Image Coding

Need data compression for:

- image transmission (communication)
- image storage (archival)
- efficient representation (feature extraction).

Applications:

- Transmission: Digital TV, picture phone, video-conferencing, tele-conferencing, tele-radiology and tele-medicine, space exploration, remote sensing, aerial reconnaissance
- Archival: Medical images, finger prints
- Features extraction: for pattern recognition and classification.

Image Coding

- A general compression system model

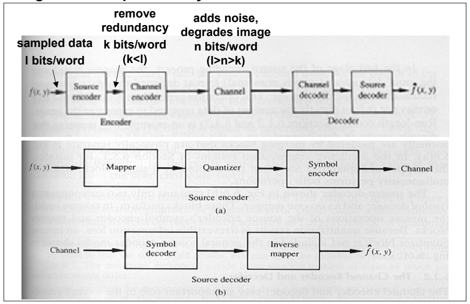


Image Coding

Concerns:

- Channel capacity, memory needs - to be minimized without sacrificing quality.

Strategies:

Image data compression possible due to:

- Coding redundancy (all codewords are not equally likely),
- Spatial redundancy or correlation in image data (pixels),
- Psychovisual redundancy present in images.

Image Coding

- Coding redundancy - Example

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$	$l_2(r_k).p_r(r_k)$
$r_0 = 0$	0.19	000	3	11	2	0.38
$r_1 = 1/7$	0.25	001	3	01	2	0.50
$r_2 = 2/7$	0.21	010	3	10	2	0.42
$r_3 = 3/7$	0.16	011	3	001	3	0.48
$r_4 = 4/7$	0.08	100	3	0001	4	0.32
$r_5 = 5/7$	0.06	101	3	00001	5	0.30
$r_6 = 6/7$	0.03	110	3	000001	6	0.18
$r_7 = 1$	0.02	111	3	000000	6	0.12
	.		$L1_{avg} = 3$			$L2_{avg} = 2.7$

$L_{avg} = \sum_{k=0}^{L-1} l(r_k)\, p_r(r_k)$ Average length of the code words

Image Coding

- Coding redundancy

$$R_D = 1 - \frac{1}{C_R} = 1 - \frac{1}{n_E / n_S}$$

$$\begin{array}{llll} n_S = n_E & \rightarrow\ C_R = 1, & R_D = 0 & 0\%\ \textit{no compression} \\ n_S << n_E & \rightarrow\ C_R \rightarrow \infty, & R_D \rightarrow 1 & 100\%\ \textit{high compression} \\ n_S >> n_E & \rightarrow\ C_R \rightarrow 0, & R_D \rightarrow -\infty & R_D < 0\ \textit{expansion} \end{array}$$

R_D: relative data redundancy of the data set 1,
C_R: compression ratio,
n_S, n_E: number of information carrying units in two data sets.

$$R_D = 1 - \frac{1}{3/2.7} = 1 - \frac{1}{1.11} = 1 - 0.90 = 0.099 \approx 0.1$$

~10% of the data using Code1 is redundant.

Image Coding

- Spatial (interpixel) redundancy

autocorrelation coefficient along one line: $\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)}$

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x,y) f(x, y+\Delta n)$$

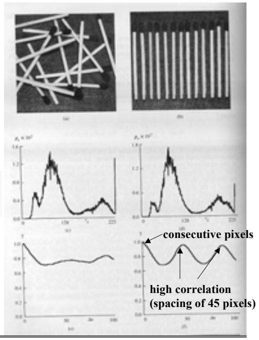
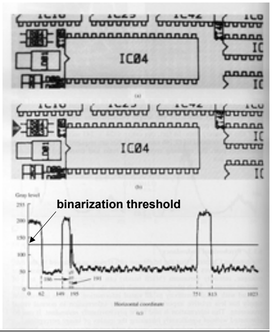


Image Coding

- **Spatial (interpixel) redundancy**



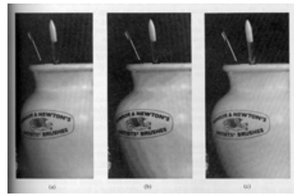
(gray level, run length)

Line 100:
(1,63)(0,87)(1,37)(0,5)(1,4)
(0,556)(1,62)(0,210)

run-length code

Image Coding

- **Psychovisual redundancy**
  - **the eye does not respond with equal sensitivity to all visual information**



Monochrome image with 256 gray leves	Uniform quantization to 16 levels	IGS quantization to 16 levels

Image Coding

Requirements:
- **Storage: coding must be reversible for exact reconstruction of image (usually).**
- **Transmission: Only reasonable fidelity expected - coding may be irreversible (minimum distortion).**
- **Feature extraction: Codes need only help in identification or classification of various types of images - image recovery not required.**

Difficulties:
- **Reversible codes can compress only ~ 2: 1;**
- **Definition of fidelity or distortion measures;**
- **Feature-space dimensionality.**

Image Coding

Information Theory Guidelines

- **A measure of information conveyed by a signal event should be related to the statistical uncertainty of the event giving rise to the information, rather than the semantic or structural content of the message or signal.**

Image Coding

Consider the occurrence of M gray levels in an image, with the probability of occurrence of the i^{th} gray level being Prob (m_i) =p_i, i = 1, 2, . . . ,M.

A measure information h(p) should be a function of p_i, satisfying the following criteria:
- **h(p) should be continuous for $0 < p < 1$,**
- **h(p) = ∞ for p = 0,**
- **h(p) = 0 for p = 1 (event allways occurs -> no uncertainty),**
- **$h(p_2) > h(p_1)$ if $p_2 < p_1$,**
- **if two independent image processes (or pixel) p_i and q_i are considered, $h(p_i, q_i) = h(p_i) + h(q_i)$.**

These requirements are met by $\boxed{h(p) = -\log_b p = \log_b 1/p}$

Image Coding

Entropy of a Source:

- **When a source generates a number of gray levels with different probabilities, we may define a measure of average information or "entropy" as the expected value of information contained in each possible level:**

$$H = \sum_{i=1}^{M} p_i h(p_i) = -\sum_{i=1}^{M} p_i \log_2 p_i \quad bits.$$

Image Coding

- **The maximum possible entropy occurs when all the gray levels occur with the same probability (= 1/M), i.e., the various gray levels are equally likely:**

$$H_{\max} = -\sum_{i=1}^{M} \frac{1}{M} \log_2 \frac{1}{M} = \log_2 M$$

- **The entropy of an image is an important measure as it gives the lower bound on the noise-free transmission rate.**

Note: If the number of gray levels M in an image is increased from 2 to 2^n bits, H_{max} increases from 1 to n bits.

Image Coding

Properties of Entropy:

$$H = -\sum_{i=1}^{M} p_i \log_2 p_i \quad bits.$$

- **$H(p) \geq 0$ with $H(p) = 0$ only for $p = 0$ or $p = 1$.**
- **$H(p_1, p_2, \ldots, p_M) \leq \log_b M$ with equality iff $p_i = 1/M$ for $i = 1, 2, \ldots, M$.**
- **Considering two pixels (or images) X with gray levels $\{x_i\}$ and Y with gray levels $\{y_i\}$, the average joint information or joint entropy is defined as**

$$H(X,Y) = -\sum_{i=1}^{M} \sum_{j=1}^{M} p(x_i, y_j) \log_2 p(x_i, y_j)$$

- **Then $H(X,Y) \leq H(X) + H(Y)$, with equality if X and Y are statistically independent.**

Image Coding

Properties of Entropy:

- **The conditional entropy of viewing one pixel X given that another element Y has been observed is**

$$H(X/Y) = -\sum_{i=1}^{M} \sum_{j=1}^{M} p(y_i) p(x_i / y_i) \log_2 p(x_i / y_j).$$

- **Then, $H(X/Y) = H(X,Y) - H(Y) \leq H(X)$ with equality if X and Y are statistically independent.**

Image Coding

- **The various pixels in an image may be considered to be symbols produced by a discrete information source with the gray levels as the states.**
- **If the successive pixels are independent, then**

$$H = -\sum_{i=1}^{M} p_i \log_2 p_i.$$

where p_i is the probability of the i^{th} gray level and $M = 2^n$.

- **While H_{max} = n bits, the actual entropy of a real image encountered in practice is considerably lesser as the pixels are not independent.**
- **Due to this reason, one must consider sequences of pixels to estimate the true entropy.**

Image Coding

Theorem:

- **Let $p(B_i)$ be the probality of a sequence B_i of the picture elements.**
- **Let**

$$H_N = -\frac{1}{N} \sum_i p(B_i) \log_2 p(B_i),$$

where the summations is over all possible sequences B_i with N pixels.

- **Then, H_N is a monotonically decreasing function of N, and**

$$\lim_{N \to \infty} H_N = H_R$$

where H_R is the real entropy of the source.

Shannon and Weaver, "*The Math. Th. Of Comm*". , B.S.T.J., 1949

Image Coding

4	4	4	4	4	4	4	0
4	5	5	5	5	5	4	0
4	5	6	6	6	5	4	0
4	5	6	7	6	5	4	0
4	5	6	6	6	5	4	0
4	5	5	5	5	5	4	0
4	4	4	4	4	4	4	0
4	4	4	4	4	4	4	0

First-Order Estimate

Gray level	Frequency	Probability	Entropy
0	8	0.125	0.375
1	0	0	0
2	0	0	0
3	0	0	0
4	31	0.484	0.507
5	16	0.250	0.500
6	8	0.125	0.375
7	1	0.016	0.096
Total	64	1.000	1.853

$H_1 = 1.853$

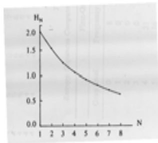
$$H = -\sum_{i=1}^{M} p_i \log_2 p_i.$$

Image Coding

Second-Order Estimate

Gray level Pairs	Frequency	Probability	Entropy
(4,4)	18	0.2813	0.515
(4,0)	8	0.1250	0.375
(0,4)	8	0.1250	0.375
(4,5)	5	0.0781	0.287
(5,4)	5	0.0781	0.287
(5,6)	3	0.0469	0.207
(6,5)	3	0.0469	0.207
(6,7)	1	0.0156	0.094
(7,6)	1	0.0156	0.094
(6,6)	4	0.0625	0.250
(5,5)	8	0.1250	0.375
Total	64	1.0000	3.066

$H_2 = \frac{1}{2}(3.066) = 1.533$



$$H_N = -\frac{1}{N}\sum_i p(B_i)\log_2 p(B_i),$$

Image Coding

A measure that is important in transmitting pictures over a communication system is "mutual information" defined as

$$\boxed{I(X, Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)}$$

This may be seen as the information transmitted:

- H (X) is the information input and
- H(X/Y) is the information about X, given that the received picture Y is known.

Then, if Y is totally correlated with X, H(X/Y) = 0, $\boxed{I(X/Y) = H(X)}$

If Y is independent of X, H(X/Y) = H(X), $\boxed{I(X/Y) = 0}$

Image Coding

- Suppose we wish to transmit an image sampled into an L x L array with B bits/pixel.
- Then a total of L^2B bits must be transmitted.
- Typically, L = 512, B = 8 $\rightarrow$ $L^2B \cong 2$ million bits!
- In high-resolution medical imaging (e.g., digital mammography), image sizes are of the order of L = 4000, B = 12 $\cong$ $L^2B \cong 200$ million bits!
- In the case of 3D medical images of size 512 x 512 x 512 voxels with 12 - 16 bits/voxel, we run into dataset sizes of 2 billion bits per image!
- This motivates one to consider efficient data compression techniques for transmission or archival of digital images.

Image Coding

Fundamental Concepts of Coding

- A *code* is a mapping of words from a source alphabet into the words of a code alphabet;
- A *word* is a finite sequence of symbols from an alphabet.
- A code is called *distinct* if each code word is distinguishable from the other code words.
- A distinct code is *uniquely decodable* if every code word is identifiable when immersed in a sequence of code words.
- A desirable property of a uniquely decodable code is that it should be decodable on a word-to-word basis. This is ensured if no code word may be a prefix to another; the code is then *instantaneously decodable*.
- A code is said to be *optimal* if it is instantaneously decodable and has the minimum average length for a given source PDF.

Image Coding

Noiseless Coding Theorem for Binary Transmission

- Given a code with an alphabet of two symbols and a source A with an alphabet of two symbols, the average length of the code words per source symbol may be made arbitrarily close to the lower bound (entropy) H(A) as desired by encoding sequences of source symbols rather than individual symbols.
- The average length L(n) of encoded n sequences is bounded

$$H_R(A) \le \frac{L(n)}{n} \le H_R(A) + \frac{1}{n}$$

- Difficulty in estimating entropy lies in the fact that pixels are statistically dependent point-to-point, line-to-line, and frame-to-frame.
- Computation of entropy requires the symbols to be considered in blocks over which the statistical dependence is negligible.

Image Coding

- If dependencies exist over T seconds and there are F frames/seconds (considering PCM digital TV, e.g.), the block to be considered has S = T.F.B.$L^2 \approx 3 \times 10^8$ bits !

  (T = 5, F = 30, B = 8, L = 512).
- Thus joint probability functions will have to be computed with a vector length of 300 million!
- As this is impossible, our approaches must be limited to single pixels or small blocks, resulting in over-estimates of H(A).
- If the blocks of pixels are chosen so that the sequence entropy estimates converge rapidly to the limit, then block-coding methods may provide results close to the minimum length.
- The entropy of most natural scenes; has been estimated to be less than 1 bit/pixel. Compression from 8 bits/pixel to 0.8 bits/pixel may be possible by optimal coding of blocks.

Image Coding

Distortion Measures and Fidelity Criteria

- The *binary symmetric* channel is characterized by a single parameter: the bit error probability *p*.
- The channel capacity is then C = 1 +*p* log p + *q log q*, where *q = 1 - p*
- Least squares single-letter fidelity criterion

$$\rho_n(x,y) = \frac{1}{n}\sum_{j=1}^{n}(x_j + y_j)^2\, 2^{j-1}$$

- where x and y the transmitted and received n-bit vectors (block or words)
- Hamming distance:

$$\rho_n(X,Y) = \frac{1}{n}\sum_{j=1}^{n}(x_j + y_j)^2$$

Image Coding

Measures of fidelity based upon whole images

$$e(x,y) = g(x,y) - f(x-y)$$

root-mean square error

$$e_{rms} = \sqrt{\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}[g(x,y) - f(x,y)]^2}$$

mean-square signal-to-noise ratio

$$SNR_{ms} = \frac{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} g^2(x,y)}{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} e^2(x,y)}$$

Image Coding

Reversible (Information-Preserving, Lossless) Coding
Huffman Coding

- The intuitive idea in Huffman coding is to use variable-length code words, and assign shorter code words to the more probable messages and longer code words to the less likely messages to obtain a code with low average length.
- The Huffman coding procedure guarantees a uniquely decodable code with the minimum average number of bits per message with H ≤ R ≤ H + 1.

Difficulties:

- Variable code word length.
- New code mapping required if source probabilities change.

Image Coding

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

Original source			Source reduction							
Sym.	Prob.	Code	1		2		3		4	
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

Image Coding

Run-length Coding

- Pixel-to-pixel correlation present in images is exploited by this simple, reversible code.
- In a digital image, a run is defined as a sequence of consecutive pixels of equal values in a specified direction.
- A reduction in average bit rate may be achieved by simply transmitting the values and lengths of such runs in the image.
- The method fails if the image is highly textured, leading to a higher bit rate than straight transmission of pixels.
- Run-length coding is best suited to binary (facsimile) images.

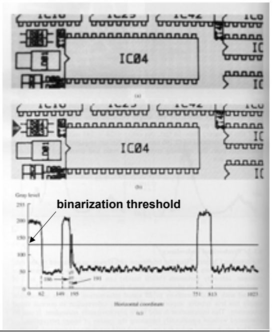
Image Coding

Run-length Coding

- The method is sensitive to errors in run lengths, leading to offsets.
- Line synchronization or standardized runs may be used to overcome this problem.
- Bit-plane images may be run-length encoded to achieve better performance than run-length encoding of the original gray-scale image.
- Two-dimensional area encoding methods may also be considered if the images have large areas of constant gray level.

Image Coding

- **Run-length Coding**



(gray level, run length)

Line 100:
(1,63)(0,87)(1,37)(0,5)(1,4)
(0,556)(1,62)(0,210)

Image Coding

Contour Encoding

- **Since a digital image has a finite number of gray levels, we may see it as being made up of a number of steps or plateaus, with the gray levels being the plateau heights.**
- **Knowledge of the heights, locations and shapes of the plateaus is equivalent to knowledge of the image.**
- **The left-most-looking rule for tracing a contour:**
  - **Look at the element to the left relative to the direction of entry; if this has the same value move to it;**
  - **else look at the element straight ahead ...**
  - **else look right ...else look back …**
- **Freeman's chain code may be used to encode the directions of movement in tracing the contour.**

Image Coding

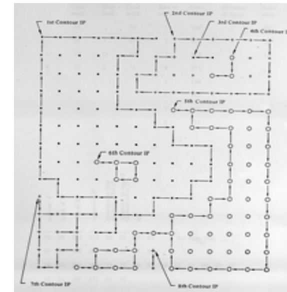


Image Coding

Contour #	Code word	Gray level	Code word	Row or column	Code word	Direction of travel	Code word
1	00	•	00	1	000	↑	00
2	01	×	01	2	001	→	01
3	10	○	10	3	010	↓	10
4	11	△	11	4	011	←	11
				5	100		
				6	101		
				7	110		
				8	111		

(a)

Image Coding

contour #	value of IP	row of IP	column of IP	direction to first element after IP	direction to second element after IP	etc.							
00	00	000	000	01	01	01	01	... (44 more bits) ...					
01	01	001	001	01	10	11	00						
10	10	001	100	10	10	01	01	10	11	11	00	00	00
11	11	101	011	01	11								

(b)

Image Coding

Lossy or Minimum Distortion Coding

- **In some image coding applications, a certain amount of distortion may be tolerable.**
- **Higher compression rates may be achieved then, subject to certain fidelity criteria.**

I. Transform Coding

- **Image transforms such as Fourier, Walsh-Hadamard, and Hotelling may be used as the mapping (decorrelating) process in the encoder.**
- **The purpose of the mapping is to reduce the correlation between pixels, and hence improve coding efficiency by processing the transform coefficients, which are independent of one another.**

Image Coding

The inverse transform may be written as $X = \sum_{k=1}^{n} \sum_{l=1}^{n} {}_{ykl} B_{kl}$,

and interpreted as a series expansion of the image X using n^2 n x n basis images:

$$B_{kl} = \begin{bmatrix} b_{kl11} & b_{kl21} & \dots & b_{kl1n} \\ b_{kl21} & b_{kl22} & \dots & b_{kl2n} \\ \vdots & & & \vdots \\ b_{k\ln 1} & b_{k\ln 2} & \dots & b_{k\ln n} \end{bmatrix}$$

with y_{kl}, K, l = 1, 2, . . . , n as the weighting coefficients of the expansion.

In the case of the Fourier transform, the basis images are 2D sinusoid.

Image Coding

II. Predictive and Interpolative Coding

- Since pixels in a local region in an image are highly correlated, one might expect that a currently considered pixel might be predicted from a knowledge of the previous pixels, or that it might be interpolated from a knowledge of a few surrounding pixels.
- Then, it should not be necessary to transmit all pixel values: a few initial pixel values and the prediction or interpolation coefficients should be adequate to facilitate reconstruction of the image.
- A simple linear predictor would produce large errors if only the starting values were available for reconstruction.
- Better results are achieved if the prediction error is also transmitted.
- The error image would require far fewer bits than the original image due to its smaller dynamic range

Image Coding

Consider forming an estimate of the pixel $\hat{f}(m,n)$ based upon the previous element f(m - 1, n) only, i.e., f(m,n) = α f(m-1, n)

$$MSE\ \bar{\varepsilon}^2 = E\{[f(m,n) - \hat{f}(m,n)]^2\}$$
$$= E\{f^2(m,n)\} - 2\alpha E\{f(m,n) f(m-1,n)\} + \alpha^2 E\{f^2(m-1,n)\}$$
$$= R(0,0) - 2\alpha R(1,0) + \alpha^2 R(0,0)$$

settings $\frac{\partial \bar{\varepsilon}^2}{\partial \alpha} = 0$, we get $\alpha_{opt} = \frac{R(1,0)}{R(0,0)}$.

The optimal first-order predictor is $\hat{f}(m,n) = \frac{R(1,0)}{R(0,0)} f(m-1,n)$.

In practice, predictor with a region of support including a few preceding pixel will be required to achieve low prediction errors.