

Solution	Marks
<p>Q.2 a. What is meant by the terms Multimedia? Explain in points the diverse uses of multimedia. (8)</p> <p>Answer: Multimedia is the field concerned with the computer-controlled integration of text, graphics, drawings, still and moving images (Video), animation, audio, and any other media where every type of information can be represented, stored, transmitted and processed digitally. Multimedia finds its application in various areas including, but not limited to, advertisements, art, education, entertainment, engineering, medicine, mathematics, business, scientific research and spatial temporal applications. Below are the several examples as follows:</p> <p>Entertainment and fine arts : Multimedia is heavily used in the entertainment industry, especially to develop special effects in movies and animations. Multimedia games are a popular pastime and are software programs available either as CD-ROMs or online. Some video games also use multimedia features. Multimedia applications that allow users to actively participate instead of just sitting by as passive recipients of information are called <i>Interactive Multimedia</i>. In the Arts there are multimedia artists, whose minds are able to blend techniques using different media that in some way incorporates interaction with the viewer. One of the most relevant could be Peter Greenaway who is melding Cinema with Opera and all sorts of digital media. Another approach entails the creation of multimedia that can be displayed in a traditional fine arts arena, such as an art gallery. For the most part these artists are using materials that will not hold up over time.</p> <p>Education : In Education, multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopaedia and almanacs. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats. Edutainment is an informal term used to describe combining education with entertainment, especially multimedia entertainment.</p> <p>Engineering : Software engineers may use multimedia in Computer Simulations for anything from entertainment to training such as military or industrial training. Multimedia for software interfaces are often done as collaboration between creative professionals and software engineers.</p> <p>Industry: In the Industrial sector, multimedia is used as a way to help present information to shareholders, superiors and coworkers. Multimedia is also helpful for providing employee training, advertising and selling products all over the world via virtually unlimited web-based technologies.</p> <p>Mathematical and Scientific Research: In Mathematical and Scientific Research, multimedia is mainly used for modelling and simulation. For example, a scientist can look at a molecular model of a particular substance and manipulate it to arrive at a new substance. Representative research can be found in journals such as the Journal of Multimedia.</p>	<p>2</p> <p>Any three points</p> <p>6</p>

Medicine: In Medicine, doctors can get trained by looking at a virtual surgery or they can simulate how the human body is affected by diseases spread by viruses and bacteria and then develop techniques to prevent it.

b. Briefly explain Gray-Level images and color images graphic/image data types. (8)

Answer:

Binary (1-bit) images : Binary images are encoded as a 2-D array, using one bit per pixel, where a 0 usually means ‘black’ and a 1 means ‘white’ (even though there is no universal agreement on that). The main advantage of this representation (Figure 1(b)) – usually suitable for images containing simple graphics, text or line art – is its small size.

Gray-level (8-bit) images: Gray-level (also referred to as *monochrome*) images are also encoded as a 2-D array of pixels, using eight bits per pixel, where a pixel value of 0 usually means ‘black’ and a pixel value of 255 means ‘white’, with intermediate values corresponding to varying shades of gray. The total number of gray-levels is larger than the human visual system requirements, making this format a good compromise between subjective visual quality and relatively compact representation and storage. An 8-bit monochrome image (Figure 1(a)) can also be thought of as a collection of bit-planes, where each plane contains a 1-bit representation of the image at different levels of detail.

Color images: Representation of color images is significantly more complex and varied. The two most common ways of storing color image contents are: *RGB* representation – in which each pixel is usually represented by a 24-bit number containing the amount of its Red®, Green (G), and Blue (B) components – and *indexed* representation – where a 2-D array contains indices to a color palette (or look-up table – LUT).

- **24-bit (RGB) color images:** Color images can be represented using three 2-D arrays of same size, one for each color channel: Red®, Green (G), and Blue (B) (Figure 2). Each array element contains an 8-bit value indicating the amount of red, green, or blue at that point, in a 0 to 255 scale. The combination of the three 8-bit values into a 24-bit number allows for 2²⁴ (16,777,216, usually referred to as 16 million or 16 M) color combinations. An alternative representation uses 32 bits and includes a fourth channel, called the *alpha channel*, which provides a measure of transparency for each pixel and is widely used in image editing effects.
- **Other color models :** The RGB color model is one of the most popular and straightforward methods for representing the contents of a digital color image, but there are several alternative models, such as: YCbCr – adopted in the JPEG standard – and the CMYK – used to encode color information for

4

4

printing purposes –, among many others.

Q.3 a. What do you understand by Gamma correction? Suppose images are not gamma corrected by a TV. Generally, how would they appear on a screen? (8)

Answer:

Gamma correction, gamma nonlinearity, gamma encoding, or often simply **gamma**, is the name of a nonlinear operation used to code and decode luminance or tristimulus values in video or still image systems. Gamma correction is, in the simplest cases, defined by the following **power-law** expression:

$$V_{\text{out}} = AV_{\text{in}}^{\gamma}$$

where A is a constant and the input and output values are non-negative real values; in the common case of $A = 1$, inputs and outputs are typically in the range 0–1. A gamma value $\gamma < 1$ is sometimes called an **encoding gamma**, and the process of encoding with this compressive power-law nonlinearity is called **gamma compression**; conversely a gamma value $\gamma > 1$ is called a **decoding gamma** and the application of the expansive power-law nonlinearity is called **gamma expansion**.

Gamma encoding of images is required to compensate for properties of human vision, to maximize the use of the bits or bandwidth relative to how humans perceive light and color. Human vision under common illumination conditions (not pitch black or blindingly bright) follows an approximate gamma or power function. If images are not gamma encoded, they allocate too many bits or too much bandwidth to highlights that humans cannot differentiate, and too few bits/bandwidth to shadow values that humans are sensitive to and would require more bits/bandwidth to maintain the same visual quality. Gamma encoding of floating point images is not required (and may be counterproductive) because the floating point format already provides a pseudo-logarithmic encoding.

A common misconception is that gamma encoding was developed to compensate for the input–output characteristic of cathode ray tube (CRT) displays. In CRT displays the electron-gun current, and thus light intensity, varies nonlinearly with the applied anode voltage. Altering the input signal by gamma compression can cancel this nonlinearity, such that the output picture has the intended luminance. However, the gamma characteristics of the display device do not play a factor in the gamma encoding of images and video — they need gamma encoding to maximize the visual quality of the signal, regardless of the gamma characteristics of the display device. The similarity of CRT physics to the inverse of gamma encoding needed for video transmission was a combination of luck and engineering which simplified the electronics in early television sets

If images are not gamma corrected by a TV, how would they appear on a

6

screen?

The gamma values on TV are the relationship between how bright a certain pixel should appear on the screen versus the actual numerical value that TV's panel enumerates. If the gamma is not correct on television, colors may seem overbearing or washed out depending on which way it is off. We can perform gamma correction on TV by synchronizing panel to target gammas, which can be found on the Internet.

2

b. What is video component of multimedia? What are the different types of video signal? Explain. (8)

Answer:

The embedding of video in multimedia applications is a powerful way to convey information which can incorporate a personal element which other media lack. Video enhances, dramatizes, and gives impact to multimedia application. Audience better understand the message of our application with the adequate and carefully planned integration of video. Video is an important way of conveying a message to the MTV generation. The advantage of integrating video into a multimedia presentation is the capacity to effectively convey a great deal of information in the least amount of time. Digital video can be divided into two main areas:

- That which has been captured/digitised from an analogue video source, e.g. VHS tape, camcorder, live video feed
- Animations created entirely in the digital domain, e.g. using software packages such as 3D Studio and Lightwave.

2

Video Signals can be classified in 3 categories as:

a) Composite Video

- Used in broadcast TV's
- Compatible with B/W TV
- Chrominance (I & Q or U & V) & Luminance signals are mixed into a single carrier wave, which can be separated at the receiver end
- Mixing of signals leads interference & create *dot crawl*

b) S-Video

- S stands *Super / Separated*
- **Uses 2 wires**, one for *luminance* & the other for *chrominance* signals

the best possible fidelity for a given bit-rate or minimizing the bit-rate to achieve a given fidelity measure but will not produce a complete facsimile of the original data.

Possible Lossless Methods:

- Zero Length Suppression
- Pattern Substitution
- Run Length Encoding
- Shano-Fannon Encoding
- Huffman Coding
- LZW/GIF Coding
- Arithmetic Coding

2

Possible Lossy Methods

- Difference Encoding/Quantisation
- Discrete Cosine Transform Coding
- Vector Quantisation
- JPEG Coding .

b. What do you understand by Huffman coding? For given a set of numbers and their frequency of use create a Huffman encoding for them: (8)

FREQUENCY	VALUE
-----	-----
5	1
7	2
10	3
15	4
20	5
45	6

Answer:

Huffman coding is an entropy encoding algorithm used for lossless data compression. The term refers to the use of a variable-length code table for encoding a source symbol (such as a character in a file) where the variable-length code table has been derived in a particular way based on the estimated probability of occurrence for each possible value of the source symbol. Huffman coding uses a specific method for choosing the representation for each symbol, resulting in a prefix code (sometimes called "prefix-free codes", that is, the bit string representing some particular symbol

3

is never a prefix of the bit string representing any other symbol) that expresses the most common source symbols using shorter strings of bits than are used for less common source symbols. For a set of symbols with a uniform probability distribution and a number of members which is a power of two, Huffman coding is equivalent to simple binary block encoding, e.g., ASCII coding.

To Create a Huffman tree we sort this list by frequency and make the two-lowest elements into leaves, creating a parent node with a frequency that is the sum of the two lower element's frequencies:

```

      12:*
     / \
    5:1 7:2

```

The two elements are removed from the list and the new parent node, with frequency 12, is inserted into the list by frequency. So now the list, sorted by frequency, is:

```

10:3
12:*
15:4
20:5
45:6

```

We then repeat the loop, combining the two lowest elements. This results in:

```

      22:*
     / \
    10:3 12:*
         / \
        5:1 7:2

```

and the list is now:

```

15:4
20:5
22:*
45:6

```

we repeat until there is only one element left in the list.

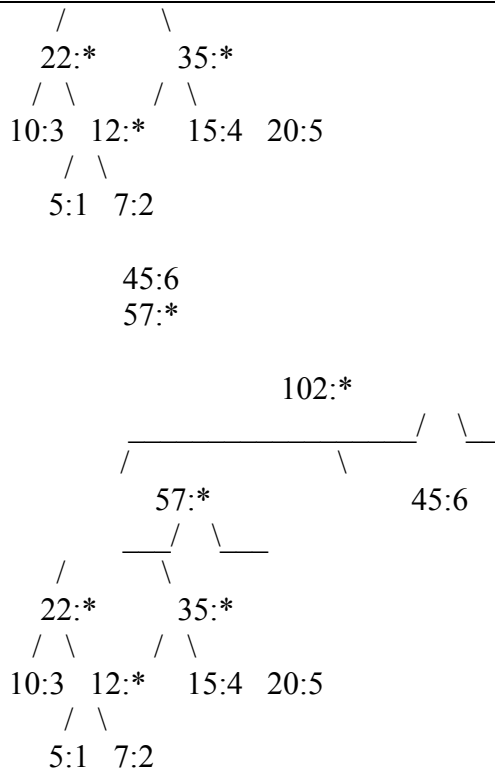
```

      35:*
     / \
    15:4 20:5

      22:*
      35:*
      45:6

      57:*
     ___/ \___

```



Now the list is just one element containing 102:*, we are done.

This element becomes the root of binary Huffman tree. To generate a Huffman code we traverse the tree to the value we want, outputting a **0** every time we take a left hand branch, and a **1** every time we take a right hand branch. (normally we traverse the tree backwards from the code we want and build the binary Huffman encoding string backwards as well, since the *first* bit must start from the top).

Q.5 a. When compression ratio is low for an image data using lossless compression technique? Also describe when Distortion measure used? (8)

Answer:

Compression ratio is low for an image data using lossless compression techniques such as Huffman coding arithmetic coding, LZW, when image histogram is relatively flat. 2

For image compression in multimedia applications where higher compression ratio is required lossy methods are usually adopted. In lossy compression the compressed image is usually not the same as the original image but it this meant to form a close approximation to the original image perpetually. To quantitatively describe how close the approximation is to the original data some form of distortion measure is required. 2

Distortion measure is a mathematical quantity that specifies how close an

more data is received, and image quality is gradually enhanced. The advantage is that the user-end has a choice whether to continue receiving image data after the first scan(s). Progressive JPEG can be realized in one of the following two ways. The main steps (DCT, quantization, etc.) are identical to those in Sequential Mode.

Q.6 a. What is the purpose of MPEG-2 standard? State its advantages. (8)

**Answer:
MPEG-2**

Development of the MPEG-2 standard started in 1990. Unlike MPEG-1, which is basically a standard for storing and playing video on the CD of a single computer at a low bit rate (1.5 Mbps), MPEG-2[7] is for higher-quality video at a bit rate of more than 4 Mbps. It was initially developed as a standard for digital broadcast TV. MPEG-2 has managed to meet the compression and bit rate requirements of digital TV/HDTV and in fact supersedes a separate standard, MPEG-3, initially thought necessary for HDTV. MPEG-2 has gained wide acceptance beyond broadcasting digital TV over terrestrial, satellite, or cable networks. Among various applications such as Interactive TV It is also adopted for digital video discs or digital versatile discs (DVDs). MPEG-2 defined seven profiles aimed at different applications (e.g., low-delay videoconferencing, scalable video, HDTV). The profiles are Simple, Main, SNR scalable, and spatially scalable, high, 4:2:2 and Multiview (where two views would refer to stereoscopic video). Within each profile, up to four levels are defined as table shows, not all profiles have four levels. For example, the Simple profile has only the Min level; whereas the High profile does not have the Low level.

Level	Maximum resolution	Maximum fps	Maximum pixels/sec	Maximum coded data rate (Mbps)	Application
High	$1,920 \times 1,152$	60	62.7×10^6	80	Film production
High 1440	$1,440 \times 1,152$	60	47.0×10^6	60	Consumer HDTV
Main	720×576	30	10.4×10^6	15	Studio TV
Low	352×288	30	3.0×10^6	4	Consumer tape equivalent

The advantages of MPEG-2 are:

- **Supporting interlaced Video:** MPEG-1 is adopted by digital broadcast TV; it must also support interlaced video, because this is one of the options for digital broadcast TV and HDTV. MPEG-2 defines frame prediction and field prediction as well as five different prediction modes, suitable for a wide range of applications where the requirement for the accuracy and speed of motion compensation vary.
- **MPEG-2 Scalabilities:** Mpeg-2 scalable coding is also known as layered coding, in which a base layer and one or more enhancement layers can be defined. MPEG-2 supports the following scalabilities:
 - **SNR scalability:** The enhancement layer provides higher SNR.
 - **Spatial scalability:** The enhancement layer provides higher spatial resolution.
 - **Temporal scalability:** The enhancement layer facilitates higher frame rate.
 - **Hybrid scalability:** This combines any two of the above three scalabilities.
 - **Data partitioning:** Quantized DCT coefficients are split into partitions.

4

b. Differentiate between I-Frame and P-Frame and their coding. (8)

Answer:

In the field of video compression a video frame is compressed using different algorithms. These different algorithms for video frames are called **picture types** or **frame types**. The major picture types used in the different video algorithms are **I** and **P**. They are different in the following characteristics:

An **I-frame** is an 'Intra-coded picture', in effect a fully specified picture, like a conventional static image file: that is it is treated as independent image. P-frames hold only part of the image information, so they need less space to store than an I-frame, and thus improve video compression rates. **I**-frames are the least compressible but don't require other video frames to decode. I-frames coding performs only spatial redundancy removal

4

A **P-frame** ('Predicted picture') are not independent holds only the changes in the image from the previous frame. They are coded by forward predictive coding method. For example, in a scene where a car moves across a stationary background, only the car's movements need to be encoded. The encoder does not need to store the unchanging background pixels in the P-frame, thus saving space. **P**-frames can use data from previous frames to decompress and are more

4

compressible than I-frames. Temporal redundancy removal is included in P-frame coding.

Q.7 a. What do you understand by VOP-based coding? Differentiate VOP-based coding with frame based coding. (8)

Answer:

VOP-based Coding

- MPEG-4 VOP-based coding also employs the Motion Compensation technique:
 - An Intra-frame coded VOP is called an I-VOP.
 - The Inter-frame coded VOPs are called *P-VOPs* if only forward prediction is employed, or *B-VOPs* if bi-directional predictions are employed.
- The new difficulty for VOPs: may have arbitrary shapes, shape information must be coded in addition to the texture of the VOP. The *texture* here actually refers to the visual content, that is the gray-level (or chroma) values of the pixels in the VOP

4

VOP-based vs. Frame-based Coding

- MPEG-1 and -2 do not support the VOP concept, and hence their coding method is referred to as frame-based (also known as Block-based coding).
- A possible example in which both potential matches yield small prediction errors for block-based coding.
- Each VOP is of arbitrary shape and ideally will obtain a unique motion vector consistent with the actual object motion.

4

b. What is the significance of MPEG-4 standard? Also explain its hierarchical structure. (8)

Answer:

MPEG-4: a newer standard. Besides compression, pays great attention to issues about user interactivities. MPEG-4 departs from its predecessors in adopting a new object-based coding:

- Offering higher compression ratio, also beneficial for digital video composition, manipulation, indexing, and retrieval.
- MPEG-4 videos can be composed and manipulated by simple operations on the visual objects.
- The bit-rate for MPEG-4 video now covers a large range between 5 kbps to 10 Mbps.

2

MPEG-4 is an entirely new standard for:

- Composing media objects to create desirable audiovisual scenes.
- Multiplexing and synchronizing the bit streams for these media data entities so that they can be transmitted with guaranteed Quality of Service (QoS).
- Interacting with the audiovisual scene at the receiving end — provides a toolbox of advanced coding modules and algorithms for audio and video compressions.

The hierarchical structure of MPEG-4 visual bit streams is very different from that of MPEG-1 and -2, it is very much video object-oriented

- a) Video-object Sequence (VS)—delivers the complete MPEG-4 visual scene, which may contain 2-D or 3-D natural or synthetic objects.
- b) Video Object (VO) — a particular object in the scene, which can be of arbitrary (non-rectangular) shape corresponding to an object or background of the scene.
- c) Video Object Layer (VOL) — facilitates a way to support (multi-layered) scalable coding. A VO can have multiple VOLs under scalable coding, or have a single VOL under non-scalable coding.
- d) Group of Video Object Planes (GOV) — groups Video Object Planes together (optional level).
- e) Video Object Plane (VOP) — a snapshot of a VO at a particular moment.

6

Q.8 a. What do you understand by MP3? Briefly explain three MPEG layers. (8)

Answer:

MP3 Stands for "MPEG-1 Audio Layer-3." MP3 is popular compressed audio file format that helped popularize digital music downloads beginning in the late 1990s. MP3 files are typically about one tenth the size of uncompressed WAVE or AIFF files, but maintain nearly the same CD-quality sound. Because of their small size and good fidelity, MP3 files have become a popular way to store music files on both computers and portable devices like the iPod.

To listen to MP3s on computer, we need an MP3 player like Nullsoft Winamp (for Windows) or Apple iTunes (for Mac and Windows). Most MP3 players also allow us to create MP3 files from CD audio tracks or other from other audio file types. Once we have converted our favorite songs to MP3 files, we can transfer them to a portable music player, like the Apple iPod, Microsoft Zune, or a music-enabled cell phone. We can also burn the MP3 files to a CD, which can be played in MP3-

2

compatible CD players.

MPEG layers :Layer 1 to 3 in MPEG Audio are compatible, because all layers include the same header file information.

- Layer 1 quality can be good provided a comparatively high bitrate is available. Digital Audio tape uses Layer1.
- Layer 2 has more complexity and was proposed for use in digital audio broadcasting.
- Layer 3 (MP3) is most complex and was originally aimed at transmission over ISDN lines.

b. State characteristics of multimedia data due to which challenges arise in multimedia network communication. (8)

Answer:

Challenges arise in multimedia network communication because of following factors:

- Multimedia systems must be *computer controlled*.
- Multimedia systems are *integrated*.
- The information they handle must be represented *digitally*. It is integrated and continuous media
- It is voluminous. It demands high data rates, possibly hundreds of MBPS.
- It is Real time and interactive. It demands low delay and synchronization between audio and video. In addition applications such as videoconferencing and interactive multimedia requires two-way traffic.
- It is sometimes bursty. Data rate fluctuate drastically for example, in video-on demand, no traffic most of the time but burst to high volume

Q.9 a. What are the various techniques of animation in multimedia? (8)

Answer:

Computer animation encompasses a variety of techniques on a computer. Figures are created and/or edited on the computer using 2D bitmap graphics or created and edited using 2D vector graphics. Animation is the rapid display of a sequence of images of 2-D artwork or model positions in order to create an illusion of movement. It is an optical illusion of motion due to the phenomenon of persistence of vision. Various techniques of **traditional Animation** and **computer animation** are:-Full Animation, Rotoscoping, -Clay/Stop -motion Animation (gumby) -Cut out animation (southpark). **Computer Animation** (computer assisted animation) encompasses a variety of techniques, the unifying factor being that the

**6
(2M
Each)**

**Any 4
points
explan
ation 2
M each
= 8 M**

4

animation is created digitally on a computer. This animation takes less time than previous traditional animation.-2d animation and 3d animation.

2D animation : In this figures are created and/or edited on the computer using 2D bitmap or created and edited using 2Dvector graphic . This includes automated computerized versions of traditional animation techniques such as of, interpolated morphing, onion skinning and interpolated rotoscoping.

3D animation: is digitally modeled and manipulated by an animator. To manipulate a mesh, it is given a digital skeletal structure that can be used to control the mesh. This process is called rigging. Various other techniques can be applied, such as mathematical functions (ex. gravity, particle simulations), simulated fur or hair, effects such as fire and water and the use of motion capture to name but a few, these techniques fall under the category of 3D dynamics. Well-made 3D animations can be difficult to distinguish from live action and are commonly used as visual effects for movies.

4

b. What is the use of DVD in Multimedia? Explain DVD+R and DVD+RW formats of DVD in multimedia. (8)

Answer:

When DVD technology first appeared in households, users were simply watching movies.. But just as CD technology evolved so that users could record and erase and re-record data onto compact discs, the same is now true of **DVDs**. There are so many different DVD formats — DVD+R, DVD+RW, DVD-RAM, DVD-R, DVD-RW, DVD-ROM.

2

DVD+R and DVD+RW

DVD+R and DVD+RW formats are supported by Philips, Sony, Hewlett-Packard, Dell, Ricoh, Yamaha and others.DVD+R is a recordable DVD format similar to CD-R. A **DVD+R** can record data only once and then the data becomes permanent on the disc. The disc cannot be recorded onto a second time. **DVD+RW** is a re-recordable format similar to CD-RW. The data on a DVD+RW disc can be erased and recorded over numerous times without damaging the medium.

6 (3 M Each)

Note: DVDs that have been made using a +R/+RW device can be read by most commercial DVD-ROM players. The crucial difference among the standards is based on which standards each manufacturer adheres to as different manufacturers support different standards. The different variations on the term DVD (e.g. +R, -R, -ROM, and so on) describe the way data is stored on or written to the disc itself. These are called physical formats.

TEXT BOOK

- I. Fundamentals of Multimedia, Ze-Nian Li and Mark S. Drew, Pentice Hall, Edition – 2007
- II. Principles of Multimedia, Ranjan Parekh, Tata McGraw-Hill, Edition 2006