MULTIMEDIA : THE INSIDE STORY

Siddhartha Bhattacharyya*, Biswanath Chakraborty# and Syamasis Majumder$
*Department of Information Technology,
Department of Computer Science & Engineering,
$ Department of Computer Applications

1. What is Multimedia?

The term "Multimedia" literally refers to a combination of several media which, taken together form the backbone of latest trends of technological advancement. In fact, the field of multimedia aims at representing computer information, both discrete (time independent) and continuous (time dependent) through various media like text, images, audio, video etc. Multimedia has evolved from such fields where either discrete or continuous data are existent. Such fields include Telecommunication, Consumer Electronics, Recording Studios and Publishing Houses.

2. Multimedia Components:

The field of multimedia involves four different domains of activities: (i) Device Domain, which deals with processing and representing information in meaningful and usable form, (ii) Application Domain, meant for presenting multimedia data, (iii) System Domain, which offers special system services for multimedia data storage and transmission and (iv) Cross Domain, which is responsible for proper synchronization between the other three domains. Multimedia has two components, viz. hardware and software components. These components make up the six different media involved in a multimedia system which are (i) the Perception Medium, (ii) the Representation Medium, (iii) the Presentation Medium, (iv) the Storage Medium, (v) the Transmission Medium and (vi) the Information Exchange Medium.

The perception medium refers to the human interpretation and understanding of the multimedia data. The representation and the presentation media refer to the standards of multimedia content management and creation. Different ways of representing and presenting multimedia information to the end user are included as multimedia document architecture for document preparation and authoring. The storage medium, as the name suggests, defines the standards and devices for storage of multimedia data. The transmission and the information exchange medium refer to those media used for disseminating multimedia information. These lay down constraints based on bandwidth requirements and common frameworks during data transfer. In the following five sections we will be focussing on the different types of multimedia information.

3. Text:

Text is a common and the most primitive form of expressing one's own feelings. It is a time independent (hence discrete) form of multimedia data. Various types of texts are in vogue. The text messages you type in the DOS prompt or in the Windows NotepadÒ window are referred to as "unformatted" or "plain" text and the files you store these messages contain ASCII (**A**merican **S**tandards **C**ode for **I**nformation **I**nterchange) codes of the same. These files are basically ASCII files and do not

contain any special information regarding the file type and size. However, if you resort to software like Microsoft Word, the texts you type in are formatted texts also referred to as "rich" texts. These texts can be formatted at will, their fonts, styles and sizes can be changed. The files used for storage of these rich texts contain special information referred to as "header information", which is typical of these types of files. The header information comprises information regarding the type, size, creator and other specifics of these files. On the other hand, whenever you go for preparing e-contents, you have to use "hypertexts", a special form of formatted texts. Hypertexts are prepared using Hypertext Markup Languages (HTML).

4. **Images and Graphics:**

Apart from the textual form of information representation, images and graphics form another discrete mode of multimedia information. Images speak eloquence. With the help of images much better can be expressed in much lesser efforts.

Images are basically a two dimensional representation of information. Images can be classified into two types, viz. range images and intensity images, depending on the method of acquisition. Range images are basically maps of relative distances between objects, while intensity images refer to maps of light intensities reflected from an illuminated object. The larger the reflectance of an object, the greater is the intensity value of the corresponding pixels to be recorded in an image and smaller the reflectance, the smaller is the intensity value.

Depending on the range and number of bits used to represent the intensity value, several types of images are there. For example, a binary/monochrome image contain only two intensity values (0 for darker regions and 1 for lighter regions) and is represented by a single bit only. Gray scale images contain 256 shades of gray intensity levels (0 for the darkest shade to 255 for the brightest shade). Since at least 8 bits are required to represent this range (0 to 255) of intensity levels, these images are also referred to as 8 bit images. Colour images are similar to these types of images with the exception that in these images, each and every piece of information (pixel) contain three colour intensity levels. These colour intensity levels are either the three primary colours (red, green and blue) or an admixture of these three primary colour components. Again, each of these three primary colour components along with their admixtures have at the most 256 shades (0 for dark colour to 255 for bright colour). If a colour image contains only the three primary colour components and their admixtures in equal amounts, that image is referred to as "pure" colour image. These images can contain pure red (0 red to 255 red) and/or pure green (0 green to 255 green) and/or pure blue (0 blue to 255 blue) colours only. These images also require 8 bits to represent each colour value, hence are 24 bit (8 bit ´ 3 colours) colour images. On the contrary, a "true" colour image contains all possible colours (either the primaries or secondary generated from the primaries like yellow, cyan, magenta etc.) in all possible shades (0 to 255). Thus the maximum number of colours possible in a "true" colour image is equal to 255x255x255 = 16777215, i.e., 16 million colours. These images are also 24 bit colour images. However, if another 8 bits (known as control bits) for controlling colour parity is also included in these images, they become 32 bit colour images.

However, though images replicate object information, they often contain redundant information. To clarify this point further, take the example of a clear sky. If you acquire a 100x100 sized image of the clear sky, most part of the information recorded in image form contain the blue colour. Other colours are also present but in small amounts. If this colour information is to be stored using 24 bits for each single pixel, you would require 24x100x100 = 240000 bits i.e. 240 kilobits. Since the blue colour occurs repetitively (about 10000 times) in the image, if you can only store the number of repetitions along with the blue colour value ignoring the other colour components (for their small contributions), then you will require only 10000x8 = 80000 bits i.e. 8 kilobits only. This mechanism of optimizing image information at the cost of loss of redundant information is referred to as compression. Several types of compression schemes are in existence and hence several image formats are in practice. Some common examples are the bitmap format (.bmp), the graphics interchange format (.gif), the animated graphics interchange format (.gif), the Joint Photographers Expert Group (JPEG) format (.jpeg) etc. However, JPEG is not an image format, it is a compression standard. The corresponding file format is a file interchange format (.jfif).

Finally, depending on the image manipulation schemes used for editing and creating images and graphics, two different types of graphics or images are there. They are the raster graphics, which solely depend on image information storage using the raster scanning principle and the vector graphics. In raster graphics, information regarding each pixel in the image requires to be stored. Hence, the memory requirements are much larger. The file formats described before use this scheme. Vector graphics systems store image information as a combination of lines, curves and closed figures. So, the memory requirements are smaller.

5. **Animation:**

The first of the types of continuous multimedia information described in this article is animation. It is time dependent information. Animation is a method of representing a sequence of related multimedia information with proper time synchronization. When several textual or graphic information is displayed frame by frame to the end user with a finite time delay between themselves, an animation is created. Animation means dynamic information. The different steps in creating an animation are :

- Creating a story of related events
- Acquiring or creating the initial frame
- Acquiring or creating the intermediate frames
- Acquiring or creating the final frame
- Adding finite delay between frames
- Synchronizing the display of frames
- Implementing value additions

The most important and laborious part of the entire procedure is the creation of the individual frames. However, of late, several software packages (e.g. Macromedia Flash, Macromedia Director) have come up, which can create the intermediate frames on their own starting form the initial frame. These packages apply interpolation schemes to generate the intermediate frames. Value additions to an animated sequence can be achieved by adding prerecorded audio files.

6. **Audio:**

Another continuous type of multimedia information is audio. Audio is air pressure fluctuations. When the frequency of these fluctuations or vibrations lies in the human audible range (20 Hertz to 20 kiloHertz), then the corresponding vibration is referred to as sound. In the external world, sound waves exist in analogue form. A computer only can handle the digital form of the analogue waveform. An analogue to digital converter (ADC) converts these waveforms into digital form for computer storage. This analogue sound is sampled, then quantized to form a digital stream. The rate or frequency of sampling is the sample rate. Sampling is done following the Nyquist sampling criteria which states that the sampling rate must be greater than or equal to the highest frequency content (bandwidth) in an waveform so that the analogue waveform can be again faithfully reconstructed from the digital stream. For CD quality audio, the rate is 44.1 kHz which is greater than 19.98 kHz (20000 - 20) Hertz. The resolution or quantization of a sample value depends on the number of bits to represent the height of the waveform. 8-bit quantization yields 256 levels. CD quality is of 16 bits. Lower the quantization the lower is the quality of the sound.

Audio signals are stored in the computer in different formats depending on the different types of audio encoding techniques, which are listed below.

- PCM encoding
- Delta Modulation
- ADPCM
- *A*-law and μ-law encoding
- Fourier Methods
- Subband Coding and MPEG (Motion Pictures Expert Group) audio encoding

These encoding schemes are different in the ways they represent the individual bits of the digital audio stream. Out of these schemes, the MPEG audio encoding scheme also compresses the audio stream. It takes human psycho-acoustics into account, which reflects the human hearing system in that weaker and unimportant signals are removed and stronger signals are given priority in the encoding procedure. Some examples of audio file formats are the resource interchange file format (.riff), the wave file format (.wav), the real audio file format (.ra), the MPEG layer 3 encoded file format (.mp3), the musical instrument digital interface file format (.mid) etc.

7. **Video:**

The final continuous type of multimedia information in this series is video. Video signals are complex waveforms comprised of signals representing a picture as well as the timing information needed to display the picture. To capture and use those complex signals, you need special electronics to do the job.

Analog video signals refer to changing the original signal acquired (in a camera) into something that represents the signal - in this case, into a waveform transferred through video cable or other transmission medium (like through air in TV broadcasts). There are also digital video signals where the picture contents are encoded in digital format (information converted to a series of bits, which represent numbers). A colour video signal is formed by the mixture of three primary colour signals - the red, the green and the blue signals. The screen of the picture tube is coated with a set of three different phosphors-one for each colour-each of which is activated by a separate electron beam. Three electron beams are scanned in unison across the screen from left to right with different resolution depending on different standards. The resolution is 525 lines

in case of NTSC (National Television System Committee) or 625 lines in case of PAL (Phase Alternating Line) and SECAM (Sequential Colour With Memory). The total screen contents are then refreshed at a rate of either 60 or 50 frames per second depending on the frequency of the mains electricity supply in different countries. Broadcast television operates differently from computer monitors in terms of scanning sequence and choice of colour signals.

In order to make black-white (monochrome) TV compatible with colour TV for displaying video signals, three main coding schemes are used. They are (i) RGB, (ii) YUV and (iii) YIQ coding schemes.

RGB coding deals with the three colour signals (Red, Green and Blue). Different colours are formed by combination of these colours, e.g. R+G+B=W produces the neutral white colour. But human eye is more sensitive to brightness than any chrominance information. So separation of brightness (luminance Y) from the chrominance (U and V) is needed.

The component division for YUV signal (used by PAL) is:

Y=0.30R+0.59G+0.11B

U=0.493(B-Y)

V=0.877(R-Y)

where, B-Y is the blue chrominance C_b and R-Y is the red chrominance C_r.

The component division for YIQ (used by NTSC) is:

Y = 0.30R+0.59G+0.11B

I = 0.60R-0.28G-0.32B

Q = 0.21R-0.52G-0.31B

Digital video is required for storage and subsequent transmission. Several digitization formats are in existence.

- 4:2:2 format
- 4:2:0 format
- HDTV format
- Source Intermediate Format (SIF)
- Common Intermediate Format (CIF)
- Quarter CIF (QCIF)

These formats differ in the relative amount of chrominance and luminance used during digitization.

A digital video can be thought of as a sequence of digitized pictures or frames. A large amount of memory is required for storage and a high bandwidth is required for faithful transmission of digital video signals. Hence proper compression techniques has to be used for optimizing the video signals. One can apply JPEG compression to each and every individual frame resulting in a moving JPEG or MJPEG compression. But the compression ratio obtained is not sufficient. Moreover, apart from spatial redundancy in a single frame, there exists other redundancy between subsequent frames. There is always slight change between frames. So, bandwidth requirements can be minimized by using this temporal difference between frames.

MPEG (Motion Pictures Expert Group) is an international body, which has set various video compression standards. The essence of their video compression technique lies in the fact that the individual frames are not transmitted in a video stream, rather only the initial frame (also referred to as the I-frame or image frame) is transmitted followed by the difference frames (also known as predictive frames or P-frames and bi-directional predictive frames or B-frames) between the succeeding frames and the first frame. At the receiving end, the intermediate frames are predicted by using motion estimation and motion compensation techniques. With time, they have evolved MPEG1, MPEG2, MPEG4, MPEG7 and MPEG21 standards for video compression.

As far as the information storage intricacies are concerned, a MPEG1 video bit stream comprises several groups of pictures (GOP), a GOP in turn contains several pictures, a picture contains several slices while a slice contains several macroblocks which are the compressed frame information.

Digital video are stored in various formats like MPEG (.mpg), Digital Audio Tape format (.dat), Audio Video Interleaved (.avi), Windows Movie format (.wmv) etc.

8. Synchronization:

As discussed before, for faithful dissemination of multimedia information, synchronization between the different units of multimedia information is a must. Whether it is an animation or an audio piece, if the atomic units of these continuous type of information are not properly synchronized, proper sequence of multimedia events expressed therein cannot be maintained and the resulting information will appear distorted and meaningless. Synchronization can be achieved at different levels of multimedia information. Intra-object synchronization is concerned with the synchronization between the individual units within the same information content. Inter-object synchronization deals with the synchronization between different information contents. These types of synchronization schemes can be achieved both in a preconceived (Synthetic Synchronization) as well as in a real time (Live Synchronization) piece of multimedia content.

9. Multimedia Applications:

The impact of multimedia technology is vast in the present era. Several fields of technology are resorting to multimedia based applications for enhanced performance. Typical examples include the Spatial Data Management System (SDMS), Movie Maps, Electronic Books, Digital Library Solutions, Video-on-Demand, Interactive Cinema, Computer Games, GIS etc. A score of multimedia application packages are now available in the market for creating, editing, formatting and storing both discrete and continuous multimedia data. Different web-editors like Microsoft FrontPage, Coffee Cup HTML etc. can be used for generating discrete textual multimedia information in a presentable form. Images and graphics can be handled by software like Microsoft Paint, Paint Shop Pro, Microsoft Imaging, Adobe photo Shop etc. Audio recording can be done by Microsoft Sound Recorder. Different audio file formats can be played using WinAmp, Microsoft Sound Player etc. Audio editing tools are now quite common and available in Sound Forge package. Ulead Video Studio Pro and Adobe Premier are well-known packages for editing and rendering video files.

10. Further Reading:

[1] Ralf Steinmetz and Klara Nahrstedt, "Multimedia:Computing, Communications and Applications"

[2] Fred Halshall, "Multimedia Communications

[3] Koegel Buford, "Multimedia Systems"

Windows Tips and Tricks :

Do you know that you can draw an arrow between your text in Microsoft Word without using your mouse. Just type "-" followed by another "-" followed by a ">". An arrow automatically appears. Isn't it simple? Send feedback to siddharthab2k@rediffmail.com or mail2arth@yahoo.com