing including digital signal processing effects.

- **Pro Tools**
  - A high-end integrated audio production and editing environment | MIDI creation and manipulation powerful audio mixing, recording, and editing software.
- **Anvil Studio**: free, for MIDI

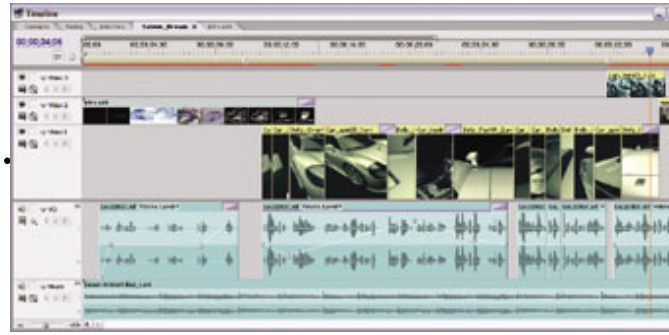
Graphics and Image/Photo Editing

- **Adobe Illustrator**
  - A powerful publishing tool from Adobe.
  - Uses vector graphics; graphics can be exported to Web.
- **Adobe Photoshop**
  - "Standard" image processing and manipulation tool.
  - Allows layers of images, graphics, and text that can be separately manipulated for maximum flexibility.
- **GIMP**: GNU Image Manipulation Program (free)



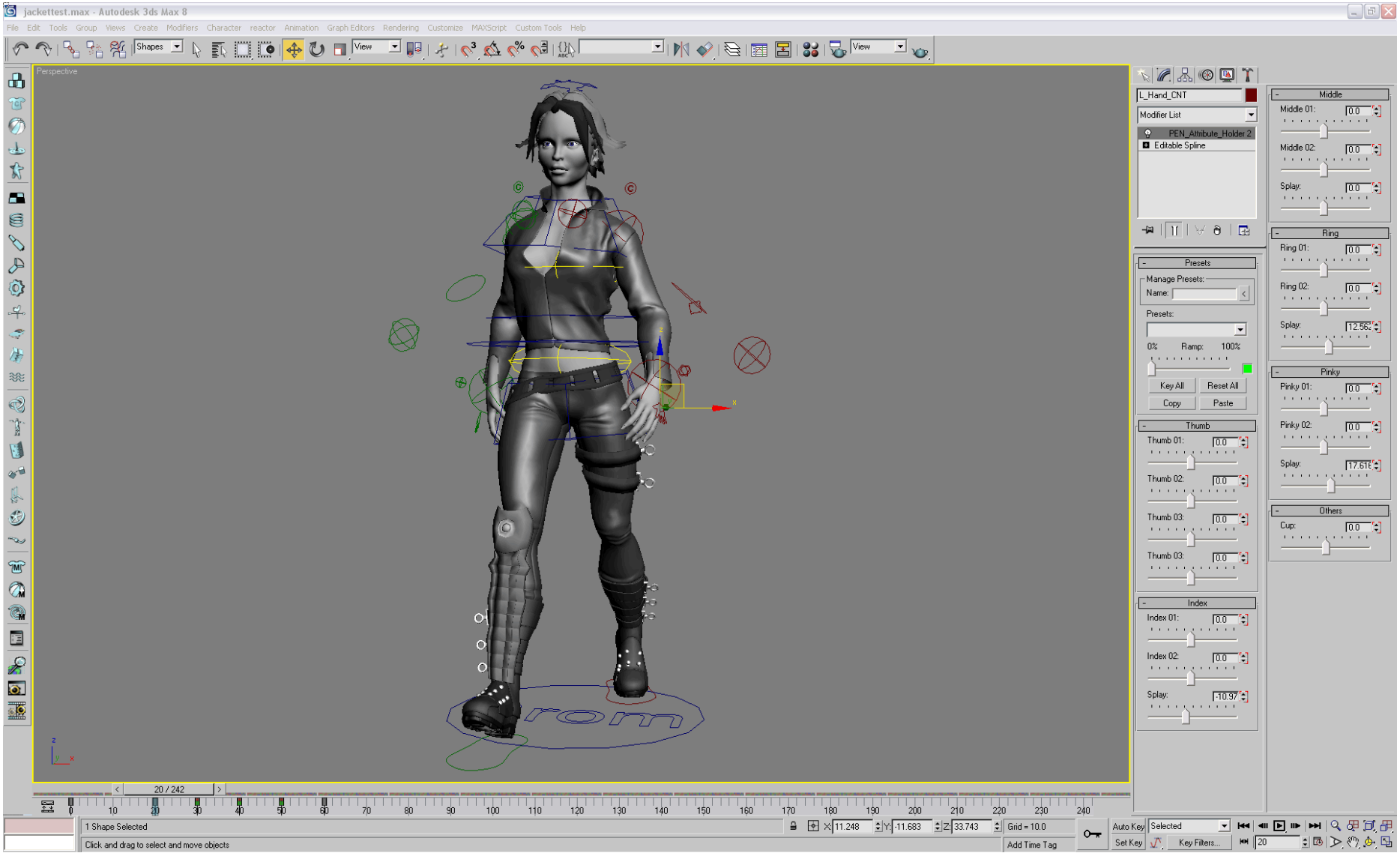
Non Linear Video Editing

- **Adobe Premiere**
  - An intuitive, simple video editing tool for **nonlinear** editing, i.e., putting video clips into any order: Video and audio are arranged in \tracks".
  - Provides a large number of video and audio tracks, superimpositions and virtual clips.
  - A large library of built-in transitions, filters and motions for clips) effective multimedia productions with little effort.
- **Adobe After Effects**
- **Final Cut Pro**
  - A video editing tool by Apple; Mac only.
- **Power Director**
  - popular and cheaper



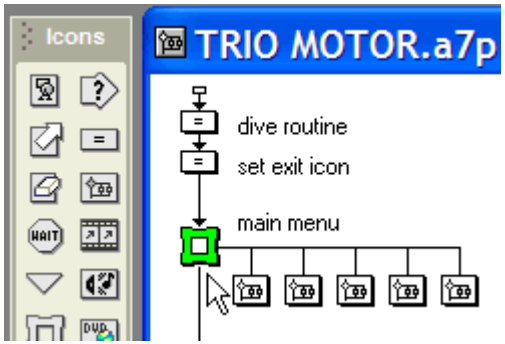
Rendering and Animation

- **Autodesk 3ds Max**
  - Rendering tool that includes a number of very high-end professional tools for character animation, game development, and visual effects production.
- **Autodesk Maya**
  - End-to-end creative workflow with comprehensive tools for animation, modeling, simulation, visual effects, rendering, match moving, and compositing on a highly extensible production platform.


jackettest.max - Autodesk 3ds Max 8
File Edit Tools Group Views Create Modifiers Character reactor Animation Graph Editors Rendering Customize MAXScript Custom Tools Help
Perspective
L_Hand_CNT
Modifier List
PEN_Attribute_Holder 2
Editable Spline
Middle
Middle 01:
Middle 02:
Splay:
Ring
Ring 01:
Ring 02:
Splay:
Pinky
Pinky 01:
Pinky 02:
Splay:
Others
Cup:
Presets
Manage Presets:
Name:
Presets:
0% Ramp: 100%
Key All
Reset All
Copy
Paste
Thumb
Thumb 01:
Thumb 02:
Thumb 03:
Thumb 03:
Index
Index 01:
Index 02:
Splay:
20 / 242
1 Shape Selected
Click and drag to select and move objects
X: 11.248
Y: -11.683
Z: 33.743
Grid = 10.0
Add Time Tag
Auto Key
Selected
Set Key
Key Filters...

Multimedia Authoring

- **Adobe Flash**
  - Allows users to create interactive movies by using the score metaphor, i.e., a timeline arranged in parallel event sequences.
- **Adobe Director**
  - Uses a movie metaphor to create interactive presentations
  - Very powerful and includes a built-in scripting language, **Lingo**, that allows creation of complex interactive movies
- **Authorware (used to be popular; but discontinued from 2003)**
  - A mature, well-supported authoring product based on the **Iconic/Flow-control** metaphor



Multimedia API

- **DirectX**
  - Windows API that supports video, images, audio and 3-D animation
- **OpenGL**
  - A highly portable, most popular 3-D API in use today.
- **Java3D**
  - API used by Java to construct and render 3D graphics, similar to the way in which the Java Media Framework is used for handling media files.
  - An abstraction layer built on top of OpenGL or DirectX (the user can select which).
  - Provides a basic set of object primitives (cube, splines, etc.) for building scenes.
- Android multimedia API/iOS multimedia API

Behind the Tools …

- Is this course about the use of these tools ?
  - No !
- What will we learn ?
  - We will learn what's behind the tools
  - That is, how to design these tools
    - (using them is then trivial)
- Computer Science vs. Computer Applications vs. Art

Grand Challenge Problems

- Social Event Detection for Social Multimedia: discovering social events planned and attended by people.
- Search and Hyperlinking of Television Content: finding relevant video segments for a particular subject and generating useful hyperlinks for each of these segments.
- Geo-coordinate Prediction for Social Multimedia: estimating the GPS coordinates of images and videos.
- Violent Scenes Detection in Film: automatic detecting.
- Preserving Privacy in Surveillance Videos: methods obscuring private information (such as faces on Google Earth).
- Spoken Term Web Search: searching for audio content within audio content by using an audio query.
- Question Answering for the Spoken Web: a variant on the above, specifically for matching spoken questions with a collection of spoken answers.
- Soundtrack Selection for Commercials: choosing the most suitable music soundtrack from a list of candidates.

Summary

- Topics to be covered:
  - Media representation
    - Audio/Image/Video
  - Media Compression:
    - Digital media signals
    - Entropy coding: Huffman, arithmetic, etc.
    - Quantization
    - Transform: KLT, DCT, Wavelet*
    - Coding standards: JPEG, MPEG, MP3, H.264
  - Multimedia Transmission*

End of Introduction