

8

WAVELET & MULTI-RESOLUTION PROCESSING

8.1 WHY WAVELET?

A time signal $\chi(t)$ contains complete information in time domain, i.e., the amplitude of the signal at any given moment t . However, no information is explicitly available in $\chi(t)$ regarding the frequency contents of the signal. On the other hand, as the spectrum $X(f)$ obtained by Fourier transform of the time signal $\chi(t)$ is extracted from the entire time duration of the signal, it contains complete information in the frequency domain in terms of the magnitude and phase angle of any frequency component f , but there is no information explicitly available in the spectrum regarding the temporal characteristics of the signal such as where in time certain frequency component appeared. Neither $\chi(t)$ in time domain nor $X(f)$ in frequency domain provides complete description of the signal.

To address this dilemma, we can truncate the signal $\chi(t)$ by a time window $\omega(t-\tau)$

$$\chi'(t) = \omega(t-\tau)\chi(t)$$

where the time window has width T and can shift in time according to τ :

$$\omega(t) = \begin{cases} 1 & 0 < t < T \\ 0 & \text{otherwise} \end{cases}$$

The spectrum of this windowed signal is for the specific time period T only. One immediate drawback of this windowed signal is the severe distortion in frequency domain caused by the sudden truncation in time domain. To reduce this distortion, the square time window can be replaced by a smooth Gaussian bell-shaped window with gradual decay:

$$\chi'(t) = g(t)\chi(t), \quad \text{where } g(t) = c \exp\left(-(t-\tau)^2/\sigma^2\right)$$

The Fourier transform of this Gaussian filtered time signal is called the Gabor transform. Although the spectral property of a windowed signal can be better localized in time, its resolution in frequency is reduced as its spectrum is blurred by the convolution

< 12 December, 2020

13:19

Wavelet and Multi-Resolution Processing

191

8.2 HAAR WAVELETS

A continuous signal can be approximated by a sequence of unit impulse functions, also called scaling functions, weighted by the sampling values of the intensity or amplitude of the signal:

$$\chi(t) = \sum_i s_i \phi_{[t_i, t_{i+1}]}(t)$$

where $\phi_{[t_i, t_{i+1}]}$ is a unit impulse with width $t_{i+1} - t_i$, defined as

$$\phi_{[t_i, t_{i+1}]}(t) = \begin{cases} 1 & t_i \leq t < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

Consider two adjacent impulse functions:

$$\phi_{[0, 1/2]}(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ 0 & \text{otherwise} \end{cases} \quad \phi_{[1/2, 1]}(t) = \begin{cases} 1 & 1/2 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

The sum of two adjacent impulse functions is a wider impulse:

$$\phi_{[0, 1]}(t) = \phi_{[0, 1/2]}(t) + \phi_{[1/2, 1]}(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

and the difference of two adjacent impulse functions is the basic wavelet, denoted by

$$\psi_{[0, 1]}(t) = \phi_{[0, 1/2]}(t) - \phi_{[1/2, 1]}(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \end{cases}$$

By solving (adding and subtracting) the two equations above, the two impulse functions can be obtained:

$$\begin{aligned} \phi_{[0, 1/2]}(t) &= (\phi_{[0, 1]}(t) + \psi_{[0, 1]}(t)) / 2 \\ \phi_{[1/2, 1]}(t) &= (\phi_{[0, 1]}(t) - \psi_{[0, 1]}(t)) / 2 \end{aligned}$$

Then any two-sample function can be written as

$$\chi(t) = s_0 \phi_{[0, 1/2]}(t) + s_1 \phi_{[1/2, 1]}(t) = s_0 \frac{\phi_{[0, 1]}(t) + \psi_{[0, 1]}(t)}{2} + s_1 \frac{\phi_{[0, 1]}(t) - \psi_{[0, 1]}(t)}{2} = \frac{s_0 + s_1}{2} \phi_{[0, 1]}(t) + \frac{s_0 - s_1}{2} \psi_{[0, 1]}(t)$$

where $(s_0 + s_1)/2$ represents the average of the function and $(s_0 - s_1)/2$ represents the change in the function. This is the Haar transform of the function. See here for more details.



< 12 December, 2020

13:20

196

8.4 WAVELET FUNCTIONS

We denote by W_j the difference between the function space V_{j+1} spanned by scaling functions $\varphi_{j+1,k}$ and the function V_j spanned by $\varphi_{j,k}$, i.e.

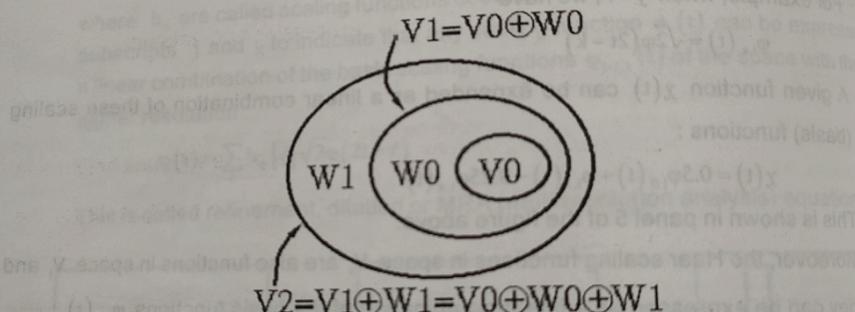
$$V_{j+1} = V_j \oplus W_j$$

(where \oplus represents the union of the two spaces). W_j is composed of all functions representable in V_{j+1} but not representable in V_j (as the scale of the basis functions of V_j is too coarse for the details of these functions). This can be carried out recursively to get:

$$V_{j+2} = V_{j+1} \oplus W_{j+1} = V_j \oplus W_j \oplus W_{j+1}$$

and finally:

$$L^2(\mathbb{R}) = V_\infty = V_0 \oplus W_0 \oplus W_{j+1} \oplus W_{j+2} \oplus \dots$$



Similar to a function space V_j spanned by the scaling functions $\varphi_{j,k}(t)$, the function space W_j is also spanned by a set of basis function, called the wavelet function

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$$

We see that the scaling sequence and the wavelet sequence correspond to low-pass filter and band-pass filter, respectively.

Also, as the wavelet functions $\psi_{j,k}(t)$ are members of the space W_j , they are also members of all the super-spaces:

$$\psi_{j,k}(t) \in W_j \subset V_{j+1} \subset \dots$$

they can be expanded in the space V_{j+1} of next higher scale with doubled resolution.



< 12 December, 2020

13:20

9 IMAGE COMPRESSION

9.1 FUNDAMENTALS

- (i) Data Compression refers to the process of reducing the amount of data required to represent a given quantity of information.
- (ii) Data and information are not the same thing; data are the means by which information is conveyed. Because various amounts of data can be used to represent the same amount of information, representations that contain irrelevant or repeated information are said to contain redundant data.

Fig. Compression as a transformation

- (iii) If we let N_1 and N_2 denote the number of bits in two representations of the same information, the relative data redundancy R of the representation with N_2 bits is

$$R = 1 - \frac{1}{C} \quad \dots(1)$$

where C = Compression Ratio $= \frac{N_1}{N_2}$ $\dots(2)$

Savings percentage This is defined as

$$\text{Savings percentage} = 1 - \frac{\text{Message (or file) size after compression}}{\text{Code (or file) size after compression}} = 1 - \frac{N_2}{N_1}$$

Bit rate: Bit rate describes the rate at which bits are transferred from the sender to the receiver and indicates the efficiency of the compression algorithm by specifying how much data is transmitted in a given amount of time. It is often given as bits per second (bps), kilobits per second (Kbps), or megabits per second (Mbps). Bit rate specifies the average number of bits per stored pixel of the image and is given as



< 12 December, 2020

13:20

Image Compression

223

$$\text{Bit rate} = \frac{\text{Size of the compressed file}}{\text{Total number of pixels in the image}} = \frac{N_2}{N} \text{ (bits per pixel)}$$

Example - 1. An image is 8 MB before compression and 2MB after compression. What are the values of compression ratio and savings percentage?

Soln. 1. Compression ratio is $8/2 = 4/1$.

2. Savings percentage is $1 - 2/8 = 6/8 = 3/4 = 75\%$.

(iv) Two-dimensional intensity arrays suffer from three principal types of data redundancies that can be identified and exploited.

(a) Coding Redundancy

(b) Spatial and temporal Redundancy

(c) Irrelevant Information

9.1.1. Coding Redundancy :

➤ A code is a system of symbols used to represent a body of information or set of events. Each piece of information or event is assigned a sequence of code symbols, called a code word.

➤ The number of symbols in each code word is its length, the 8-bit codes that are used to represent the intensities in most 2D intensity arrays contain more bits than are needed to represent the intensities.

➤ Assume that a discrete random variable r_k in the interval $[0, L-1]$ is used to represent the intensities of an $M \times N$ image and that each r_k occurs with probability $p_r(r_k)$,

$$P_r(r_k) = \frac{n_k}{MN}, k = 0, 1, 2, \dots, (L-1) \quad (1)$$

where L = The number of intensity level or values.

n_k = The number of times that the k^{th} intensity appear in the image

➤ If the number of bits used to represent each value of r_k is $\ell(r_k)$, then the average number of bits required to represent each pixel is,

$$L_{\text{avg}} = \sum_{k=0}^{L-1} \ell(r_k) p_r(r_k) \quad (2)$$

➤ The average length of the code words assigned to the various intensity value is found by summing the product of number of bits used to represent each intensity values is found by summing the products of the number of bits used to represent each intensity and the probability that the intensity occurs. The total number of bits required to represent an $M \times N$ image. = $M \times N \times L_{\text{avg}}$



9.2.1. Information Content of a Symbol

Let us consider a Discrete Memory less source (DMS) denoted by X and having alphabet $\{x_1, x_2, x_3, \dots, x_m\}$.

The information content of a symbol x_i , denoted by $I(x_i)$, is defined by

$$I(X_i) = \log_2 \left(\frac{1}{P(x_i)} \right)$$

Where $P(X_i)$ = probability of occurrence of symbol x_i

9.2.2. Average Information

Suppose M different and independent messages m_1 and m_2 , probabilities of occurrence P_1, P_2, \dots . Suppose that during a long period of transmission a sequence of length L has been generated, if L is very large we may expect that in the L message sequence we transmitted $P_1 L$ messages $m_1, P_2 L$ messages of m_2 , etc.

Total information in such sequence will be,

$$I(\text{total}) = P_1 L \log_2 \left(\frac{1}{P_1} \right) + P_2 L \log_2 \left(\frac{1}{P_2} \right) + \dots \rightarrow (1)$$

The average information per message interval = $I(\text{total})/L$

$$= P_1 \log_2 \left(\frac{1}{P_1} \right) + P_2 \log_2 \left(\frac{1}{P_2} \right) + \dots$$

$$= \sum_{i=1}^M P_i \log_2 \left(\frac{1}{P_i} \right) \rightarrow (2)$$

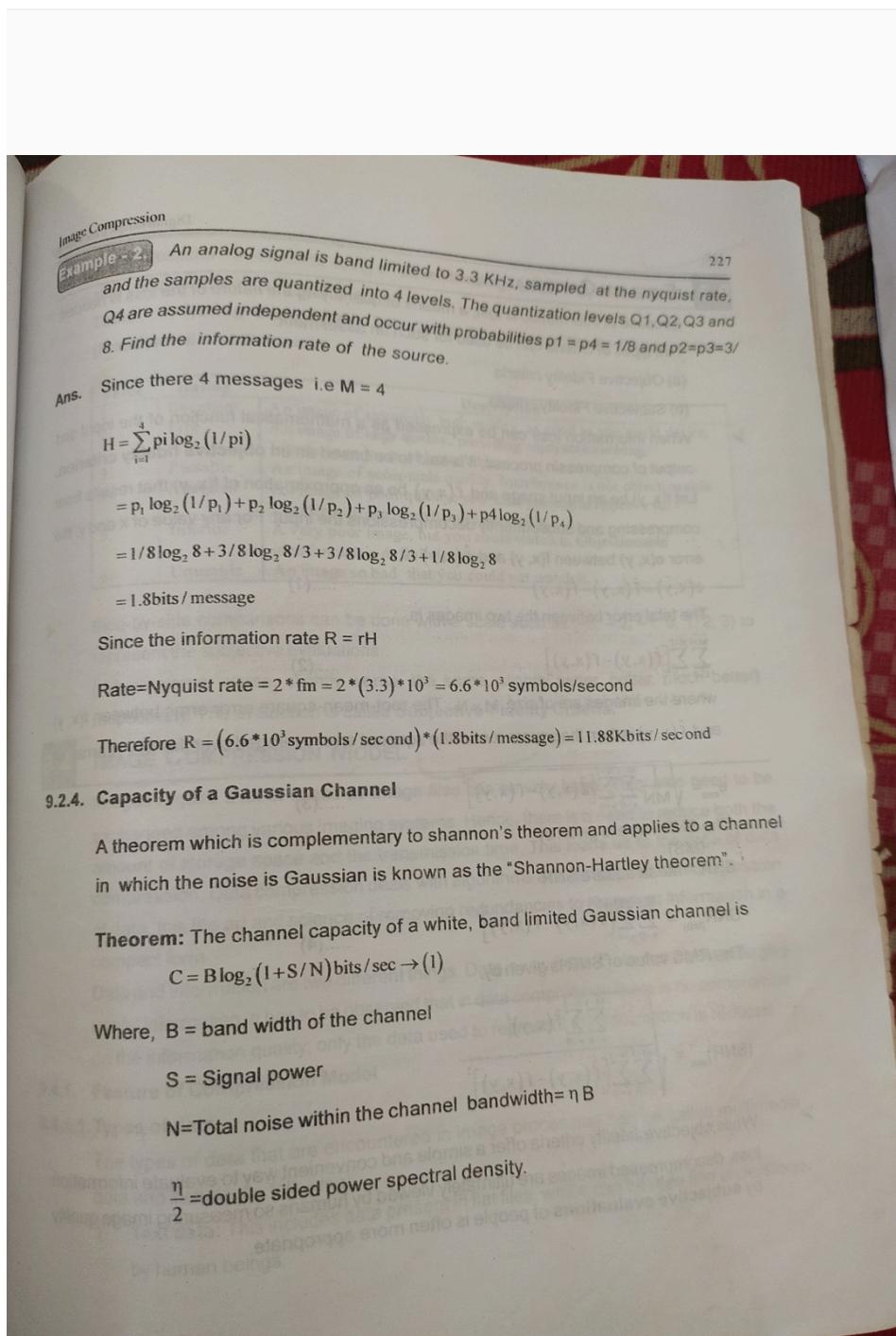
9.2.3. Information Rate

- If the source of the messages at the rate r messages per second, then the information rate is defined to be $\boxed{\text{Information rate} = rH}$
- It is designated by "R".
- The unit of R is "average number of bits of information /second"



< 12 December, 2020

13:21



< 12 December, 2020

13:21

Image Compression
229

The evaluations may be made using an absolute ratings scale or by means of side-by-side comparisons of $f(x, y)$ and $\hat{f}(x, y)$

Table - 1. Rating scale of the Television Allocations Study Organization. (Frendendall and Behrend.)

Value	Rating	Description
1.	Excellent	An image of extremely high quality, as good as you could desire.
2.	Fine	An image of high quality, providing enjoyable viewing. Interference is not objectionable.
3.	Passable	An image of acceptable quality. Interference is not objectionable.
4.	Marginal	An image of poor quality; you wish you could improve it. Interference is somewhat objectionable.
5.	Inferior	A very poor image, but you could watch it. Objectionable interference is definitely present.
6.	Unusable	An image so bad that you could not watch it.

- Side-by-side comparisons can be done with a scale such as $\{-3, -2, -1, 0, 1, 2, 3\}$ to represent the subjective evaluations.
 { Much worse, worse, slightly worse, the same, slightly better, better, much better} respectively. The evaluation is called **Subjective-Fidelity Criteria**.

9.4 IMAGE COMPRESSION MODEL

Images require a lot of space as image files can be very large. They also need to be exchanged among various imaging systems. Hence, there is a need to reduce both the amount of storage space and the transmission time. This leads us to the area of data compression. Data compression deals with algorithms and techniques for compression of data. It is the art and science of removing redundancies to represent information in a compact form.

Data and information are two different things. Data is raw facts. Data is processed to give useful information. It has to be observed that in data compression there is no compromise on the information quality; only the data used to represent the information is reduced.

9.4.1. Feature of Compression Model

9.4.1.1. Types of data

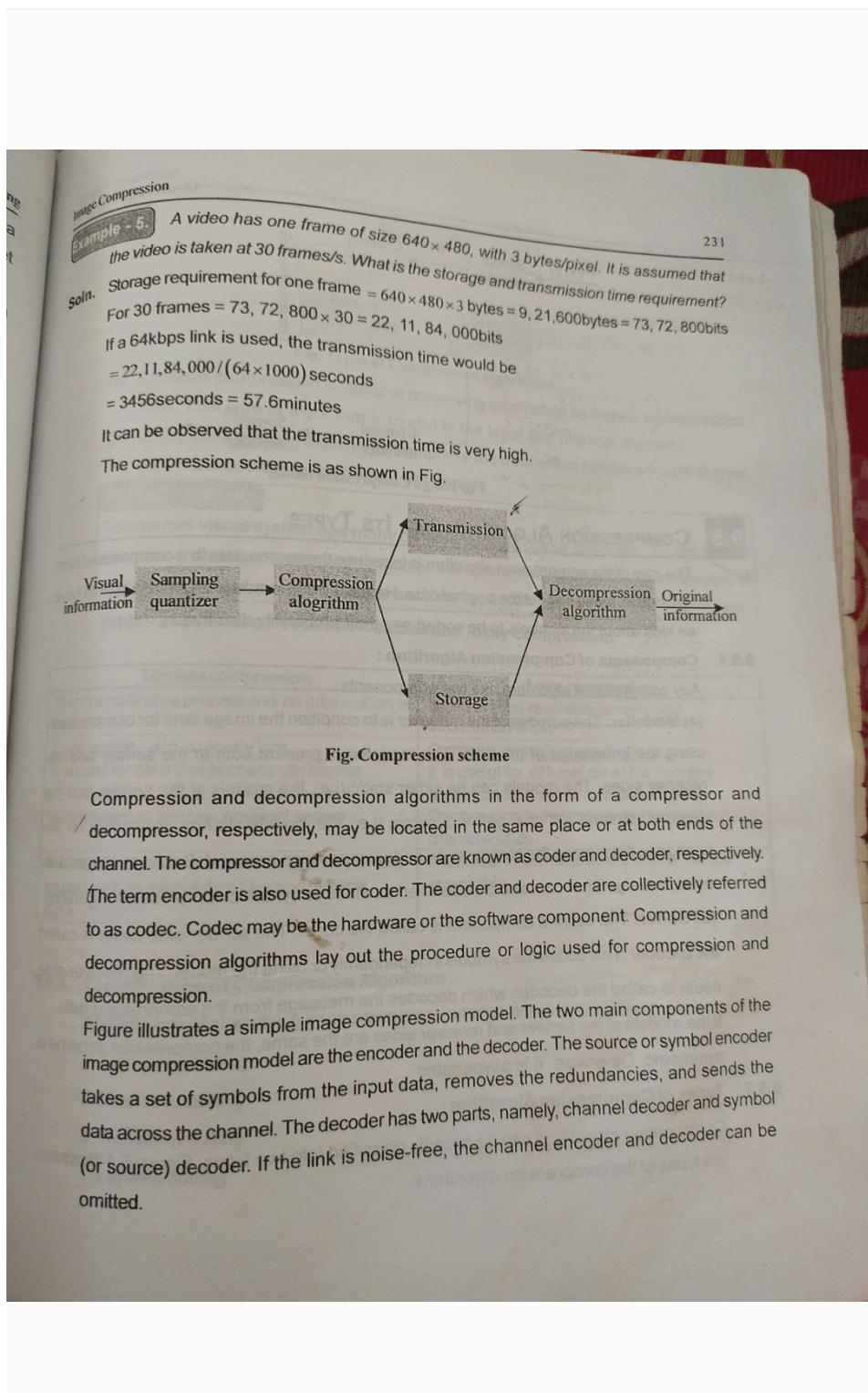
The types of data that are encountered in image processing can be called multimedia data and include the following:

1. **Text data:** This includes data present in flat files, which can be read and understood by human beings.



< 12 December, 2020

13:21



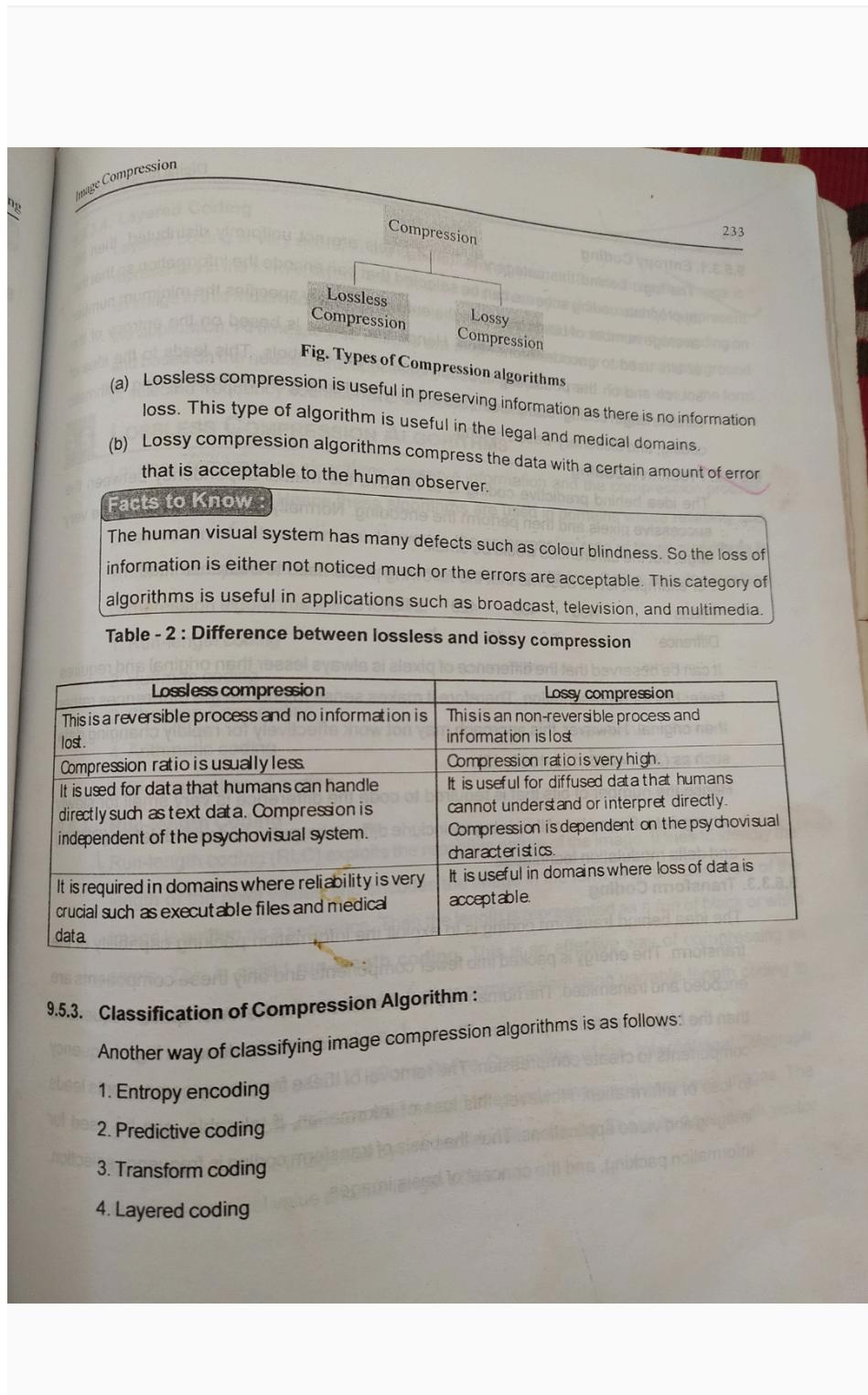
Compression and decompression algorithms in the form of a compressor and decompressor, respectively, may be located in the same place or at both ends of the channel. The compressor and decompressor are known as coder and decoder, respectively. The term encoder is also used for coder. The coder and decoder are collectively referred to as codec. Codec may be the hardware or the software component. Compression and decompression algorithms lay out the procedure or logic used for compression and decompression.

Figure illustrates a simple image compression model. The two main components of the image compression model are the encoder and the decoder. The source or symbol encoder takes a set of symbols from the input data, removes the redundancies, and sends the data across the channel. The decoder has two parts, namely, channel decoder and symbol (or source) decoder. If the link is noise-free, the channel encoder and decoder can be omitted.



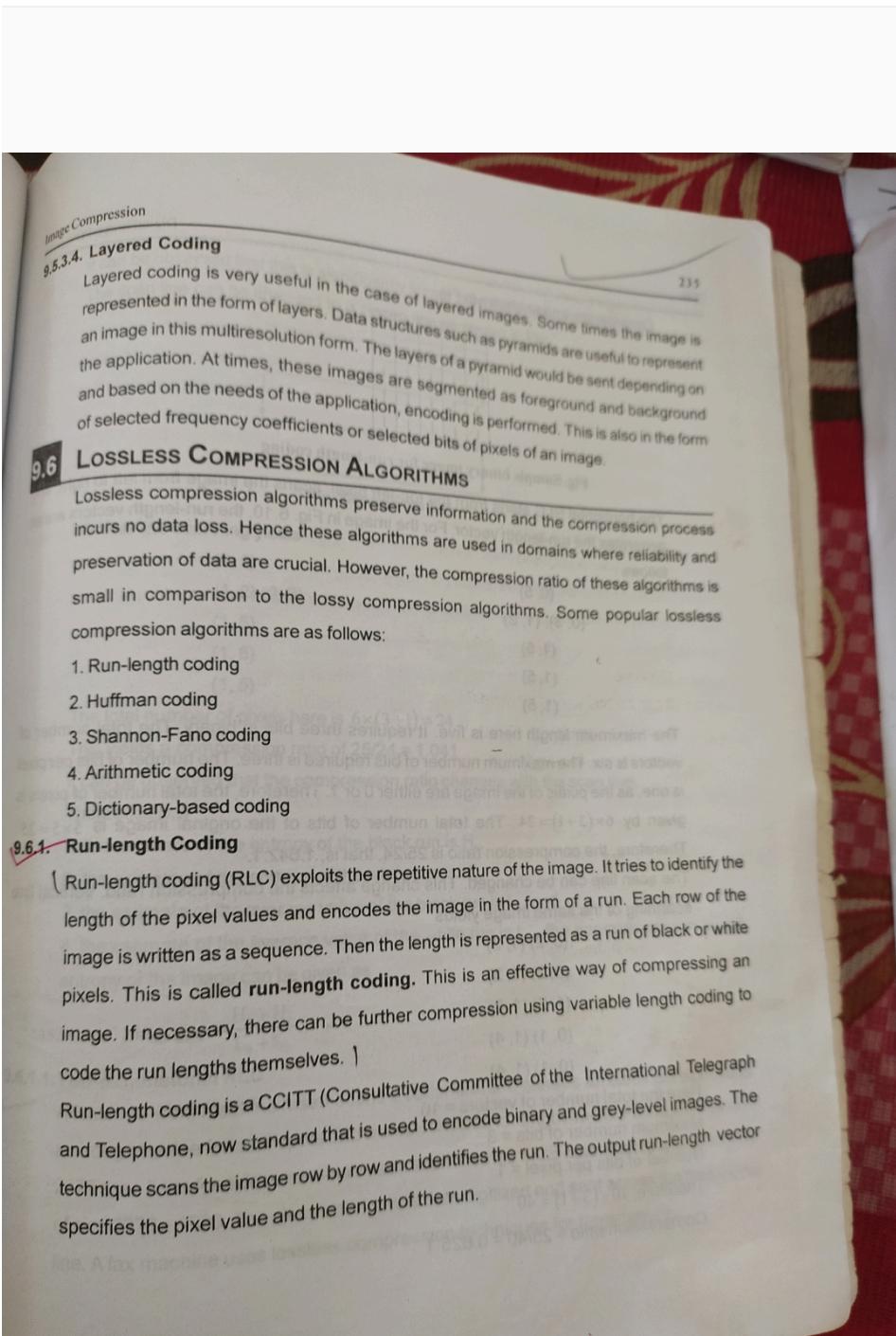
< 12 December, 2020

13:21



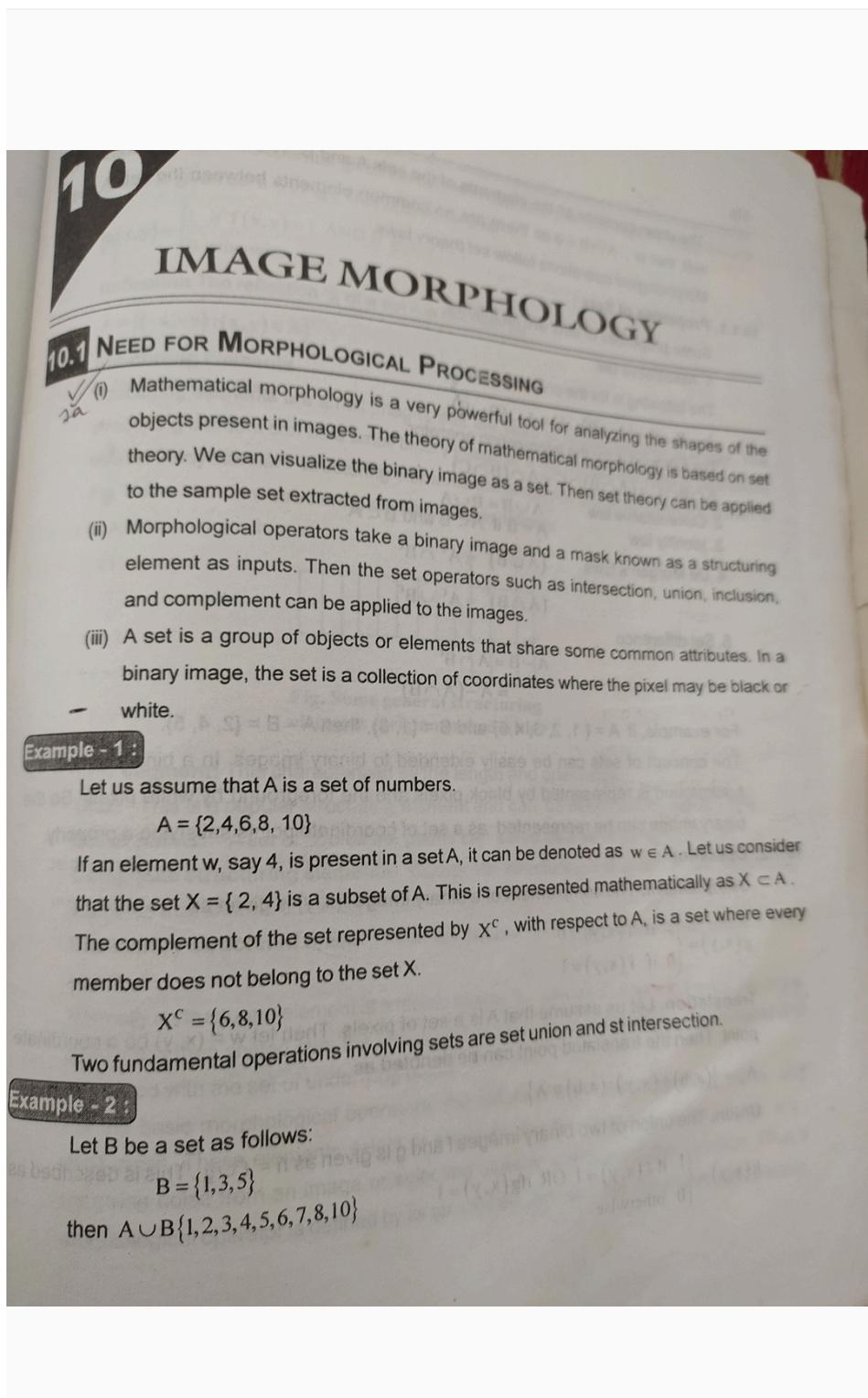
< 12 December, 2020

13:21



< 12 December, 2020

13:22



< 12 December, 2020

13:22

276

Digital Image Processing

The union contains all the elements of the sets A and B. The intersection of A and B is null, that is, $A \cap B = \emptyset$ as there are no common elements between the sets A and B.

Morphological operations follow set theory laws.

10.1.1. Properties of Sets :

Some properties of sets are as follows:

1. Associative law

The following is for the union operator:

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$$

The following is for the intersection operator:

$$A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$$

2. Commutative law

$$A \cup B = B \cup A$$

$$A = B \text{ if } A \subset B \text{ and } B \subset A$$

3. Identity law

$$(A \cup B)^c = A^c \cap B^c$$

4. De Morgan's laws

$$(A \cap B)^c = A^c \cup B^c$$

5. Set difference

$$A - B = A \cap B^c$$

$$= A - (A \cap B)$$

For example, if $A = \{1, 2, 3, 4, 5\}$ and $B = \{1, 3\}$, then $A - B = \{2, 4, 5\}$.

The concept of sets can be easily extended to binary images. In a binary image, the background is represented by black pixels and the foreground by white pixels. So the binary image can be represented as a set of coordinate points satisfying a property.

10.1.2. Operation :

Complement. The complement of a binary image $f(x, y)$ is given as

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) = 0 \\ 0 & \text{if } f(x, y) = 1 \end{cases}$$

Translation. Let us assume that A is a set of pixels. Then let $w = (x, y)$ be a coordinate point. Then the translated point can be denoted as

$$A_w = \{(a, b) + (x, y) : (a, b) \in A\}$$

Union. The union of two binary images f and g is given as $h = f \cup g$. This is described as

$$h(x, y) = \begin{cases} 1 & \text{if } f(x, y) = 1 \text{ OR } g(x, y) = 1 \\ 0 & \text{otherwise} \end{cases}$$

