IMPERIAL COLLEGE LONDON

E4.08
SO14
ISE4.33

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2006

MSc and EEE/ISE PART IV: MEng and ACGI

DIGITAL IMAGE PROCESSING

Corrected Copy

Friday, 5 May 10:00 am

Time allowed: 3:00 hours

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

© University of London 2006

1. (a) Consider an $M \times N$ -pixel grayscale image $f(x,y)$ which is zero outside $0 \le x \le M-1$ and $0 \le y \le N-1$. In transform coding, we discard the transform coefficients with small magnitudes and code only those with large magnitudes. Let $F(u,v)$ denote the $M \times N$ -point Discrete Fourier Transform (DFT) of $f(x,y)$. Let $G(u,v)$ denote $F(u,v)$ modified by

$$G(u,v) = \begin{cases} F(u,v), & \text{when } |F(u,v)| \text{ is large} \\ 0, & \text{otherwise} \end{cases}$$

Let

$$\frac{\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}|G(u,v)|^2}{\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}|F(u,v)|^2} = \frac{9}{10}$$

We reconstruct an image $g(x,y)$ by computing the $M \times N$ -point inverse DFT of $G(u,v)$. Express

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}|f(x,y)-g(x,y)|^2$$

in terms of

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}|f(x,y)|^2$$

[10]

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

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

Each component $f_i(x,y), i = 1,2,3$ represents an $M \times N$ -pixel grayscale image. The population arises from the formation of the vectors across the entire collection of pixels. Consider now a population of vectors $\underline{g}$ of the form

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

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

The covariance matrix of the population $\underline{f}$ calculated as part of the transform is

$$\underline{C}_{\underline{f}} = \begin{bmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{bmatrix}$$

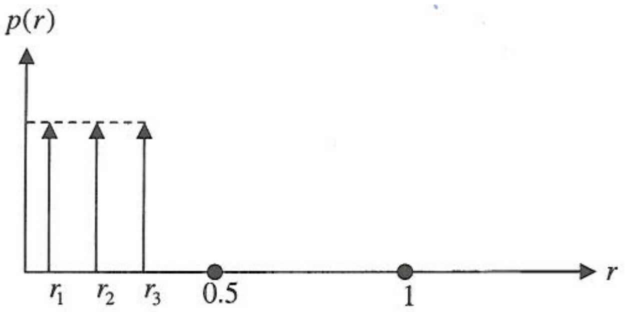
(i) Suppose that a credible job could be done of reconstructing approximations to the three original images by using one principal component image. What would be the mean square error incurred in doing so, if it is known that $c < a - b$ and $b > 0$?

[5]

(ii) Suppose that a credible job could be done of reconstructing approximations to the three original images by using two principal component images. What would be the mean square error incurred in doing so, if it is known that $c > a + b$ and $b > 0$?

[5]

2. (a) Consider an $M \times N$-pixel grayscale image $f(x, y)$ which is zero outside $0 \leq x \leq M-1$ and $0 \leq y \leq N-1$. The histogram $p(r)$ of $f(x, y)$ is sketched below. (Intensities r have been normalized so that they take values from 0 to 1.)



(i) What can we say about $f(x, y)$? **[1]**

(ii) Propose an intensity transformation function which will improve the contrast of the image when it is used to modify the intensity of the image. **[3]**

(iii) Sketch the histogram of the transformed intensity. **[3]**

(iv) Calculate the mean and the variance of the two images (original and modified). **[3]**

(b) Propose a method for spatially adaptive noise reduction in images, that exploits jointly the following two observations:

1. Noise is typically less visible to human viewers in image regions of high detail, than in image regions of low detail.
2. Noise is typically less visible to human viewers in bright areas than in dark areas.

[6]

(c) (i) Find the impulse response $h(m,n)$ of the Laplacian spatial mask. Show that this impulse response can be written as $h(m,n) = g(m)\delta(n) + \delta(m)g(n)$ where $g(n)$ is a one-dimensional discrete function and $\delta(n)$ is the discrete unit impulse function defined as

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the function $g(n)$. **[2]**

(ii) Show that the Laplacian spatial mask is an approximation of the local second derivative of the image intensity. **[2]**

3. We are given the blurred and noisy version $g(x,y)$ of an image $f(x,y)$ such that

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n}$$

where $\mathbf{f}$ and $\mathbf{g}$ are the lexicographically ordered original and degraded image respectively, $\mathbf{H}$ is the degradation matrix, and $\mathbf{n}$ is the lexicographically ordered noise term which is assumed to be zero mean, independent from the original image and white.

(a) (i) Consider the Wiener filtering image restoration technique. Write down without proof the expressions for both the Wiener filter estimator and the restored image both in the spatial domain and the frequency domain. Explain all symbols used. **[5]**

(ii) Discuss the relation between Wiener filtering and both inverse and pseudo-inverse filtering. **[5]**

(b) Let $g(x,y)$ be a degraded only by noise image that can be expressed as

$$g(x,y) = f(x,y) + n(x,y)$$

where $f(x,y)$ is the original image and $n(x,y)$ is a background noise. In Wiener filtering we assume that the Discrete Space Fourier Transform $S_{ff}(u,v)$ of the autocorrelation function of the original image is available.

One method of estimating $S_{ff}(u,v)$ is to model the autocorrelation function as follows

$$R_{ff}(k,l) = \underset{(x,y)}{E}\{f(x,y)f(x+k,y+l)\} = \rho^{|k|+|l|}$$

with ρ a real unknown parameter with $0 < \rho < 1$.

(i) Assuming $n(x,y)$ is a zero mean, white noise with unknown variance σ_n^2 and independent of $f(x,y)$, write down without proof the expressions for the Wiener filter estimator and the restored image in the frequency domain as functions of ρ and σ_n^2. **[5]**

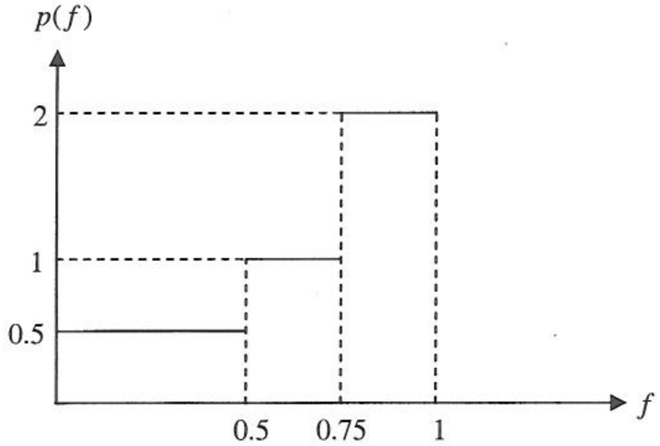
(ii) Develop a method of estimating ρ using the autocorrelation function samples of $g(x,y)$. **[5]**

Hints:

1. The Discrete Space Fourier Transform $S_{ff}(u,v)$ of the autocorrelation function is defined as $S_{ff}(u,v) = \sum_{k=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty} R_{ff}(k,l)e^{-j(uk+vl)}$.
2. The following result holds: $\sum_{k=0}^{+\infty} a^k = \frac{1}{1-a}, |a| < 1$.

---

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. (Intensities f have been normalized so that they take values from 0 to 1).



(i) Suppose four reconstruction levels are assigned to quantize the intensity $f(x, y)$. Determine these reconstruction levels using a uniform quantizer. **[3]**

(ii) Explain when uniform quantization of an image is optimal in terms of the mean square error of quantization. **[3]**

(iii) 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 f. **[4]**

(iv) Determine the entropy, the redundancy and the coding efficiency of the Huffman code for this example. **[4]**

(b) Consider an image with intensity $f(x, y)$ that can take three possible values r_1, r_2 and r_3 with probabilities shown in Figure 4.1 below. We wish to compress the image using Huffman coding.

Intensities	***Probabilities***
r_1	0.95
r_2	0.02
r_3	0.03

Figure 4.1

(i) Determine the redundancy of the Huffman code for this example. Comment on the result. **[3]**

(ii) Suppose that extended-by-2 Huffman coding is applied in the above set of intensities, achieving finally 1.222 bits/new symbol. Is extended Huffman code effective in this example? Justify your answer. **[3]**

DIGITAL IMAGE PROCESSING
SOLUTIONS 2006

E4.08
ISE 4.33
SO14

1/5

1. (a)

The signal $f(x,y)-g(x,y)$ is obtained by the Inverse DFT of the signal $F(u,v)-G(u,v)$. Therefore, according to Parseval's theorem the energy of the signal $f(x,y)-g(x,y)$ is equal to the energy of the signal $F(u,v)-G(u,v)$. The signal $F(u,v)-G(u,v)$ consists of the DFT samples of $F(u,v)$ which were excluded in forming $G(u,v)$. Since, $G(u,v)$ captures 0.9 of the energy of $F(u,v)$, the signal $F(u,v)-G(u,v)$ will capture 0.1 of the energy of $F(u,v)$. Therefore,

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(f(x,y)-g(x,y))^2 = 0.1\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)^2$$

(b)

The eigenvalues of the covariance matrix $\underline{C}_f = \begin{bmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{bmatrix}$ are found by the following relationship:

$$\det\begin{bmatrix} a-\lambda & b & 0 \\ b & a-\lambda & 0 \\ 0 & 0 & c-\lambda \end{bmatrix} = (c-\lambda)[(a-\lambda)^2-b^2] = (c-\lambda)[(a-\lambda)-b](c-\lambda)[(a-\lambda)+b]$$

$$=0 \Rightarrow \lambda_1 = c, \lambda_2 = a-b, \lambda = a+b$$

(i) If $b<a-c$ then because $c \geq 0$ since it represents variance of an image the eigenvalues will be sorted according to magnitude as $a+c \geq a-c > b$ and therefore by using only one principal component the error of reconstruction will be $a-c+b$

(ii) If $b > a+c$ the eigenvalues will be sorted according to magnitude as $b > a+c \geq a-c$ and therefore by using only two principal components the error of reconstruction will be $a-c$

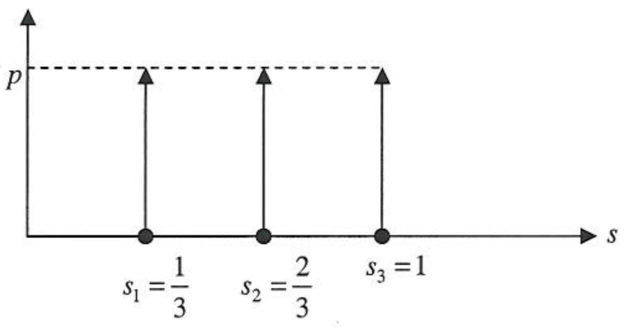
---

2. (a)

(i) $f(x,y)$ will be a dark image since the intensities are concentrated in the lower half of the intensity range. Moreover, it will consists of three intensities only with equal probabilities, and therefore, $p=\frac{1}{3}$.

(ii) We can use histogram equalization. By doing so, the three intensities are mapped to the following: $s_1 = T(r_1) = \frac{1}{3}$, $s_2 = T(r_2) = \frac{2}{3}$, $s_3 = T(r_3) = 1$

(iii)



2/5

(iv) For the original image we have:

Mean: $m_1 = \frac{1}{3}(r_1 + r_2 + r_3)$

Variance: $\sigma_1^2 = \frac{1}{3}(r_1^2 + r_2^2 + r_3^2) - \frac{1}{9}(r_1 + r_2 + r_3)^2$

For the equalised image we have:

Mean: $m_2 = \frac{1}{3}(s_1 + s_2 + s_3) = \frac{1}{3}(\frac{1}{3} + \frac{2}{3} + 1) = \frac{2}{3}$

Variance: $\sigma_2^2 = \frac{1}{3}(s_1^2 + s_2^2 + s_3^2) - \frac{4}{9} = \frac{1}{3}(\frac{1}{9} + \frac{4}{9} + 1) - \frac{4}{9} = \frac{2}{27}$

(b)

Since dark areas are represented by small intensities we can calculate the local mean around each pixel in order to decide whether this belongs to a dark or to a bright area. We denote the mean around a pixel with co-ordinates (x, y) as $m_{(x,y),W}$. The subscript W denotes the neighborhood used in order to calculate this mean. If this is a rectangular window of size $(2W+1)(2W+1)$ symmetrically placed around the pixel of interest, then

$$m_{(x,y),W} = \frac{1}{(2W+1)^2} \sum_{i=-W}^{i=+W} \sum_{j=-W}^{j=+W} f(x+i, y+j)$$

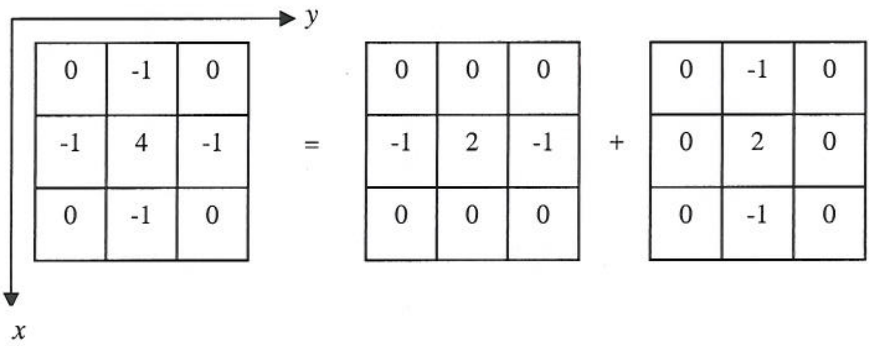
We can calculate the local variance around each pixel in order to decide whether this belongs to a high detail or to a low detail area. We denote the variance around a pixel with co-ordinates (x, y) as $\sigma^2_{(x,y),W}$. The subscript W denotes the neighborhood used in order to calculate this variance. If this is a rectangular window of size $(2W+1)(2W+1)$ symmetrically placed around the pixel of interest, then

$$\sigma^2_{(x,y),W} = \frac{1}{(2W+1)} \sum_{i=-W}^{i=+W} \sum_{j=-W}^{j=+W} f(x+i, y+j)^2 - m^2_{(x,y),W}$$

Therefore, if $m_{(x,y),W} > T_1$ or $\sigma^2_{(x,y),W} > T_2$ then use a small low-pass spatial mask for noise removal. Otherwise, if $m_{(x,y),W} < T_1$ and $\sigma^2_{(x,y),W} < T_2$ then use a large low-pass spatial mask for noise removal.

(c)

(i) The 2-D impulse response is separable since it can be written as follows.



$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

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

$g(x, y) = f(x, y) - f(x, y-1)$

7/5

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

We therefore see that the Laplacian operator is the first derivative of the first derivative and therefore, is the second derivative.

---

3. (a) (i) Wiener filter in space

$$\mathbf{W}=\mathbf{R}_{\mathbf{fy}}\mathbf{R}_{\mathbf{yy}}{}^{-1}=\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathrm{T}}+\mathbf{R}_{\mathbf{nn}})^{-1}$$

Estimate for the original image is

$$\hat{\mathbf{f}}=\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathrm{T}}+\mathbf{R}_{\mathbf{nn}})^{-1}\mathbf{y}$$

Note that knowledge of $\mathbf{R}_{\mathbf{ff}}=E\{\mathbf{ff}^{\mathrm{T}}\}$ and $\mathbf{R}_{\mathbf{nn}}=E\{\mathbf{nn}^{\mathrm{T}}\}$ is assumed.

In frequency domain

$$W(u,v)=\frac{S_{ff}(u,v)H^{*}(u,v)}{S_{ff}(u,v)|H(u,v)|^{2}+S_{nn}(u,v)}$$

$$\hat{F}(u,v)=\frac{S_{ff}(u,v)H^{*}(u,v)}{S_{ff}(u,v)|H(u,v)|^{2}+S_{nn}(u,v)}Y(u,v)$$

The noise variance has to be known, otherwise it is estimated from a flat region of the observed image.

In practical cases where a single copy of the degraded image is available, it is quite common to use $S_{yy}(u,v)$ as an estimate of $S_{ff}(u,v)$. This is very often a poor estimate.

(ii) If $S_{nn}(u,v)=0\Rightarrow W(u,v)=\dfrac{1}{H(u,v)}$ which is the inverse filter. If $S_{nn}(u,v)\to 0$

$$\lim_{S_{nn}\to 0}W(u,v)=\begin{cases}\dfrac{1}{H(u,v)} & H(u,v)\neq 0\\[2ex] 0 & H(u,v)=0\end{cases}$$

which is the pseudo-inverse filter.

(b) (i) The noise is white with variance σ_n^2 and therefore the autocorrelation function of the noise is

$$R_{nn}(k,l)=\underset{(x,y)}{E}\{n(x,y)n(x+k,y+l\}=\sigma_n^2\delta(k,l)$$

and the power spectrum of the noise is σ_n^2.

$$S_{ff}(u,v)=\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}R_{ff}(k,l)e^{-j(uk+vl)}=\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\rho^{|k|+|l|}e^{-j(uk+vl)}=\sum_{k=-\infty}^{\infty}\rho^{|k|}e^{-juk}\sum_{l=-\infty}^{\infty}\rho^{|l|}e^{-jvl}$$

We find that

$$\sum_{n=-\infty}^{\infty}r^{|n|}e^{-j\omega n}=\sum_{n=-\infty}^{-1}r^{-n}e^{-j\omega n}+\sum_{n=0}^{\infty}r^{n}e^{-j\omega n}=\sum_{n=1}^{\infty}r^{n}e^{j\omega n}+\sum_{n=0}^{\infty}r^{n}e^{-j\omega n}=\sum_{n=0}^{\infty}r^{n}e^{j\omega n}+\sum_{n=0}^{\infty}r^{n}e^{-j\omega n}-1$$

$$=\sum_{n=0}^{\infty}(re^{j\omega})^{n}+\sum_{n=0}^{\infty}(re^{-j\omega})^{n}-1=\frac{1}{1-re^{j\omega}}+\frac{1}{1-re^{-j\omega}}-1=\frac{2-2r\cos\omega}{1+r^{2}-2r\cos\omega}-1=\frac{1-r^{2}}{1+r^{2}-2r\cos\omega}$$

Therefore,

$$S_{ff}(u,v)=\frac{(1-\rho^{2})^{2}}{(1+\rho^{2}-2\rho\cos u)(1+\rho^{2}-2\rho\cos v)}$$

4/5

and

$$W(u,v)=\frac{(1-\rho^2)^2}{(1-\rho^2)^2+(1+\rho^2-2\rho\cos u)(1+\rho^2-2\rho\cos v)\sigma_n^2}$$

(ii) Develop a method of estimating ρ_1 and ρ_2 from $g(x,y)$.

$$R_{gg}(k,l)=R_{ff}(k,l)+R_{nn}(k,l)=\rho^{|k|+|l|}+\sigma_n^2\delta(k,l)$$

If we choose two pairs of values for the parameters k,l as $(k,l)=(1,0)$ and $(k,l)=(0,0)$ we have the following relationships:

For $(k,l)=(1,0)$: $R_{gg}(1,0)=\rho$ where $R_{gg}(1,0)$ is estimated from the available data.

For $(k,l)=(0,0)$: $R_{gg}(0,0)=1+\sigma_n^2$ where $R_{gg}(0,0)$ is estimated from the available data

---

4. (a) (i) The four reconstruction levels are:

$r_1=1/8$ with probability $1/8$

$r_2=3/8$ with probability $1/8$

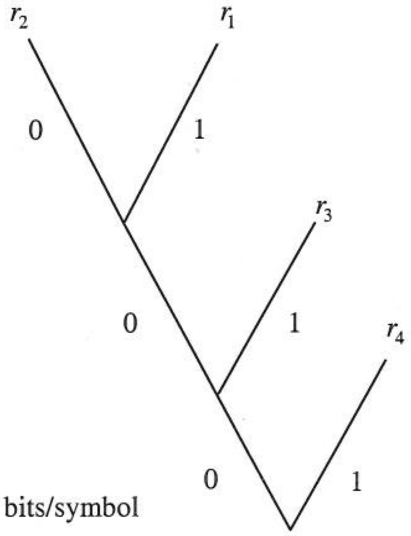
$r_3=5/8$ with probability $1/4$

$r_4=7/8$ with probability $1/2$

(ii) Uniform quantization does not exploit the pdf of the alphabet to be quantized.

(iii) The Huffman code is found below.

Step 1		**Step 2**		**Step 3**	
r_4	1/2	r_4	1/2	r_4	1/2
r_3	1/4	r_3	1/4	$\{r_3,\{r_2,r_1\}\}$	1/2
r_2	1/8	$\{r_2,r_1\}$	1/4		
r_1	1/8				



Symbol	*Codeword*
r_1	001
r_2	000
r_3	01
r_4	1

Average number of bits to represent f

$$l_{avg}=\frac{1}{2}+2\frac{1}{4}+6\frac{1}{8}=\frac{7}{4}\text{ bits/symbol}$$

(iv) Entropy $H(s)=-\sum_{i=1}^{3}p_i\log_2(p_i)=-(\frac{1}{2}\log_2\frac{1}{2}+\frac{1}{4}\log_2\frac{1}{4}+\frac{1}{8}\log_2\frac{1}{8}+\frac{1}{8}\log_2\frac{1}{8})$

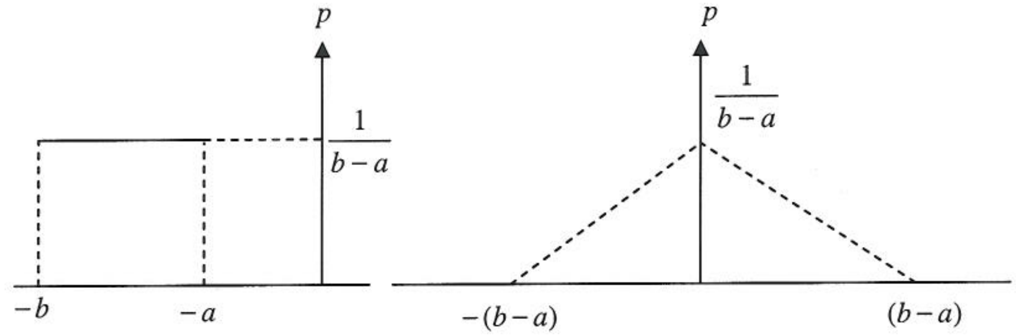
5/5

$$=\frac{1}{2}+\frac{2}{4}+\frac{3}{8}+\frac{3}{8}=\frac{4}{8}+\frac{4}{8}+\frac{3}{8}+\frac{3}{8}=\frac{14}{8}=\frac{7}{4} \text{ bits/symbol}$$

Redundancy $=0$ bits/symbol

Coding efficiency $H(s)/l_{avg}=100\%$

(b)



The pdf of $-f(x-1,y-1)$ is shown on the left and the pdf of $g(x,y)$ is shown on the right (convolution of the 2 pdf's). The pdf of $g(x,y)$ is more skewed and therefore more appropriate for symbol encoding.