# Chapter 1
# Introduction to Multimedia

*Li, Drew, & Liu   © Springer 2021*

---

# 1.1  What is Multimedia?

- When different people mention the term multimedia, they often have quite different, or even opposing, viewpoints.

  - A consumer entertainment vendor: interactive TV with hundreds of digital channels available, or a cable TV-like service delivered over a high-speed Internet connection; a smartphone.
  - A Computer Science (CS) student: applications that use multiple modalities, including text, images, drawings (graphics), animation, video, sound including speech, and interactivity.
  - Graphics, visualization, HCI, artificial intelligence, computer vision, data compression, graph theory, networking, database systems all have contributions to make in multimedia.

*Li, Drew, & Liu   © Springer 2021*

# 1.1.1 Components of Multimedia

- Multimedia involves multiple modalities of text, audio, images, drawings, animation, and video.

Examples of how these modalities are put to use:

1. Video-conferencing
2. Tele-medicine
3. A web-based video editor that lets anyone create a new video by editing, annotating, and remixing editable professional videos on the cloud
4. Geographically-based, realtime augmented-reality, massively multiplayer online video games
5. Shapeshifting TV, where viewers vote on the plot path
6. A camera that suggests what would be the best type of next shot for developing good storyboards

7. Cooperative education environments that allow schoolchildren to share a single educational game using two mice at once

8. Searching (very) large video and image databases for target visual objects, using semantics of objects

9. Compositing of artificial and natural video into hybrid scenes

10. Visual cues of video-conferencing participants, taking into account gaze direction and attention

11. Making multimedia components editable — allowing the user side to decide what components, video, graphics, and so on are actually viewed — making components distributed

12. Building "inverse-Hollywood" applications that can recreate the process by which a video was made.

# 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, multi-modal interaction and integration: "ubiquity" — web-everywhere devices, multimedia education including Computer Supported Collaborative Learning, and design and applications of virtual environments.

*Li, Drew, & Liu  © Springer 2021*

# 1.2  Multimedia: Past and Present

## 1.2.1  Early History of Multimedia

1. **Newspaper:** perhaps the *first* mass communication medium, uses text, graphics, and images.

2. **Motion pictures:** conceived of in 1830's in order to observe motion too rapid for perception by the human eye.

3. **Wireless radio transmission:** Guglielmo 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.

*Li, Drew, & Liu  © Springer 2021*

*Fundamentals of Multimedia 3$^{rd}$ ed., Chapter 1*

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.

➔ Link to full V. Bush 1945 Memex article, "As We May Think"

- 1965 – 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.

7                                                     *Li, Drew, & Liu  © Springer 2021*

---

*Fundamentals of Multimedia 3$^{rd}$ ed., Chapter 1*

- 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.

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

8                                                     *Li, Drew, & Liu  © Springer 2021*

- 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 program.

- 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.

# 1.2 Hypermedia, WWW, and Internet

- A hypertext system: meant to be read nonlinearly, by following links that point to other parts of the document, or to other documents (Fig. 1.1)

- HyperMedia: not constrained to be text-based, can include other media, e.g., graphics, images, and especially the continuous media, sound and video.

    - The World Wide Web (WWW) — the best example of a hypermedia application.

- Multimedia means that computer information can be represented through audio, graphics, images, video, and animation in addition to traditional media.
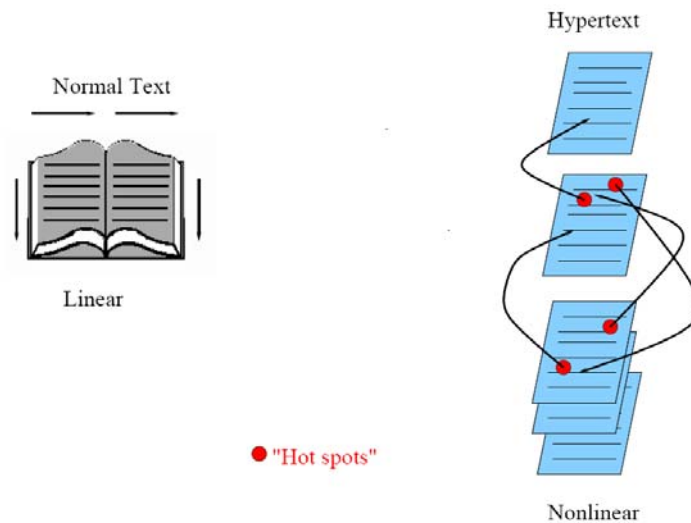
---

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*



Fig 1.1: Hypertext is nonlinear

*Li, Drew, & Liu  © Springer 2021*

---

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

# HTML (HyperText Markup Language)

- HTML: a language for publishing Hypermedia on the WWW — defined using SGML (Standard Generalized Markup Language):

    1. HTML uses ASCII, it is portable to all different (possibly binary incompatible) computer hardware.

    2. The current version of HTML is version 4.01.

    3. The next generation of HTML, HTML5, is still under development.

- HTML uses tags to describe document elements:
    - `<token params>` — defining a starting point.
    - `</token>` — the ending point of the element.
    - Some elements have no ending tags.

*Li, Drew, & Liu  © Springer 2021*

- A very simple HTML page is as follows:

```html
<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>
```

- Naturally, HTML has more complex structures and can be mixed in with other standards.

---

# XML (Extensible Markup Language)

- XML: a markup language for the WWW in which there is modularity of data, structure and view so that user or application can be able to define the tags (structure).

- Example of using XML to retrieve stock information from a database according to a user query:
  1. First use a global Document Type Definition (DTD) that is already defined.
  2. The server side script will abide by the DTD rules to generate an XML  document according to the query using data from your database.
  3. Finally send user the *XML Style Sheet* (XSL) depending on the type of  device used to display the information.

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

- In addition to XML specifications, the following XML-related specifications are standardized:

  - XML Protocol. Used to exchange XML information between processes. It is meant to supersede HTTP and extend it as well as to allow interprocess communications across networks.

  - XML Schema. A more structured and powerful language for defining XML data types (tags). Unlike a DTD, XML Schema uses XML tags for type definitions.

  - XSL. This is basically CSS for XML. On the other hand, XSL is much more complex, having three parts: *XSL Transformations* (XSLT), *XML Path Language* (XPath), and *XSL Formatting Objects*.

15                                    *Li, Drew, & Liu  © Springer 2021*

---

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

- An example of an XML document structure — the definition for a small XHTML document:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE htmlPUBLIC "- //W3C//DTD XHTML 1.0"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-
transition.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
      ... [html that follows the above
      mentioned XML rules]
</html>
```

16                                    *Li, Drew, & Liu  © Springer 2021*

## 1.2.3  Multimedia in the New Millennium

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

- 2001  The first peer-to-peer file sharing system, Napster, was shut down by court order. Coolstreaming was the first large-scale peer-to-peer streaming system, attracting over 1 million users by 2004. 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 .

*Li, Drew, & Liu  © Springer 2021*

- 2005  YouTube created.
   Google launched online maps

- 2006 Twitter created: 500 million users in 2012, 340 million tweets per 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.
  - Google launched Android mobile operating system.

- 2009 - The first LTE (Long Term Evolution) network was set, an important step toward 4G wireless networking.
  - James Cameron's film, Avatar, created a surge on the interest in 3D video.

*Li, Drew, & Liu  © Springer 2021*

- 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. HTML5 is a W3C "Candidate Recommendation"; it is able to run on low powered devices such as smartphones and tablets.

*Li, Drew, & Liu © Springer 2021*

- 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.

- 2015 YouTube launched support for publishing and viewing 360-degree videos, with playback on its website and its Android mobile apps.
  - AlphaGo, a computer program that plays the board game Go, became the first program to beat a human professional player. Its core technology Deep Learning attracted significant attention and have seen success in multimedia content understanding and generation.

*Li, Drew, & Liu © Springer 2021*

- 2016 HoloLens, a pair of mixed reality smartglasses developed and manufactured by Microsoft, started to be available in the market.

  Pokémon Go, an augmented reality (AR) mobile game, was released and credited with popularizing location-based and AR technologies.

  Netflix completely migrated to the Amazon AWS cloud platform, and Skype moved to the Microsoft Azure platform.

- 2017 TikTok, a video-sharing social networking service for creating and sharing short lip-sync, comedy, and talent videos, was launched for the global market (it's Chinese version, Douyin, was launched in 2016).

- 2018 The world's first 16K Ultra High Definition (UHD) short video film, Prairie Wind, was created.

  5G cellular systems started deployment, providing enhanced mobile broadband and ultra low latency access.

  The WiFi 6 (802.11ax) standard was released, offering theoretical maximum throughput of 1 Gbps.

- 2020 Due to the outbreak of corona virus (COVID-19) around the world, work/study from home became a norm in early 2020. Multimedia-empowered online meeting and teaching tools, e.g., Zoom, Google Class, and Microsoft Teams, saw booming use during this period.

## 1.3 Multimedia Software Tools: A Quick Scan

- 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

  7. Multimedia Broadcasting

*Li, Drew, & Liu © Springer 2021*

## 1.3.1  Music Sequencing and Notation

- Cakewalk by Bandlab.
  - The term sequencer comes from older devices that stored sequences of notes ("events", in MIDI).
  - Can insert digital-audio WAV files as well.

- Finale, Sibelius.
  - Composer-level notation systems; these programs likely set the bar for excellence, but their learning curve is fairly steep.

*Li, Drew, & Liu © Springer 2021*

8/18/2020

Fundamentals of Multimedia 3rd ed., Chapter 1

# 1.3.2  Digital Audio

- Digital Audio tools deal with accessing and editing the actual sampled sounds that make up audio:

- **Adobe Audition (formerly Cool Edit):** a very powerful and popular digital audio toolkit; emulates a professional audio studio — multitrack productions and sound file editing, along with digital signal processing effects.

- **Sound Forge:** a sophisticated PC-based program for editing audio WAV files.

- **Avid Pro Tools:** a high-end integrated audio production and editing environment — MIDI creation and manipulation; powerful audio mixing, recording, and editing software. Fill effects depend on purchasing a dongle.

25    Li, Drew, & Liu  © Springer 2021

Fundamentals of Multimedia 3rd ed., Chapter 1

# 1.3.3  Graphics and Image Editing

- **Adobe Illustrator:** a powerful publishing tool from Adobe. Uses vector graphics; graphics can be exported to Web.

- **Adobe Photoshop:** the standard in a tool for graphics, image processing and manipulation.  Allows layers of images, graphics, and text that can be separately manipulated for maximum flexibility, and its set of filters permits creation of sophisticated lighting effects.

- **GIMP:** a free and open-source graphics editor alternative to Photoshop. It supports many bitmap formats, such as GIF, PNG, and JPEG. It also supports *vector-based* formats.

26    Li, Drew, & Liu  © Springer 2021

13

# 1.3.4  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.

27                                                   *Li, Drew, & Liu   © Springer 2021*

# Video Editing (cont'd)

- CyberLink PowerDirector:  Another popular nonlinear video editing software; provides a rich set of audio and video features and special effects and is easy to use. Not as "programmable" as Premiere.

- Adobe After Effects: A powerful video editing tool that enables users to add and change existing movies. Can add many effects: lighting, shadows, motion blurring, layers.

28                                                   *Li, Drew, & Liu   © Springer 2021*

# Video Editing (cont'd)

- **iMovie**: a video editing tool for MacOS and iOS devices. It is versatile, convenient for video editing and creation of movie trailers. iMovie on iPhones is especially handy and popular. Later versions of iMovie also support 4K UHD video editing.

- **Final Cut Pro**: a video editing tool by Apple for the MacOS. It allows the input of video and audio from numerous sources, and provides a complete environment, from editing and color correction to the final output of a video file.

*Li, Drew, & Liu  © Springer 2021*

# 1.3.5  Animation

- **Multimedia APIs:**
  - 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.
    1. Provides a basic set of object primitives (cube, splines, etc.) for building scenes.
    2. It is an abstraction layer built on top of OpenGL or DirectX (the user can select which).
  - DirectX: Windows API that supports video, images, audio and 3-D animation.
  - OpenGL: created in 1992, highly portable; is still most popular 3D API.

*Li, Drew, & Liu  © Springer 2021*

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

- **Animation Software:**
  - Autodesk 3ds Max (formerly 3D Studio Max): rendering tool that includes a number of high-end professional tools for character animation, game development, and visual effects production, e.g., for Sony Playstation.
  - Autodesk Maya: it is a complete modeling, animation, and rendering package; runs on Windows, MacOS and Linux.
  - Blender: a free and open-source alternative to the paid Autodesk suite of tools. It also offers a complete modeling, animation, and rendering feature set, as well as python scripting capabilities.

- **GIF Animation Packages:** a simpler approach to animation — looping through several images, allows very quick development of effective small animation for the web.

31                                                      *Li, Drew, & Liu  © Springer 2021*

---

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

# 1.3.6  Multimedia Authoring

Tools that provide the capability for creating a complete multimedia presentation, including interactive user control, are called authoring programs.

- **Adobe Animate (formerly Adobe Flash):** allows users to create interactive presentations for many different platforms in many different formats, such as HTML5 and WebGL. The content creation process in Animate follows the score metaphor — a timeline arranged in parallel event sequences, much like a musical score consisting of musical notes.

32                                                      *Li, Drew, & Liu  © Springer 2021*

- Adobe Director (formerly Macromedia Director): a multimedia application authoring platform, uses movie metaphor. It includes a built-in scripting language, Lingo. Although not supported by Adobe since 2017, it is still being used to date.

- Adobe Dreamweaver: web page authoring tool that allows users to produce multimedia presentations without learning any HTML.

- Software Development Kits:

  Unity Engine and Unreal Engine

---

# 1.3.7 Multimedia Broadcasting

- OBS, XSplit: two widely used broadcasting tools. OBS is free and open-source, while XSplit is proprietary and paid.

  These tools can be thought of as an entire broadcasting production studio in digital form.

  They offer built-in support for switching between different cameras and other multimedia sources for real-time broadcasting. Users can broadcast live video feeds to websites like YouTube Live, Mixer, Twitch, and various other live streaming websites.

# 1.4  Multimedia in the Future

- *Innovations now or in the near future:*

  - Better camera-based object tracking technology

  - Video shot detection and video classification for online video

  - 3D capture technology for acquiring dynamic facial expression, and synthesizing realistic facial animation

  - Multimedia applications aimed at handicapped persons

  - Crowdsourcing -- Amazon's "Mechanical Turk"

  - Deployment of "Digital fashion" + Wearable computing

# 1.4  Multimedia in the Future (cont'd)

"Grand challenge" problems, which act as a type of state-of-the-art for multimedia interests:

- Social Event Detection for Social Multimedia: discovering social events planned and attended by people.

- Sports Video Annotation: using video classification to label video segments with certain actions such as strokes in table tennis, penalty kicks in soccer games, etc.

- GameStory: a video game analytics challenge in which e-sport games often involving millions of players and viewers are analyzed.

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

- Live Video Streaming: requiring ultra low end-to-end latency. The main challenge is the QoE (Quality of Experience), due to the latency constraint.
- Violent Scenes Detection in Film: automatic detection.
- Preserving Privacy in Surveillance Videos: methods obscuring private information (such as faces on Google Earth).
- Deep Video Understanding: understanding the relationships between different entities from a long duration movie. The relations can be family, work, social and other types.

37                                    *Li, Drew, & Liu  © Springer 2021*

---

*Fundamentals of Multimedia 3ʳᵈ ed., Chapter 1*

- Large-scale Human-centric Video Analysis: analyzing various crowd and complex events such as getting off a train, dining in a busy restaurant, earthquake escape, etc.
- Searching and Question Answering for the SpokenWeb: searching for audio content within audio content by using an audio query, matching spoken questions with a collection of spoken answers.
- Multimedia Recommender Systems: improving the quality of recommender systems to produce items more relevant to users' interests. Applications include movie/news recommendation, etc.

38                                    *Li, Drew, & Liu  © Springer 2021*