INFO 430 MULTIMEDIA

INFO430

Dr Siba Haidar

Course Learning Objectives

- present concepts in multimedia

and

- acquire the management techniques of the media in diverse applications

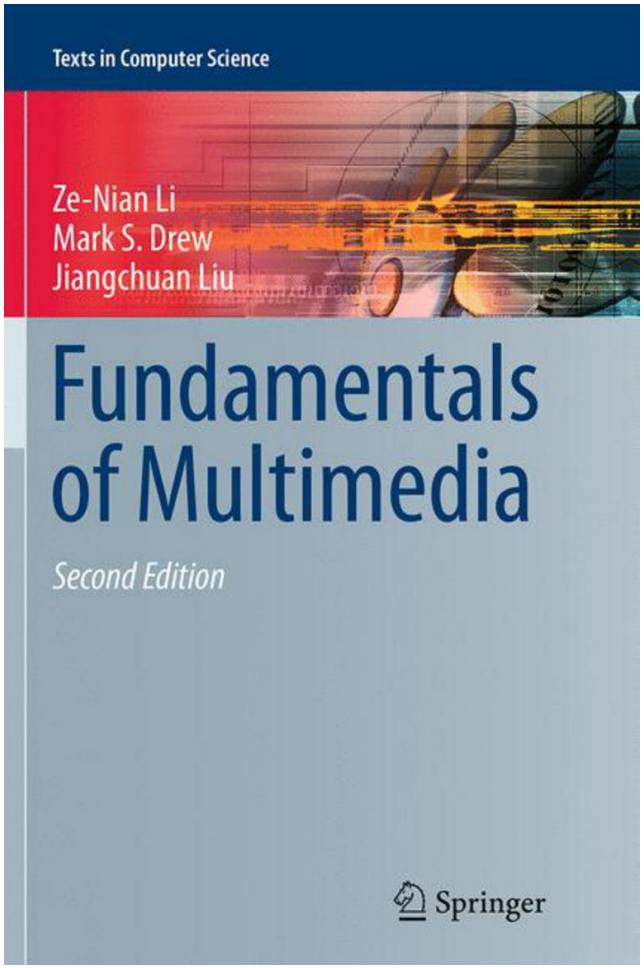
- prerequisites
  - good programming skills
  - clear math mind

Course Outline

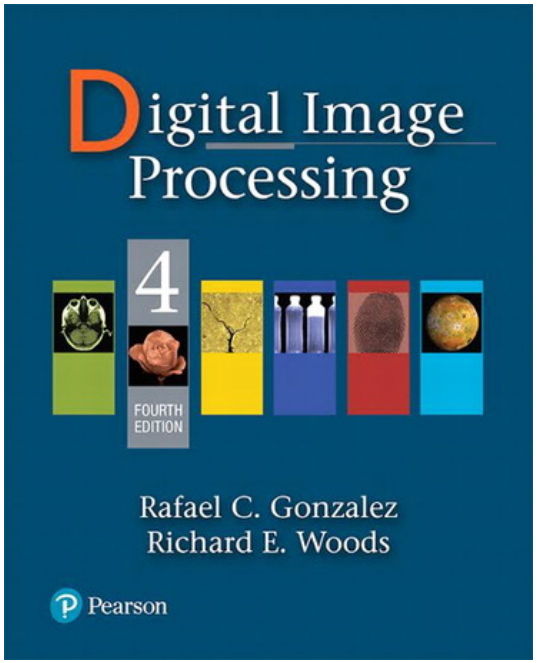
- Chapter 1 - Introduction to multimedia
- Chapter 2 - Digital representation of graphics and images
- Chapter 3 - Colors in images and video
- Chapter 4 - Fundamentals video
- ~~Chapter 5 - Digital Audio~~
- Chapter 6 - Lossless compression algorithms
- Chapter 7 - Lossy compression algorithms (including JPEG standard)
- Chapter 8 - Video Coding (MPEG)
- Chapter 9 - Image Processing

- Tutorials: Mastery of techniques for handling and processing of media streams with MATLAB.

Textbooks



- then for the last lectures :



About this course

- 2 lectures weekly
  - Tuesday 10:00 → 11:50 room 40 or lab3
  - Thursday 10:00 → 11:50 room 40 or lab3
- lectures, labs, resources available through
  - Moodle
  - Google Classroom
- labs → using Matlab

INTRODUCTION TO MULTIMEDIA

INFO430

Chapter 1

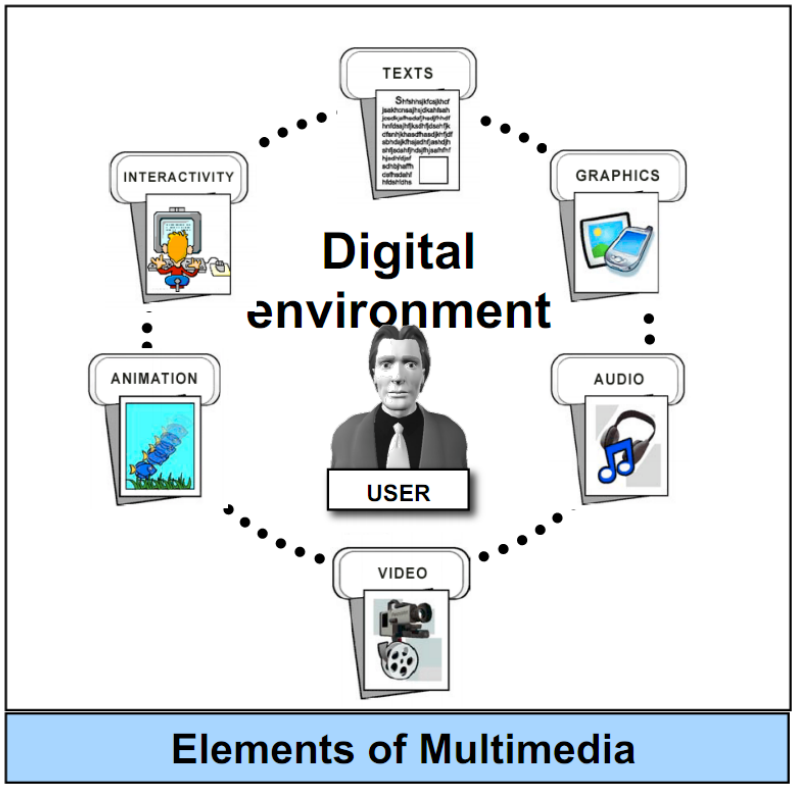
Lecture outline

- What is Multimedia?
- Brief Look at Multimedia Data
- Multimedia: Past, Present and Future

What is Multimedia?

What is Multimedia?

- is the field concerned with computer controlled integration of
  - text
  - graphics
  - drawings
  - still and moving images (video)
  - animation
  - audio …
- where every type of information can be ... digitally
  - represented
  - stored
  - transmitted
  - processed



Elements of Multimedia

Definitions

- multimedia application (MA)
  - uses a collection of multiple media sources: text, graphics, images, sound/audio, animation and/or video
- hypertext
  - text containing links to other texts (term by Ted Nelson 1965)
- hypermedia
  - not only text → other media: graphics, images, sound & video
  - examples: WWW, PowerPoint, Adobe Acrobat
- examples of MAs
  - multimedia authoring (Adobe/Macromedia Director), hypermedia courseware, video-on-demand, interactive TV, computer games, virtual reality, digital video editing, multimedia database systems

Multimedia System (MS)

- system capable of processing multimedia data and applications
- characterized by
  - processing
  - storage
  - generation
  - manipulation
  - rendition … of Multimedia information
- 4 basic characteristics
  - computer controlled
  - integrated
  - information represented digitally
  - interactive interface

Components of a MS

- capture devices
  - video camera, video recorder, audio microphone, keyboards, mice, graphics tablets, 3D input devices, tactile sensors, digitizing hardware
- storage devices
  - hard disks, CD-ROMs, DVD-ROM …
- communication networks
  - local networks, intranets, internet …
- computer systems
  - multimedia desktop machines, workstations, MPEG/VIDEO/DSP hardware
- display devices
  - CD-quality speakers, HDTV, SVGA, Hi-Res monitors, color printers …

Challenges for a MS

- distributed networks
- temporal relationship between data
  - render different data at same time — continuously
  - sequencing **within** the media
    - playing frames in correct order/time frame in video
  - synchronization — **inter-media** scheduling
    - video & audio — Lip synchronization!

Desirable Features for a MS

- very high processing power → to deal with
  - large data processing
  - real time delivery of media
  - special hardware commonplace
- multimedia capable file system → to deliver
  - real-time media
    - video/audio streaming
- efficient and high I/O
- special OS
  - access to file system & process data efficiently & quickly
- storage and memory
  - + large caches & high speed buses
- network support
  - client-server systems
  - distributed systems
- data representations → file formats
  - easy to handle
  - allow for compression/decompression in real-time
- software tools
  - user friendly to design & develop

Reasons of Expension

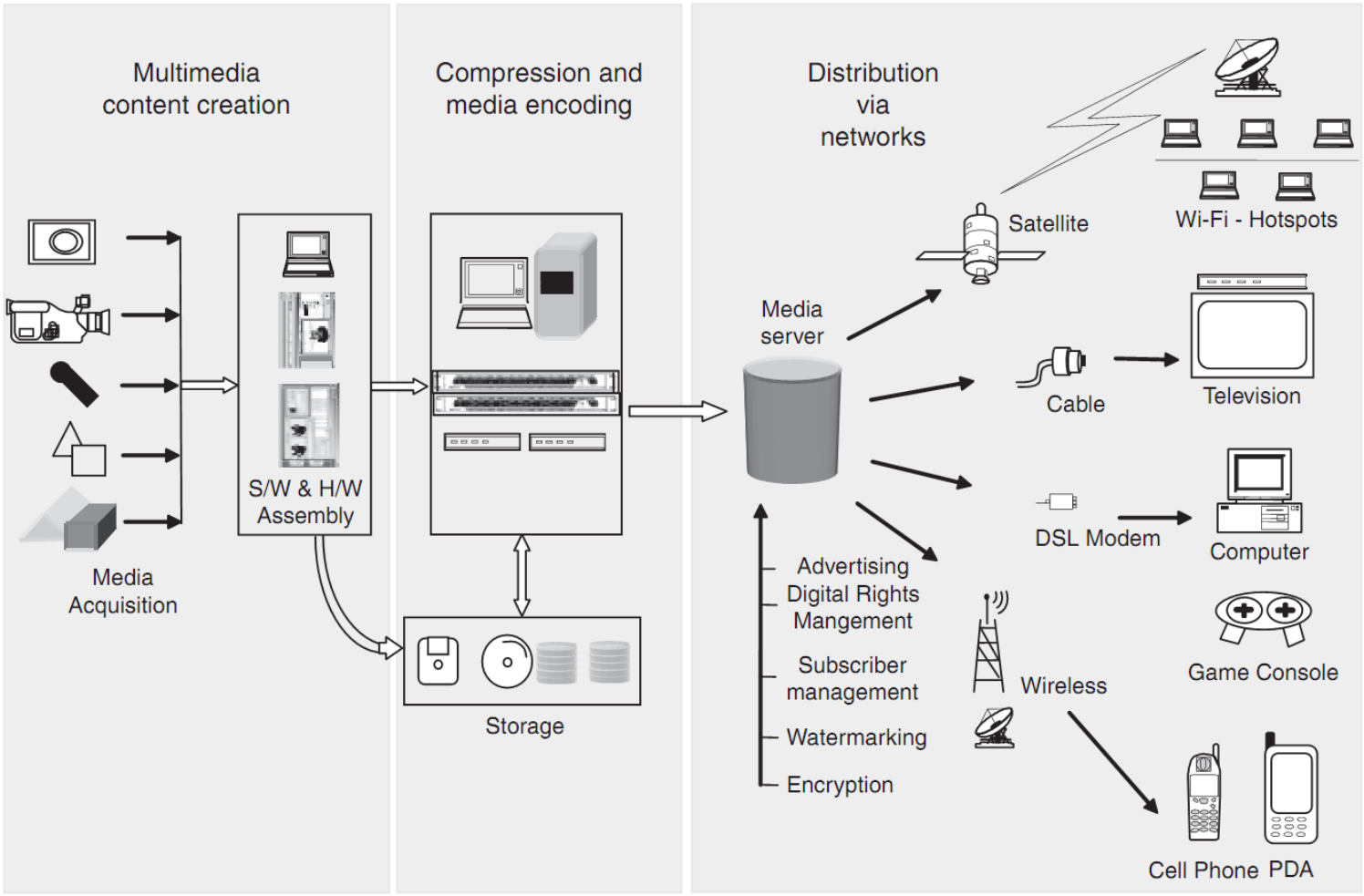
- The multimedia started to change our ways of communication with:
  - availability of low-cost capture devices, rendering devices, and smarter software to create content;
  - larger, less expensive storage devices along with research in better compression of media content;
  - technological advances in digital networks and standardization of distribution protocols

Processes in MS

- Starting from these three points → three processes in multimedia systems:

- Multimedia content creation or multimedia authoring
  - involves digitizing media (audio, images, video) using capture devices and assembling/processing them using smart software and hardware
- Storage and compression
  - significant memory requirements and has to be engineered so as to minimize necessities for storage and distribution
  - state-of-the-art compression algorithms and standards for audio, video, images, and graphics
- Distribution
  - content distributed via various media, such as wired cables, optical networks, satellite, wireless networks, or any combination thereof, to specific platforms ranging from television, computers, personal digital assistants (PDAs), and so on

A Multimedia System today: 3 parts



Brief Look at Multimedia Data

Input and format

Text and Static Data


```
This is an unformatted
text also called plain
text.

All the characters have
the same style and font
and their pitch is the
same.

An advantage is that
  v e r t i c a l
  a l i g n m e n t
is easier.
```

This is an example of
rich text.

Characters can have different *styles*
and their `pitch` is variable.

Βυτ αγαιν τηε μοστ ιμπορταντ

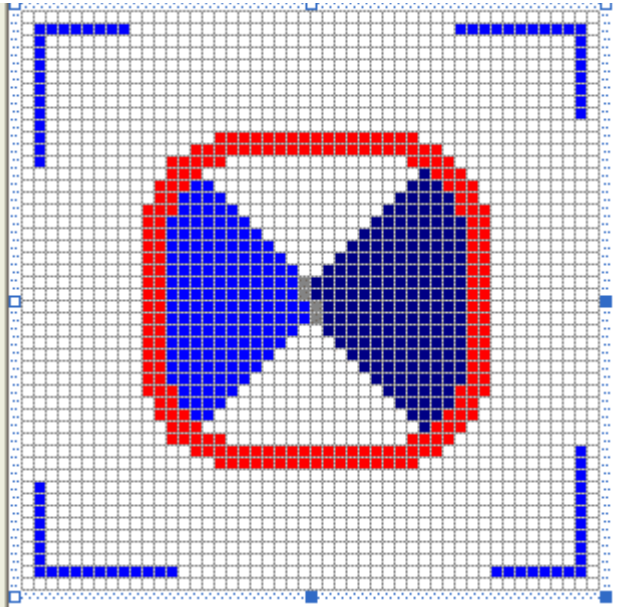
They may also
respect certain
formatting rules.

Graphics

- format
  - composition of primitive objects : lines, polygons, circles, curves, arcs
- input
  - generated by a program (Illustrator)
  - automatically (Postscript)
- editable or revisable (unlike images)
- input devices
  - keyboard (text & cursor control)
  - mouse, trackball, graphics tablet
- standards: OpenGL, PHIGS, GKS
  - OpenGL (Open Graphics Library)
    - cross-language, cross-platform API for writing applications that produce 2D and3D computer graphics
  - PHIGS (Programmer's Hierarchical Interactive Graphics System)
    - for rendering 3D computer graphics, in 1990.
  - GKS (Graphical Kernel System)
    - first ISO standard for low-level computer graphics 1977
- files
  - store primitive assembly
- storage size
  - do not take up a very high storage overhead

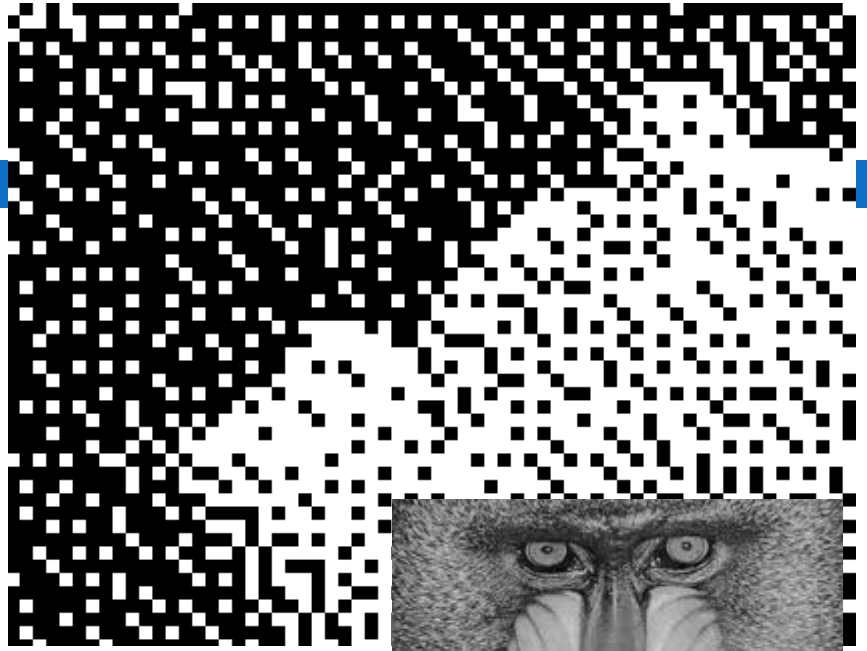
Images

- still pictures 
- uncompressed →
  - bitmap = a grid of pixels
  - pixel = picture element
- input
  - digitally scanned
  - from digital camera
  - by programs similar to graphics or animation programs



Images

- storage
  - black & white
    - 1 bit/pixel
  - grayscale/color map
    - 8 bpp
  - true color
    - 24 bpp



Bits and Bytes

- Which units are there even, and how big are they?
- Refer to this short list:
  - 1 B = 1 byte = 8 bits
  - 1 kB = 1.000 bytes
  - 1 MB = 1.000 kB or 1.000.000 bytes
  - 1 GB = 1.000 MB, 1.000.000 kB, or 1.000.000.000 bytes

Images

- size
  - 512x512 grayscale
    - 8x512x512=262144 bits/8=32768 bytes → 1/4Mb
  - 512x512 (24 bit) → 3/4 Mb with no compression
  - modern high digital camera 10+ Megapixels → 29Mb uncompressed!
  - → compression is commonly applied

Audio

- audio signals → continuous analog signals
- input
  - microphones then
  - digitized & stored
- CD quality audio → 16-bit sampling at 44.1 KHz (even higher audiophile rates (24-bit, 96 KHz))
- 1 min mono CD quality (uncompressed) → 5 Mb
- 1 min stereo CD quality (uncompressed) → 10 Mb
- usually compressed: pcm, mp3, aac, …

Video

- input
  - analog video captured by video camera then
  - digitized
- formats → variety (analog and digital)
- raw video →
  - series of single images
  - 25, 30 or 50 frames per second (fps)
- storage example
  - 512×512 size monochrome video images → 25 × 0,25 = 6.25Mb for 1 sec uncompressed
  - typical PAL digital video (720 × 576 pixels per color frame → 1.2 × 25 = 30Mb for 1 sec uncompressed
  - high definition DVD (1440×1080 = 1.5 Megapixels per frame) → 4.5 × 25 = 112.5Mb for 1 sec uncompressed
- digital video clearly needs to be compressed

Common Properties

- Digital
  - always
  - underlying representation of the information → bits and bytes
- Voluminous
  - size of data → understandably large
  - storage and transmission bandwidth limitations
  - → data must be compressed
- Interactive
  - content can be interacted with
  - ex: choose a video to watch or a set of images to browse
  - ex: click on areas of an image causing an action
- Real-time and synchronization
  - real-time requirements and resulting synchronization
  - → crucial role in the system's architecture and design

- → core to design good working multimedia applications

Data Compression

- evident necessity
- problems for storage and networks
- → we need excellent and fast data compression

- Table 2.1 : some values
  - for standard-definition, and
  - for HD broadcast video

Standard definition video		
640×480 full color	=	922 kB/frame
@ 30 frames/s	=	28 MB/s
	=	221 Mb/s
× 3,600 s/h	=	100 GB/h
High definition video		
1,920×1,080 full color	=	6.2 MB/frame
@ 30 frames/s	=	187 MB/s
	=	1.5 Gb/s
× 3,600 s/h	=	672 GB/h

How much compression is required?

- depends on
  - application
  - capability of viewing and display
  - bandwidth (in bits per second) to stream and view decompression
- in JPEG amount of compression controlled by a value Q in range 0–100
  - image quality best for Q = 100 and worst for Q = 0

Table 2.2 JPEG file sizes (bytes) and percentage size of data for JPEG compression with Quality Factor Q = 75, 25, and 5

Quality factor	Compressed file size	Percentage of original (%)
–	529,620	100
75	37,667	7.11
25	16,560	3.13
5	5,960	1.13

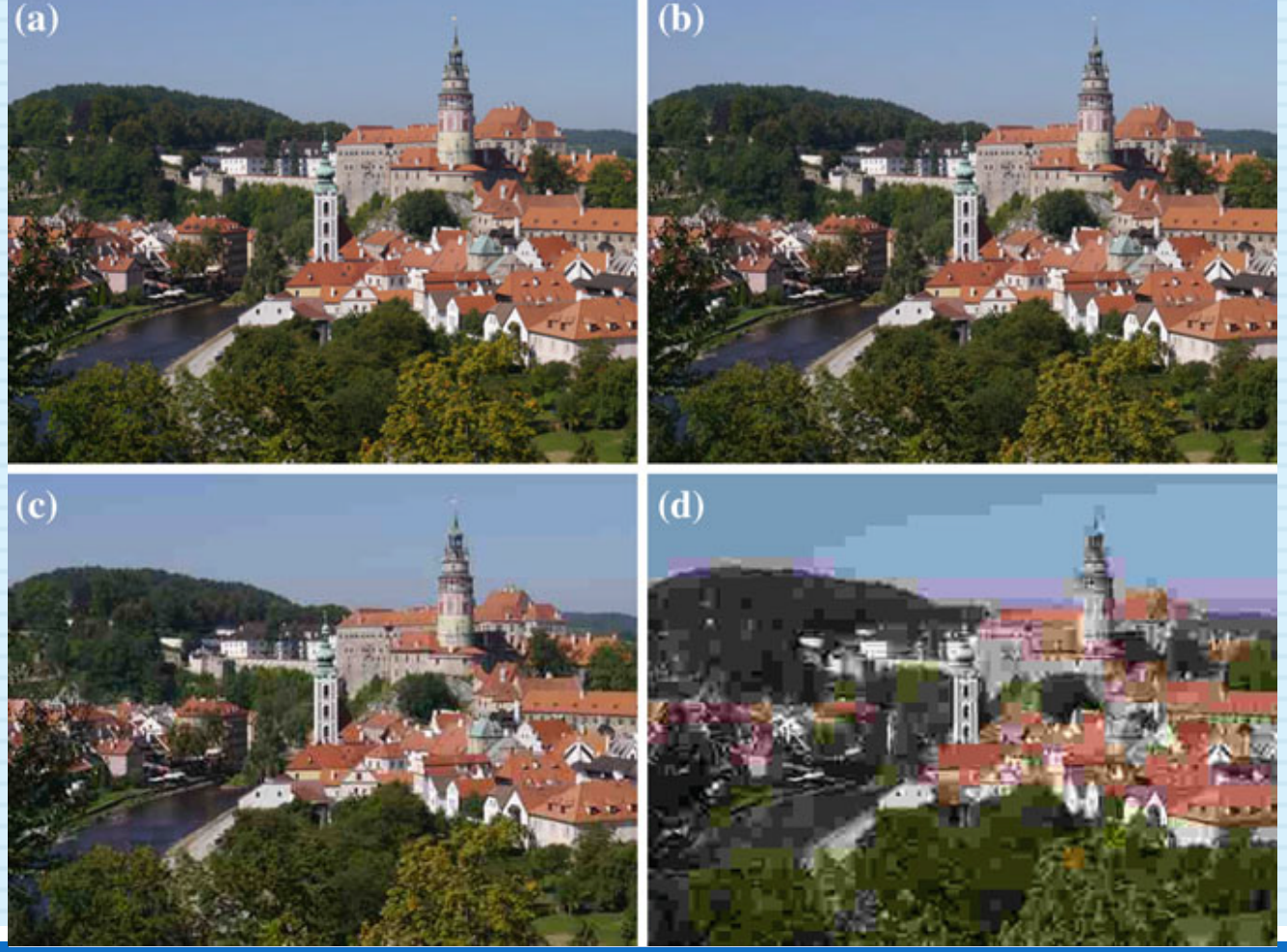


Fig. 2.9 JPEG compression

a original uncompressed image; b JPEG compression with Quality Factor Q = 75 (the typical default); c, d Factors Q = 25 and Q = 5

Reducing File Sizes

- image files →cropped
  - photographs for web or Facebook do not have to be big
  - buttons, lines, clipart → formats smaller size → PNG or GIF
  - RAW camera images (CR2 | NEF) 2-6 times bigger than JPG
  - TIFF and BMP converted to JPG
- document files → online slimmed down → PDF
- video files → converted to compressed formats MPEG
- audio files → converted to MP3, loss of quality to some degree but marginal to listening

Compression

- Still Picture Coding: JPEG 2000
- Video Coding: MPEG-4
- Multimedia Content Description Interface: MPEG-7
- Speech Coding Standards: ITU G.729
- Audio Coding Standards: MPEG-4
- Text compression standards: most popular Lempel-Ziv (LZ) family

Average File Sizes

- Images
  - PNG ~ 2 – 4 kB
  - GIF ~ 6 – 8 kB
  - JPG ~ 9 – 12 kB
  - TIFF ~ 900 – 1.000 kB
  - BMP ~ 900 – 1.000 kB
- Documents
  - DOCX ~4 – 8 kB
  - PDF ~ 18 – 20 kB
  - ODT ~ 80 – 90 kB
- Media Files
  - eBook ~ 1 – 5 MB
  - MP3 song ~ 3 – 4 MB
  - DVD Movie ~ 4 GB
  - HD Movie ~ 5 – 8 GB
  - Blu-Ray Movie ~ 20 – 25 GB

Multimedia: Past, Present and Future

Early History of Multimedia

- newspapers → 1st mass communication medium
  - hand-drawn graphics and images
- Joseph Nicéphore Niépce captured 1st natural image from his window in 1826
- Alphonse Giroux built 1st commercial camera with a double-box design
- by the end of 19th century → film-based cameras

Fig. 1.1 A vintage dry-plate camera.

E&H T Anthony model Champion, circa 1890

Early History of Multimedia

- 1877 →
  - Thomas Alva Edison's phonograph
    - 1st device able to record and reproduce sound
  - Alexander Graham Bell improved it later
  - then Emile Berliner
- 20th century
  - → quality sound close to origin
- mid-1980s
  - → audio tapes, CD, …
- 1830
  - → motion pictures conceived to observe motion too rapid for perception by the human eye.
- 1887
  - → Edison → motion picture camera
- 1910 → 1927
  - silent feature films
- 1927
  - → release of the Jazz Singer

Fig. 1.2 An Edison phonograph, model GEM.

Fig. 1.3 Evolution of audio storage media.

Left to right an Edison cylinder record, a flat vinyl record, a reel-to-reel magnetic tape, a cassette tape, and a CD

Early History of Multimedia

- 1895
  - Guglielmo Marconi →1st wireless radio transmission at Pontecchio, Italy
- 1901
  - radio waves beamed across the Atlantic
  - invented for telegraph → radio now a major medium for audio broadcasting
- 1909
  - Marconi shared the Nobel Prize for Physics
- 1884
  - Paul Gottlieb Nipkow, a 23-year-old university student in Germany, patented the first electromechanical television system
- 1907
  - cathode ray tube (CRT) became available
- 1920
  - CRT-based TV established video as a commonly available medium and has since changed the world of mass communication

From Analog to Digital Medias

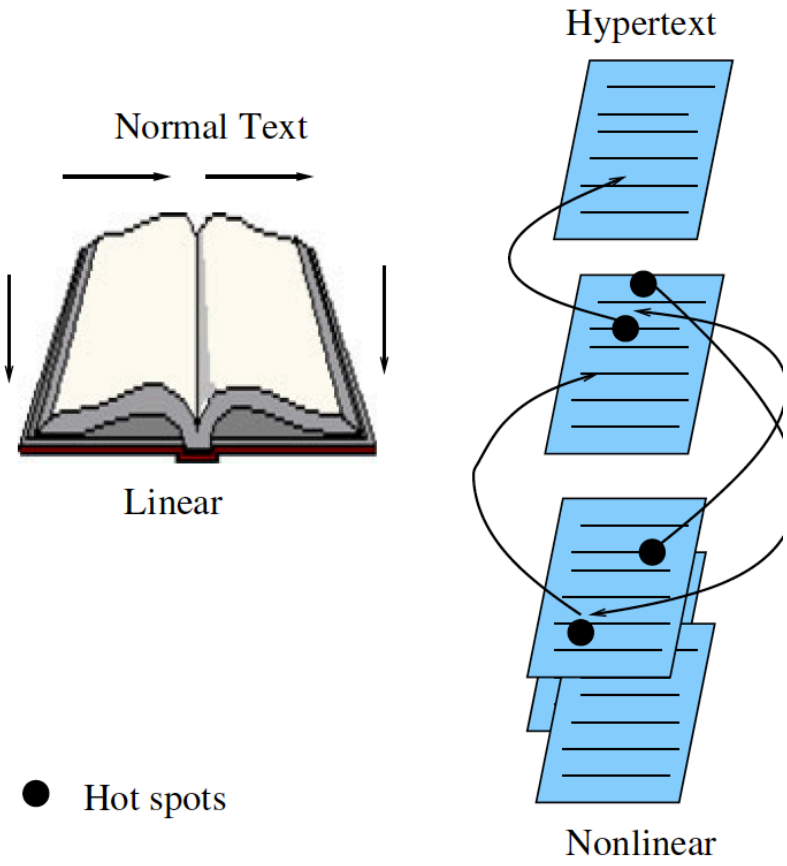
- all of the previous
  - analog format
  - → time (variable) of signal is continuous representation of the input
  - analogous to the input audio, image, or video signal
- connection between computers and digital media
  - media data represented using discrete binary format
  - emerged over a short period

Digital Media

- 1967
  - Nicholas Negroponte → Architecture Machine Group at MIT
- 1969
  - Nelson and van Dam at Brown University → early hypertext editor called FRESS
- 1976
  - MIT Architecture Machine Group proposed "Multiple Media"
- 1978
  - 1st videodisk
- 1982
  - Compact Disc (CD) available by Philips and Sony
  - standard for digital audio data → replace analog magnetic tape
- 1985
  - Negroponte and Wiesner co-founded the MIT Media Lab
- 1990
  - Kristina Hooper Woolsey headed the Apple Multimedia Lab, with a staff of 100
- 1991
  - MPEG-1was approved as an international standard for digital video.
    - then MPEG-2, MPEG-4, and further MPEGs, in the 1990s
  - introduction of PDAs (personal digital assistant)
    - 1996 → PDA with no keyboard
- 1992
  - JPEG accepted as the international standard for digital image compression
  - 1st audio multicast on the multicast backbone (MBone)
- 1995
  - JAVA language created for platform-independent application development
  - widely used for developing multimedia applications
- 1996
  - DVD video; high-quality, full-length movies
- 1998
  - MP3audio players with 32 MB of flash memory

Hypermedia, WWW and Internet

- Multimedia
  - integration to enable rich interaction amongst medias
  - and between media and human beings
- In 1945, as part of MIT's postwar
  - → Vannevar Bush → Memex : forerunner of the WWW
- after World War II → 6,000 scientists →
- 1960
  - Ted Nelson → Xanadu project : term hypertext
  - book: linear medium
  - hypertext system: read nonlinearly, following links



Hypermedia, WWW and Internet

- MAs roots in nuclear physics!
- 1990
  - Tim Berners-Lee proposed the WWW to CERN
  - CERN = European Center for Nuclear Research
  - to organize and share their work and experimental results
- His team invented HTML and HTTP for this purpose too
- then came XML, XSL and family

Multimedia in the New Millennium

- new generation of social, mobile, and cloud computing → MM processing & sharing
- role of Internet evolved
  - from original use as communication tool
  - to provide easier and faster sharing of an infinite supply of information
  - multimedia content enriched
- HD (High definition) videos and 3D multi-view videos
  - captured and browsed by pcs
  - stored and processed with remote cloud resources
- users engaged to be part of a social ecosystem not anymore passive
- revolution
  - driven by penetration of 3G/4G wireless networks
  - and smart mobile devices
  - high intuitive interfaces + rich multimedia functionalities
  - integrated with online social networking
  - → instant media content generation and sharing

Milestones in the New Millennium

- 2000
  - WWW size > 1 billion pages
  - Sony → Blu-ray Disc prototypes
- 2001
  - Napster : 1st p2p (peer-to-peer) file sharing (MP3 music)
  - NTT DoCoMo in Japan launched the first commercial 3G wireless network
- 2003
  - Skype p2p voice over the Internet
- 2004
  - Web 2.0 : "virtual community"
  - Ex: social networking, blogs, wikis, etc.
  - Facebook by Mark Zuckerberg
  - Flickrc by Ludicorp
- 2005
  - YouTube for video sharing purchased by Google in 2006
  - Google launched online map service + satellite imaging, real-time traffic, and Streetview later
- 2006
  - Twitter
    - > 500 million registered users in 2012
    - 340 million tweets per day
  - Amazon's Web Services (AWS)
  - Nintendo Wii home video game console

Milestones in the New Millennium

- 2007
  - Apple iPhone running iOS + touch screen + App Store mobile apps
  - Goolge Android mobile os + Open Handset Alliance
    - consortium of hardware, software, and telecommunication companies devoted to advancing open standards for mobile devices
- 2008
  - 1st Android-powered phone
  - Google Play
  - tablet computers iOS, Android, and Windows with larger touch screens
- 2009
  - 1st LTE (Long Term Evolution) network set up in Oslo, Norway, and Stockholm, Sweden
  - making an important step toward 4G wireless networking
  - James Cameron's film, Avatar, created a surge on the interest in 3D video
- 2010
  - Netflix
    - previously a DVD rental service provider
    - migrates to Amazon AWS cloud computing platform
    - Master copies of digital films from movie studios are stored on Amazon S3
    - each film is encoded into over 50 different versions based on video resolution, audio quality
    - > 1 petabyte of data stored on Amazon's cloud
  - Microsoft Kinect
    - horizontal bar with full-body 3D motion capture, facial recognition, and voice recognition capabilities, for its game console Xbox 360.

Milestones in the New Millennium

- 2012
  - HTML5 is a W3C "Candidate Recommendation
  - provide support for latest multimedia formats
  - ability to run on low-powered devices such as smartphones and tablets
- 2013
  - Sony PlayStation 4
    - video game console 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.

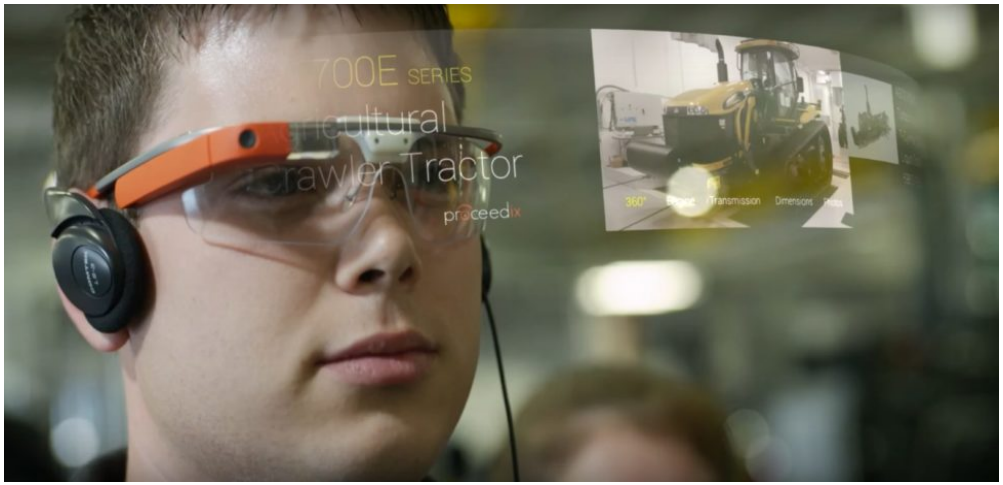
Multimedia in the Future

- multimedia research → young + vigorously growing
- many exciting topics together
- old interests
  - shot detection
  - video classification
- new challenges due to the abundance of online video
  - camera-based object tracking
  - face detection and recognition
  - event detection for security
    - person leaving a bag unattended in an airport

Advanced Modelling

- 2D video → 3D capture technology
  - → acquire dynamic characteristics of human facial expression during speech
  - to synthesize highly realistic facial animation from speech for low-bandwidth applications
- multiple views from several cameras or from a single camera under differing lighting can accurately acquire data
  - give both the shape and surface properties of materials
  - generate synthetic graphics models
  - photo-realistic (video-quality) synthesis of virtual actors
- handicapped persons
  - poor vision
  - elderly

Google Glass



- Google Glass
  - optical head-mounted display
  - enables interactive display for its users
  - communicate using natural language voice commands
- → wearable computing of great potentials

Online Social Media

- YouTube, Facebook, Twitter
  - appeared past decade
  - changed information generation and sharing
- Research on social media is likely one of the most important areas under scrutiny
  - 100,000 academic articles per year
- topics:
- Crowdsourcing for multimedia
  - input of a large number of human contributors → large attention
- example
  - having people provide tags → aid in understanding visual content of images and video
  - Amazon's "Mechanical Turk" to outsource such time-consuming tasks as semantic video annotation to a large number of workers who are willing to work for small reward or just for fun
  - analyze "sentiment" such as the popularity of a particular brand-name as evidenced by reading several thousand tweets on the subject.
  - "Digital fashion" aims to develop smart clothing that can communicate certain thoughts and feelings to be broadcast automatically, for exchange with others equipped with similar technology

Executable Academic Papers

- in science and engineering
- traditional way →
  - publication of papers in academic journals
  - datasets and programming code typically not supplied
- new idea →
  - results discussed in a published paper are often difficult to reproduce
  - allow the “reader” to interact with and interactively manipulate data and code
  - further understand findings presented
  - rerun code, change parameters or upload different data

Other Applications

- Animated Lifelike Virtual Agents
  - virtual educators for special needs children
- model interaction between people
  - enable natural interaction by virtual characters
  - augmented interaction used to develop interfaces between real and virtual humans
    - storytelling

Some Multimedia Interests

- Social Event Detection for Social Multimedia
  - as indicated by collections of multimedia content that was captured by people and uploaded to social-media sites
- Search and Hyperlinking of Television Content:
  - for a particular subject and generating useful hyperlinks
  - automated intelligently
- Geo-coordinate Prediction for Social Multimedia:
  - estimating the GPS coordinates of images and videos using tags, audio, and users
- Violent Scenes Detection in Film
- Preserving Privacy in Surveillance Videos
  - obscuring private information (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
  - variant on the above
  - match spoken questions with collection of spoken answers
- Soundtrack Selection for Commercials

Exercise

- to be done in group of 3
- Identify three novel multimedia applications.
- Discuss why you think these are novel and their potential impact.
- Assignment due next week to upload
  - PowerPoint
  - pdf
  - docx
  - or any other electronic format