Teaching Multimedia – from Multimedia Signals, Audio and Visual Processing, to Multimedia Networks

Iliya Georgiev

Abstract: *Multimedia processing is becoming mainstream in today's computing environment. Teaching and learning multimedia basics are relevant to a diversity of scientific areas: signals and systems, data compression, computer graphics, image processing and understanding, digital audio and video, networking. The paper describes a generalized approach to multimedia teaching that is designed to present a multimedia-processing model and the basics of the related disciplines.*

Key words: *Teaching Multimedia, Multimedia Signals and Systems, Data Compression, Computer Graphics, Image Processing, Multimedia Networks*

INTRODUCTION

Digital multimedia courses are not often included in many computer science and electrical engineering curricula. This situation is going to be changed for several reasons. First, multimedia has become a pervasive part of modern computing especially with the "picture and sound" nature of WWW. Beyond the obvious usage in scientific visualization, information industry and entertainment, multimedia has become a central component of net-centred computing, human-computer interfaces, different Web applications, as well as data analysis for domains such as medical and environmental studies. Scientific fundamental concepts to multimedia now need to be introduced as soon as possible in the curriculum to give students the background they need for today's computing. Second, multimedia can provide other disciplines with practical, real-world exercises. Students are always interested in how the concepts they are learning will benefit them when they get a job. Showing them how the theory they just leaned can be directly applied to most interesting and practically oriented applications, such as those that multimedia can provide, is clearly important.

A mayor challenge for the teachers in multimedia is to cover a large body of basic theory with different mathematical background. Manipulating multimedia components and sequences uses methods of several scientific disciplines like signals and systems, digital imagery (image processing, image understanding, computer graphics, computer animation, geometric modelling), audio processing, speech processing, data compression and networking.

It is impossible to cover all topics in a single course and often depth must be sacrificed by breath. Traditionally, there are two complementary approaches to teach such theoretical- intensive courses. The first is to organize a review course, which explains the structure of multimedia documents, methods for their manipulating, multimedia networks. Here students are often expected to implement multimedia manipulation in software to develop skills for performing well in the industry. In addition to the abundance of diverse material, the students encounter implementation hurdles in form of miscellaneous programming details, which reduces the time they spend on the road of algorithm understanding. The second approach is to choose a subset of topics that gives students a skewed view of the field of multimedia. They learned only the selected subset of basic methods but ended up believing that they had learned the most important ones. In multimedia, which is a relatively immature subject, most problems are still open for research; such approach gives the impression that most problems can be solved using the set of just learned techniques.

This paper documents a model of multimedia processing that has been used to form teaching and research activities at university level. The model has been successfully used giving broad knowledge of the relationship between different classical techniques and problems, which remain unsolved or only partly solved.

OBJECTIVES

Our objectives correlate with the understanding what a successful students should be able to:

- Read multimedia literature, critically analyze recent methods and determine their applicability for solving a problem;
- Implement techniques reported on in recent multimedia literature;
- Understand the problems open for research;
- Follow the broad standardization process of different multimedia topics.

Most of the courses teach students to know some classical techniques. Although classical techniques are important, it is not feasible to cover all of them, or to decide which one will be most significant in future research. Therefore, we decided to spend 30 % of the teaching time to present a main approach to create, compress, and transfer multimedia documents. The model of multimedia processing offers a comprehensive abstraction of scientific basics of the multimedia-computing environment.

MODEL OF MULTIMEDIA PROCESSING

Figure 1 presents the model we used as a teaching platform for multimedia course. The model shows the multimedia system properties and functions as well as the sciences that provide corresponding methods, algorithms and techniques.

Multimedia System Properties

The model was used to explain several important multimedia characteristics.

The diversity of the digital representation of multimedia components. Three presentations with different scientific foundations are analyzed.

Multimedia presentation as signals evolving in time. Media can be divided into two types: time-independent or discrete media (text, generated graphics, still images) and time-dependent (audio, video). In time-dependent media, the information is generated continuously and is passed directly to the destination, and at the destination, the information stream is played directly as it is received. The scientific discipline *signal and systems* provides several methods for uniform processing of different media as signals.

Static digital presentation for transformation, manipulation and compression. The media comprises a single block of information that is often stored as a file. Here the media is manipulated and structured using the methods of the *traditional computer sciences* and several specific like *image processing and understanding*, *audio and speech processing*, *computer graphics,* and *geometric modelling*.

Bitstream presentation for networking transfer and reproduction. In term of the bit rate at which the source information stream is generated, this may be either constant or variable bit rate. The digitized audio stream is generated at a constant bit rate. In the case of video, although the individual pictures that make up the video are generated at a constant rate, after compression, the amount of information associated with each frame varies. Multimedia specifics have modified the functionality of computer communications and now *multimedia networks* are considered as a specific domain.

The duality between integration and independence of media. Multimedia systems include two or more media, at least one continuous (time dependent) and one discrete (time independent) media. A high level of integration ensures that changing the content of one media causes corresponding changes in other media. On the other hand, different media should have a high degree of independence. This is an important criterion for a multimedia system, as it enables independent processing and provides the flexibility of combining media in arbitrary forms. Some of the most conventional information sources that include two or more media are not considered digital multimedia systems, because these systems do not satisfy the independence criteria. For example, text and images in a newspaper are tightly coupled.

The domains of multimedia. Multimedia can be divided into four domains: device, system, application and cross-domains. The *device domain* includes storage media, networks and the basic concepts such as audio, video, graphics, and images. Conversely,

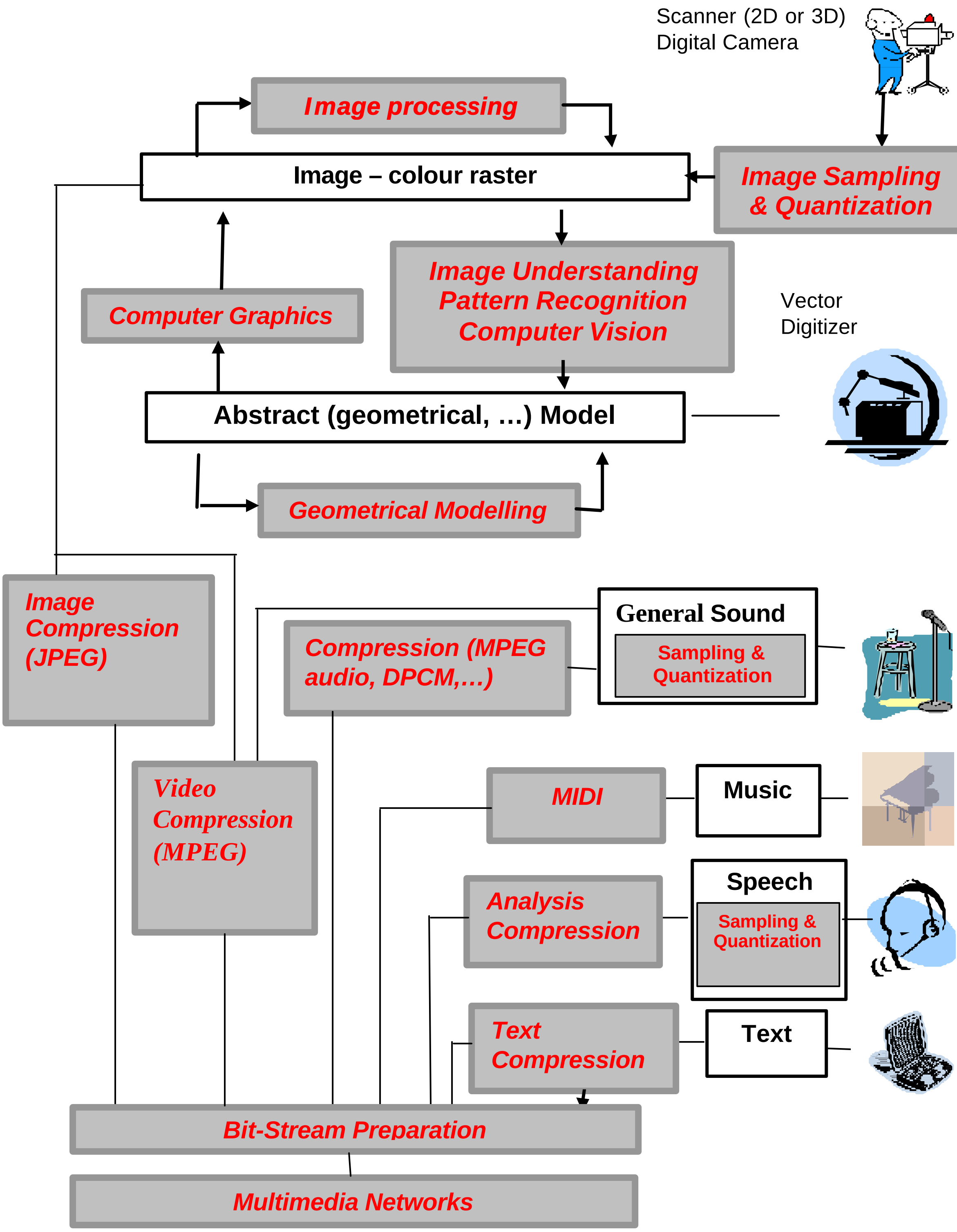

Figure 1. A model of multimedia processing

the *system domain* includes multimedia database libraries, operating systems and communication systems. The *application domain* includes the user interface through which various tools, applications, and documents are made accessible. The *cross domain* includes the integration and synchronization of various media.

Multimedia System Functionality and Related Scientific Disciplines

We start with the nature of video and audio signals and their transmission over the communication channel. Signals of this kind are analogue signals that we consider as information carriers. For students it is important to understand that some analogue signals cannot technically be transferred over all types of channels. Transformation into digital form makes the transmission possible. This in turn requires using proper scientific methods of *sampling and quantization* that convert the analogue signals to discrete signals. This conversion changes the information into signals that could be analysed with the *basic concepts of information theory. Topics learned:* the human auditory and visual system; audio and video signal representation in time and frequency domain; sampling, alias signals; quantization, quantization errors and noise.

Further, we present the scientific foundation of the different media: visual, audio, and text. The processing of the *visual information* we explain with a processing model (upper part of the Fig.1) emphasizing on the difference of images, taken from the real life, and computer-synthesized images. On the one side of the processing model, abstract geometrical models present man-made objects. G*eometrical modeling* provides the required transformations of the abstract models. *Computer graphics* generates from abstract models realistically looking images using *computational geometry* (Euclid geometry for man-made objects, fractal geometry for nature). *Image processing* transforms the taken from real-life images. *Image understanding and computer vision* extract specific image properties (contours detection, pattern recognition, surface recognition, etc.) and generate abstract model of the image. *Topics learned:* geometrical models (2D, 2 ½ D, 3D: wire frame representation, boundary representation, constructive solid geometry, surface modelling); image representation and transformations: full-colour raster image, colour image with reduced number of colours, chain differential encoding of edges, vector representation; computer graphics visualization pipeline (coordinate systems, clipping and windows transformations, coordinate transformations, hidden lines analysis, shading, textures); image processing stages: image transformation, image restoration, segmentation, shape representation).

Three different form of *audio* are considered digital media. *General audio* is processed by *signal and system methods* for filtering, noise removal. *Music* is represented by a sequence that specifies how the musical instruments are to be played at different time instances. *Speech processing* decomposes different components of the human speech and groups the phonemes. *Topics learned:* general audio, MIDI; speech representation.

Although computers become faster and data storage becomes less expensive and more efficient, the increased importance of sound and video necessitates of *data compression* due to vast storage and transmission requirements. By using the model we present the representatives methods of *lossless compression* of text, and *lossy compression* of audio, speech, images and video. *Topics learned:* entropy encoding (run-length, statistical), source encoding (differential, transform).

Text data are truly digital data and can be assumed to be created by a discrete information source that emits symbols corresponding to an alphabet. There is no time-domain sampling or quantization. Typical text data have redundant information. We emphasize on the structured text in XML format. The t*ext compression* reduce or eliminate this redundancy. There are primary two types of redundancies present in text data: statistical and knowledge redundancy. Statistical redundancy refers to non-uniform probabilities of the symbols occurrences. The symbols that occur more frequently should be given a larger number of bits than the less probable symbols. Knowledge redundancy works over text in limited scope. A common knowledge can be associated with it with both the encoder and decoder. The encoder can then transmit only the necessary information. *Topics learned:* Huffman coding (static, dynamic), dictionary encoding (LZW algorithm).

Audio compression is diverse according to the type (general sound, speech, music) and to the requirements to the transmission channel. *General sound* requires a large number of bits for representation. For example, CD quality stereo audio requires 176.4 Kbytes/sec data rate. This bandwith is too large for many applications like transmission over Internet. We present used techniques that achieve superior compression performance (good quality at low bit-rate) by removing redundancies in an audio signal. Statistical redundancy removal is based on the probability of occurrence of audio samples. The sound signal is time varying and there is a strong corellation between the neighboring sample values. This temporal inter-state redundancy is removed by differential predictive coding and transform coding (based on discrete cosine transformation or wavelet transformation. Analysing the whole scope of the sound signal, a common knowledge can be associated with it and to transfer only the necessity information required for reconstruction. For example, the pieces of the signal we can not hear (frequency masking and loudness coma). *Music compression* is based on MIDI representation that exploit knowledge redundancy and encodes the names of instrument and how they are played. *Speech compression* is based on speech decomposition in white noice (produced by mouth) and sinwave signals (produced by voice cords) and is coded by prediction encoding. *Topics learned:* transform encoding, differential encoding, linear predictive code, perceptual coding, and MPEG audio.

Image compression uses almost all encoding techniques to remove redundancy. Statistical redundancy refers to the correlation between neighboring pixels. It is removed by exploiyng prediction and transform encoding. Further we note that the image is originally a projection of 3D objects onto 2D plane. Therefore, if the image is encoded using structural image models that take into account the 3D properties of the scene, a high compression ratio can be achieved. For example, a segmentation coding approach that considers an image as an assembly of many regions and encodes the contour and texture of each region separately, can efficiency explit the structural redundancy. The following specific properties (psychovisual model) of the human visual system are used in compression: greater senstivity to distortion in smooth areas compared to areas with sharp changes; greater sensitivity to signal changes in the luminance component compared to the chrominance in colour pictures. *Topics learned:* image compression: graphics interchange format, JPEG principles; video compression principles (frame types, motion estimation and compensation), MPEG video.

Multimedia bitstreams are transferred over *multimedia networks*. If the transmitted bitstream is the same as the received signal, we call the network noiseless. The real networks are noisy. There is a need to detect the bitstream distortion to rectify the problem either by retransmitting it or by fixing it on the receiving side by *error-correcting codes.* The operational parameters associated with a network are known as the network *Quality of Service* (QoS). *Topics learned:* basic multimedia networks (telephone networks, data networks, broadcast television networks, integrated services digital networks, broadcast multi-service networks), polynomial error correcting codes, network QoS.

COURSE ORGANIZATION

A *60-hour course* we organized in following sequence: multimedia processing model and the basics of the related sciences – 20 hours; solving problems from sampling and quantization – 4 hours; selected compression methods (Huffman tree construction, LZW dictionary compression, JPEG sequence, MPEG frames) – 6 hours; selected algorithms in computer graphics (Bresenham algorithm, fill area algorithm, clipping algorithm, z-buffer hidden surfaces algorithm) – 8 hours; image processing selected algorithms (brightness transformation, image restoration) – 6 hours; computer animation – 2 hours; multimedia networks – 6 hours; practical issues – 6 hours; projects presentations and discussions – 2 hours.

Students have to choose from *two alternative projects*. The *first assignment* is oriented to students who want to gain some skills in usage of some multimedia tools (authoring and simulation – Macromedia Director, HTML editor – Macromedia Dreamweaver, vector graphics and drawings – Macromedia Freehand and Flash, video editor – Adobe Premiere, sound editor – Sound Forge) and to create several minutes multimedia presentation. The *second project* is to learn fundamentals of the development of multimedia programs (interfaces and libraries) and to implement a Media Player by using JAVA Media Framework or Visual C++ Media Control Interface that should be able to play WAV, MPEG, MIDI and AVI files with a user interface and reasonable controls.

RELATED WORKS

Multimedia sciences are highly mathematical subjects. No one teaching approach can expect all students to experience the process of self-teaching of unfamiliar mathematics and there is no practical way to use it in teaching. This is the main reason for our approach to organize multimedia courses around a comprehensive explanation of the multimedia processing.

For the same reason most courses incorporate only two or three topics. Here we can enumerate some typical representatives: courses that master multimedia tools and their application to engineering (eml.on.edu/papers2000/julian.pdf at the University of Oklahoma); courses that concentrate on multimedia signals and systems (www.cs. ualberta.ca/courses/Cmput414.html at University of Alberta); computer science courses emphasising on multimedia digital representation and compression (www.cogs.susx.ac.uk/ users/andyh/lastyear/multimedia-systems/ms.html at University of Sussex); computer science courses emphasising on multimedia compression and multimedia networks (eeclass.stanford.edu/ee384b/ at Stanford University).

CONCLUSIONS AND FUTURE WORK

We have presented multimedia course with a scientific orientation based on a model of multimedia processing. It deals in a systematic manner with all the issues that come from the lack of sufficiently accurate knowledge, both instantaneous and predictive, of the state of a multimedia system. The success of the course has inspired our conviction that it does not seem like a hodgepodge of topics. In the future, we are going experimentally to reorganize the course following way. The first part explaining the multimedia-processing model will remain unchanged. All the rest will be changed toward problem-based learning. The covered set of topics will be reduced to some units. The students will focus on their own research into recent primary literature and discover new techniques to solve particular problems.

REFERENCES (recommended books in alphanumerical order)

[1] Chapman N., J. Chapman. Digital Multimedia. John Wiley & Sons, 2002.

[2] Drozdek A. Elements of Data Compression. Brooks Cole, 2002.

[3] Foley J., A. van Dam, S. Feiner, J.Hughes, R.Phillips. Introduction to Computer Graphics. Addison-Wesley, 2002.

[4] Halsall F. Multimedia Communications. Addison-Wesley, 2001.

[5] Mrinal Kr. Mandal , Multimedia Signals and Systems, Kluwer, December 2002.

[6] Solary S. Digital Video and Audio Compression. McGraw-Hill, 2001.

[7]Sonka M., V.Hlavac, R.Boyle. Image Processing, Analysis, and Machine Vision. ITP, 1999.

[8] Stockman G., L. Shapiro. Computer Vision. Prentice-Hall, 2001.

ABOUT THE AUTHOR

Iliya Georgiev, Prof. PhD, Department of Mathematical and Computer Sciences, Metro State College of Denver, Campus box 38, P.O. Box 173362, Denver 80217, USA Phone: +303 673 9403, ?-mail: gueorgil@mscd.edu, URL: http://clem.mscd.edu/~gueorgil/