Chapter 1
Introduction to Multimedia

Li, Drew, and Liu

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**.

- Multimedia and Computer Science:

  -Graphics, HCI, visualization, computer vision, data compression, graph theory, networking, database systems.

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. Geographically-based, realtime augmented-reality, massively multiplayer online video games.
2. Shapeshifting TV, where viewers vote on the plot path.
3. Tele-medicine.
4. A camera that suggests what would be be best type of next shot.
5. A web-based video editor that lets anyone create a new video by editing, annotating, and remixing editable professional videos on the cloud.
6. Cooperative education environments that allow schoolchildren to share a single educational game using two mice at once.

Li, Drew, and Liu

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

8.   Compositing of artificial and natural video into hybrid scenes.

9.   Visual cues of video-conference participants, taking into account gaze direction and attention.

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

11.   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.

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

Current Multimedia Projects

- Many exciting research projects are currently underway. Here are a few of them:

1. **Camera-based object tracking technology**: tracking of the control objects provides user control of the process.

2. **3D motion capture**: used for multiple actor capture so that multiple real actors in a virtual studio can be used to automatically produce realistic animated models with natural movement.

3. **Multiple views**: allowing photo-realistic (video-quality) synthesis of virtual actors from several cameras or from a single camera under differing lighting.

4. **3D sentiment- and speech-capture technology**: allow synthesis of highly realistic facial animation from speech.

5. **Specific multimedia applications**: aimed at handicapped persons with low vision capability and the elderly — a rich field of endeavor.

6. **Digital fashion**: aims to develop smart clothing that can communicate with other such enhanced clothing using wireless communication, so as to artificially enhance human interaction in a social setting.

7. **Distributed medical care**: an initiative for providing interactive health monitoring services to patients in their homes

8. **Augmented Interaction applications**: used to develop interfaces between real and virtual humans for tasks such as augmented storytelling.

Current Multimedia Projects

- 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.

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"

- 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.

- 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** 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.

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

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.
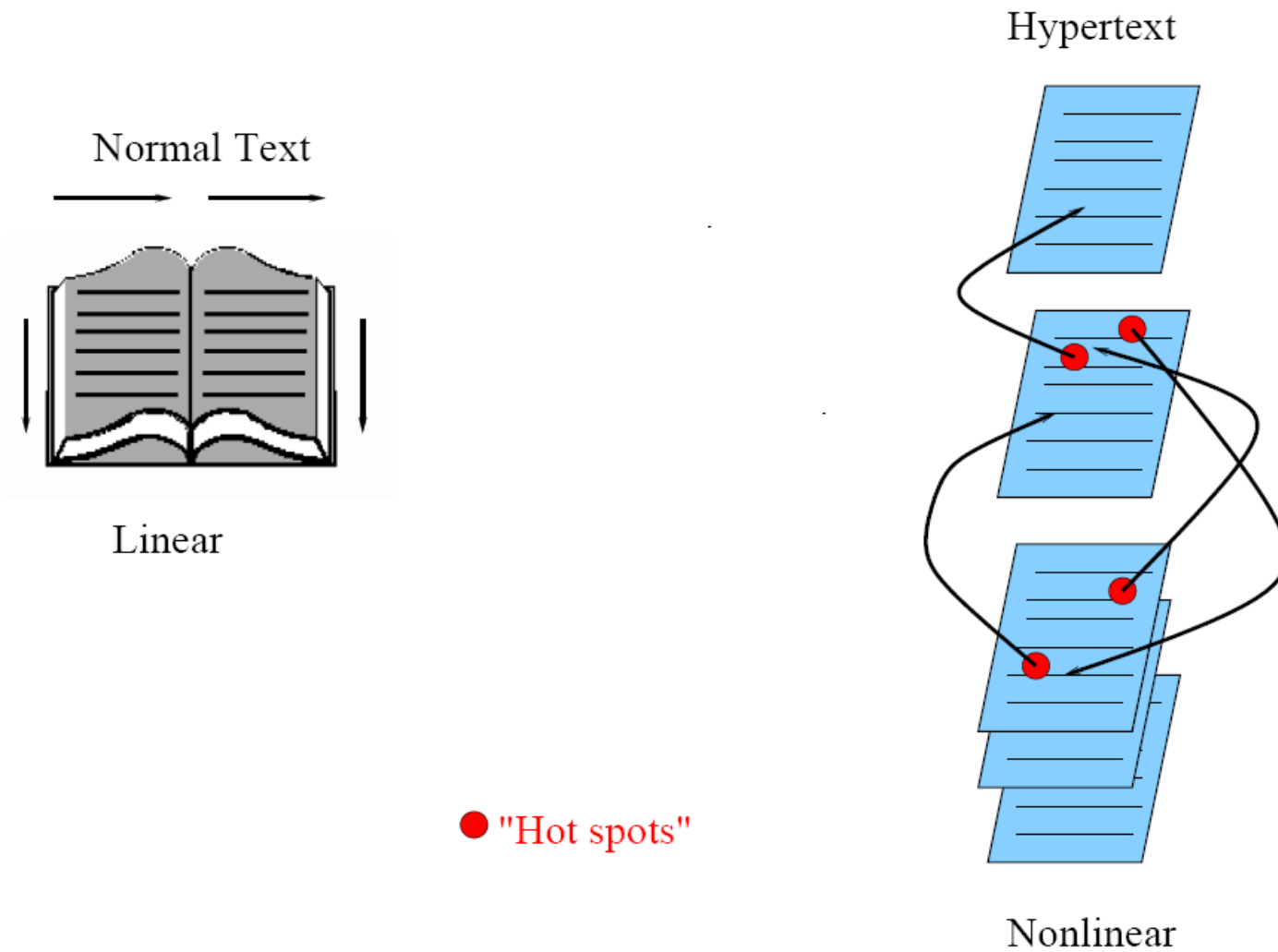
Fig 1.1: Hypertext is nonlinear

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.

- A very simple HTML page is as follows:

```
<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.

- 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*.

- 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>
```

Multimedia in the New Millennium

- **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 .

- **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.

- **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.

- **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.

Li, Drew, and Liu

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**

Li, Drew, and Liu

Music Sequencing and Notation

- **Cakewalk Pro Audio.**

  – 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 sequencing/editing program, with capabilities similar to those of Cakewalk. Include some digital audio editing tools.

Digital Audio

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

- **Adobe Audition**: a very powerful and popular digital audio toolkit; emulates a professional audio studio — multitrack productions and sound file editing including digital signal processing effects.

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

- **Pro Tools**: a high-end integrated audio production and editing environment — MIDI creation and manipulation; powerful audio mixing, recording, and editing software. Popular with a professional niche market on Macintosh.

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 graphics, image processing and manipulation tool.
  - Allows layers of images, graphics, and text that can be separately manipulated for maximum flexibility.

- **Adobe Fireworks**: software for making graphics specifically for the web.

- **Adobe Freehand**: a text and web graphics editing tool that supports many bitmap formats such as GIF, PNG, and JPEG.

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

Video Editing

- **CyberLink PowerDirector:** The most 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.

- **Final Cut Pro**: a video editing tool by Apple; Macintosh only. Final Cut is the definitive high-end video editing tool.

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, and Liu

- **Animation Software**:

- **Autodesk 3ds Max** (formerly 3D Studio Max): rendering tool that includes a number of very high-end professional tools for character animation, game development, and visual effects production.

- **Autodesk Softimage** (previously called Softimage XSI): a powerful modeling, animation, and rendering package used for animation and special effects in films and games.

- **Autodesk Maya**: competing product to Softimage; as well, it is a complete modeling package.

- **RenderMan**: rendering package created by Pixar.

- **GIF Animation Packages**: a simpler approach to animation, allows very quick development of effective small animations for the web.

Multimedia Authoring

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

- **Adobe Flash**: allows users to create interactive movies by using the score metaphor, i.e., a timeline arranged in parallel event sequences. Flash movies are commonly used to show movies or games on the web.

- **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.

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

- **Authorware**: legacy flowchart-based authoring system.

1.4 Multimedia in the Future

- *Innovations now or in the near future:*

  - Better camera-based object tracking technology

  - Video shot detection—finding where scene changes exist in video—and video classification

  - 3D capture technology; multiple views from several cameras; or a single camera under differing lighting

  - Multimedia applications aimed at handicapped persons

  - Crowdsourcing -- Amazon's "Mechanical Turk"

  - Deployment of "Digital fashion" + Wearable computing

1.4 Multimedia in the Future

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

- 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.

All these use extra features ("meta-data") such as text, descriptive features calculated for audio and video, web pages, and social tags to help in the task.

Li, Drew, and Liu