Solutions

1. (a) For an image given by the function $f(x,y) = (x+y)^3$ where x, y are continuous variables, evaluate $f(x,y)\delta(x-1,y-2)$ (2 marks) and $f(x,y) * \delta(x-1,y-2)$ (2 marks) where $\delta(x,y)$ is the two-dimensional Dirac Delta function. **[4]**

Solution

$\delta(x-1,y-2) = 0$ unless $x = 1$ and $y = 2$, hence the product $f(x,y)\delta(x-1,y-2)$ is also zero unless both $x = 1$ and $y = 2$. The product is therefore also a delta function at the same position. However, the size of the delta function is multiplied by the value of $f(x,y) = (x+y)^3$ at $x = 1$ and $y = 2$. Hence the final answer is $27\delta(x-1,y-2)$; this means that the area under the delta function is now 27. **[2]**

$f(x,y) * \delta(x-1,y-2) = f(x-1,y-2) = (x+y-3)^3$ **[2]**

(b) Three binary images (with value 255 in black areas and value 0 elsewhere) are shown below. Describe roughly on the form of the amplitude of the 2D DFT of these images by using properties of the DFT.

[3]

Solution

The amplitude of the DFT of the circular pillbox is a circularly symmetric function. **[1]**

The DFT of the straight line is also a straight line which is perpendicular to the original line. **[1]**

The third image consists of two displaced replications of the first image. Shifting a function causes change only to the phase of the DFT. Therefore, its DFT's amplitude is the same as the DFT amplitude of the original image multiplied by 2. **[1]**

(c) Given the filter impulse response $h(m,n)$

$$h(m,n) = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(i) Calculate the DFT of the given filter impulse response $h(m,n)$. **[4]**

(ii) Based on the result from (c) (i), determine if the given filter is a high-pass or a low-pass filter. Explain your reasoning. **[2]**

Solution

(i)

$$H(u,v) = \frac{1}{9}\sum_{m=-1}^{m=1}\sum_{n=-1}^{n=1} h(m,n)e^{-j2\pi(um/3+vn/3)} = \frac{1}{9}\sum_{m=-1}^{m=1}\sum_{n=-1}^{n=1} e^{-j2\pi(um/3+vn/3)}$$

$$\frac{1}{9}(e^{-j2\pi\left(-\frac{u}{3}-\frac{v}{3}\right)}+e^{-j2\pi\left(-\frac{u}{3}\right)}+e^{-j2\pi\left(-\frac{u}{3}+\frac{v}{3}\right)}+e^{-j2\pi\left(-\frac{v}{3}\right)}+1+e^{-j2\pi\left(\frac{v}{3}\right)}$$
$$+e^{-j2\pi\left(\frac{u}{3}-\frac{v}{3}\right)}+e^{-j2\pi\left(\frac{u}{3}\right)}+e^{-j2\pi\left(\frac{u}{3}+\frac{v}{3}\right)})=$$
$$\frac{1}{9}(2\cos\left[2\pi\left(\frac{u}{3}+\frac{v}{3}\right)\right]+2\cos\left[2\pi\left(\frac{u}{3}-\frac{v}{3}\right)\right]+2\cos\left[2\pi\frac{u}{3}\right]+2\cos\left[2\pi\frac{v}{3}\right]+1)=$$

$$\frac{1}{9}(4\cos\left[2\pi\left(\frac{u}{3}\frac{v}{3}\right)\right]+2\cos\left[2\pi\frac{u}{3}\right]+2\cos\left[2\pi\frac{v}{3}\right]+1)$$
$$\frac{1}{9}(2\cos\left[2\pi\frac{u}{3}\right]+1)(2\cos\left[2\pi\frac{v}{3}\right]+1)$$

(ii) $u=v=0, H(0,0)=1$
$u=v=\pi, H(\pi,\pi)=0$
We see from the above that $H(u,v)$ behaves like a lowpass filter.

(d) 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:

$$f_1(x,y)=\begin{cases} r_1 & 1\le x\le M, 1\le y\le \frac{M}{2}\\ r_2 & 1\le x\le M, \frac{M}{2}< y\le M\end{cases}$$
$$f_2(x,y)=r_3,\ 1\le x\le M,\ 1\le y\le M$$
$$f_3(x,y)=r_4,\ 1\le x\le M,\ 1\le y\le M$$

Consider now a population of 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. **[5]**

(ii) Comment on whether you could obtain the result of (b)-(i) above using intuition rather than by explicit calculation. **[2]**

Solution

(i) 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

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

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) 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) The following figures show (3 marks): a 3-bit image $f(x,y)$, $x, y \in [0,4]$ of size 5-by-5, a Laplacian filter and a low-pass filter

$$\begin{bmatrix} 3 & 7 & 6 & 2 & 0 \\ 2 & 4 & 6 & 1 & 1 \\ 4 & 7 & 2 & 5 & 4 \\ 3 & 0 & 6 & 2 & 1 \\ 5 & 7 & 5 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0.01 & 0.1 & 0.01 \\ 0.1 & 0.56 & 0.1 \\ 0.01 & 0.1 & 0.01 \end{bmatrix}$$

Compute the following:

(i) The output of a 3×3 mean (average) filter at (2,2). **[1]**

(ii) The output of the 3×3 Laplacian filter shown above at (2,2). **[1]**

(iii) The output of the 3×3 low-pass filter shown above at (2,2). **[1]**

Solution

(i) $\frac{4+6+1+7+2+5+0+6+2}{9} = \frac{33}{9}$

(ii) $6 \cdot 1 + 6 \cdot 1 + 7 \cdot 1 + 5 \cdot 1 - 4 \cdot 2 = 16$

(iii) $(4 + 1 + 2) \cdot 0.01 + (7 + 6 + 6 + 5) \cdot 0.1 + 2 \cdot 0.56 = 0.07 + 2.4 + 1.12 = 3.59$

(b) Show that median filters are non-linear filters. **[2]**

Solutions

We can use a counter example to show that.

$x = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ $y = \begin{bmatrix} 0 \\ -4 \\ 4 \end{bmatrix}$ $x + y = \begin{bmatrix} 2 \\ 0 \\ 10 \end{bmatrix}$

We see that $\text{median}(x) = 4$, $\text{median}(y) = 0$,

$\text{median}(x + y) = 2 \neq \text{median}(x) + \text{median}(y)$

(c) Consider an image with 8 grey levels $r \in [0,7]$ with the histogram $(2, 2, 4, 8, 16, 32, 64, 128)$.

(i) Apply histogram equalization to the given image. If the resulting intensities are not integers you can round them to the nearest integer. **[3]**

(ii) Explain using the result how histogram equalization enhances the contrast of an image. **[3]**

Solution

(i) By applying histogram equalisation, we transform the original intensity r to a new intensity s as follows:

$0 \to 2, 1 \to 4, 2 \to 8, 3 \to 16, 4 \to 32, 5 \to 64, 6 \to 128, 7 \to 256$

In order to normalise the new intensities within the range $[0, 7]$, we multiply them with $\frac{7}{256}$. Therefore, we obtain

$0 \to 2 \cdot \frac{7}{256} = 0.054 \to 0, 1 \to 4 \cdot \frac{7}{256} = 0.109 \to 0, 2 \to 8 \cdot \frac{7}{256} = 0.21875 \to 0$

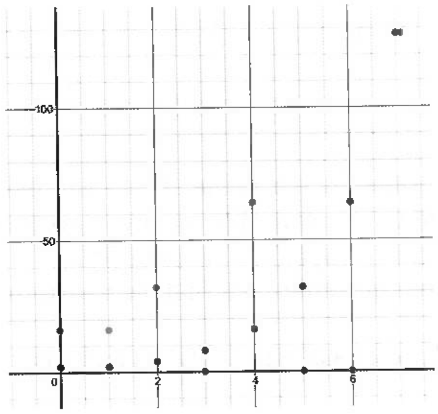
$3 \to 16 \cdot \frac{7}{256} = 0.4375 \to 0, 4 \to 32 \cdot \frac{7}{256} = 0.875 \to 1, 5 \to 64 \cdot \frac{7}{256} = 1.75 \to 2$

$6 \to 128 \cdot \frac{7}{256} = 3.5 \to 4, 7 \to 256 \cdot \frac{7}{256} = 7 \to 7$

The new image has intensities $(0, 1, 2, 4, 7)$. We see that the intensities $3, 5, 6$ are missing.

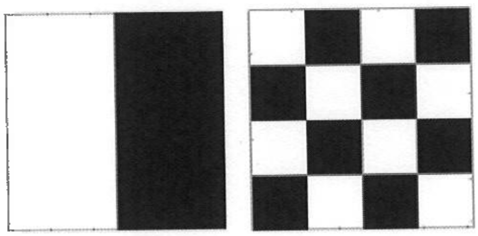
The histogram of the new image is $(16, 16, 32, 0, 64, 0, 0, 128)$.

(ii) By comparing the two histograms we verify the well-known universal observation that in case of discrete histogram equalisation, the new image has a "flatter" histogram (shown in orange below) but is definitely not flat.



(d) Consider the two binary images of the same size shown below. The pixels shown with black have 0 intensity and the pixels shown with white have intensity equal to 1. The images shown are quite different, but their histograms are the same. Suppose that each image is blurred with a 3×3 averaging kernel. Would the histograms of the blurred images still be equal? Explain. If your answer is no, sketch the two histograms. In order to convolve the outer pixels with the 3×3 kernel you can assume zero values outside the images.

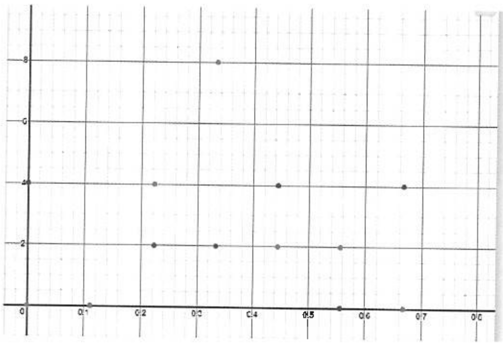
[4]



Solution

The two images which are formed by convolving the original images with a 3×3 kernel are shown below. The two histograms are not the same. Their plots are given below with blue (image on the left) and orange (image on the right).

$$\begin{bmatrix} \frac{4}{9} & \frac{4}{9} & \frac{2}{9} & 0 \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{4}{9} & \frac{4}{9} & \frac{2}{9} & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{9} & \frac{1}{3} & \frac{1}{3} & \frac{2}{9} \\ \frac{1}{3} & \frac{5}{9} & \frac{4}{9} & \frac{1}{3} \\ \frac{1}{3} & \frac{4}{9} & \frac{5}{9} & \frac{1}{3} \\ \frac{2}{9} & \frac{1}{3} & \frac{1}{3} & \frac{2}{9} \end{bmatrix}$$



(e) An image with intensities in the range [0, 1] has the pdf of the form $p_r(r) = 2 - 2r$, $r \in [0,1]$. It is desired to transform the intensity levels of this image so that they will have the specified pdf as given by $p_z(z) = 3z^2$, $z \in [0,1]$.

(i) Assume continuous quantities and find the transformation (in terms of r and z) that will accomplish this. **[3]**

(ii) Repeat Part (a) for the case when $p_z(z)$ is uniform. **[2]**

Solution

(i) $s = T(r) = (L-1)\int_0^r p_r(w)dw = (L-1)\int_0^r (2-2w)dw = 2r - r^2$

$$v = G(z) = (L-1)\int_0^z 3w^2 dw = z^3$$

$$2r - r^2 = z^3 \Rightarrow z = (2r - r^2)^{\frac{1}{3}}$$

(ii) In case that the desired pdf is uniform we simply carry histogram equalisation, and therefore, $z = 2r - r^2$.

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

$$g = Hf + n$$

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. All images involved have size $N \times N$ after extension and zero-padding.

(a) (i) Describe how inverse filtering can be used to restore the degraded image above. [4]

(ii) Given knowledge of the exact degradation function, under what assumption can we perfectly restore the image? (2 marks in total) How can we avoid erratic behaviour when the assumption is not met? (2 marks in total) [4]

(iii) A particular image $f(x, y)$ is distorted by convolution with either the space invariant point function $h_1(x, y)$ or the function $h_2(x, y)$ where:

$h_1(x,y) = \delta(x,y) + \delta(x-1,y) + \delta(x+1,y) + \delta(x,y-1) + \delta(x,y+1)$

$h_2(x,y) = 5\delta(x,y) + \delta(x-1,y) + \delta(x+1,y) + \delta(x,y-1) + \delta(x,y+1)$

Assuming that the distorted images do not contain random additive noise then in one image the distortion can be effectively removed using an Inverse Fourier filter, while in the other a Pseudo Inverse filter must be used. Which image must use the Pseudo Inverse filter and why? [4]

Solution

(i) The objective is to minimize $J(\mathbf{f}) = \|\mathbf{n}(\mathbf{f})\|^2 = \|\mathbf{y} - \mathbf{Hf}\|^2$.

We set the first derivative of the cost function equal to zero

$$\frac{\partial J(\mathbf{f})}{\partial \mathbf{f}} = 0 \Rightarrow -2\mathbf{H}^{\mathrm{T}}(\mathbf{y} - \mathbf{Hf}) = \mathbf{0}.$$

By solving that system we obtain $\mathbf{H}^{\mathrm{T}}\mathbf{Hf} = \mathbf{H}^{\mathrm{T}}\mathbf{y}$.

If $M = N$ and $\mathbf{H}^{-1}$ exists then $\mathbf{f} = \mathbf{H}^{-1}\mathbf{y}$.

According to the previous analysis if $\mathbf{H}$ (and therefore $\mathbf{H}^{-1}$) is block circulant the above problem can be solved as a set of $M \times N$ scalar problems as follows

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

Bookwork

(ii) 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. Therefore, only under the assumption that the image is noiseless we can perfectly restore it.

In order to avoid erratic behaviour we 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:

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

or

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

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. The parameter ε (called **threshold**) is a small number chosen by the user.

Bookwork

(iii) The DFT of the first point spread function given is

$$H_1(u,v)=\frac{1}{N^2}\sum_{x=-1}^{x=1}\sum_{y=-1}^{y=1}h(x,y)e^{-j2\pi\frac{(ux+vy)}{N}}$$
$$=\frac{1}{N^2}\left(1+e^{-j2\pi\frac{v}{N}}+e^{j2\pi\frac{v}{N}}+e^{-j2\pi\frac{u}{N}}+e^{j2\pi\frac{u}{N}}\right)$$
$$=\frac{1}{N^2}[1+2\cos(2\pi\frac{u}{N})+2\cos\left(2\pi\frac{v}{N}\right)]$$

In the above relationship we can easily see that there are frequency pairs for which $H_1(u,v)=0$, as for example $u=2N/3$ and $v=N/4$ in the case where both $2N/3$ and $N/4$ are integers. Therefore, a Pseudo Inverse Filter must be used.

For the second point spread function we see that

$$H_2(u,v)=\frac{1}{N^2}[5+2\cos(2\pi\frac{u}{N})+2\cos\left(2\pi\frac{v}{N}\right)]$$

Due to the first term of 5 inside the brackets, the above function is always positive and therefore Inverse Filtering can be applied.

(b) (i) Explain why, when an image is corrupted by additive random noise, the noise is more visible in smooth regions with low contrast than in regions of high contrast and texture. **[2]**

(ii) 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. **[2]**

(iii) Explain the role of the regularization parameter α in the Constrained Least Squares image restoration method. **[2]**

(iv) Propose a spatially adaptive Constrained Least Squares image restoration method which employs a spatially varying α in order to restore image g defined above. **[2]**

Solution

(i) This is because, if we calculate a local Signal-to-Noise-Ratio instead of a global one, this will be smaller in background areas where the image intensity is slowly varying. In other words, the power of the original signal will be negligible compared to the power of the noise in slowly varying areas.

(ii) CLS restoration refers to the solution of the following constrained minimization problem.

Minimize: $J(\mathbf{f}) = \|\mathbf{n}(\mathbf{f})\|^2 = \|\mathbf{y} - \mathbf{Hf}\|^2$

Subject to: $\|\mathbf{Cf}\|^2 < \varepsilon$

where $\mathbf{Cf}$ is a high pass filtered version of the image.

The idea behind the above constraint is that the highpass version of the image contains a considerably large amount of noise.

The minimization of the above leads to the following estimate for the original image

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

where α is the so-called regularization parameter that controls the contribution between the two terms in the cost function.

In frequency and under the presence of noise we have:

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

Bookwork

(iii) For large regularization parameter, we achieve strong elimination of high frequencies. This is desirable if the image of interest is heavily corrupted by noise. For small regularization parameter the CLS restoration tends to imitate Inverse Filtering.

(iv) For each pixel we can calculate the local variance. If this is larger than a pre-specified threshold we use a small regularization parameter. Otherwise we use a large one.

4. (a) The characters a to h have the set of frequencies based on the first 8 Fibonacci numbers as follows: **[5]**

symbol	frequency
a	1
b	1
c	2
d	3
e	5
f	8
g	13
h	21

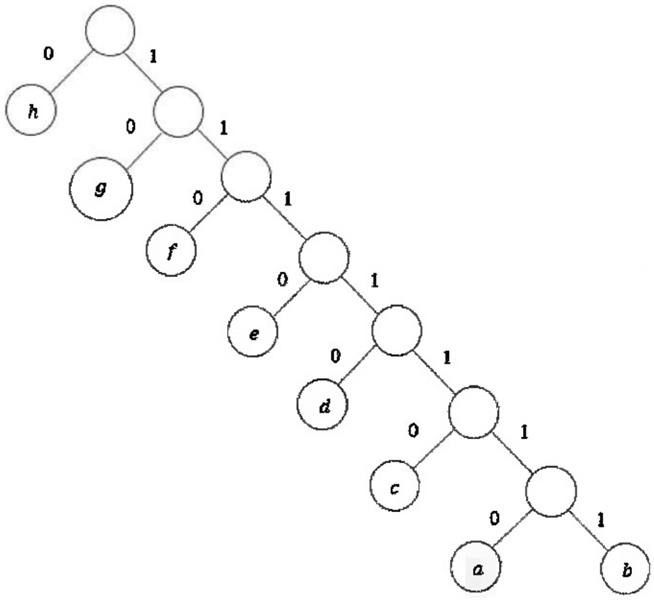
Figure 4.1

(i) A Huffman code is used to represent the characters. What is the sequence of characters corresponding to the code 110111100111010?

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

Solution

(i) The Huffman code is shown below.



Character	Code
a	1111110
b	1111111
c	111110
d	11110
e	1110
f	110
g	10
h	0

Figure 4.1

(ii) The entropy of the source is 2.37.
Average number of bits per symbol is 3.
Redundancy $3 - 2.7 = 0.3$.

(b) A networking company uses a compression technique to encode the message before transmitting over the network. Suppose the message contains the following characters with their frequency:

symbol	frequency
a	5
b	9
c	12
d	13
e	16
f	45

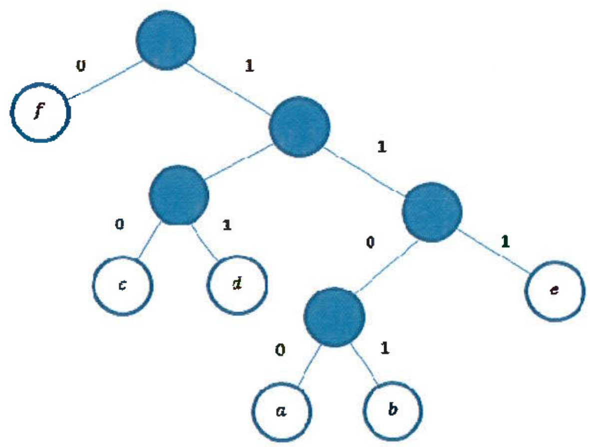
Figure 4.2

(i) Determine the Huffman code for this example. **[5]**

(ii) Note that each character in the input message takes 1 byte. If the compression technique used is Huffman Coding, how many bits will be saved in the message? **[5]**

Solutions

(i) The Huffman code is shown below.



(ii) Total number of characters = sum of frequencies = 100
size of 1 character = 1 byte = 8 bits
Total number of bits = $8 \cdot 100 = 800$
Using Huffman Encoding, Total number of bits needed can be calculated as:

$$5 \cdot 4 + 9 \cdot 4 + 12 \cdot 3 + 13 \cdot 3 + 16 \cdot 3 + 45 \cdot 1 = 224$$

Bits saved $= 800 - 224 = 576$.