Paper Number(s): **E4.08**
SO14
ISE4.33

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2001

MSc and EEE/ISE PART IV: M.Eng. and ACGI

DIGITAL IMAGE PROCESSING

Friday, 18 May 10:00 am

There are FOUR questions on this paper.

Answer THREE questions.

Time allowed: 3:00 hours

Corrected Copy

Examiners: Stathaki,T. and Clarke,T.J.W.

Special instructions for invigilators:	None
Information for candidates:	None

1. (a) Let $f(x,y)$ denote an $N \times N$-point 2-D sequence, that has zero value outside $0 \le x \le N-1$, $0 \le y \le N-1$, where N is an integer power of 2. In implementing the standard Walsh transform of $f(x,y)$, we relate $f(x,y)$ to a new $N \times N$-point sequence $W(u,v)$.

(i) Define the sequence $W(u,v)$ in terms of $f(x,y)$. **[5]**

(ii) For the case $N = 2$ and $f(x,y) = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ calculate the forward Walsh transform coefficients. **[20]**

(b) Consider the population of random vectors $\underline{f}$ of the form

$$\underline{f} = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \\ \vdots \\ f_N(x,y) \end{bmatrix}, N \ge 2.$$

Each component $f_i(x,y)$ represents an image and (x,y) denotes a randomly chosen pixel. The population arises from their formation across the entire collection of pixels.
Consider now a population of random vectors of the form

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

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

(i) Prove and explain the relationship between the covariance matrix of $\underline{g}$ and the covariance matrix of $\underline{f}$. What is the structure of the covariance matrix of $\underline{g}$? **[15]**

(ii) Can the elements of the covariance matrix of $\underline{g}$ be negative? Justify your answer. **[10]**

(iii) Suppose that $N = 8$ and the eigenvalues of the covariance matrix of $\underline{f}$ are

$$[6.1 \ 168 \ 0.08 \ 13 \ 64 \ 214 \ 1.2 \ 0.2].$$

What will be the mean square error if we use principal component images associated with the largest eigenvalues for 2:1 and 4:1 data compression? **[20]**

(c) Consider again Question 1(b) in the case of two images ($N = 2$). The covariance matrix of the population is $\underline{C}_f$ with elements defined as $C_{m,n} = \rho^{|m-n|}, 1 \le m,n \le 2, 0 < \rho < 1$.

(i) What are the variances of the images $g_1(x,y)$ and $g_2(x,y)$? **[20]**

(ii) What will be the mean square error as a function of ρ if we use the principal component image from the set of images $g_1(x,y)$ and $g_2(x,y)$ to reconstruct the original images? **[10]**

2. (a) (i) Sketch a possible histogram of a low contrast image and a possible histogram of a high contrast image. **[10]**

(ii) Describe the technique of histogram equalisation. **[10]**

(iii) Explain why the histogram of a discrete image is not flat after histogram equalisation. **[5]**

(iv) An image has the gray level probability density function $p_r(r) = \dfrac{e^r - 1}{e^1 - 2}, 0 \le r \le 1$. It is desired to transform the gray levels of this image so that they will have the specified probability density function $p_z(z) = 2z, 0 \le z \le 1$. Assume continuous quantities and find the transformation (in terms of r and z) that will accomplish this. **[25]**

(b) (i) Explain why smoothing typically blurs image edges. **[10]**

(ii) Show using a simple example that median filters are non-linear filters. **[20]**

(c) Let $f(x, y)$ denote an image of size $N \times N$ corrupted by zero mean random noise. Describe a spatially adaptive low-pass filtering technique that attempts to reduce the amount of noise without destroying the edges. **[20]**

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, independent and white.

(a) Describe the image restoration technique of pseudoinverse filtering. **[30]**

(b) Explain the effect of the regularization term in CLS filtering. Explain the advantages and drawbacks that arise from the use of (i) a large regularization parameter and (ii) a small regularization parameter. **[30]**

(c) Describe a spatially adaptive image restoration technique that attempts to eliminate the problem of noise amplification without destroying the edges of the original image. **[40]**

4. (a) (i) Consider an image $f(x, y)$ with intensity of each pixel r that can be modelled as a sample obtained from the probability density function $p_r(r) = 2r, 0 \le r \le 1$.

Suppose five reconstruction levels are assigned to quantize the intensity r. Determine these reconstruction levels using a uniform quantizer. **[10]**

(ii) Determine the codeword to be assigned to each of the five reconstruction levels using Huffman coding. Specify what the reconstruction level is for each codeword. **[20]**

(iii) For your codeword assignment in (ii), determine the average number of bits required to represent r. **[10]**

(iv) Determine the entropy and the redundancy of the Huffman code for this example. **[10]**

(b) Suppose that we have a binary representation for the Huffman codewords of a set of symbols.

(i) Explain why the Huffman code is not unique. **[10]**

(ii) Suppose that the Huffman codewords of the set of symbols of Question 4(a) are to be transmitted over a noisy channel where the probability of error of a one being received as a zero is higher than the probability of error of a zero being received as a one. Choose the Huffman codeword set with the smallest error rate under these conditions. **[20]**

(c) Describe the technique of run-length coding for bitonal images. **[20]**

E 4.08

SOLUTIONS - DIGITAL IMAGE PROCESSING May 01

QUESTION 1

(a)

(i) Forward

$$W(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}\right] \text{ or}$$

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

Inverse

$$f(x,y)=\frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}W(u,v)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}\right] \text{ or}$$

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

[5]

(ii)

$$W(u,v)=\frac{1}{2}\sum_{x=0}^{1}\sum_{y=0}^{1}f(x,y)\left[(-1)^{(b_0(x)b_0(u)+b_0(y)b_0(v))}\right]$$

$$=\frac{1}{2}\left(f(0,0)\left[(-1)^{(b_0(0)b_0(u)+b_0(0)b_0(v))}\right]+f(0,1)\left[(-1)^{(b_0(0)b_0(u)+b_0(1)b_0(v))}\right]\right)$$

$$+\frac{1}{2}\left(f(1,0)\left[(-1)^{(b_0(1)b_0(u)+b_0(0)b_0(v))}\right]+f(1,1)\left[(-1)^{(b_0(1)b_0(u)+b_0(1)b_0(v))}\right]\right)$$

$$=\frac{1}{2}(f(0,0)+f(0,1)(-1)^{b_0(v)}+f(1,0)(-1)^{b_0(u)}+f(1,1)(-1)^{b_0(u)+b_0(v)})\Rightarrow$$

$$W(0,0)=\frac{1}{2}(f(0,0)+f(0,1)+f(1,0)+f(1,1))=4$$

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

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

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

[20]

(b)

(i) The mean vector of the population is defined as

$$\underline{m}_f=E\{\underline{f}\}\Rightarrow\begin{bmatrix}m_1\\m_2\\\vdots\\m_n\end{bmatrix}=\begin{bmatrix}E\{f_1\}\\E\{f_2\}\\\vdots\\E\{f_n\}\end{bmatrix}$$

The covariance matrix of the population is defined as

$$\underline{C}_f=E\{(\underline{f}-\underline{m}_f)(\underline{f}-\underline{m}_f)^T\}$$

For M vectors from a random population, where M is large enough, the mean vector and covariance matrix can be approximately calculated by summations

$$\underline{m}_f=\frac{1}{M}\sum_{k=1}^{M}\underline{f}_k,\quad \underline{C}_f=\frac{1}{M}\sum_{k=1}^{M}\underline{f}_k\underline{f}_k^{\,T}-\underline{m}_f\underline{m}_f^{\,T}$$

Very easily it can be seen that $\underline{C}_f$ is real and symmetric. In that case a set of n orthonormal eigenvectors always exists.

Let $\underline{A}$ be a matrix whose rows are formed from the eigenvectors of $\underline{C}_f$, ordered so that the first row of $\underline{A}$ is the eigenvector corresponding to the largest eigenvalue, and the last row the eigenvector corresponding to the smallest eigenvalue.
The Karhunen-Loeve transform maps the vectors $\underline{f}$'s into vectors $\underline{g}$'s with the relationship

$$\underline{g} = \underline{A}(\underline{f} - \underline{m}_f)$$

The mean of the $\underline{g}$ vectors resulting from the above transformation is zero ($\underline{m}_g = \underline{0}$) and the covariance matrix is $\underline{C}_g = \underline{A}\underline{C}_x\underline{A}^T$, where $\underline{C}_g$ is a diagonal matrix whose elements along the main diagonal are the eigenvalues of $\underline{C}_f$

$$\underline{C}_g = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

The off-diagonal elements of the covariance matrix are 0, so the elements of the $\underline{g}$ vectors are uncorrelated. **[15]**

(ii) The element λ_i represents the variance of the image $g_i(x,y)$. Therefore it cannot be negative. If λ_i has very small value then the variance of the image $g_i(x,y)$ and the image $g_i(x,y)$ carries no information. **[10]**

(iii) Mean square error for compression 2:1 is $\sum_{j=5}^{8} \lambda_j = 7.58$.

Mean square error for compression 4:1 is $\sum_{j=3}^{8} \lambda_j = 84.58$. **[20]**

(c)

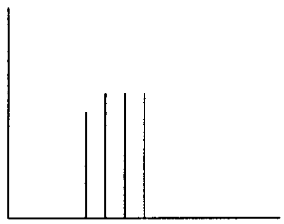
(i) The covariance matrix of $\underline{f}$ is $\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ and the eigenvalues of this matrix which are the variances of the images $g_1(x,y)$ and $g_2(x,y)$ are $\lambda_1 = 1+\rho, \lambda_2 = 1-\rho$. **[20]**

(ii) The error is $1-\rho$. **[10]**

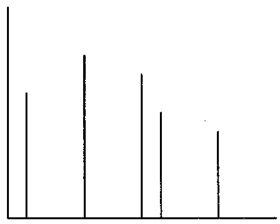
QUESTION 2

(a)

(i)



Low contrast



High contrast

[10]

(ii) Suppose we divide the grey levels in the image with the maximum value $L-1$.
Let the variable r represent the new grey levels in the image, where now $0 \le r \le 1$.
We apply the following transformation function:

$$s = T(r) = \int_0^r p_r(w)dw\,,\ 0 \le r \le 1$$

If $P_r(r)$, $P_s(s)$ are the probability distribution functions of r, s respectively, then

$$P_r[r, r+dr] \cong p_r(r)dr$$

$$P_s[s, s+ds] \cong p_s(s)ds$$

$$p_r(r)dr \overset{r=T^{-1}(s)}{=} p_s(s)ds \Rightarrow p_s(s) = p_r(r)\frac{dr}{ds}\bigg|_{r=T^{-1}(s)}$$

$$\frac{ds}{dr} = p_r(r) \Rightarrow p_s(s) = \left[p_r(r)\frac{1}{p_r(r)}\right]_{r=T^{-1}(s)} = 1,\ 0 \le s \le 1$$

$p_s(s)$ is a uniform density in the interval of definition of the transformed variable s.
The discrete form of histogram equalisation is given by the relation

$$s_k = T(r_k) = \sum_{j=0}^{k}\frac{n_j}{N^2} = \sum_{j=0}^{k} p_r(r_j),\ 0 \le r_k \le 1,\ k = 0,1,\ldots,L-1$$ **[10]**

(iii) The improvement over the original image is quite evident. Note that the new histogram is not flat because of the discrete approximation of histogram equalisation. Note, however, that the grey levels of an image that has been subjected to histogram equalisation are spread out and always reach white. This process increases the dynamic range of grey levels and produces an increase in image contrast. **[5]**

(iv) In the specific example $p_r(r) = \dfrac{e^r - 1}{e-2}, 0 \le r \le 1$ and $p_z(z) = 2z, 0 \le z \le 1$.

$$s = G(z) = \int_0^z p_z(w)dw = z^2$$

$$s = T(r) = \int_0^r \frac{e^w - 1}{e-2}dw = \int_0^r \frac{e^w}{e-2}dw - \int_0^r \frac{1}{e-2}dw = \frac{e^r - 1}{e-2} - \frac{r}{e-2}$$

$z_{1,2} = G^{-1}(s) = \pm\sqrt{s}$. Because $0 \le z \le 1$ we keep $z = \sqrt{s} = \sqrt{\dfrac{e^r - 1}{e-2} - \dfrac{r}{e-2}}$. **[25]**

(b)

(i) Because it eliminates high frequency components which are related to abrupt changes within an image, i.e. the edges of the image. **[10]**

(ii) median{2,3,8}=3
median{8,4,2}=4
median{10,7,10}=10$\neq$7=median{2,3,8}+median{8,4,2} **[20]**

(c)

To protect the edges from blurring while smoothing, a directional averaging filter can be useful. Spatial averages $g(x, y:\theta)$ are calculated in several selected directions (for example could be horizontal, vertical, main diagonals)

$$g(x, y:\theta) = \frac{1}{N_\theta} \sum_{(k,l) \in W_\theta} \sum f(x-k, y-l)$$

and a direction θ^* is found such that

$|f(x, y) - g(x, y:\theta^*)|$ is minimum. (Note that W_θ is the neighborhood along the direction θ and N_θ is the number of pixels within this neighborhood).

Then by replacing $g(x, y:\theta)$ with $g(x, y:\theta^*)$ we get the desired result. **[20]**

QUESTION 3

(a)
In inverse filtering the restored image is given by

$$F(u,v)=\frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2}\Rightarrow f(i,j)=\Im^{-1}\left[\frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2}\right]$$

Suppose first that the additive noise $n(i,j)$ is negligible. A problem arises if $H(u,v)$ becomes very small or zero for some point (u,v) or for a whole region in the (u,v) plane. In that region inverse filtering cannot be applied. Note that in most real applications $H(u,v)$ drops off rapidly as a function of distance from the origin.
In the presence of external noise we have that

$$\hat{F}(u,v)=\frac{H^*(u,v)(Y(u,v)-N(u,v))}{|H(u,v)|^2}=$$

$$\frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2}-\frac{H^*(u,v)N(u,v)}{|H(u,v)|^2}\Rightarrow$$

$$\hat{F}(u,v)=F(u,v)-\frac{N(u,v)}{H(u,v)}$$

If $H(u,v)$ becomes very small, the term $N(u,v)$ dominates the result.
The solution is to carry out the restoration process in a limited neighborhood about the origin where $H(u,v)$ is not very small.
This procedure is called pseudoinverse filtering.
In that case we set

$$\hat{F}(u,v)=\begin{cases}\dfrac{H^*(u,v)Y(u,v)}{|H(u,v)|^2} & H(u,v)\neq 0\\ 0 & H(u,v)=0\end{cases}$$

In general, the noise may very well possess large components at high frequencies (u,v), while $H(u,v)$ and $Y(u,v)$ normally will be dominated by low frequency components. **[30]**

(b)
Constrained least squares (CLS) restoration can be formulated by choosing an $\mathbf{f}$ to minimize the Lagrangian

$$\min\left(\|\mathbf{y}-\mathbf{Hf}\|^2+\alpha\|\mathbf{Cf}\|^2\right)$$

α represents either a Lagrange multiplier or a fixed parameter known as regularisation parameter.
α controls the relative contribution between the term $\|\mathbf{y}-\mathbf{Hf}\|^2$ and the term $\|\mathbf{Cf}\|^2$.
The minimization of the above leads to the following estimate for the original image

$$\mathbf{f}=\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}+\alpha\mathbf{C}^{\mathrm{T}}\mathbf{C}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y}$$

With larger values of α, and thus more regularisation, the restored image tends to have more ***ringing***.
With smaller values of α, the restored image tends to have more ***amplified noise effects***. **[40]**

(c)
The functional to be minimized takes the form

$$M(\mathbf{f},\alpha)=\|\mathbf{y}-\mathbf{Hf}\|^2_{\mathrm{W}_1}+\alpha\|\mathbf{Cf}\|^2_{\mathrm{W}_2}$$

where

$\|\mathbf{y}-\mathbf{Hf}\|^2_{\mathrm{W}_1}=(\mathbf{y}-\mathbf{Hf})^{\mathrm{T}}\mathbf{W}_1(\mathbf{y}-\mathbf{Hf})$

$\|\mathbf{Cf}\|^2_{\mathbf{W}_2} = (\mathbf{Cf})^{\mathrm{T}} \mathbf{W}_2 (\mathbf{Cf})$

$\mathbf{W}_1, \mathbf{W}_2$ are diagonal matrices, the choice of which can be justified in various ways. The entries in both matrices are non-negative values and less than or equal to unity.

In that case

$\Phi(\mathbf{f}) = \nabla_{\mathbf{f}} M(\mathbf{f}, \alpha) = (\mathbf{H}^{\mathrm{T}} \mathbf{W}_1 \mathbf{H} + \alpha \mathbf{C}^{\mathrm{T}} \mathbf{W}_2 \mathbf{C})\mathbf{f} - \mathbf{H}^{\mathrm{T}} \mathbf{W}_1 \mathbf{y}$

A more specific case is

$$M(\mathbf{f}, \alpha) = \|\mathbf{y} - \mathbf{Hf}\|^2 + \alpha \|\mathbf{Cf}\|^2_{\mathbf{W}}$$

where the weighting matrix is incorporated only in the regularization term. This method is known as **weighted regularised image restoration**. The entries in matrix $\mathbf{W}$ will be chosen so that the high-pass filter is only effective in the areas of low activity and a very little smoothing takes place in the edge areas. **[40]**

QUESTION 4

(a)

(i) Reconstruction levels are at $r_0 = (0 + \frac{1}{5})\frac{1}{2} = \frac{1}{10}$, $r_1 = (\frac{1}{5} + \frac{2}{5})\frac{1}{2} = \frac{3}{10}$, $r_2 = (\frac{2}{5} + \frac{3}{5})\frac{1}{2} = \frac{5}{10}$, $r_3 = (\frac{3}{5} + \frac{4}{5})\frac{1}{2} = \frac{7}{10}$ and $r_4 = (\frac{4}{5} + \frac{5}{5})\frac{1}{2} = \frac{9}{10}$. **[10]**

(ii) The probabilities of the five symbols are

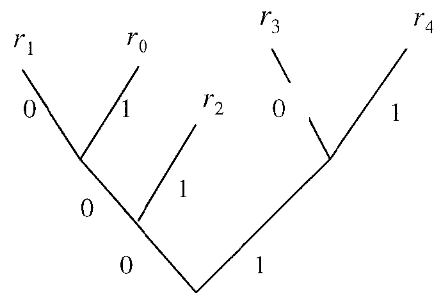
$$p_0 = \int_0^{\frac{1}{5}} 2r dr = r^2\Big|_0^{\frac{1}{5}} = \frac{1}{25},\ p_1 = \int_{\frac{1}{5}}^{\frac{2}{5}} 2r dr = r^2\Big|_{\frac{1}{5}}^{\frac{2}{5}} = \frac{3}{25},\ p_2 = \int_{\frac{2}{5}}^{\frac{3}{5}} 2r dr = r^2\Big|_{\frac{2}{5}}^{\frac{3}{5}} = \frac{5}{25},$$

$$p_3 = \int_{\frac{3}{5}}^{\frac{4}{5}} 2r dr = r^2\Big|_{\frac{3}{5}}^{\frac{4}{5}} = \frac{7}{25},\ p_4 = \int_{\frac{4}{5}}^{1} 2r dr = r^2\Big|_{\frac{4}{5}}^{1} = \frac{9}{25}$$

The Huffman code is found below. Probabilities for each r_i are found by evaluating the integral of the PDF over the relevant decision region. The result is shown below.

Symbol	*Probability*
r_0	1/25
r_1	3/25
r_2	5/25
r_3	7/25
r_4	9/25

Step 1	**Step 2**	**Step 3**	**Step 4**
r_4 9/25	r_4 9/25	$\{\{r_1, r_0\}, r_2\}$ 10/25	$\{r_3, r_4\}$ 16/25
r_3 7/25	r_3 7/25	r_4 9/25	$\{\{r_1, r_0\}, r_2\}$ 10/25
r_2 5/25	r_2 5/25	r_3 7/25	
r_1 3/25	$\{r_1, r_0\}$ 5/25		
r_0 2/25			



Symbol	*Codeword*
r_0	001
r_1	000
r_2	01
r_3	10
r_4	11

[20]

(iii) Average number of bits to represent f

$$l_{avg} = 3\cdot\frac{1}{25} + 3\cdot\frac{3}{25} + 2\cdot\frac{5}{25} + 2\cdot\frac{7}{25} + 2\cdot\frac{9}{25} = \frac{54}{25} = 2.16 \text{ bits/word}$$

[10]

(iv) For the above example we have:

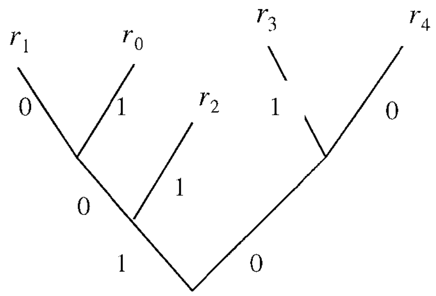
Entropy $H(s) = -\sum_{i=1}^{5} p_i \log_2(p_i) = 2.062$ bits/symbol

Redundancy $l_{avg} - H(s) = 0.098$ bits/symbol **[10]**

(b)

(i) Because in the process of merging two symbols it does not make any difference in which branch of the tree we assign the one and in which we assign the zero. **[10]**

(ii) The zeros are preferred to ones. Therefore, in the process of merging two symbols we assign the one to that branch that leads to the symbol that has the smallest probability between the two. **[20]**



(c)

In every bitonal image there are large regions that are either all white or all black. For instance, in Figure 4.1, we show a few pixels of a line in a bitonal image. Note that, the six contiguous pixels of the same color can be described as a run of six pixels with value 0. Thus, if each pixel of the image is remapped from say, its (position, value) to a **run** and **value**, then a more compact description can be obtained. In our example, no more than four bits are needed to describe the six-pixel run. In general, for many document type images, significant compression can be achieved using such preprocessing. Such a mapping scheme is referred to as a run-length coding scheme. **[20]**