IMPERIAL COLLEGE LONDON

E4.08
SO14
ISE4.33

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2005

MSc and EEE/ISE PART IV: MEng and ACGI

Corrected Copy

DIGITAL IMAGE PROCESSING

Thursday, 12 May 10:00 am

Time allowed: 3:00 hours

Q4. (c) → (b)

There are FOUR questions on this paper.

Answer THREE questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : P.T. Stathaki
Second Marker(s) : P.L. Dragotti

1. (a) (i) Explain why the Fourier transform amplitude of an image alone often does not capture the intelligibility of the image.

[2]

(ii) In a specific experiment it is observed that the amplitude response of an image exhibits energy concentration along a straight line in the 2-D frequency plane. What are the implications of this observation as far as the original image is concerned? Your answers must be fully justified.

[4]

(b) Let $f(x,y)$ denote the following 4×4 digital image that is zero outside $0\le x\le 3$, $0\le y\le 3$, with r a constant value.

$$\begin{bmatrix} r & r & r & r \\ r & r & r & r \\ r & r & r & r \\ r & r & r & r \end{bmatrix}$$

(i) Give the standard Hadamard Transform of $f(x,y)$ without carrying out any mathematical manipulations. Justify your answer.

[4]

(ii) Comment on the energy compaction property of the standard Hadamard Transform.

[2]

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

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

Each component $f_i(x,y)$ represents an image. The population arises from their formation across the entire collection of pixels. Suppose that $n > 2$, i.e. you have at least three images. 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) Suppose that a credible job could be done of reconstructing approximations to the n original images by using only the two principal component images associated with the largest eigenvalues. Find a lower bound for the mean square error incurred, as a function of $C_{\underline{g}}(n,n)$, where $C_{\underline{g}}$ is the covariance matrix of $\underline{g}$.

[4]

(ii) Suppose that the covariance matrix of $\underline{g}$ turns out to be the identity matrix. Is the Karhunen-Loeve transform useful in that case? Justify your answer.

[4]

2. (a) (i) Knowing that adding uncorrelated images convolves their histograms, how would you expect the contrast of the sum of two uncorrelated images to compare with the contrast of its component images? Justify your answer.

[4]

(ii) Consider an $N \times N$ image $f(x,y)$. From $f(x,y)$ create an image

$$g(x,y) = 2f(x,y) - f(x,y-1) - f(x,y+1).$$

Comment on the histogram of $g(x,y)$ in relation to the histogram of $f(x,y)$.

[4]

(b) Propose a method that uses variable size spatial filters to reduce background noise without blurring the image significantly.

[4]

(c) Why are bandpass filters useful in image processing? Justify your answer. Propose a method to obtain a bandpass filtered version of an image using spatial masks.

[4]

(d) Propose a method that detects edges in an image along the directions $\pm 45^\circ$.

[4]

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, statistically independent from the image and white.

(a) The autocorrelation matrix of the additive noise n is necessary in order to restore the image using Wiener Filtering.

(i) Since n is unknown, suggest a way to obtain an estimate of its autocorrelation matrix. Justify fully your answer using mathematical expressions.

[3]

(ii) State a possible scenario where the method proposed in (i) would give poor results.

[3]

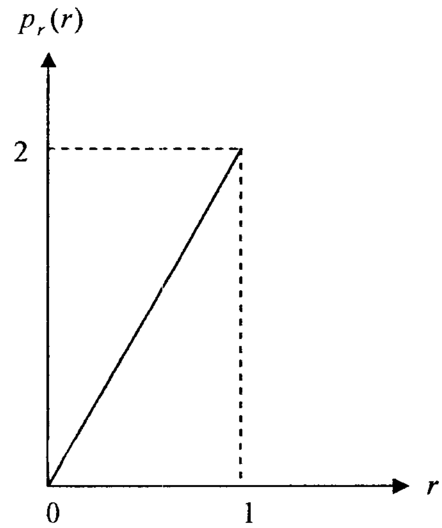
(b) Would the technique of Inverse Filtering yield better results in the case of restoration of an image that contains mainly low frequencies or in the case of an image that contains low and high frequencies? Assume that the spectrum of the degradation is of lowpass form. Justify your answer.

[6]

(c) Propose a technique to restore an image using a spatially adaptive constrained least squares filter. A full mathematical analysis is required.

[8]

4. (a) Consider an image with intensity $f(x,y)$ that can be modeled as a sample obtained from the probability density function sketched below:



(i) Suppose four reconstruction levels are assigned to quantize the intensity $f(x,y)$. Determine these reconstruction levels using a uniform quantizer.

[4]

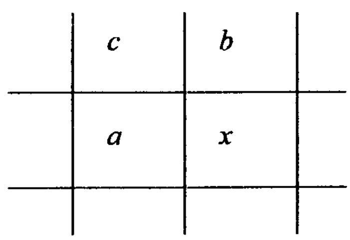
(ii) Determine the codeword to be assigned to each of the four reconstruction levels using Huffman coding. Specify what the reconstruction level is for each codeword. For your codeword assignment, determine the average number of bits required to represent r.

[4]

(iii) Determine the entropy, the redundancy and the coding efficiency of the Huffman code for this example.

[4]

(c) In lossless JPEG, one forms a prediction residual using previously encoded pixels in the current line and/or the previous line. Suppose that the prediction residual for pixel with intensity x in the following figure is defined as $r = y - x$ where y is the function $y = a + b - c$.



(i) Consider the case with pixel values $a = 101$, $b = 191$, $c = 190$ and $x = 180$. Find the codeword of the prediction residual r, knowing that the Huffman code for seven is 11110.

[4]

(ii) What is the key issue in achieving compression in lossless JPEG?

[4]

Solutions - DIGITAL IMAGE PROCESSING 2005

1 (a) (i) In viewing a picture, some of the most important visual information is contained in the edges and regions of high contrast. Intuitively, regions of maximum and minimum intensity in a picture are places at which complex exponentials at different frequencies are in phase. Therefore, it seems plausible to expect the phase and NOT the amplitude of the Fourier transform of a picture to contain much of the information in the picture, and in particular, the phase should capture the information about the edges.

(ii) The original image will possess a pattern of straight lines perpendicular to the line in the frequency domain.

(b) (i)

$$\begin{bmatrix} r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the image is constant only the zero frequency pair component of the Hadamard transform will be non zero and equal to the mean of the image.

(ii) The standard Hadamard transform does not have the advantageous property of energy compaction since as the index of the transform increases the frequency of the corresponding basis vector of the same index does not increase. Alternatively, we can say that the Hadamard matrix is non ordered.

(c) (i) Since the mean square error is equal to the sum of the eigenvalues that we discard we can say that the mean square error is at least equal to $(n-2)C_{\underline{g}}(n,n)$.

(ii) No because in that case all the new component images will have the same variance so there is no point of excluding some of the images from the set.

2\. (a) (i) Convolution increases the range of values so the new histogram will occupy a higher range of values and hence the new image will be of higher contrast.

(ii) The image $g(x,y)$ is a high pass filtered version of the image $f(x,y)$. Therefore it should contain large flat areas. Therefore the histogram of $g(x,y)$ should look like the following figure on the right.

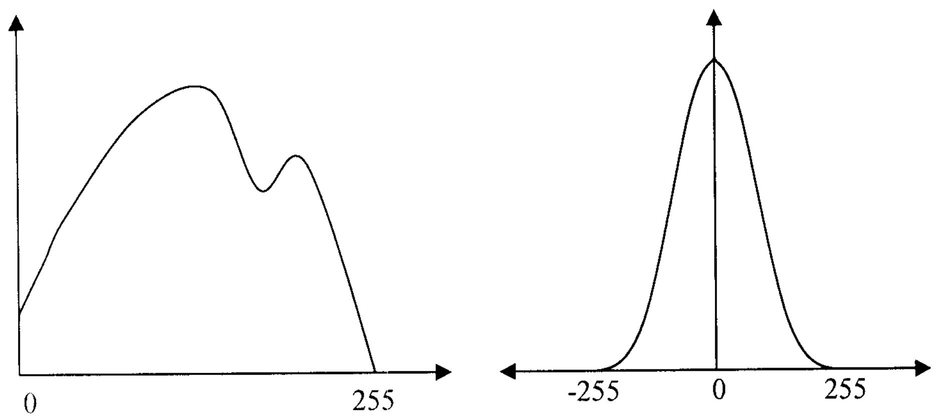


Figure: Original histogram (left) and new histogram (right)

(b) The background noise is more visible within flat areas instead of detailed areas. This is because the local Signal-to-Noise Ratio (SNR) is lower within flat areas. Therefore we can use larger spatial lowpass filtering type of masks within flat areas to reduce the amount of noise and smaller masks for high activity areas. An area can be classified as low activity or high activity depending on the local variance of the pixel of interest.

(c) Bandpass filters are useful for removing background noise without completely eliminating the background information. A bandpass filter can be implemented by a spatial mask as follows: The original image $f(x,y)$ is first convolved with a spatial mask of size $N_1 \times N_1$ to produce an output $g_1(x,y)$. Then is convolved with a spatial mask of size $N_2 \times N_2$ with $N_2 > N_1$ to produce an output $g_2(x,y)$. The final output is obtain as the difference $g_1(x,y) - g_2(x,y)$ and this is the bandpass filtered version of the image.

(d) Propose a method that detects edges in an image along the directions $\pm 45°$.

1	1	0
1	0	-1
0	-1	-1

-1	-1	0
-1	0	1
0	1	1

3 (a) (i) The additive noise in usually assumed to be white. In that case its autocorrelation matrix is a diagonal matrix of the form $\sigma_n^2 I$ where I is the identity matrix. Therefore there is only one unknown to be estimated and this is the variance σ_n^2 of the additive noise n. This can be estimated from a flat region of the available distorted image. This is because for a flat region we might write $\sigma_g^2 = \sigma_{Hf}^2 + \sigma_n^2 = \sigma_n^2$. A flat region can be selected visually.

(ii) The method proposed in (i) would give very poor results if there isn't any flat region available, i.e., the image is too detailed.

(b) In the presence of external noise we have that

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

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

If $H(u,v)$ becomes very small, the term $N(u,v)$ dominates the result.

(c) The functional to be minimized takes the form

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

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

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

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. Each pixel (x,y) is classified as a low activity pixel or a high activity (edge) pixel. This is done by calculating the local variance $\sigma_{(x,y)}^2$ around this pixel. Then the corresponding location in the matrix $\mathbf{W}$ can be set as $c/\sigma_{(x,y)}^2$ with c a constant value calculated so that the element of $\mathbf{W}$ sum up to 1.

4 (a) (i) $p_r(r) = 2r$

Reconstruction level 1 $r = \frac{1}{2}(\frac{1}{4} + 0) = \frac{1}{8}$ with probability $\int_0^{\frac{1}{4}} 2r dr = \frac{1}{16}$

Reconstruction level 1 $r = \frac{1}{2}(\frac{1}{4} + \frac{2}{4}) = \frac{3}{8}$ with probability $\int_{\frac{1}{4}}^{\frac{1}{2}} 2r dr = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$

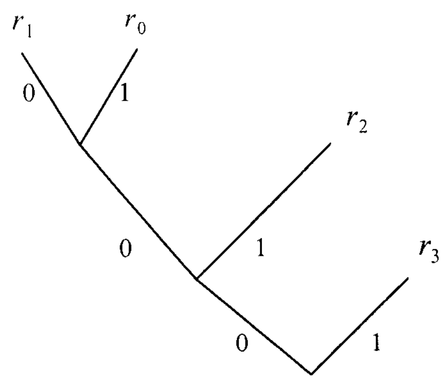
Reconstruction level 1 $r = \frac{1}{2}(\frac{2}{4} + \frac{3}{4}) = \frac{5}{8}$ with probability $\int_{\frac{1}{2}}^{\frac{3}{4}} 2r dr = \frac{9}{16} - \frac{1}{4} = \frac{5}{16}$

Reconstruction level 1 $r = \frac{1}{2}(\frac{3}{4} + 1) = \frac{7}{8}$ with probability $\int_{\frac{3}{4}}^{1} 2r dr = 1 - \frac{9}{16} = \frac{7}{16}$

(ii) The Huffman code is found below.

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

Symbol	Codeword
r_0	001
r_1	000
r_2	01
r_3	1



Average number of bits to represent f

$$l_{avg} = 3\frac{4}{16} + 2\frac{5}{16} + \frac{7}{16} = \frac{29}{16} = \text{ bits/symbol}$$

(iii) Entropy $H(s) = -\sum_{i=1}^{4} p_i \log_2(p_i) = 1.7604$ bits/symbol

Redundancy $= 0.9896$ bits/symbol

Coding efficiency $H(s)/l_{avg} = 64\%$

(c) (i) $r = a + b - c - x = -78$. Since $(-78)_{10} = (1001110)_2$, in lossless JPEG -78 will be represented by 0110001. Therefore, the representation of -78 will be 111100110001.

(ii) The fact that small categories occur very frequently.