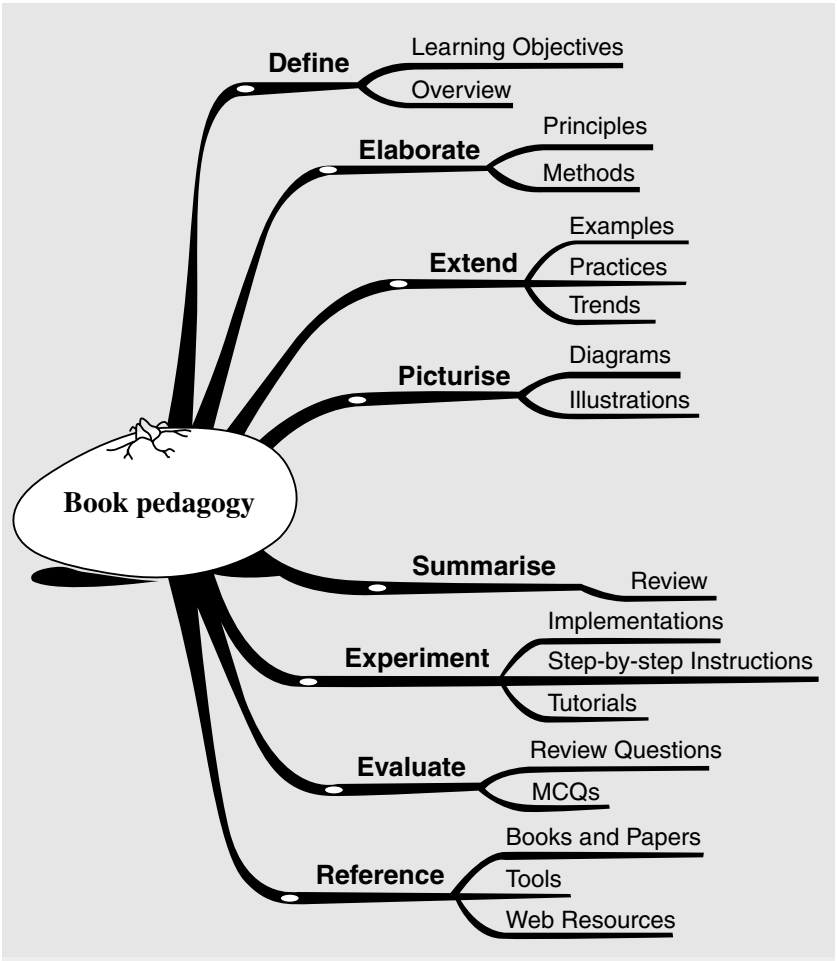
KNOW YOUR BOOK VISUALLY

Fundamentals of Multimedia Technology
DEEP SEER X model!



eXpand: Online Learning Centre for Students and Faculty - http://www.mhhe.com/banerji/fmt

Each chapter begins with learning objectives and an overview.
Objectives *provide a concise statement of expected learning outcomes.* ***Overview*** *gives a brief summary of the concepts discussed and their relevance.*

DEFINE

DIGITAL AUDIO

CHAPTER 8

OBJECTIVES

After reading this chapter and completing the learning activities for it, you should be able to:
- Describe advantages of representing sound in digital form.
- Elaborate the process of producing digital audio.
- Describe the principle of psychoacoustics and its applications.
- Describe different formats of audio files.
- Demonstrate and describe the basic audio recording and editing process.

8.1 OVERVIEW

In the previous three chapters of this book (Chapters 5, 6 and 7) we discussed text, pictures and graphics. These are the media for two of our principal channels of communication in the visual mode. Although, text and graphics are the predominant modes of human-computer communication, the media through which human—human communication started was audio. Drawings and pictures were the main vehicles of human communication before written languages. But even earlier, it was sound through which we started to communicate naturally. Sound is an auditory impression produced by sensations perceived by our ears. We are constantly immersed in the world of sound in the form of speech, music and other naturally produced sounds (including noise).

It is obvious, therefore, why audio is an important component of multimedia. It is a crucial element for the entertainment industry. Sounds are used to enhance multimedia games through background music and sound effects. Sounds can be combined with other multimedia components to create moods, enhance understanding, and reinforce concepts. The scopes of digital audio applications are many. Some major application areas are: computer generated sound, sound storage and processing, digital communications, answering service, speech synthesis, speech recognition, computerised call centre, and presentation of data as sound. If nothing else, sound can enable programs to have a more user friendly interface, with the ability to have a person talk to the user helping them to use a program. Basically, the study of digital audio involves the processes of generation, propagation, amplification and transformation, all of which lead to the understanding of information (Fig. 8.1). This chapter will provide the foundation concepts and theoretical background for the representation, storage and processing of digital audio for use in multimedia applications.

Many commercial organisations are exploring the Internet and intranet as a delivery channel for multimedia information and services, and are offering multimedia services on the World Wide Web (WWW). Electronic commerce (called e-commerce) via multimedia information systems on the Internet has become a reality. For example, see Amazon.com (*www.amazon.com*). Apple computer's virtual music store iTunes has already become a phenomenon due to iPods, its MP3 music player.

Future multimedia technologies include many innovative systems. For example, technologies like interactive television in which the television will become a two-way communication device by adding a 'set-top-box'. This will allow the telecaster to build direct links to clients. Potential customers can be contacted by via a TV, which will, in fact, be a computer. The possibilities are fascinating.

2.4 APPLICATION EXAMPLES

The real impact of multimedia is achieved when the application goes beyond textual display and effectively integrates sound, animation and graphics. For example, animation can clearly demonstrate the

4.5.4 Discrete Cosine Transform

The Discrete Cosine Transform (DCT) is a key mathematical transformation method for powerful graphics and image compression. Therefore, it is discussed in detail here. DCT falls in the class of mathematical operations that include well-known Fourier transform and Fast Fourier transform (FFT). The basic operation performed by such mathematical functions is to transform the data into a frequency domain. For example, when we take a set of samples from an incoming audio signal we end up with a representation

Since the 2D DCT can be computed by applying 1D transforms separately to the rows and then the columns, we can say that the 2D DCT is separable in the two dimensions. We can use the following equation to get the basis matrix:

$$C = k * \cos((2m + 1)\, n\pi / 2N)$$

where

$$k = \sqrt{(1/N)} \text{ if } n = 0$$

$$k = \sqrt{(2/N)} \text{ otherwise}$$

$m, n = 0, 1, 2, \ldots (N - 1)$.

Thus, a two dimensional DCT is defined as:

$$Y = C * X * C'$$

Where X is an N N image block (represent the pixels of a given image, $\mathbf{A}_{mn}$), Y contains the computed $N \times N$ DCT coefficients (represent the frequency domain coefficients, $\mathbf{B}_{pq}$), C is an $N \times N$ matrix and C' is the transpose of matrix C. Here, matrix X is first multiplied on the left by the basis matrix C. This transforms the rows. The columns are then transformed by multiplying on the right by the transpose of the C. This block matrix now consists of N^2 DCT coefficients. The top-left coefficient correlates to the low frequencies of the original image block. As we move away from this in all directions, the DCT coefficients correlate to higher and higher frequencies of the image block. It is important to note that the human eye

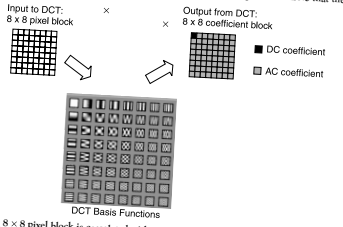
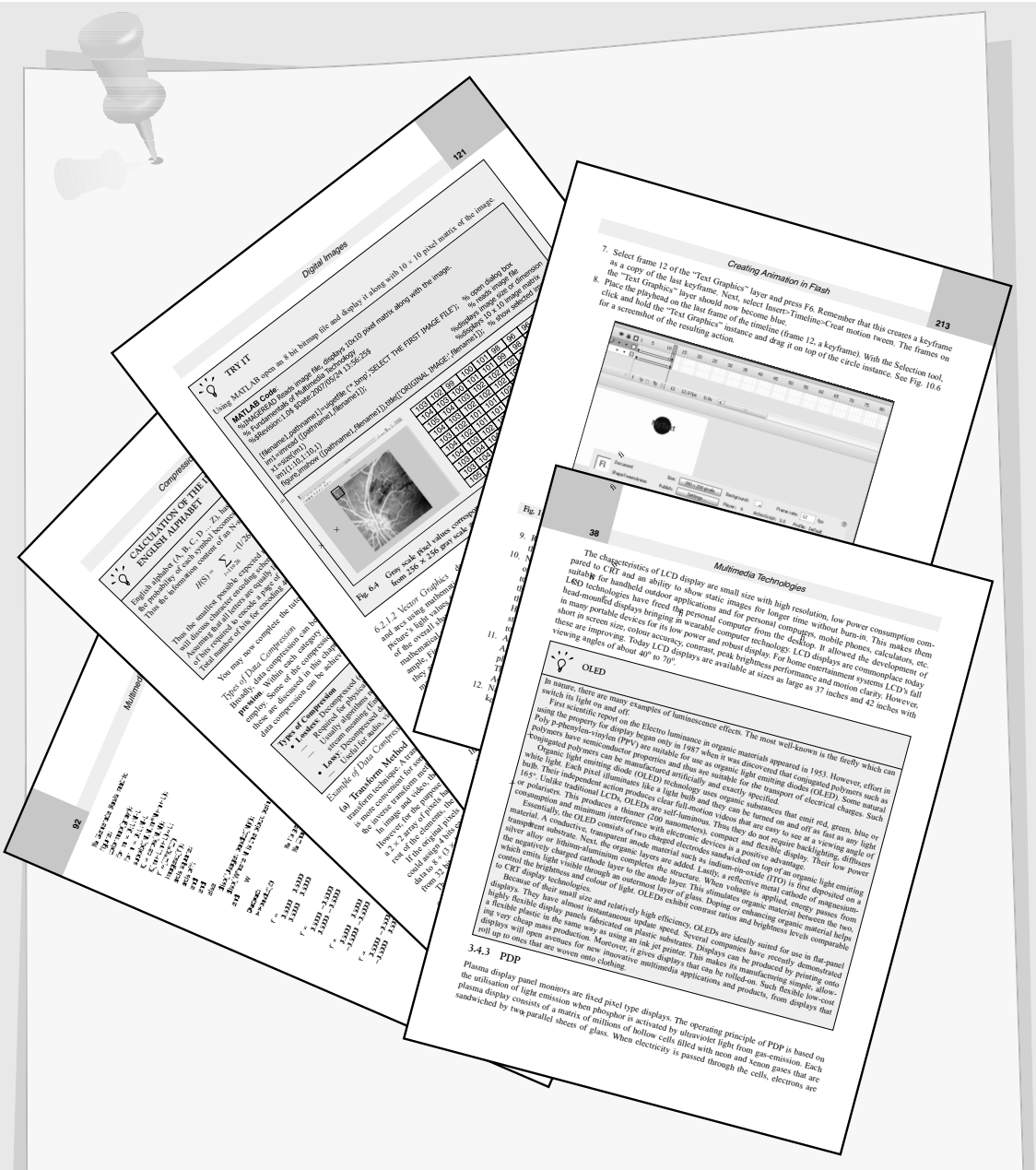


Fig. 4.7 Each 8 × 8 pixel block is correlated with 64 DCT basis functions. The result from each correlation produces an output coefficient.

To support concept building, each chapter has been neatly divided into ***sections*** *and* ***sub-sections*** *so that the subject matter is studied in a logical progression of ideas and concepts.*
Concept building is strengthened by describing Applications and elaborating Principles and Methods.

ELABORATE



Adequate features are provided to contextualise concept building and help in understanding by experimentation. These include: implementations, worked examples, exploration and step-by-step instructions. Total ***45 worked examples*** *of implementations are provided in the book.*
Step-by-step instructions are provided to guide the students in developing experimental applications.
Latest technologies, trends and research directions are also provided in appropriate sections.

EXTEND

*Ample illustrations (**123 figures** and **28 tables**) are used in the book to explain the complex concepts, functional relationships, and provide sketches for algorithmic models.*

PICTURISE

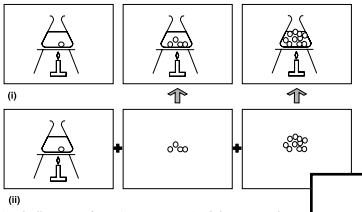


Fig. 9.5 Simple illustration of MPEG compression. (i) If shown in quick succe three images will produce animation of boiling water in a beaker. (ii) Sa transferred by the 'I' frame (key frame) and the difference information b frames (B frames). Intra frame redundancies are removed

between adjacent pictures. These are shown in Fig. 9.5(ii). Illusion of motion is difference between the frames, coupled with our persistence of vision. Figure 9.5(i sible to create the second and third frame simply by transmitting the differences. be constructed from the first frame by adding only the extra objects. Similarly, th be constructed from the first frame. In both cases, the data load for the subsequen is greatly reduced.

At the cost of higher complexity, the result of applying MPEG video compress of data transmitted across the network is less than that of Motion JPEG. This is ill where only information about the differences in the second and third frames is t the key frame data.

MPEG is in fact far more complex than that indicated above. Parameters such a in a scene and identifying objects are additional techniques or tools used within ent applications can make use of different tools, for example: a real-time surveill animated movie.

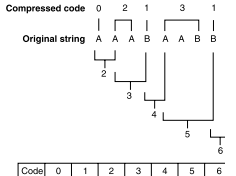


Fig. 4.4 LZW compression explained diagrammatically

Originally, the size of the string AAABAABBB is 9 × 8 bits (size of ASCII) = 72 bits.

For the compressed string, if we define the code size as 4 bits, the size of the compressed string 0 2 1 3 1 1 will be 6 × 4 bits = 24 bits. Thus with this simple method, we can achieve a compression ratio of 3:1.

*A **review**, at the end of each chapter, gives the essence of the topics covered and will be helpful for a quick summary.*

SUMMARISE

compression techniques including DCT and DWT transforms, for different partitions as well as at different levels of encoding and decoding can be incorporated.

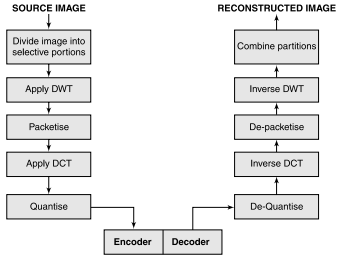


Fig. 4.10 Block Diagram of a Hybrid Compression Scheme

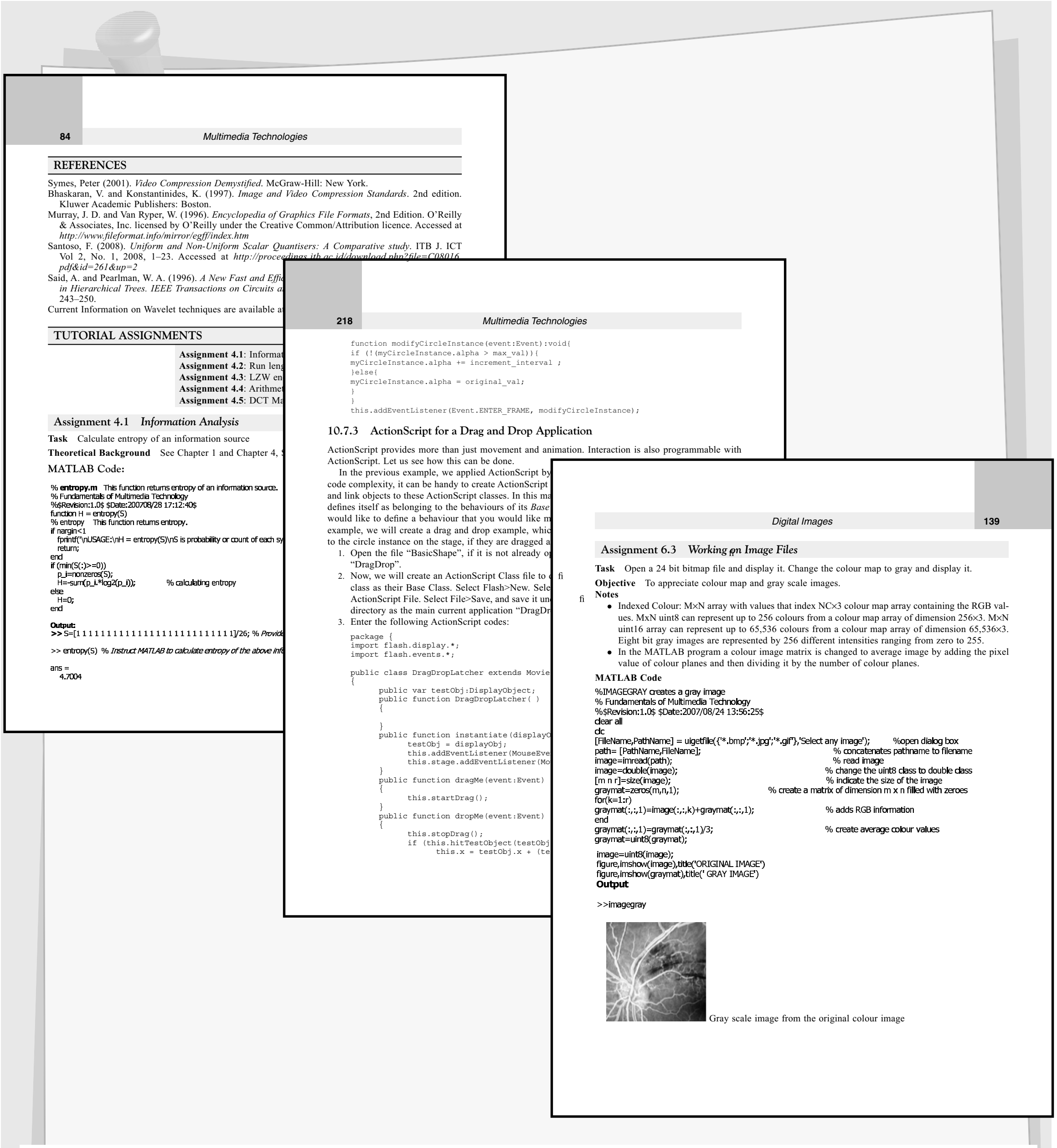
4.6 REVIEW

In this chapter we covered some of the basic techniques of data compression that are relevant for multimedia. First, we examined the issue of 'why compression is required'. Then we moved on to understand the fundamentals of compression and the possible approaches. There are two broad classes of data compression techniques—lossless and lossy compression. Within each class various methods are available. In order to understand the basis of compression we examined some of the techniques and their examples. These include **lossless**—Run length encoding, dictionary based LZW compression, variable length based Huffman and Arithmetic Coding; **lossy**—DCT, Wavelet and DWT transformation. The importance of using hybrid compression techniques for remote sensing, medical image processing and integration has also been briefly mentioned.

The subject of data compression is huge and extremely interesting. By no way we could have discussed all of them within the confines of this book. Nevertheless, after reading this chapter you will gain a fairly good foundation. With the appreciation of the broader framework of multimedia data compression, we are now ready to discuss the various multimedia components in the subsequent chapters.

REVIEW QUESTIONS

1. Why data compression is needed for multimedia components like video and audio, images and texts?



84 Multimedia Technologies

REFERENCES

Symes, Peter (2001). *Video Compression Demystified*. McGraw-Hill: New York.
Bhaskaran, V. and Konstantinides, K. (1997). *Image and Video Compression Standards*. 2nd edition. Kluwer Academic Publishers: Boston.
Murray, J. D. and Van Ryper, W. (1996). *Encyclopedia of Graphics File Formats*, 2nd Edition. O'Reilly & Associates, Inc. licensed by O'Reilly under the Creative Common/Attribution licence. Accessed at *http://www.fileformat.info/mirror/egff/index.htm*
Santoso, F. (2008). *Uniform and Non-Uniform Scalar Quantisers: A Comparative study*. ITB J. ICT Vol 2, No. 1, 2008, 1–23. Accessed at *http://proceedings.itb.ac.id/download.php?file=C08016.pdf&id=261&up=2*
Said, A. and Pearlman, W. A. (1996). *A New Fast and Effic... in Hierarchical Trees. IEEE Transactions on Circuits a...* 243–250.
Current Information on Wavelet techniques are available a...

TUTORIAL ASSIGNMENTS

Assignment 4.1: Informat...
Assignment 4.2: Run leng...
Assignment 4.3: LZW en...
Assignment 4.4: Arithmet...
Assignment 4.5: DCT Ma...

Assignment 4.1 *Information Analysis*

Task Calculate entropy of an information source

Theoretical Background See Chapter 1 and Chapter 4, ...

MATLAB Code:

```
% entropy.m   This function returns entropy of an information source.
% Fundamentals of Multimedia Technology
%$Revision:1.0$ $Date:2007/08/28 17:12:40$
function H = entropy(S)
% entropy    This function returns entropy.
if nargin<1
   fprintf('\nUSAGE:\nH = entropy(S)\nS is probability or count of each sy...
   return;
end
if (min(S(:)>=0))
   p_i=nonzeros(S);
   H=-sum(p_i.*log2(p_i));            % calculating entropy
else
   H=0;
end
```

Output:

```
>> S=[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1]/26; % Provid...

>> entropy(S)  % Instruct MATLAB to calculate entropy of the above inf...

ans =
   4.7004
```

218 Multimedia Technologies

```
function modifyCircleInstance(event:Event):void{
if (!(myCircleInstance.alpha > max_val)){
myCircleInstance.alpha += increment_interval ;
}else{
myCircleInstance.alpha = original_val;
}
}
this.addEventListener(Event.ENTER_FRAME, modifyCircleInstance);
```

10.7.3 ActionScript for a Drag and Drop Application

ActionScript provides more than just movement and animation. Interaction is also programmable with ActionScript. Let us see how this can be done.

In the previous example, we applied ActionScript by... code complexity, it can be handy to create ActionScript... and link objects to these ActionScript classes. In this ma... defines itself as belonging to the behaviours of its *Base*... would like to define a behaviour that you would like m... example, we will create a drag and drop example, whic... to the circle instance on the stage, if they are dragged a...

1. Open the file "BasicShape", if it is not already o... "DragDrop".
2. Now, we will create an ActionScript Class file to ... class as their Base Class. Select Flash>New. Sele... ActionScript File. Select File>Save, and save it un... directory as the main current application "DragDr...
3. Enter the following ActionScript codes:

```
package {
import flash.display.*;
import flash.events.*;

public class DragDropLatcher extends Movie...
{
      public var testObj:DisplayObject;
      public function DragDropLatcher( )
      {

      }
      public function instantiate(displayO...
            testObj = displayObj;
            this.addEventListener(MouseEve...
            this.stage.addEventListener(Mo...
      }
      public function dragMe(event:Event)
      {
            this.startDrag();
      }
      public function dropMe(event:Event)
      {
            this.stopDrag();
            if (this.hitTestObject(testObj...
                  this.x = testObj.x + (te...
```

Digital Images 139

Assignment 6.3 *Working on Image Files*

Task Open a 24 bit bitmap file and display it. Change the colour map to gray and display it.

Objective To appreciate colour map and gray scale images.

Notes

- Indexed Colour: M×N array with values that index NC×3 colour map array containing the RGB values. MxN uint8 can represent up to 256 colours from a colour map array of dimension 256×3. M×N uint16 array can represent up to 65,536 colours from a colour map array of dimension 65,536×3. Eight bit gray images are represented by 256 different intensities ranging from zero to 255.
- In the MATLAB program a colour image matrix is changed to average image by adding the pixel value of colour planes and then dividing it by the number of colour planes.

MATLAB Code

```
%IMAGEGRAY creates a gray image
% Fundamentals of Multimedia Technology
%$Revision:1.0$ $Date:2007/08/24 13:56:25$
clear all
clc
[FileName,PathName] = uigetfile({'*.bmp';'*.jpg';'*.gif'},'Select any image');      %open dialog box
path= [PathName,FileName];                          % concatenates pathname to filename
image=imread(path);                                 % read image
image=double(image);                                % change the uint8 class to double class
[m n r]=size(image);                                % indicate the size of the image
graymat=zeros(m,n,1);                % create a matrix of dimension m x n filled with zeroes
for(k=1:r)
graymat(:,:,1)=image(:,:,k)+graymat(:,:,1);         % adds RGB information
end
graymat(:,:,1)=graymat(:,:,1)/3;                    % create average colour values
graymat=uint8(graymat);

image=uint8(image);
figure,imshow(image),title('ORIGINAL IMAGE')
figure,imshow(graymat),title(' GRAY IMAGE')
```

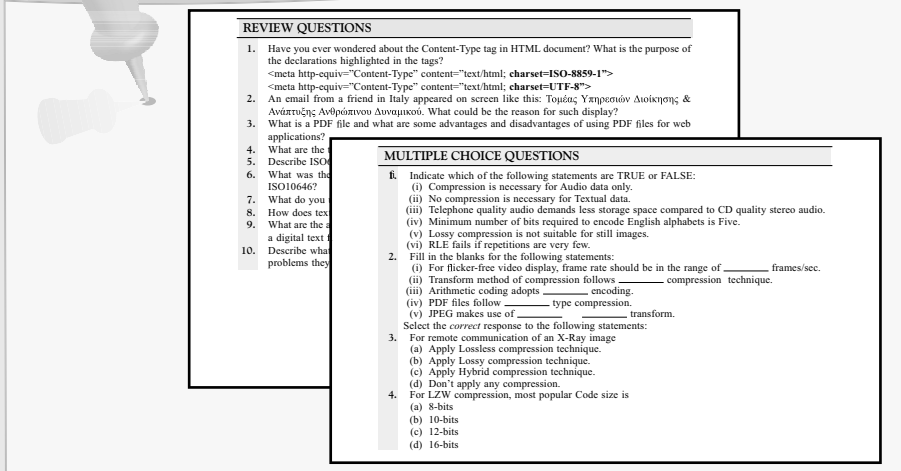
Output

```
>>imagegray
```

Gray scale image from the original colour image

*Each chapter contains a set of problems to practice in laboratory, totalling to **37 tutorial assignments** in the book. These will complement learning through hands-on lab experiments and problem solving.*
Solution to the tutorials requires not only application of the material covered in the book but also enables the student to strive towards good comprehension of the subject matter. Answers, steps or hints have been provided for all the problem sets as appropriate.

EXPERIMENT



REVIEW QUESTIONS

1. Have you ever wondered about the Content-Type tag in HTML document? What is the purpose of the declarations highlighted in the tags?
   <meta http-equiv="Content-Type" content="text/html; **charset=ISO-8859-1"**>
   <meta http-equiv="Content-Type" content="text/html; **charset=UTF-8"**>
2. An email from a friend in Italy appeared on screen like this: Τομέας Υπηρεσιών Διοίκησης & Ανάπτυξης Ανθρώπινου Δυναμικού. What could be the reason for such display?
3. What is a PDF file and what are some advantages and disadvantages of using PDF files for web applications?
4. What are the
5. Describe ISO
6. What was the ISO10646?
7. What do you
8. How does tex
9. What are the a a digital text f
10. Describe what problems they

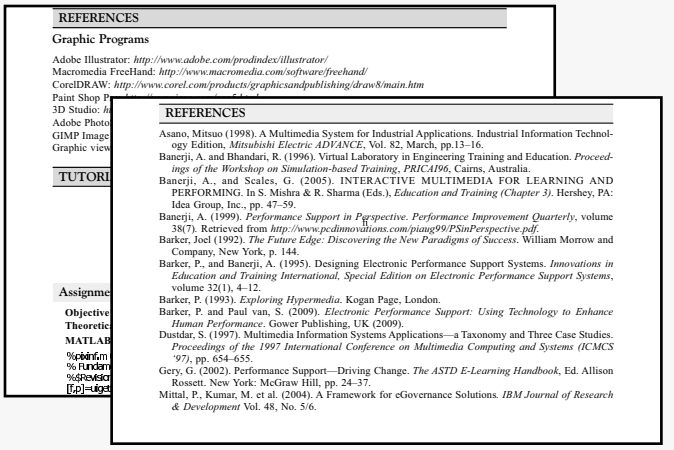
MULTIPLE CHOICE QUESTIONS

fi. Indicate which of the following statements are TRUE or FALSE:
   (i) Compression is necessary for Audio data only.
   (ii) No compression is necessary for Textual data.
   (iii) Telephone quality audio demands less storage space compared to CD quality stereo audio.
   (iv) Minimum number of bits required to encode English alphabets is Five.
   (v) Lossy compression is not suitable for still images.
   (vi) RLE fails if repetitions are very few.
2. Fill in the blanks for the following statements:
   (i) For flicker-free video display, frame rate should be in the range of _____ frames/sec.
   (ii) Transform method of compression follows _____ compression technique.
   (iii) Arithmetic coding adopts _____ encoding.
   (iv) PDF files follow _____ type compression.
   (v) JPEG makes use of _____ transform.

   Select the *correct* response to the following statements:
3. For remote communication of an X-Ray image
   (a) Apply Lossless compression technique.
   (b) Apply Lossy compression technique.
   (c) Apply Hybrid compression technique.
   (d) Don't apply any compression.
4. For LZW compression, most popular Code size is
   (a) 8-bits
   (b) 10-bits
   (c) 12-bits
   (d) 16-bits

*To evaluate learning, each chapter contains Review Questions for the students to respond and reflect. More than **170 review questions** in the text are included to hone the problem-solving skills.*
*In addition, each chapter contains a set of Objective Questions, totalling to **128 multiple choice questions** in the book. This enables the user to obtain clear comprehension of the subject matter.*

EVALUATE



REFERENCES

Graphic Programs

Adobe Illustrator: *http://www.adobe.com/prodindex/illustrator/*
Macromedia FreeHand: *http://www.macromedia.com/software/freehand/*
CorelDRAW: *http://www.corel.com/products/graphicsandpublishing/draw8/main.htm*
Paint Shop P
3D Studio: *h*
Adobe Photo
GIMP Image
Graphic view

TUTORI

Assignme

Objective
Theoretic
MATLAB

```
%pbinfm
% Fundam
%%Revisio
[F,p]=uige
```

REFERENCES

Asano, Mitsuo (1998). A Multimedia System for Industrial Applications. Industrial Information Technology Edition, *Mitsubishi Electric ADVANCE*, Vol. 82, March, pp.13–16.

Banerji, A. and Bhandari, R. (1996). Virtual Laboratory in Engineering Training and Education. *Proceedings of the Workshop on Simulation-based Training*, *PRICAI96*, Cairns, Australia.

Banerji, A., and Scales, G. (2005). INTERACTIVE MULTIMEDIA FOR LEARNING AND PERFORMING. In S. Mishra & R. Sharma (Eds.), *Education and Training (Chapter 3)*. Hershey, PA: Idea Group, Inc., pp. 47–59.

Banerji, A. (1999). *Performance Support in Perspective. Performance Improvement Quarterly*, volume 38(7). Retrieved from *http://www.pcdinnovations.com/piaug99/PSinPerspective.pdf*.

Barker, Joel (1992). *The Future Edge: Discovering the New Paradigms of Success*. William Morrow and Company, New York, p. 144.

Barker, P., and Banerji, A. (1995). Designing Electronic Performance Support Systems. *Innovations in Education and Training International, Special Edition on Electronic Performance Support Systems*, volume 32(1), 4–12.

Barker, P. (1993). *Exploring Hypermedia*. Kogan Page, London.

Barker, P. and Paul van, S. (2009). *Electronic Performance Support: Using Technology to Enhance Human Performance*. Gower Publishing, UK (2009).

Dustdar, S. (1997). Multimedia Information Systems Applications—a Taxonomy and Three Case Studies. *Proceedings of the 1997 International Conference on Multimedia Computing and Systems (ICMCS '97)*, pp. 654–655.

Gery, G. (2002). Performance Support—Driving Change. *The ASTD E-Learning Handbook*, Ed. Allison Rossett. New York: McGraw Hill, pp. 24–37.

Mittal, P., Kumar, M. et al. (2004). A Framework for eGovernance Solutions. *IBM Journal of Research & Development* Vol. 48, No. 5/6.

*At the end of each chapter, a comprehensive list of **references** is provided to support further studies.*
Relevant web addresses of different multimedia tools and useful websites are also provided.

REFERENCES