ANSWERS. 4.08 Digital Image Proc.

1. a) Let $f(x,y)$ denote a digital image of size 256×256. In order to compress this image, we take its Discrete Cosine Transform $C(u,v), u,v=0,\ldots,255$ and keep only the Discrete Cosine Transform coefficients for $u,v=0,\ldots,n$ with $0\le n<255$. The percentage of total energy of the original image that is preserved in that case is given by the formula $an+b+85$ with a,b constants. Furthermore, the energy that is preserved if $n=0$ is 85%. Find the constants a,b. **[4]**

ANSWER

For $n=0$ it is given that the preserved energy is 85%. This is the case where only the (0,0) frequency pair is kept. Therefore,

$a\cdot0+b+85=85\Rightarrow b=0$

In case where the entire DCT is kept we have $n=255$ and the preserved energy should be 100%. In that case:

$$a\cdot255+85=100\Rightarrow 255a=15\Rightarrow a=\frac{1}{17}\Rightarrow b=0$$

b) Let $f(x,y)$ denote an $M\times N$-point 2-D sequence that is zero outside $0\le x\le M-1$, $0\le y\le N-1$, where M and N are integers and powers of 2. In implementing the standard Discrete Walsh Transform of $f(x,y)$, we relate $f(x,y)$ to a new $M\times N$-point sequence $H(u,v)$.

(i) State the main disadvantage of the Discrete Hadamard Transform. **[2]**

ANSWER

Bookwork. The main disadvantage of Hadamard Transform is the fact that it doesn't possess the property of energy compaction. This is because the basis functions of the Hadamard matrix are not exhibiting increasing sequency as the value of their index increases.

(ii) In the case of $M=N=2$ and $f(x,y)=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}$ calculate the Hadamard transform coefficients. **[4]**

ANSWER

Bookwork, use the definition of Hadamard Transform $H(u,v)=\begin{bmatrix}4 & -1\\ -1 & 0\end{bmatrix}$.

$$H(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)(-1)^{\sum_{i=0}^{n-1}(b_i(x)b_i(u)+b_i(y)b_i(v))}\quad \text{with } n=1$$

$$H(0,0)=\frac{1}{2}\sum_{x=0}^{1}\sum_{y=0}^{1}f(x,y)(-1)^{\sum_{i=0}^{0}(b_i(x)b_i(u)+b_i(y)b_i(v))}=\frac{1}{2}\sum_{x=0}^{1}\sum_{y=0}^{1}f(x,y)(-1)^{xu+yv}$$

$$=\frac{1}{2}\sum_{x=0}^{1}\sum_{y=0}^{1}f(x,y)(-1)^{0}=4$$

$$H(0,1)=\frac{1}{2}\sum_{x=0}^{1}\sum_{y=0}^{1}f(x,y)(-1)^{\sum_{i=0}^{0}(b_i(x)b_i(u)+b_i(y)b_i(v))}$$

$$=\frac{1}{2}f(0,0)(-1)^{0\cdot0+0\cdot1}+\frac{1}{2}f(0,1)(-1)^{0\cdot0+1\cdot1}+\frac{1}{2}f(1,0)(-1)^{1\cdot0+0\cdot1}+\frac{1}{2}f(1,1)(-1)^{1\cdot0+1\cdot1}$$

$$=\frac{1}{2}f(0,0)+\frac{1}{2}f(0,1)(-1)+\frac{1}{2}f(1,0)+\frac{1}{2}f(1,1)(-1)=\frac{1}{2}1+\frac{1}{2}2(-1)+\frac{1}{2}2+\frac{1}{2}3(-1)=-1$$

$$H(1,0)=\frac{1}{2}f(0,0)(-1)^{0\cdot 1+0\cdot 0}+\frac{1}{2}f(0,1)(-1)^{0\cdot 1+1\cdot 0}+\frac{1}{2}f(1,0)(-1)^{1\cdot 1+0\cdot 0}+\frac{1}{2}f(1,1)(-1)^{1\cdot 1+1\cdot 0}$$
$$=\frac{1}{2}f(0,0)+\frac{1}{2}f(0,1)+\frac{1}{2}f(1,0)(-1)+\frac{1}{2}f(1,1)(-1)$$
$$=\frac{1}{2}1+\frac{1}{2}2+\frac{1}{2}2(-1)+\frac{1}{2}3(-1)=-1$$
$$H(1,1)=\frac{1}{2}f(0,0)(-1)^{0\cdot 1+0\cdot 1}+\frac{1}{2}f(0,1)(-1)^{0\cdot 1+1\cdot 1}+\frac{1}{2}f(1,0)(-1)^{1\cdot 1+0\cdot 1}+\frac{1}{2}f(1,1)(-1)^{1\cdot 1+1\cdot 1}$$
$$=\frac{1}{2}f(0,0)+\frac{1}{2}f(0,1)(-1)+\frac{1}{2}f(1,0)(-1)+\frac{1}{2}f(1,1)(-1)$$
$$=\frac{1}{2}1+\frac{1}{2}2(-1)+\frac{1}{2}2(-1)+\frac{1}{2}3=0$$

Therefore, $H(u,v)=\begin{bmatrix}4 & -1\\ -1 & 0\end{bmatrix}$

c) Consider the population of vectors $\underline{f}$ of the form

$$\underline{f}(x,y)=\begin{bmatrix}f_1(x,y)\\ f_2(x,y)\\ f_3(x,y)\end{bmatrix}.$$

Each component $f_i(x,y)$, $i=1,2,3$ represents an image of size $M\times M$ where M is even. The population arises from the formation of the vectors $\underline{f}$ across the entire collection of pixels (x,y). The three images are defined as follows:

$$f_1(x,y)=\begin{cases}r_1 & 1\le x\le \frac{M}{2}, 1\le y\le M\\ r_2 & \frac{M}{2}<x\le M, 1\le y\le M\end{cases}$$

$$f_2(x,y)=r_3,\ 1\le x\le M,\ 1\le y\le M$$

$$f_3(x,y)=r_4,\ 1\le x\le M,\ 1\le y\le M$$

Consider now a population of random vectors of the form

$$\underline{g}(x,y)=\begin{bmatrix}g_1(x,y)\\ g_2(x,y)\\ g_3(x,y)\end{bmatrix}$$

where the vectors $\underline{g}$ are the Karhunen-Loeve (KL) transforms of the vectors $\underline{f}$.

(i) Find the images $g_1(x,y)$, $g_2(x,y)$ and $g_3(x,y)$ using the Karhunen-Loeve (KL) transform. **[8]**

ANSWER

$$f_1(x,y)=\begin{cases}r_1 & 1\le x\le \frac{M}{2}, 1\le y\le M\\ r_2 & \frac{M}{2}<x\le M, 1\le y\le M\end{cases}$$

Mean value of $f_1(x,y)$ **is** $m_1=\frac{r_1}{2}+\frac{r_2}{2}$. **Zero-mean version of** $f_1(x,y)$ **is**

$$f_1(x,y)-m_1=\begin{cases}\frac{r_1}{2}-\frac{r_2}{2} & 1\le x\le\frac{M}{2},1\le y\le M \\ \frac{r_2}{2}-\frac{r_1}{2} & \frac{M}{2}<x\le M,1\le y\le M\end{cases}$$

Mean value of $f_2(x,y)$ is r_3. Zero-mean version of $f_2(x,y)$ is $f_2(x,y)-m_2=0$.
Mean value of $f_3(x,y)$ is r_4. Zero-mean version of $f_3(x,y)$ is $f_3(x,y)-m_3=0$.
Variance of $f_1(x,y)-m_1$ is $\frac{1}{2}\frac{1}{4}(r_1-r_2)^2+\frac{1}{2}\frac{1}{4}(r_1-r_2)^2=\frac{1}{4}(r_1-r_2)^2$.
Variance of $f_2(x,y)-m_2$ is 0.
Variance of $f_3(x,y)-m_3$ is 0.
Covariance between $f_1(x,y)-m_1$ and $f_2(x,y)-m_2$ is 0. Therefore, the covariance matrix is $\begin{bmatrix}\frac{1}{4}(r_1-r_2)^2 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$ with eigenvalues $\frac{1}{4}(r_1-r_2)^2$ and 0. Therefore, by using the Karhunen Loeve transform we produce three new images, with two of them being 0 and the other being $f_1(x,y)-m_1$.

(ii) Comment on whether you could obtain the result of c)-i) above using intuition rather than by explicit calculation.

[2]

ANSWER
The above result is expected since two of the given images are constant and therefore they don't carry any information. This means that there is only one principal component in the given set.

2. a) A 3×3 two dimensional filter is separable if its response can be written in the form:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}\begin{bmatrix} d & e & f \end{bmatrix}$$

where a,b,c,d,e,f are constant coefficients. Consider the 3×3 two dimensional filters with coefficients shown below:

$$\frac{1}{9}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9A-1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

with A a constant real number slightly greater than 1.

For each filter given above, answer the following questions.

(i) Is it a separable filter? If yes, present the vertical and horizontal filters. **[3]**

ANSWER

The first filter is not separable and the second filter is separable since it can be written as $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$

(ii) What is the functionality of the filter? Explain your reasoning. **[3]**

ANSWER

The first filter is a High Boost Filter which enhances the edges of an image without eliminating the background completely.

The second filter is a derivative filter that detects horizontal edges.

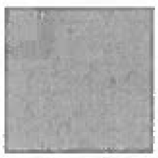
b) Consider a grey level image $f(x,y)$ of size 256×256 with $1\le x,y\le 256$, which has the following intensities:

$$f(x,y)=\begin{cases} r+1 & 1\le x\le 12 \text{ and } 1\le y\le 12 \\ r & 13\le x\le 16, 1\le y\le 16 \text{ and } 1\le x\le 12, 13\le y\le 16 \\ r+2 & \text{elsewhere} \end{cases}$$

with $0\le r\le 253$.

(i) Sketch the image and comment on its visual appearance. Justify your answer. **[3]**

ANSWER



The three intensities are very close to each other so their differences are not large enough to be perceived by the human eye. Therefore, the above image should appear constant with a grey level around r.

(ii) Apply global histogram equalisation on the above image. Comment on the visual appearance of the resulting equalised image. **[3]**

ANSWER

$$p(r)=\frac{256-144}{256\cdot 256}=\frac{112}{256\cdot 256}=\frac{7}{16\cdot 256}$$

$$p(r+1)=\frac{144}{256\cdot 256}=\frac{9}{16\cdot 256}$$

$$p(r+2)=\frac{256\cdot 256-256}{256\cdot 256}=\frac{255}{256}$$

Therefore we get:

$$T(r) = p(r) = \frac{7}{16 \cdot 256}$$

$$T(r+1) = \frac{7+9}{16 \cdot 256} = \frac{1}{256}$$

$$T(r+2) = 1$$

By multiplying with 255 we get the new image intensities as follows:

$$r \to \frac{7}{16}\frac{255}{256} = 0$$

$$(r+1) \to \frac{255}{256} = 1$$

$$(r+2) \to 255$$

In the resulting image we will still not be able to distinguish the two new intensities which arise from r and $r+1$.

(iii) Apply local histogram equalisation on the above image using non-overlapping image patches of size 16×16. Comment on the visual appearance of the resulting equalised image. **[3]**

For the top left part of the image of size 16×16 we will get:

$$p(r) = \frac{256-144}{256} = \frac{112}{256} = \frac{7}{16}$$

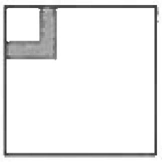
$$p(r+1) = 1$$

By multiplying with 255 we get the new image intensities as follows:

$$r \to \frac{7}{16} 255 = 112$$

$$(r+1) \to 255$$

The rest of the image will turn white, i.e., it will be of intensity 255.
Therefore, the locally equalised will look as below:



(iv) Based on the above observations, which of the two types of equalisations processes would you choose for the visual improvement of the particular image? Justify your answer. **[2]**

ANSWER

Obviously local histogram equalisation is able to extract the local pattern on the top left part and therefore, it is preferable.

c) Propose a method that uses variable sizes spatial filters to reduce background noise without blurring the image significantly.

[3]

We can apply low pass filters only within low contrast areas (uniform backgrounds and weak texture) and not high contrast areas (edge areas). We can measure the activity around each pixel using the local variance of a neighborhood around the pixel.

3. We are given the degraded version g of an image f such that in lexicographic ordering

$$g = Hf + n$$

where H is the degradation matrix which is assumed to be block-circulant and n is the noise term which is assumed to be zero-mean, white and independent of the image f. The images have size $N \times N$.

a) (i) Consider the Inverse Filtering image restoration technique. Give the general expressions for both the Inverse Filtering estimator and the restored image in both spatial and frequency domains and explain all symbols used. **[5]**

ANSWER

Bookwork

(ii) In a particular scenario the degradation process can be modelled as a linear filter with the two dimensional impulse response given below:

$$h(x,y) = \begin{cases} 1 & -1 \le x \le 1,\, y = 0 \\ 0 & \text{elsewhere} \end{cases}$$

Estimate the frequency pairs for which Inverse Filtering cannot be applied. **[5]**

ANSWER

Inverse Filtering cannot be applied if $H(u,v) = 0$.

$$H(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} h(x,y)e^{-j2\pi(ux/M+vy/N)} = \frac{1}{MN}(e^{-j2\pi u/M} + e^{j2\pi u/M} + 2)$$

$$= \frac{2}{MN}[\cos(2\pi u/M) + 1]$$

$$\cos(2\pi u/M) + 1 = 0 \Rightarrow 2\pi u/M = (2k+1)\pi \Rightarrow u = \frac{M(2k+1)}{2},\, k = 0,1,2,\ldots$$

b) (i) Consider the Constrained Least Squares (CLS) Filtering image restoration technique. Give the general expressions for both the CLS filter estimator and the restored image in both spatial and frequency domains and explain all symbols used. **[5]**

ANSWER

Bookwork

(iii) In a particular scenario, the degradation process can be modelled as a linear filter with the transfer function given below:

$$H(u,v) = \sqrt{2\pi}\sigma(u^2+v^2)e^{-j2\pi^2\sigma^2(u^2+v^2)}$$

In the above formulation σ is a constant parameter. Generate the expression of the CLS filter in frequency domain by assuming that the high pass filter used in CLS is a Laplacian filter.

[5]

ANSWER

$$H^*(u,v) = \sqrt{2\pi}\sigma(u^2+v^2)e^{j2\pi^2\sigma^2(u^2+v^2)}$$

$$|H(u,v)|^2 = H(u,v)H^*(u,v) = 2\pi\sigma^2(u^2+v^2)^2$$

The CLS filter in frequency domain is $\dfrac{\sqrt{2\pi}\sigma(u^2+v^2)e^{j2\pi^2\sigma^2(u^2+v^2)}}{2\pi\sigma^2(u^2+v^2)^2 + \alpha|C(u,v)|^2}$ **where** $C(u,v)$ **is the Laplacian filter in frequency domain.**

4. a) (i) Give the definition of a Discrete Memoryless Source (DMS). **[3]**

ANSWER

The discrete memoryless source (DMS) has the property that its output at a certain time does not depend on its output at any earlier time.

(ii) Consider a set of symbols generated from a DMS. Give the minimum number of bits per symbol that we can achieve if we use Huffman coding for the binary representation of the symbols. Explain what type of probabilities the symbols must possess in order to achieve the minimum number of bits per symbol. **[3]**

ANSWER

The minimum number of bits per symbol we can achieve is the entropy of the source. In order to achieve this, the probabilities of the symbols must be negative powers of 2.

(iii) Provide a scenario where Huffman coding would not reduce the number of bits per symbol from that achieve using fixed number of bits per symbol. **[2]**

ANSWER

This happens when the probabilities of the symbols are equal.

b) The following Figure 4a shows a list 7 symbols and their probabilities. It is assumed that this set of symbols are generated by a Discrete Memoryless Source (DMS).

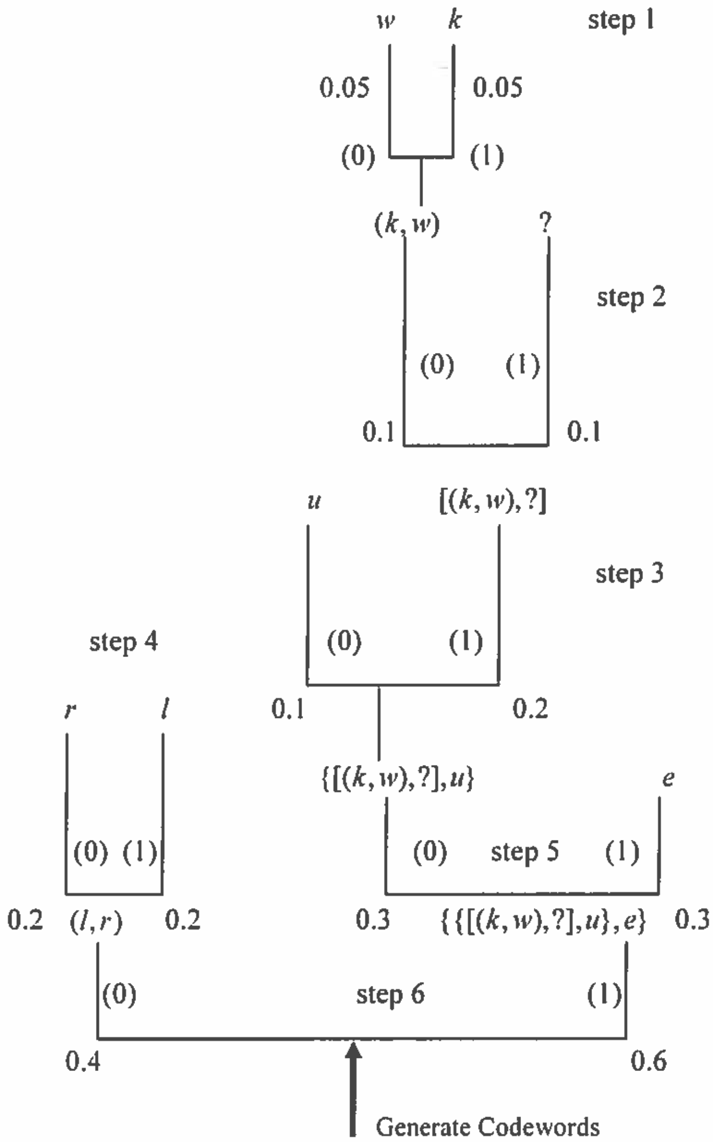
Symbol	**Probability**
k	0.05
l	0.2
u	0.1
w	0.05
e	0.3
r	0.2
?	0.1

Figure 4a

(i) Derive a Huffman code taking into consideration that the probability of a 1 being transmitted as 0 is zero and the probability of a 0 being transmitted as a 1 is 0.05. **[4]**

ANSWER

While merging branches 1 is preferred to 0 since it has zero probability of being wrongly transmitted and therefore, we assign 1s to branches that correspond to symbols with higher probabilities.



	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
k 0.05	e 0.3	e 0.3	e 0.3	e 0.3	(l,r) 0.4	$\{\{[(k,w),?],u\},e\}$ 0.6
l 0.2	l 0.2	l 0.2	l 0.2	$\{[(k,w),?],u\}$ 0.3	e 0.3	(l,r) 0.4
u 0.1	r 0.2	r 0.2	r 0.2	l 0.2	$\{[(k,w),?],u\}$ 0.3	
w 0.05	u 0.1	u 0.1	$[(k,w),?]$ 0.2	r 0.2		
e 0.3	? 0.1	? 0.1	u 0.1			
r 0.2	k 0.05	(k,w) 0.1				
? 0.1	w 0.05					

Symbol	Probability	Codeword
k	0.05	10101
l	0.2	01
u	0.1	100
w	0.05	10100
e	0.3	11
r	0.2	00
?	0.1	1011

(ii) Calculate the compression ratio. [4]

ANSWER

In this example, the average codeword length is 2.6 bits per symbol. In general, the average codeword length is defined as

$$l_{avg} = \sum l_i p_i$$

where l_i is the codeword length (in bits) for the codeword corresponding to symbol s_i. The average codeword length is a measure of the compression ratio. Since our alphabet has seven symbols, a fixed-length coder would require at least three bits per codeword. In this example, we have reduced the representation from three bits per symbol to 2.6 bits per symbol; thus the corresponding compression ratio can be stated as 3/2.6=1.15.

(ii) In the particular transmission system described in b)(i) above, find the probability of a codeword equal to 100 being transmitted wrongly. [4]

ANSWER

This can happen if one of the 0 is transmitted as 1 or both of them. Therefore, the probability of this event is 0.05+0.05+0.05*0.05=0.1+0.0025=0.1025.