
KTUNOTES

Digital Image Processing
MODULE 6

By
Shyno K G
(Asst. Professor, Dept. Of ECE)

Image Compression

Image Compression-Definition

"It is the process of mapping an image from a higher dimensional space to a lower dimensional space"

Image compression algorithms utilizes the following strategies:

- Removal of statistical redundancy
- Exploiting perceptual irrelevancy

Need for Image compression

1. **Reduction of storage space**: Development in multimedia systems demands usage of high quality images which requires large amount of space. Compression reduces the storage space

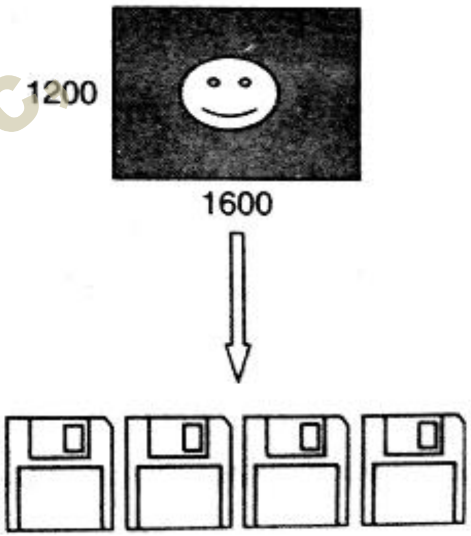
For example, if we want to store a 1600 × 1200 colour image then the space required to store the image is

$$1200 \times 1600 \times 8 \times 3 = 46{,}080{,}000 \text{ bits}$$
$$= 5{,}760{,}000 \text{ bytes}$$
$$= 5.76 \text{ Mbytes}$$

Image Compression

Need for Image compression

The maximum space available in one floppy disk is 1.44 Mb. If we have three floppies then the maximum space is $1.44 \times 4 = 5.76$ Mbytes. That is, a minimum of four floppies are required to store an RGB image of size 1600×1200.



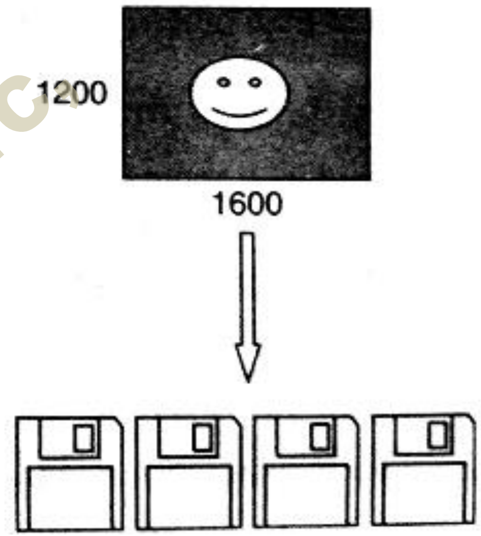
2. **Reduction of transmission bandwidth**: Amount of data transmitted through the internet doubles every year, and most of them are images. Compressed images require lesser bandwidth for transmission
3. **Improvement in processing speed**: Advanced technologies supports processing of compressed data; which improves higher processing speed.

E.g.: Compressed sensing

Image Compression

Need for Image compression

The maximum space available in one floppy disk is 1.44 Mb. If we have three floppies then the maximum space is $1.44 \times 4 =$ 5.76 Mbytes. That is, a minimum of four floppies are required to store an RGB image of size 1600×1200.



2. **Reduction of transmission bandwidth**: Amount of data transmitted through the internet doubles every year, and most of them are images. Compressed images require lesser bandwidth for transmission
3. **Improvement in processing speed**: Advanced technologies supports processing of compressed data; which improves higher processing speed.

E.g.: Compressed sensing

Image Compression

Redundancy

- Images are highly ***coherent***(contains redundant information)
- It is otherwise called duplicate information

Redundancy can be classified into:

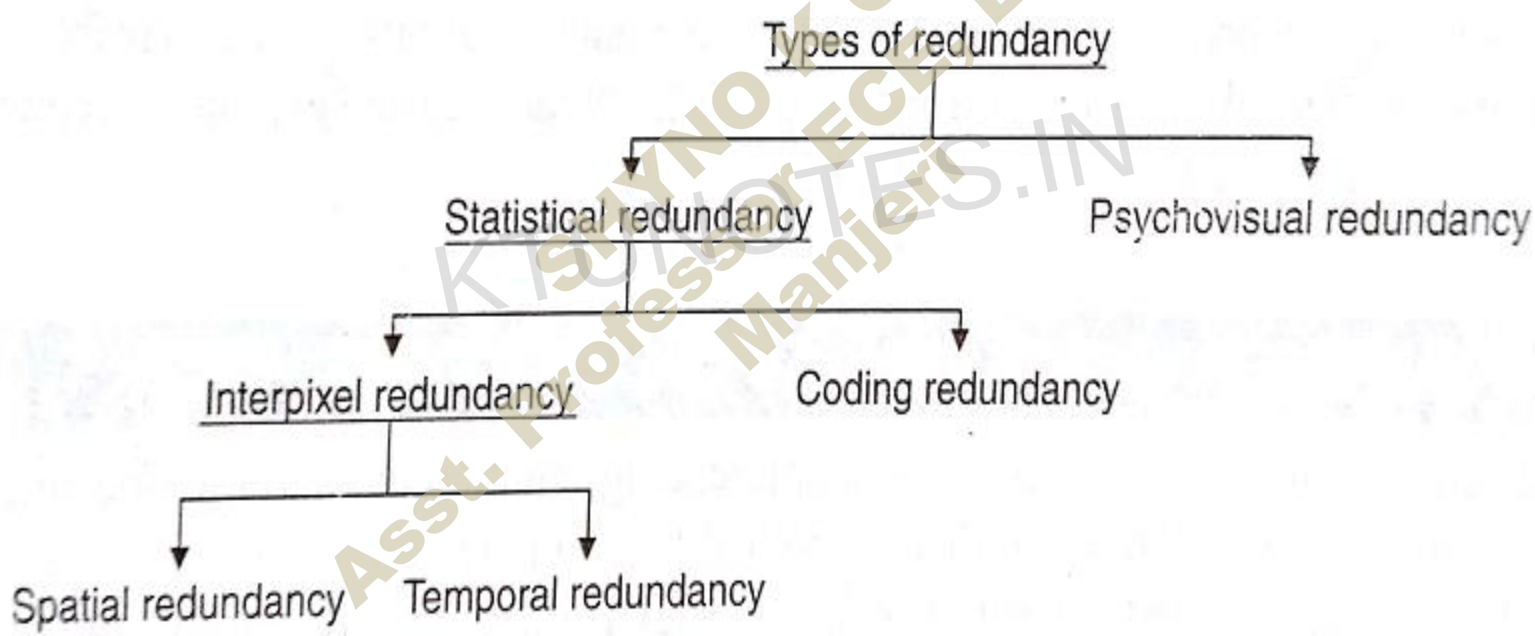


Image Compression

Redundancy

1. **Statistical Redundancy**: It is obtained from statistical analysis of an image. Two types are there:

a) **Inter-pixel redundancy**: Occurs when the neighboring pixels are statistically dependent (i.e. they possess a high correlation). Inter-pixel redundancy is again classified to:

i. **Spatial redundancy**: Here all the pixels are not necessarily represented, instead a pixel can be predicted from its neighbor using correlation property.

E.g.: Differential coding

ii. **Temporal redundancy(Inter frame redundancy)** : It occurs when the pixels of different frames of video shows a correlation. i.e. a time dependant correlation. It is used in video compression. E.g.: Motion compensated predictive coding

b) **Coding redundancy**: Occurs when the image is encoded. Then the repetition of codes is called code redundancy (E.g.: Redundancy in Huffman/Arithmetic codes)

Image Compression

Redundancy

2. **Psycho-visual Redundancy**: It is associated with characteristics of **HVS** (Human Visual System). If less data is used to represent less important visual information, perception will not be effected.

- Visual information is psycho-visually redundant

Image compression schemes

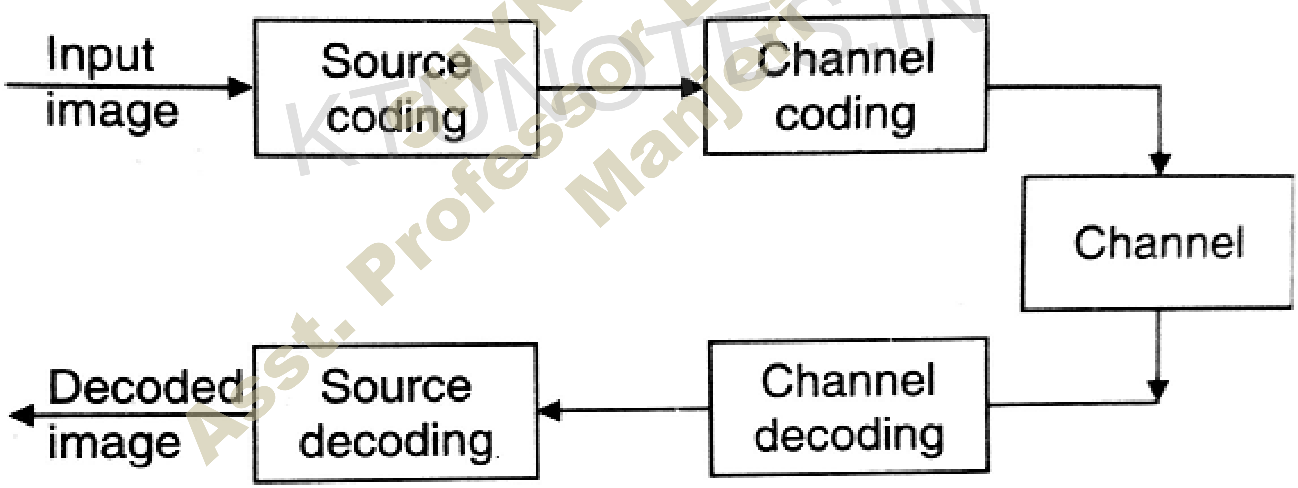


Image-compression scheme

Image Compression

Image compression schemes

- The source encoder and decoder pair is known as ***source codec module***
- The Channel encoder and decoder pair is known as ***channel codec module***
- The design of codec is based on Shannon's theorem
- **Source coding** : It is used for efficient conversion of source data into a sequence of bits. Source coding reduces entropy (Average number of bits required to represent the image).
- **Source decoding** : It is the inverse operation of Source coding
- **Channel** : It represents the mathematical model of the medium over which communication occurs
- **Channel coding** : It introduces a controlled redundancy to the compressed output of the source encoder. Which gives immunity against noise and transmission errors.
- **Channel decoder** : It exploits the redundancy in the bit sequence to reproduce the compressed bits (Inverse of channel coding)

Image Compression

Classification of Image compression schemes

1. **Lossless compression (Reversible Compression)**

- ➢ The image after decompression is identical to original image
- ➢ Modest compression rate
- ➢ Used where no loss of image data is compromised
- ➢ E.g.: Medical image compression, PNG, GIF

2. **Lossy compression (Irreversible Compression)**

- ➢ The decompressed image contains degradations with respect to original image
- ➢ Higher compression rate
- ➢ No visible degradation under a set of designated viewing conditions
- ➢ E.g.: Multimedia applications, TIFF, MNG

Image Compression

Fundamentals of Information Theory

Entropy

If 'x' → random variable generated by a discrete memory less source 'X' is a finite alphabet set

$$X = \{x_1, x_2 \dots x_N\}$$

If $P_x(x)$ → is the PDF of x

Shannon defined a measure of information content of 'X' (The minimum possible data rate by which X can be transmitted without any loss) is called entropy of X given by:

$$H(X) = \sum_{x \in X} P_x(x) \log \frac{1}{P_x(x)}$$

Image Compression

Fundamentals of Information Theory

If the conditional probability of X w.r.t Y is given as P_{XY} (where X & Y are random variables

Then $$H(X|Y) = \sum_{x \in X, y \in Y} P_{XY} \log \frac{1}{P_{XY}}$$

where $H(X|Y)$ = is the conditional entropy.

Mutual Information

$$I(X:Y) = H(X) - H(X|Y)$$

or

$$I(X:Y) = H(Y) - H(Y|X)$$

Image Compression

Fundamentals of Information Theory

Shannon's Source Coding Theorem

According to shannon's source coding theorem

" Let X be the random variable.

$X = \{x_1, x_2 \ldots x_N\}$ from a discrete memoryless source with entropy $H(X)$

Let blocks of J symbols from the source are encoded into code words of the length N from a binary alphabet. Then the probability of a block-decoding failure is given by

$$R \equiv \frac{N}{J} \geq H(X) + \varepsilon .$$

Image Compression

Fundamentals of Information Theory

If J is sufficiently large

$$R \leq H(X) - \varepsilon.$$

if $J \to \infty$

$R \to 1$

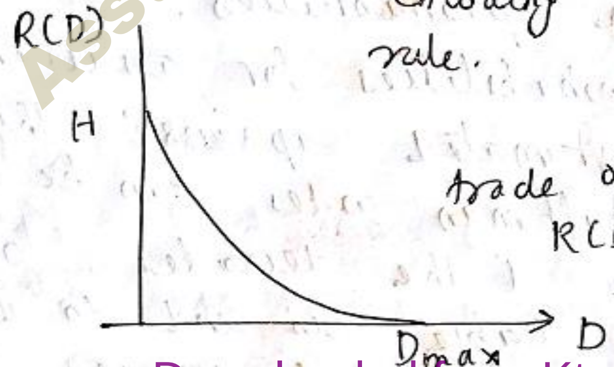
Rate Distortion theorem.

$$R(D) = \inf_{P(Y|X),\ MSE\ \mid\ P(Y|X) \leq D} I(X;Y)$$

← mutual info..

↑ $P(Y|X)$: source coding rule

↑ $\leq D$: MSE of encoding rule.



trade off b/n R(D) & D.

Image Compression

Fundamentals of Information Theory

$R(D) \rightarrow$ minimum rate that is required to encode the source with maximum distortion.

$\inf_{a} f(x) = y$ → min of 'f(x)' with max 'a'.

Image Compression

Huffman Coding

- By David Huffman in 1950.
- It assigns a binary code for each intensity value.
- Use shorter codes for intensities with higher probabilities.
- If probabilities for available intensities are are estimated apriori then the table of Huffman codes can be fixed in both encoder & the decoder. Otherwise the coding table is sent to the decoder along with compressed image.

Image Compression

Huffman Coding

The main parameters for Huffman coding are:

a) Entropy

b) Average length

c) Efficiency

d) Variance

Huffman code is a Pre-fix code (unique mapping).

Image Compression

Huffman Coding

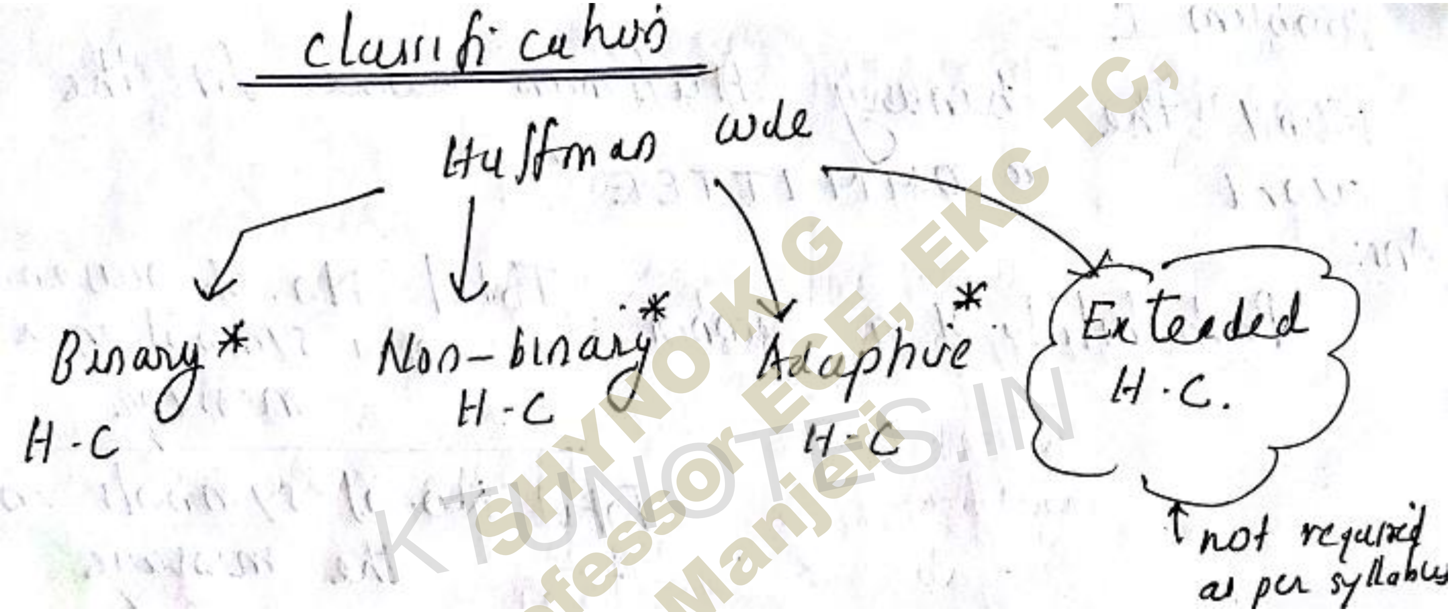


Image Compression

Huffman Coding

I. Binary Huffman Code

- Each symbol (Pixel value) is represented by a binary code
- It can be represented by a binary tree where the leaves will be the symbols. The code for any symbol is obtained by traversing the tree from the root node to the leaf corresponding to the symbol by assigning a 0 to the code word every time the traversal takes us over an upper branch & assigning a '1' every time the traversal takes us over a lower branch.

*(For Problems Refer Notebook)**

Image Compression

Huffman Coding

II. Non-Binary Huffman Code

$$N = r + \alpha(r-1). \rightarrow (1)$$

Here N = Number of probabilities

r = Base (eg: binary 2,
ternary 3,
HEX 16 etc).

To code a word with 'N' symbols using an index (base) 'r' Huffman code it should satisfy eqn (1). If not add a dummy symbol with probability '0' and discard the code at the end.

*(For Problems Refer Notebook)**

Image Compression

Huffman Coding

III. Adaptive Huffman Code

- Discovered by ***Newton Faller***
- It learns the symbol probabilities by dynamically using the symbol counts to adjust the code tree
- Used in coding live streaming where there is no knowledge about the statistics of source data
- It uses learning techniques to encode

i.e. taking into account the impact of the previous symbol to the probability of the current symbol (e.g., "qu" often come together, ...)

Limitations of Huffman Coding

- Probability has the property ;when one of the symbol has a very high probability (P_{max}) It can suppress the other probabilities (i.e . $\sum P_i = 1$)
- E.g.: Consider the sequence '***qqqqqrv***' the high probability of q suppress that of ***r*** and ***v*** (Here P_{max}=5/7 i.e very high)

Image Compression

Huffman Coding

Limitations of Huffman Coding

- Huffman code satisfies the above inequality

$$H(S) \leq L_{avg} < H(S) + P_{max} + 0.086$$

- I.e. When P_{max} is very high the average code length also very high this makes Huffman coding inefficient since it is a probability dependent coding scheme
- So to keep Huffman coding efficient the input data should be random enough

Image Compression

Arithmetic Coding

- Efficient code :- It can efficiently code non-random data also.
- It is a stream based coding scheme rather than symbol base. [a sequence forms a unique code]
- works on block of data : a word is coded to a "tag".
- Works well with low entropy sequences.

*(Problems and encoding/decoding procedure refer notebook)**

Image Compression

Limitations of Arithmetic Coding

- Use of finite precision arithmetic
- We need to put an additional end of message symbol to understand that the decoding has been finished

Image Compression

Dictionary based Compression

This uses dictionary or look up table. The original word is compared with available patterns in the dictionary (LUT).

Dictionary can be static or dynamic (adaptive)

Adaptive dictionaries update the dictionary each time a new pattern is encountered.

Dictionary based compression uses Lempel Ziv algorithm developed by Jacob Ziv & Abraham Lempel.

eg: LZ77 , LZ78

1977 Jacob ziv — 1978 Abraham Lempel.

Image Compression

Dictionary based Compression

another dictionary based compression is LZW (Lempel-Ziv-Welch)

LZW is utilized in GIF, TIFF, & PDF.

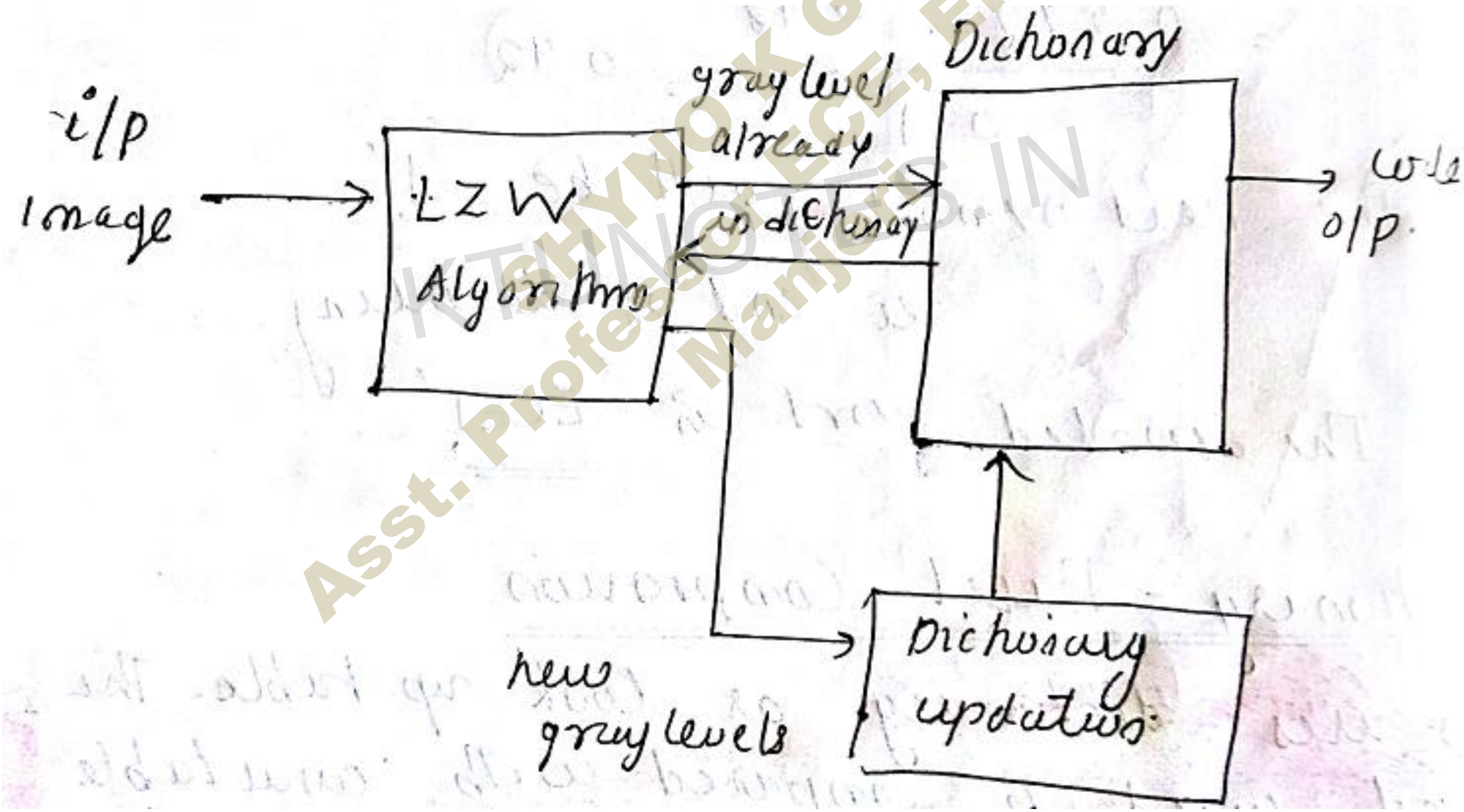


Image Compression

Dictionary based Compression

when a new intensity level is detected it is assigned the next free location in the dictionary.

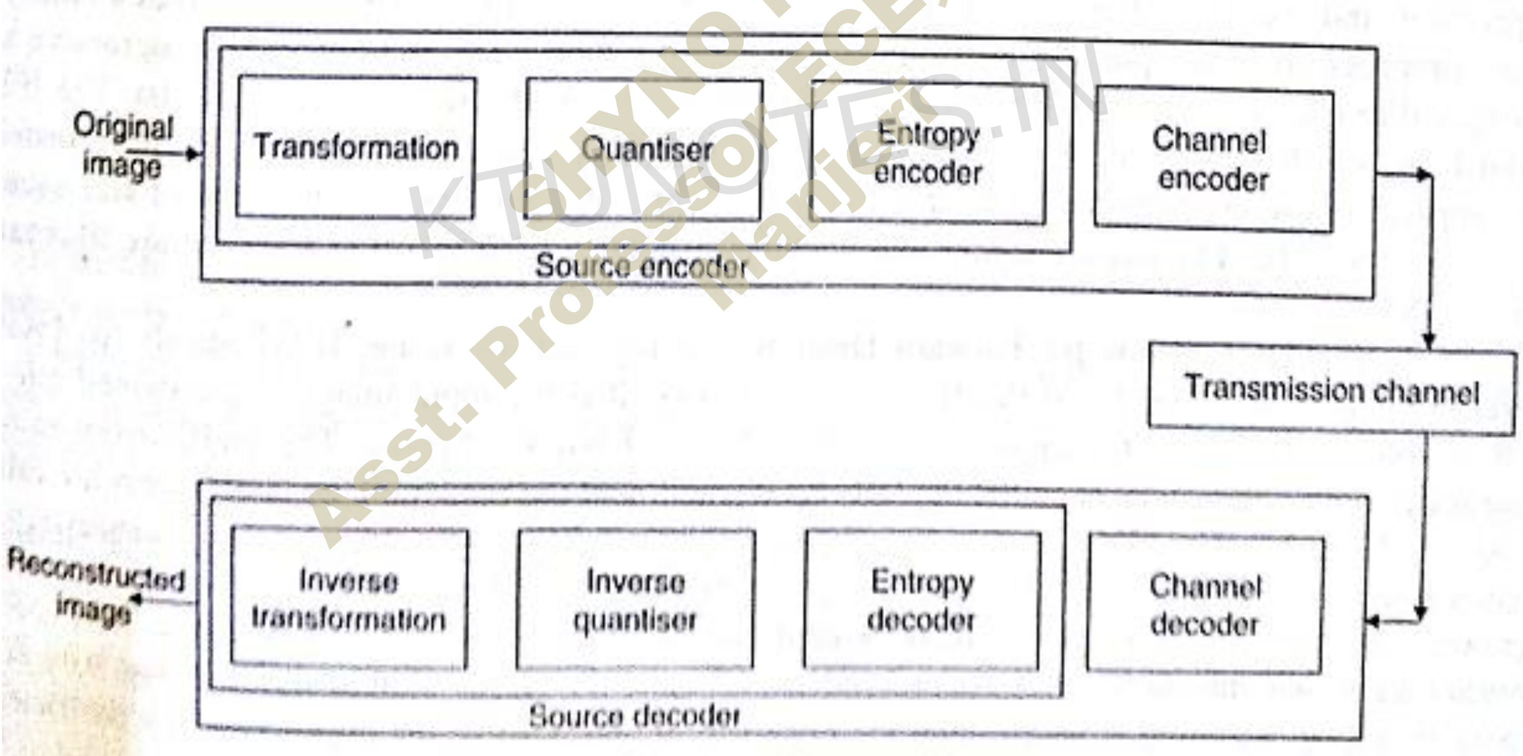
Transform based Compression

- Transform based coding uses a ***linear transformation*** (DFT, DCT etc) to convert an image containing highly correlated pixels into a few non-correlated transform coefficients
- It is based on the ***correlation of pixels*** as well as the ***energy compaction capability*** of transform
- If we use a frequency transformation (e.g. DCT) we can reduce the spatial redundancies in time domain
- After transform the useful data (Energy) will be concentrated around low frequencies

Image Compression

Transform based Compression

- Because of the above property we can reconstruct the image without much change using the low frequency transform coefficients. And this gives good compression ratio
- Transform coder segments the image into non overlapping blocks, each represented as ***weighted sum of discrete basis functions***



Transform-based image-coding scheme

Image Compression

Transform based Compression

Transformation

- It reorganizes the gray values to change the correlations of original pixels. It produces transform coefficients which are highly un correlated
- An efficient transformation should possess the following properties:

i. *De-correlation*: The transform coefficients should be highly uncorrelated

ii. *Linearity*: Which ensures one-to-one mapping and make reconstruction easy

iii. *Orthogonality*: Which ensures elimination of redundancy

Quantization

- It is a non-linear irreversible process
- Here the transform coefficients are truncated/round off to achieve the required word-length of the processor

Entropy Encoder

- The quantized data is encoded using any of the Entropy based coding schemes
- E.g. : Huffman coding, Arithmetic coding, RLE (Run Length Encoding)

Image Compression

Transform based Compression

Entropy Encoder

- Which further improves efficiency since all these techniques are lossless and utilizes the entropy levels to set the code length

Channel Encoder

- It uses the error checking schemes such as CRC, Parity etc

Decoder Section

- It performs all the above encoding process in the reverse order and finally the inverse transform will give you the reconstructed image

E.g. for transform based compression are JPEG and MPEG

Image Compression Standards

Image compression standards

"It defines procedures for compressing and de-compressing images"

- Different standardization authorities such as ISO, ITU-T etc. defined many standardizations
- We are focusing on two of them
- JPEG (Image compression standard)
- MPEG (Video compression standard)

I. JPEG (Joint Photographic Experts Group)

- Used for compressing Continuous tone still images
- Supported by ISO/ITU-T/IEC (International Electro-technical Commission)
- Normal JPEG is a lossy baseline coding system
- Uses quantized DCT coefficients on 8X8 image blocks
- Coding is done using Huffman coding/RLE
- Most widely used over Internet

Image Compression Standards

JPEG (Joint Photographic Experts Group)

JPEG-LS

- It's a JPEG standard supported by ISO/IEC/ITU-T
- Lossless to nearly lossless
- Use adaptive prediction, context modeling and Golomb coding

JPEG-2000

- It's a JPEG standard supported by ISO/IEC/ITU-T
- Can be Lossy/Lossless
- Use Arithmetic coding and DWT (Discrete Wavelet Transform)

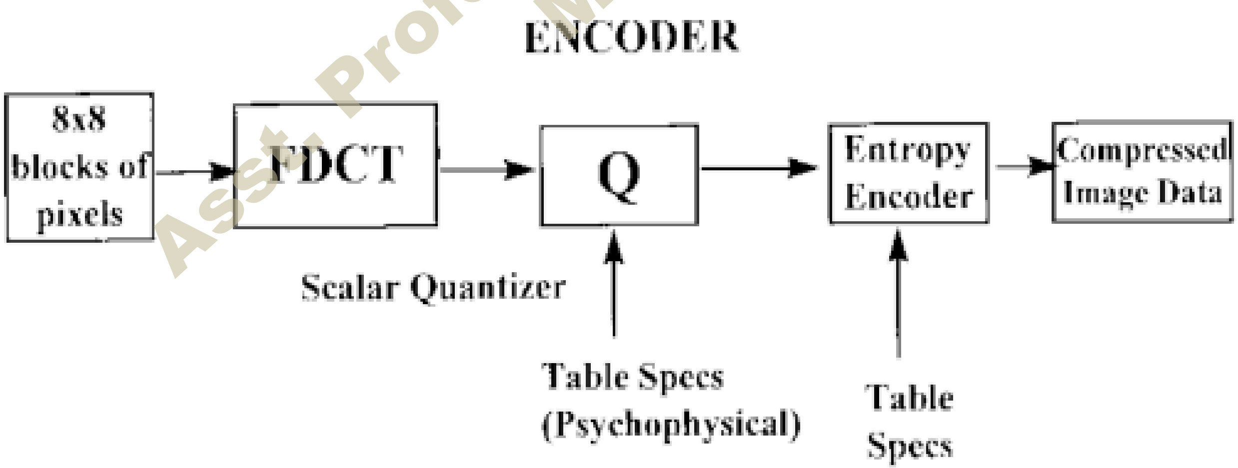


Image Compression Standards

JPEG (Joint Photographic Experts Group)

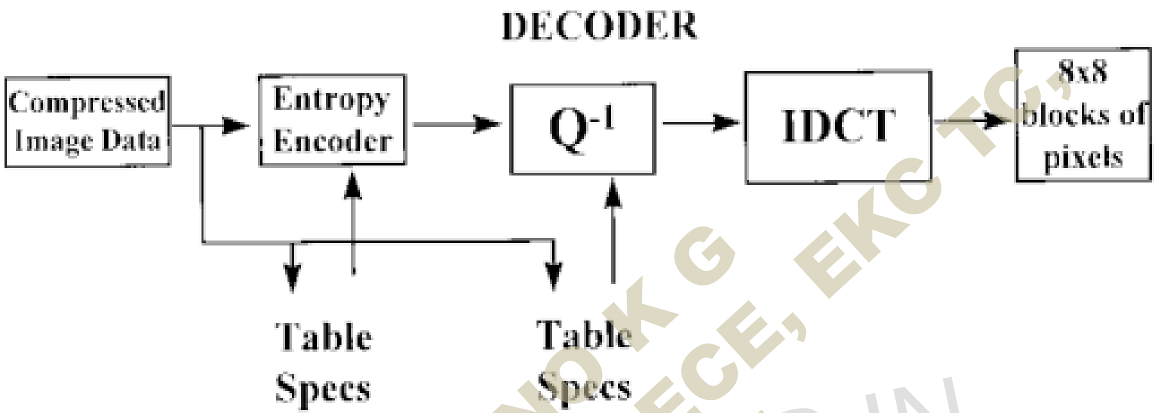


Table specification includes Probability of symbols, Symbols used, Word length etc.

Q^{-1} is the inverse quantizer

(*Each blocks in the JPEG are familiar to you ; Explained in class so write a brief description for them in exams)**

Image Compression Standards

II. MPEG/MPEG-1 (Motion Pictures Experts Group)

- Used for video compression
- Supported by ISO/IEC
- Standard for CD ROM applications with non-interlaced video up to 1.5 Mbps
- Frame predictions are based on Previous frame, next frame or an interpolation of both
- Supported by almost all computers and DVD layers

MPEG-2

- Supported by ISO/IEC
- Extension of MPEG-1 for DVDs with transfer rate 15 Mbps
- Supports interlaced video and HDTV
- Most successful video standard

MPEG-4

- Supported by ISO/IEC
- Extension of MPEG-2 supports variable block sizes and differential prediction within frames

Image Compression Standards

II. MPEG/MPEG-1 (Motion Pictures Experts Group)

MPEG-4-AVC

- Supported by ISO/IEC
- Extension of MPEG-4 supports Advanced Video Coding (AVC)

MPEG-2 ENCODER

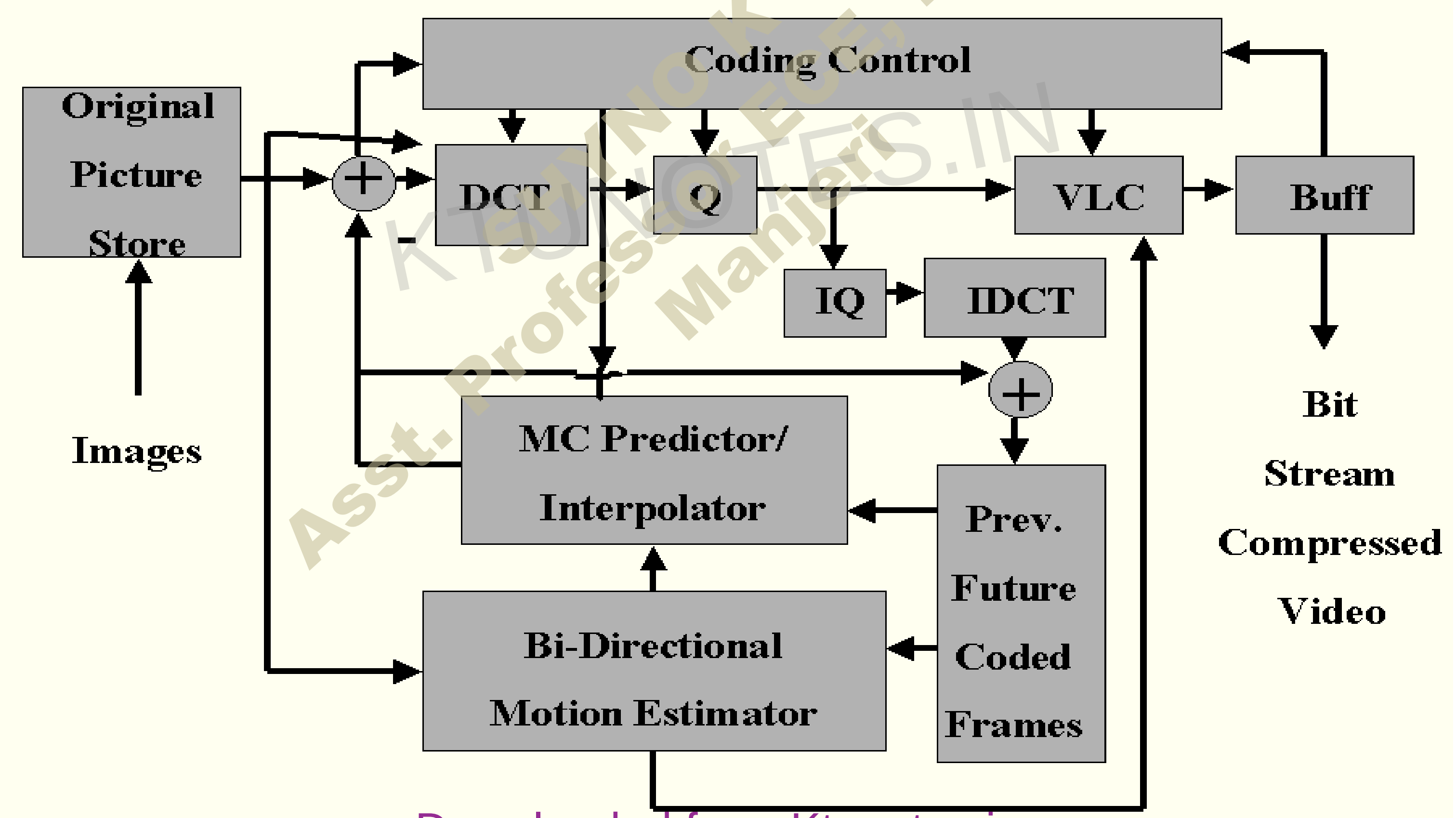


Image Compression Standards

II. MPEG/MPEG-1 (Motion Pictures Experts Group)

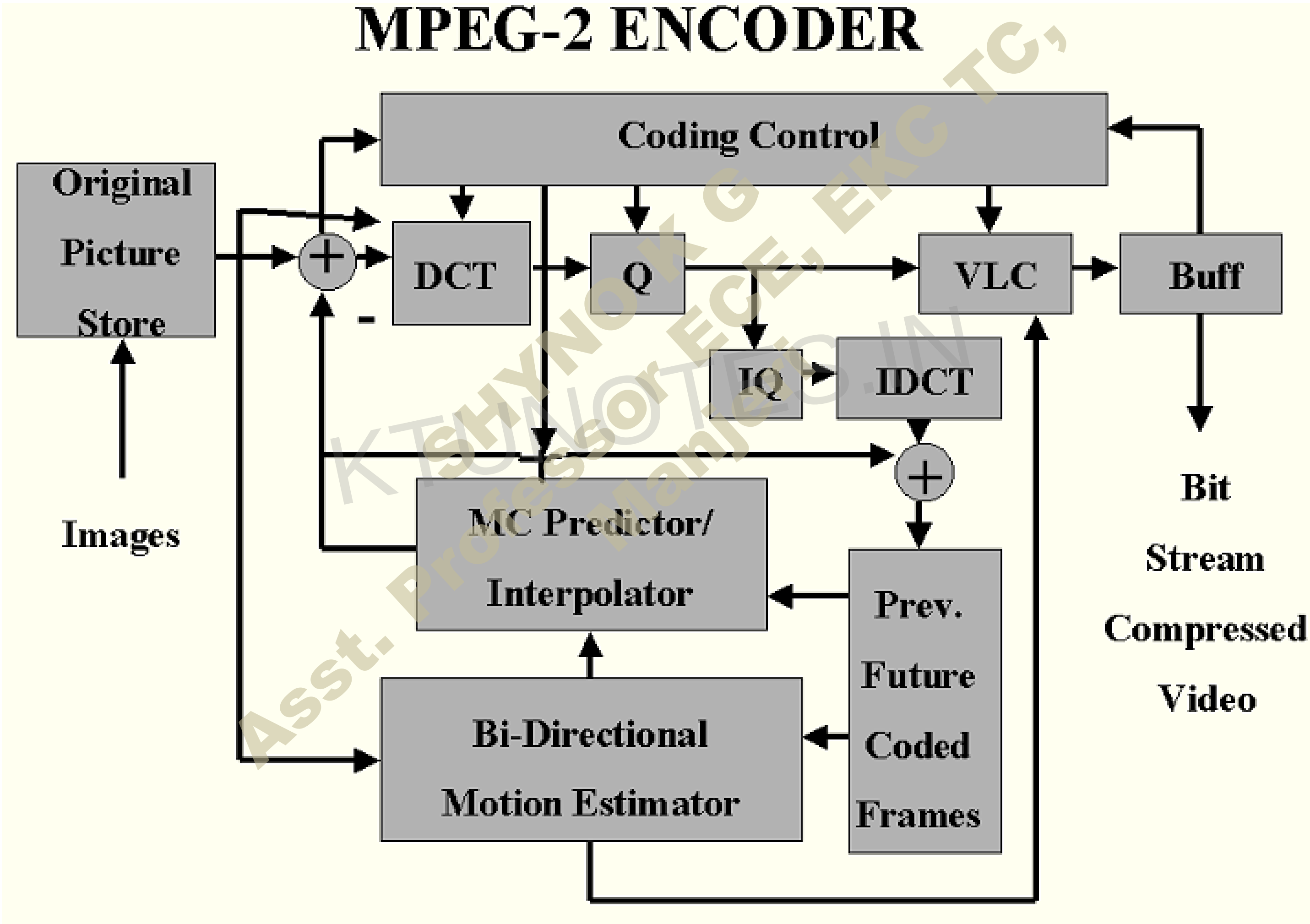


Image Compression Standards

II. MPEG (Motion Pictures Experts Group)

- **MC Predictor (Interpolator)** : is the Motion Compensator Predictor. Since MPEG is a video CODEC it utilizes both spatial and temporal redundancies. MC predictor estimate the temporal redundancies required for compression using a Motion vector
- **Bi-directional Motion estimator**: used to estimate the motion vector based on previous/next/interpolated frames
- **VLC**: is the Variable Length Coding block (Entropy coding)
- The entire system utilizes a feedback mechanism

(*Video compression is a vast topic and cannot be covered in the syllabus in detail. So simple explanation about the block are required for image processing point of view . Hope other blocks are familiar to you*)*

Image Compression

Wavelet based image compression

Wavelets

- Comes with the importance of Time freq. Analysis.
- A **wavelet** is a wave-like oscillation with an amplitude that begins at zero, increases, and then decreases back to zero.
- Gives an infinite set of bases to represent a function.

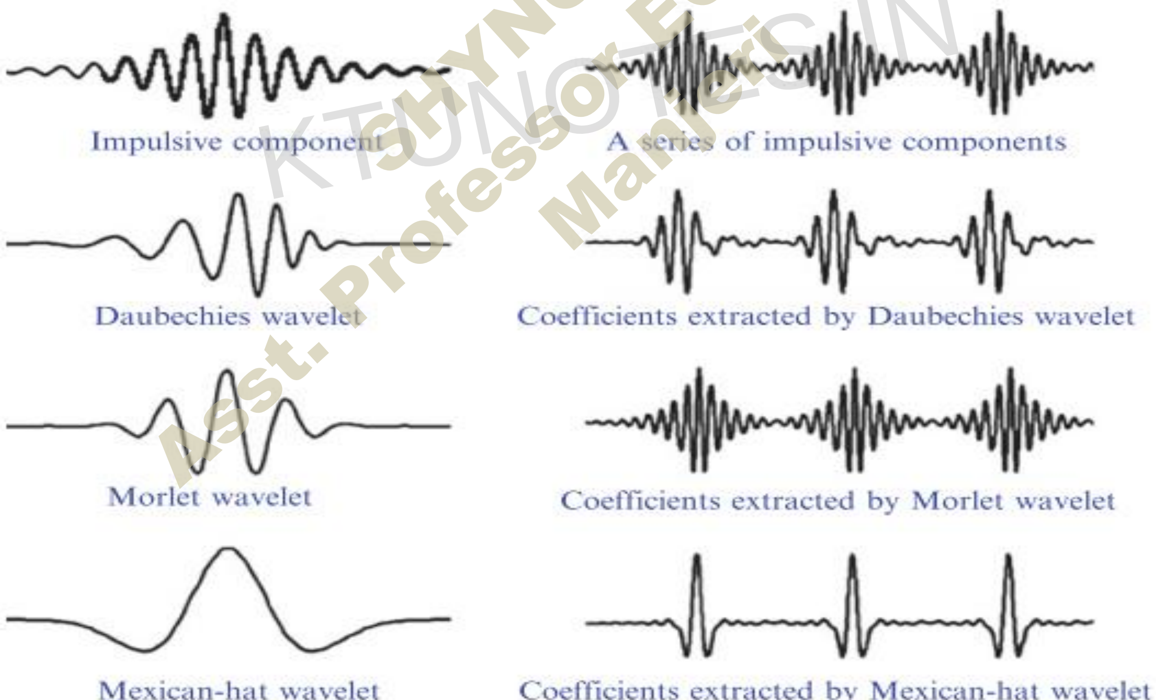


Image Compression

Wavelet based image compression

Wavelet Transform

Continuous Time WT (CWT)

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\, \Psi^* \left(\frac{t-b}{a}\right) dt$$

Inverse Continuous Time WT (ICWT)

$$f(t) = \frac{1}{\sqrt{|a|}} \int_{0}^{\infty} \int_{-\infty}^{\infty} W(a,b) \Psi \left(\frac{t-b}{a}\right) dbda$$

Discrete Time WT (DWT)

$$W(j,k) = 2^{-j/2} \int_{-\infty}^{\infty} f(t) \Psi^* \left(2^{-j} t - k\right) dt$$

Inverse Discrete Time WT (IDWT)

$$f(t) = \sum_{j} \sum_{k} w(j,k) 2^{-j/2} \Psi(2^{-j} t - k)$$

Image Compression

Wavelet based image compression

Wavelet Transform

- Where ψ(.) represents the ***Mother Wavelet***
- ***a*** and ***j*** are called the ***Dilation parameter*** or ***Scaling parameter*** in analog and discrete domains respectively
- ***b*** and ***k*** are called the ***Translation parameter*** or ***Shifting parameter*** in analog and discrete domains respectively
- Since we are focused on discrete images we will consider the DWT

Features

- Dynamic Resolution with varying Window size.
- Time frequency analysis.
- Sub band decomposition and **MRA**.
- Suitable for natural existing signals.
- Signal Compression is achieved.
- Better performance with Wavelet packet transform
- Computationally efficient & inherently local

Image Compression

Wavelet based image compression

Selection of Wavelet

1. **Qualitative Approach**

- Orthogonality (Non overlapping Freq. bands)
- Symmetry(Linear Phase Filter)
- Compact Support (Within finite interval basis fn. is non zero)
- Shape matching.

Thus efficiently represents s/l that have localized features.

E.g.: Coiflects(EMG),Morlet and Gaussian (ECG).

- The morphology and latency of bio signal peaks are preserved by symmetric wavelets.
- Daubechies Wavelets are suited for motor unit potentials hidden in EMG.

Image Compression

Wavelet based image compression

Selection of Wavelet

1. Quantitative Approach

- Uncertainty Model
- Energy of Wavelet Coefficients

$$E(a) = \int |W(a,b)|^2 db$$

- Shannon Entropy (Lower Entropy Higher Energy Conc.)

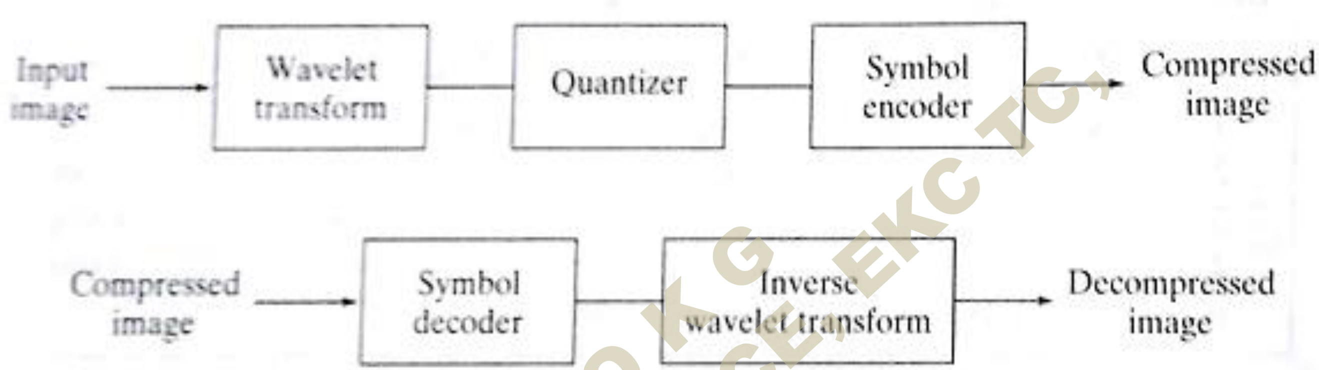
$$H(a) = -\sum P_i log_2 P_i$$

$$0 \leq H(a) \leq log_2 N$$

- E.g.: DB-8 for ECG (Max. Cross Correlation Coefficient)

Image Compression

Wavelet based image compression



- Symbol Encoder/Decoder can be any of the entropy coding schemes

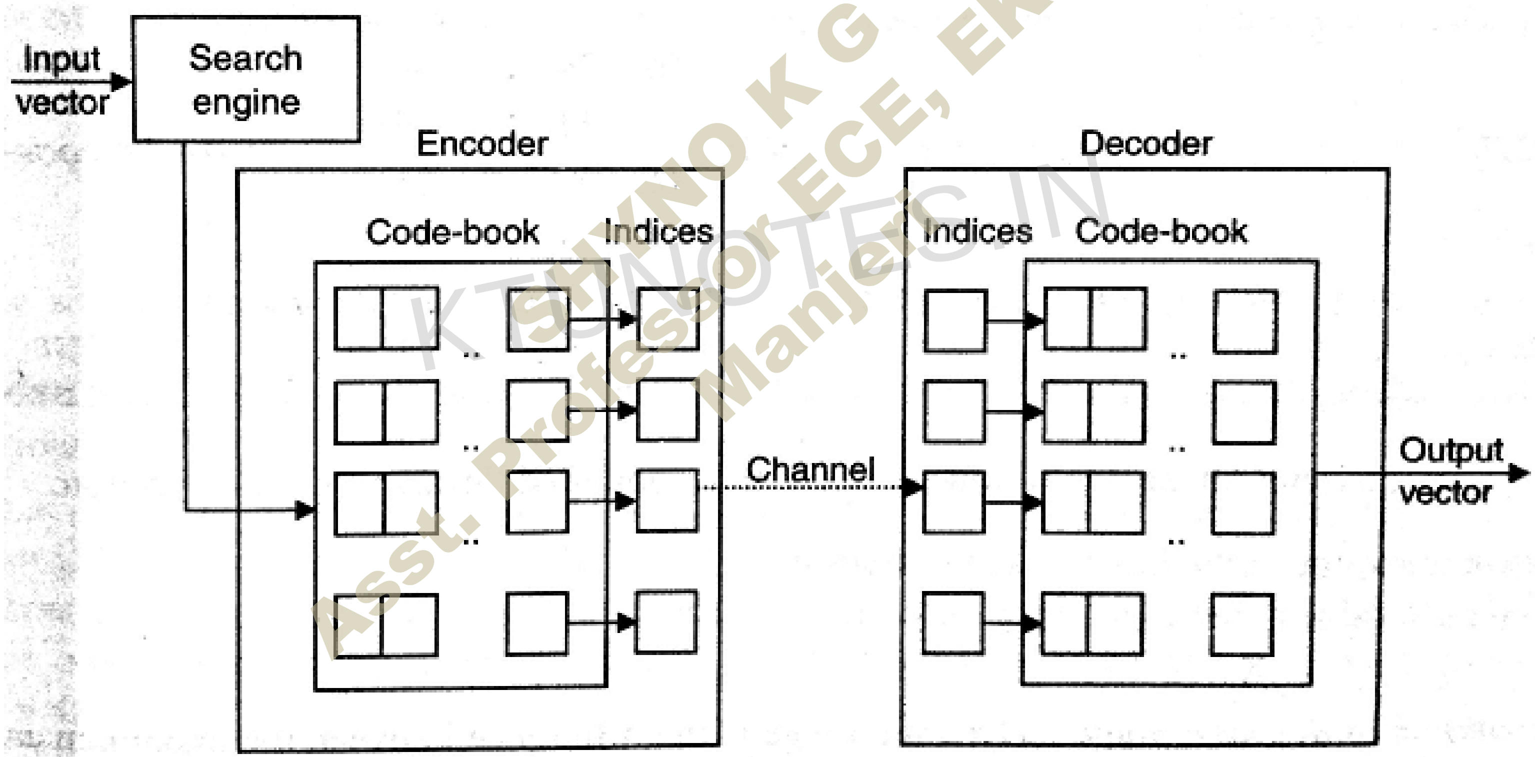
(*Other blocks are well known for you. Please describe in your on language*)*

- E.g. for Wavelet based image compression is JPEG 2000

Image Compression

Vector Quantization (VQ)

- VQ is a block coding technique which utilizes the correlation between neighboring pixels



Vector quantisation scheme

Image Compression

Vector Quantization (VQ)

- At the encoder side the input image is partitioned into a set of non overlapping image blocks
- The closest code word in the code book is found for each image block
- Closest code word for a given block is the one in the code book that has the minimum squared Euclidean distance from the input block
- Compression is achieved by sending the indices corresponding to a particular code book entry
- At the decoder side the same code book is used such that received indices can regenerate the code words
- VQ code book generation is the process of finding an optimal code book that yields the lowest possible distortion when compared to all other code books of the same size
- Computational complexity increases exponentially with the size of vector blocks
- VQ blocks are usually small

Image Compression

Vector Quantization (VQ)

- The encoder searches the code book by minimizing the distortion between the original image block and the chosen vector from the codebook according to some distortion metric
- Search complexity increases with number of vectors in the code book
- VQ can be used in both spatial and frequency domains
- Lossy compression
- Each code word is called code vector and set of all code vectors is the code book
- VQ utilizes the two techniques:

1. Mapping technique (Grouping the values)
2. Designing a code book (Mechanism f mapping or entries of code words)

- Let the number of vectors is ***N*** and ***R***(bits/Pixel) be the rate and ***L*** (grouping) be the dimension then:

$$N=2^{RL}$$

(For problems refer notebook)*