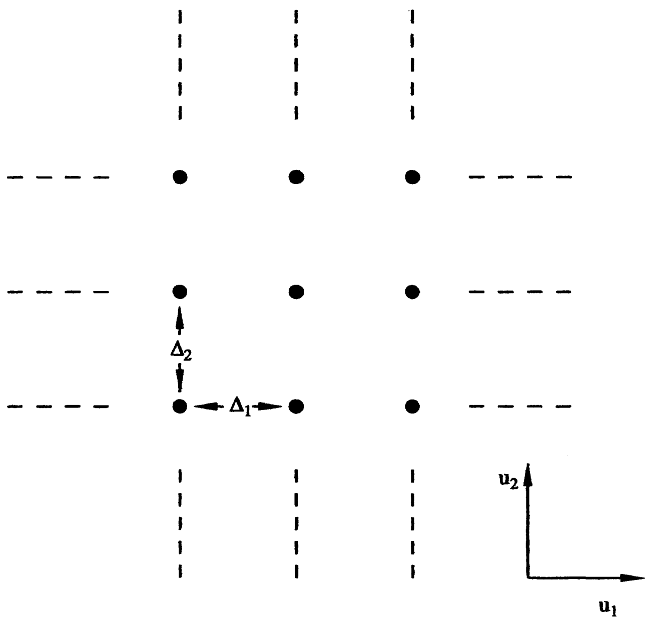
Module 4F8, April 2005 – IMAGE PROCESSING AND IMAGE CODING – Solutions

1 (a) **Sampled Spectrum**

Consider the sampling grid shown here:



The sampling function for this grid is clearly

$$s(u_1, u_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \delta(u_1 - n_1\Delta_1, u_2 - n_2\Delta_2)$$

The sampled image $g_s(u_1, u_2)$ can then be written as

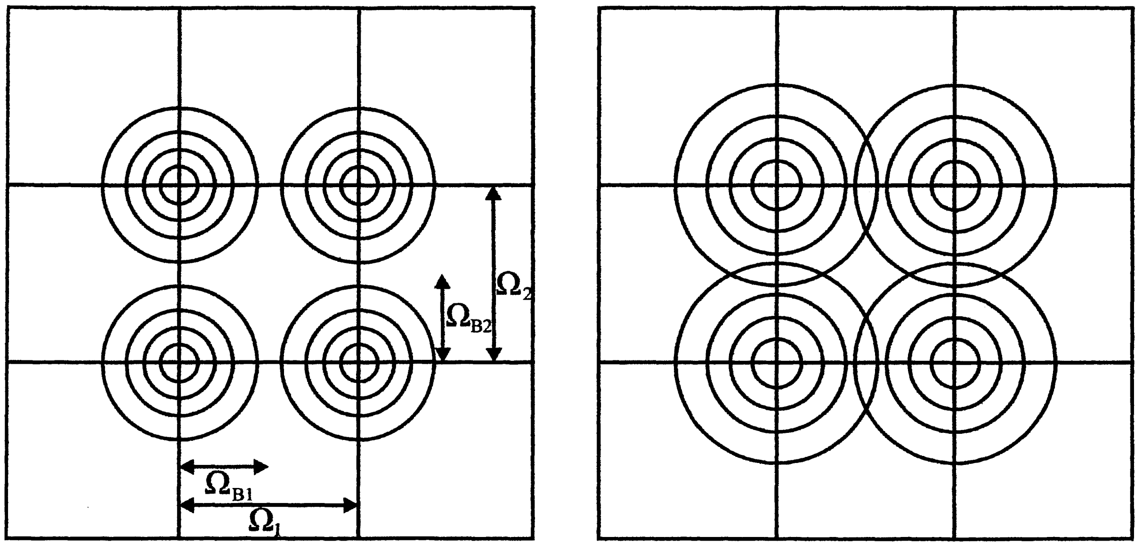
$$g_s(u_1, u_2) = s(u_1, u_2)\, g(u_1, u_2)$$

Because $s(u_1, u_2)$ is periodic, we can write it in terms of a Fourier series and using this we arrive at the following result: the fourier transform of $g_s(u_1, u_2)$ is $G_s(\omega_1, \omega_2)$ where

$$\boxed{G_s(\omega_1, \omega_2) = \frac{1}{\Delta_1 \Delta_2} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} G(\omega_1 - p_1\Omega_1, \omega_2 - p_2\Omega_2)}$$

$\Omega_i = \frac{2\pi}{\Delta_i}, i = 1, 2$. [There is no need to prove this, the question only asks you to write it down].

We can see from this expression that the FT of the sampled signal (up to a scale factor) is the spectrum of the original signal reproduced at each grid point.



It can be seen in left of the above figure that if the image is spatially bandlimited to Ω_{B1} in the u_1 direction and Ω_{B2} in the u_2 direction, then the original continuous image can be recovered from the sampled image by ideal low-pass filtering at Ω_{B1}, Ω_{B2}, provided that the samples are taken such that $\Omega_{B1} < \frac{1}{2}\Omega_1$ and $\Omega_{B2} < \frac{1}{2}\Omega_2$, so that the periodic repeats of the spectrum do not overlap.

This can also be written as:

$$\Omega_1 = \frac{2\pi}{\Delta_1} > 2\Omega_{B1}$$
$$\Omega_2 = \frac{2\pi}{\Delta_2} > 2\Omega_{B2}$$

$2\Omega_{B1}$ and $2\Omega_{B2}$ are known as the **2D Nyquist frequencies**. Thus the *2-D sampling theorem* states that a bandlimited image sampled at or above its u_1 and u_2 Nyquist rates can be recovered without error by low-pass filtering the sampled spectrum. If we sample below the Nyquist rates, as in the right of the above figure, the spectrum will display *aliasing* which, as in the 1D case, occurs when we have overlap of the repeated unsampled spectra in the frequency plane. If aliasing occurs we are not able to recover the true spectrum without error. [25%]

(b) Hexagonal sampling

By inspection, the sampling function, s_1, for the grid marked by squares is given by

$$s_1(u_1,u_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \delta(u_1 - n_1\Delta_1, u_2 - n_2\Delta_2)$$

where $\Delta_1 = 2s$ and $\Delta_2 = \sqrt{3}s$.

Similarly, the sampling function, s_2, for the grid marked by circles can be seen to be

$$s_2(u_1,u_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \delta(u_1 - [n_1 + \tfrac{1}{2}]\tilde{\Delta}_1, u_2 - [n_2 + \tfrac{1}{2}]\tilde{\Delta}_2)$$

where $\tilde{\Delta}_1 = s$ and $\tilde{\Delta}_2 = \sqrt{3}s$.

Writing s_1 and s_2 as fourier series we have

$$s_1(u_1,u_2) = \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} c_1(p_1,p_2)\, e^{j(p_1\Omega_1 u_1 + p_2\Omega_2 u_2)}$$

and

$$s_2(u_1,u_2) = \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} c_2(p_1,p_2)\, e^{j(p_1\tilde{\Omega}_1 u_1 + p_2\tilde{\Omega}_2 u_2)}$$

where:

$$\Omega_1 = \frac{2\pi}{\Delta_1}, \quad \Omega_2 = \frac{2\pi}{\Delta_2} \quad \text{and} \quad \tilde{\Omega}_1 = \frac{2\pi}{\tilde{\Delta}_1}, \quad \tilde{\Omega}_2 = \frac{2\pi}{\tilde{\Delta}_2}$$

c_1 and c_2 are given, so we can write the sampled function $g_s(u_1,u_2)$ as

$$g_s(u_1,u_2) = \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \quad g(u_1,u_2)\, c_1(p_1,p_2)\, e^{j(p_1\Omega_1 u_1 + p_2\Omega_2 u_2)} + g(u_1,u_2)\, c_2(p_1,p_2)\, e^{j(p_1\tilde{\Omega}_1 u_1 + p_2\tilde{\Omega}_2 u_2)}$$

and therefore, using the shift theorem, the FT of g as

$$G_s(\omega_1,\omega_2) = \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \quad c_1(p_1,p_2)\, G(\omega_1 - \Omega_1 p_1,\ \omega_2 - \Omega_2 p_2) + c_2(p_1,p_2)\, G(\omega_1 - \tilde{\Omega}_1 p_1,\ \omega_2 - \tilde{\Omega}_2 p_2)$$

Substituting the given values for c_1 and c_2, we have

$$G_s(\omega_1,\omega_2) = \frac{1}{2\sqrt{3}s^2} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \quad G\left(\omega_1 - \frac{\pi p_1}{s},\ \omega_2 - \frac{2\pi p_2}{\sqrt{3}s}\right) + 2G\left(\omega_1 - \frac{2\pi p_1}{s},\ \omega_2 - \frac{2\pi p_2}{\sqrt{3}s}\right) e^{-j(p_1+p_2)\pi}$$

To get this into the given form note that we can write the above equation as

$$G_s(\omega_1, \omega_2) = \frac{1}{2\sqrt{3}\,s^2} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \quad G(\omega_1 - \frac{\pi p_1}{s},\ \omega_2 - \frac{2\pi p_2}{\sqrt{3}s})$$
$$+ \ 2\alpha\, G(\omega_1 - \frac{\pi p_1}{s},\ \omega_2 - \frac{2\pi p_2}{\sqrt{3}s})\, e^{-j(p_1/2+p_2)\pi}$$

where, in the second term, we have replaced p_1 by $p_1/2$ and chosen $\alpha = 0$ if p_1 is odd and $\alpha = 1$ if p_1 is even in order to correct for this change. Hence the second term only contributes when p_1 is even and thus we cope with the difference between Ω_1 and $\tilde{\Omega}_1$. The form of $f(p_1, p_2)$ then follows straightforwardly from this and f is given by

$$f(p_1, p_2) = \frac{1}{2\sqrt{3}\,s^2}\left[1 + 2\alpha \exp\left(-j(\frac{p_1}{2} + p_2)\pi\right)\right]$$

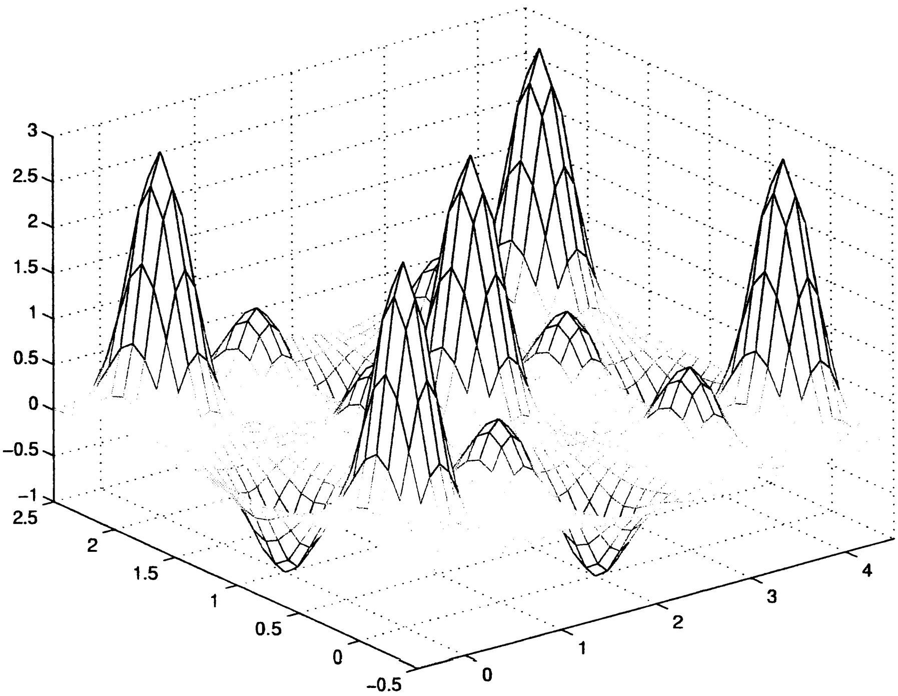
with α as above. [60%]

(c) **Contour plots of sampled spectrum**

A matrix of values for $f(p_1, p_2)$ for $p_1 = 0, 1, 2, 3, 4\ldots$ (horizontally) and $p_2 = 0, 1, 2\ldots$ (vertically upwards) is:

$$f(p_1, p_2) = \frac{1}{2\sqrt{3}\,s^2}\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \\ 3 & 1 & -1 & 1 & 3 & \cdots \\ -1 & 1 & 3 & 1 & -1 & \cdots \\ 3 & 1 & -1 & 1 & 3 & \cdots \end{bmatrix}$$

Hence the form of G_s will be G repeated at intervals of $\frac{\pi}{s}$ horizontally and $\frac{2\pi}{\sqrt{3}s}$ vertically with amplitude scaling factors given by $f(p_1, p_2)$ above. If G is a 2-D Gaussian lowpass function, then G_s will be of the form shown below (a sketch of a contour plot with labelled peak values, would be adequate in the exam, as mesh plots are difficult to draw). [15%]



2 (a) **Bandpass filter**

Impulse response of the lowpass filter, H_s, described is given by

$$\begin{aligned} h(n_1,n_2) &= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_s(\omega_1,\omega_2)\, e^{j(\omega_1 n_1\Delta_1+\omega_2 n_2\Delta_2)}\, d(\Delta_1\omega_1)\, d(\Delta_2\omega_2) \\ &= \frac{\Delta_1\Delta_2}{(2\pi)^2}\int_{-\pi/\Delta_2}^{\pi/\Delta_2}\int_{-\pi/\Delta_1}^{\pi/\Delta_1} H_s(\omega_1,\omega_2)\, e^{j(\omega_1 n_1\Delta_1+\omega_2 n_2\Delta_2)}\, d\omega_1\, d\omega_2 \end{aligned}$$

putting in the form of H_s gives

$$\begin{aligned} h(n_1,n_2) &= \frac{\Delta_1\Delta_2}{(2\pi)^2}\int_{-\Omega_2}^{\Omega_2}\int_{-\Omega_1}^{\Omega_1} e^{j(\omega_1 n_1\Delta_1+\omega_2 n_2\Delta_2)}\, d\omega_1\, d\omega_2 \\ &= \frac{\Delta_1\Delta_2}{(2\pi)^2}\int_{-\Omega_1}^{\Omega_1} e^{j\omega_1 n_1\Delta_1}\, d\omega_1 \int_{-\Omega_2}^{\Omega_2} e^{j\omega_2 n_2\Delta_2}\, d\omega_2 \\ &= \frac{\Delta_1\Delta_2\Omega_{U1}\Omega_{U2}}{(\pi)^2}\mathrm{sinc}(\Omega_{U1}n_1\Delta_1)\,\mathrm{sinc}(\Omega_{U2}n_2\Delta_2) \end{aligned}$$

which is the result given, using the following working:

$$\int_{-\Omega}^{\Omega} e^{j\omega n\Delta} d\omega = \left[\frac{e^{j\omega n\Delta}}{jn\Delta}\right]_{-\Omega}^{\Omega} = \frac{e^{j\Omega n\Delta}-e^{-j\Omega n\Delta}}{jn\Delta} = \frac{2\sin(\Omega n\Delta)}{n\Delta} = 2\Omega\,\mathrm{sinc}(\Omega n\Delta)$$

The easiest way of finding impulse response of the filter in fig.2 is by considering a linear combination of lowpass filters, but can also do it by considering the given filter to be the difference of the two standard bandpass filters covered in lectures.

First way:

If we define the following lowpass filters:

$$H_1(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{U1} \text{ and } |\omega_2| < \Omega_{L2} \\ 0 & \text{otherwise} \end{cases}$$

$$H_2(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{L1} \text{ and } |\omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

$$H_3(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{L1} \text{ and } |\omega_2| < \Omega_{L2} \\ 0 & \text{otherwise} \end{cases}$$

The filter, H, in fig 2 can be formed as $H = (H_1 - H_3) + (H_2 - H_3) = H_1 + H_2 - 2H_3$.

Thus the impulse response, h, is given by the corresponding linear combination of impulse responses: $h = h_1 + h_2 - 2h_3$. But the h_i can easily be written down from result in first part of the question. Therefore

$$\begin{aligned} h(n_1,n_2) &= \frac{\Delta_1\Delta_2}{\pi^2}\left[\Omega_{U1}\,\Omega_{L2}\,\mathrm{sinc}(\Omega_{U1}n_1\Delta_1)\,\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)\right. \\ &\quad + \Omega_{L1}\,\Omega_{U2}\,\mathrm{sinc}(\Omega_{L1}n_1\Delta_1)\,\mathrm{sinc}(\Omega_{U2}n_2\Delta_2) \\ &\quad \left. - 2\,\Omega_{L1}\,\Omega_{L2}\,\mathrm{sinc}(\Omega_{L1}n_1\Delta_1)\,\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)\right] \end{aligned}$$

Second way:

This is more involved, but some candidates may write down straightway the bandpass filters covered in the lectures. Consider the bandpass filter given in the figure – one way to construct this is to say that the ideal frequency response of this filter, $H(\omega_1, \omega_2)$, can be written as

$$H(\omega_1, \omega_2) = H_1(\omega_1, \omega_2) - H_2(\omega_1, \omega_2)$$

where H_1 is a rectangular bandpass filter given by $H_{1a} - H_{1b}$

$$H_{1a}(\omega_1, \omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{U1} \text{ and } |\omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

$$H_{1b}(\omega_1, \omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{L1} \text{ and } |\omega_2| < \Omega_{L2} \\ 0 & \text{otherwise} \end{cases}$$

and H_2 is the separable ideal bandpass filter representing just the small corner rectangles, given by

$$H_2(\omega_1, \omega_2) = \begin{cases} 1 & \text{if } \Omega_{L1} < |\omega_1| < \Omega_{U1} \text{ and } \Omega_{L2} < |\omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

We know that since H_2 is separable, we can write it as the product of two 1d filters, i.e.

$$H_2(\omega_1, \omega_2) = H_a(\omega_1)\, H_b(\omega_2)$$

where $H_a(\omega_1)$ is an ideal 1-D bandpass filter with a lower cut-off frequency of Ω_{L1} and an upper cut-off frequency of Ω_{U1}. Similarly $H_b(\omega_2)$ is an ideal 1-D bandpass filter with cut-off frequencies Ω_{L2} and Ω_{U2}. More explicitly we have

$$H_a(\omega_1) = \begin{cases} 1 & \text{if } \Omega_{L1} < |\omega_1| < \Omega_{U1} \\ 0 & \text{otherwise} \end{cases}$$

$$H_b(\omega_2) = \begin{cases} 1 & \text{if } \Omega_{L2} < |\omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

Thus, we can now work out the ideal impulse response of the filter from the impulse responses of H_1 and H_2. We have (where $h(n_1, n_2) \equiv h(n_1\Delta_1, n_2\Delta_2)$)

$$h(n_1, n_2) = \frac{\Delta_1\Delta_2}{(2\pi)^2} \int_{-\pi/\Delta_2}^{\pi/\Delta_2} \int_{-\pi/\Delta_1}^{\pi/\Delta_1} H_s(\omega_1, \omega_2) e^{j(\omega_1 n_1 \Delta_1 + \omega_2 n_2 \Delta_2)} d\omega_1 d\omega_2$$

from first part of question we can write down the impulse responses for H_{1a} and H_{1b}:

$$h_{1a}(n_1, n_2) = \frac{\Delta_1\Delta_2\Omega_{U1}\Omega_{U2}}{(\pi)^2} \text{sinc}(\Omega_{U2} n_2 \Delta_2)\, \text{sinc}(\Omega_{U1} n_1 \Delta_1)$$

Similarly, for H_{1b} we have

$$h_{1b}(n_1, n_2) = \frac{\Delta_1\Delta_2\Omega_{L1}\Omega_{L2}}{(\pi)^2} \text{sinc}(\Omega_{L2} n_2 \Delta_2)\, \text{sinc}(\Omega_{L1} n_1 \Delta_1)$$

The impulse response for H_2 is similarly given by

$$h(n_1, n_2) = \frac{\Delta_1}{2\pi} \int_{-\pi/\Delta_1}^{\pi/\Delta_1} H_a(\omega_1)\, e^{j\omega_1 n_1 \Delta_1}\, d\omega_1 \cdot \frac{\Delta_2}{2\pi} \int_{-\pi/\Delta_2}^{\pi/\Delta_2} H_b(\omega_2) e^{j\omega_2 n_2 \Delta_2}\, d\omega_2$$

$$= \frac{\Delta_1 \Delta_2}{(2\pi)^2} \left[\int_{-\Omega_{U1}}^{\Omega_{U1}} e^{j\omega_1 n_1 \Delta_1}\, d\omega_1 - \int_{-\Omega_{L1}}^{\Omega_{L1}} e^{j\omega_1 n_1 \Delta_1}\, d\omega_1 \right] \left[\int_{-\Omega_{U2}}^{\Omega_{U2}} e^{j\omega_2 n_2 \Delta_2}\, d\omega_2 - \int_{-\Omega_{L2}}^{\Omega_{L2}} e^{j\omega_2 n_2 \Delta_2}\, d\omega_2 \right]$$

where we have formed H_a and H_b as the difference of two lowpass filters in each case.

Thus we have

$$\begin{aligned} h_2(n_1, n_2) &= \frac{\Delta_1 \Delta_2}{(\pi)^2} \left[\Omega_{U1} \operatorname{sinc}(\Omega_{U1}\, n_1 \Delta_1) - \Omega_{L1} \operatorname{sinc}(\Omega_{L1}\, n_1 \Delta_1)\right] \\ &\quad \times \left[\Omega_{U2} \operatorname{sinc}(\Omega_{U2}\, n_2 \Delta_2) - \Omega_{L2} \operatorname{sinc}(\Omega_{L2}\, n_2 \Delta_2)\right] \end{aligned}$$

Now, forming $h(n_1, n_2)$ from the difference of $h_1(n_1, n_2)$ and $h_2(n_1, n_2)$ we have

$$\begin{aligned} h(n_1, n_2) &= \frac{\Delta_1 \Delta_2}{\pi^2} [\Omega_{U1}\, \Omega_{L2} \operatorname{sinc}(\Omega_{U1} n_1 \Delta_1) \operatorname{sinc}(\Omega_{L2} n_2 \Delta_2) \\ &\quad + \Omega_{L1}\, \Omega_{U2} \operatorname{sinc}(\Omega_{L1} n_1 \Delta_1) \operatorname{sinc}(\Omega_{U2} n_2 \Delta_2) \\ &\quad - 2\Omega_{L1}\, \Omega_{L2} \operatorname{sinc}(\Omega_{L1} n_1 \Delta_1) \operatorname{sinc}(\Omega_{L2} n_2 \Delta_2)] \end{aligned}$$

[60%]

b) **Histogram equalisation**

The differential equation we start with is

$$p_Y(y)dy = p_X(x)dx$$

or

$$\frac{dy}{dx} = \frac{p_X(x)}{p_Y(y)}$$

It is required that the output image probability density $p_Y(y)$ be constant over the grey level range O to L.

$$p_Y(y) = \frac{1}{L}$$

$$\frac{dy}{dx} = L\,p_X(x)$$

$$y = g(x) = \int_0^x L\,p_X(u)du$$

(We assume that the lowest value of input image pixels is 0).

In practice the input image probability density is not known and is approximated by the image histogram and the integral is approximated by a sum.

Let the input image be quantised into N_L levels x_i spaced by Δx_i then $N_L \Delta x_i = L$ and

$$y_k = \sum_{i=0}^{k} L\,p_X(x_i)\Delta x_i \quad \text{for} \quad k = 0, .., N_L - 1$$

Now

$$p_X(x_i)\Delta x_i = Pr\left\{x_i \leq X \leq x_i + \Delta x_i\right\}$$

so if the histogram of the image has N_i occurrences in the bin x_i to $x_i + \Delta x_i$ then

$$p_X(x_i)\Delta x_i = Pr\left\{x_i \leq X \leq x_i + \Delta x_i\right\} = \frac{N_i}{N \times M}$$

where N and M are the dimensions of the image in pixels. thus the mapping rule becomes:

$$y_k = \sum_{i=0}^{k} L\frac{N_i}{NM}$$

We can see that if $k = N_L - 1$, $y_{N_L-1} = L$ as required. This may be considered as a look-up table – i.e. the above values of y_k are formed and stored so that we can scan our image x and when the value of pixel i in x falls within the kth greylevel bin, we map it to y_k. [40%]

3 (a) **DCT matrix**

For the forward DCT in 2-D, we may transform first the columns of **X**, by premultiplying by **T**, and then the rows of the result by postmultiplying by $\mathbf{T}^T$. Hence:

$$\mathbf{Y} = \mathbf{T}\,\mathbf{X}\,\mathbf{T}^T$$

Since **T** is orthonormal, $\mathbf{T}^{-1} = \mathbf{T}^T$ and $\mathbf{T}^T\mathbf{T} = \mathbf{I}$

Hence, to invert the DCT and get **X** on the RHS we premultiply the above equation by $\mathbf{T}^T$, and postmultiply by **T** to get

$$\mathbf{T}^T\,\mathbf{Y}\,\mathbf{T} = \mathbf{T}^T\,\mathbf{T}\,\mathbf{X}\,\mathbf{T}^T\,\mathbf{T} = \mathbf{I}\,\mathbf{X}\,\mathbf{I} = \mathbf{X}$$

This is the required operation. [20%]

(b) **Sum of basis functions**

The n^2 basis functions represent the contributions of each of the n^2 elements (or coefficients) of **Y** to the output image **X**.

In the expression $\mathbf{T}^T\ \mathbf{Y}\ \mathbf{T}$ above, we may consider the effect of each element $y_{i,j}$ in **Y** separately, by evaluating the case when **Y** is a matrix of zeros apart from a single non-zero element $y_{i,j}$ at location (i, j).

In this case we see that the contribution of $y_{i,j}$ to **X** is given by:

$$\mathbf{X}(i,j) = \mathbf{t}_i^T\, y_{i,j}\, \mathbf{t}_j = y_{i,j}\, \mathbf{t}_i^T\, \mathbf{t}_j = y_{i,j}\, \mathbf{T}(i,j)$$

where $\mathbf{t}_i$ and $\mathbf{t}_j$ are the i^{th} and j^{th} row vectors from **T**, and $\mathbf{T}(i,j)$ is the $n \times n$ basis function matrix for the coefficient $y_{i,j}$.

Hence the 2-D basis functions $\mathbf{T}(i,j)$, from which **X** may be reconstructed, are given by the n^2 separate products of the column vectors $\mathbf{t}_i^T$ with the row vectors $\mathbf{t}_j$:

$$\mathbf{T}(i,j) = \mathbf{t}_i^T\, \mathbf{t}_j \qquad \text{so that} \qquad \mathbf{X} = \sum_{i=1}^{n}\sum_{j=1}^{n} y_{i,j}\, \mathbf{T}(i,j)$$

[25%]

(c) **Coding of a large image**

For the case where the DCT size $n = 4$, there are 16 subbands but only 7 separate values for $\sigma_{i,j}$, corresponding to

$$p = i + j = 2 \ldots 8 \qquad \text{where} \qquad \sigma_p = \frac{A}{p}$$

When $Q = A/2$, the entropy of the $(p-1)^{th}$ diagonal row of subbands is then given by

$$
\begin{aligned}
H_p &= \tfrac{1}{2}\log_2\left(\frac{2\,e^2\,(A/p)^2}{(A/2)^2}+1\right) = \tfrac{1}{2}\log_e\left(\frac{8\,e^2}{p^2}+1\right)/\log_e 2 \\
&= 1.9899 \quad \text{when} \;\; p=2 \\
&= 1.4600 \quad \text{when} \;\; p=3 \\
&= 1.1155 \quad \text{when} \;\; p=4 \\
&= 0.8752 \quad \text{when} \;\; p=5 \\
&= 0.7008 \quad \text{when} \;\; p=6 \\
&= 0.5708 \quad \text{when} \;\; p=7 \\
&= 0.4719 \quad \text{when} \;\; p=8
\end{aligned}
$$

This represents the number of bits per coefficient required for each subband.

The 4×4 subbands occur in diagonal rows for a given p, so the number of subbands for $p = 2 \ldots 8$ is $\{1,2,3,4,3,2,1\}$.

Each subband contains $k^2 = 256^2 = 64$ K coefficients.

Hence the total number of bits for the image is

$$
\begin{aligned}
N_{\text{bits}} &= (H_2 + 2H_3 + 3H_4 + 4H_5 + 3H_6 + 2H_7 + H_8)\,.\,k^2 \\
&= 15.4732\,.\,64\text{ K} = 990.28\text{ K} = 1014048 \text{ bits}
\end{aligned}
$$

[35%]

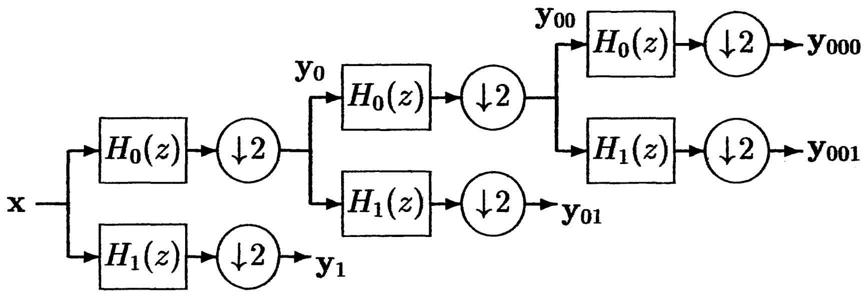
(d) **Coarser quantisers and colour**

DCT coefficients representing higher frequency basis functions can usually be coded more coarsely than those for lower frequencies because the higher frequency basis functions are less visible to the human eye. This is because the contrast sensitivity of the eye falls off at high frequencies.

It is also found that the sensitivity of the eye to colour difference (chrominance) components falls off at much lower frequencies than the sensitivity to brightness (luminance) components. Hence the red and blue chrominance images can be subsampled by at least 4:1 (2:1 in each direction) relative to the luminance image sample rate, so the total area of the two chrominance images is only half that of the luminance image. In addition the contrast sensitivity to chrominance components is lower than for luminance so the chrominance coefficients can be quantised more coarsely and the increase in bit rate will then be less than 50% . [20%]

4 (a) **Three-level Wavelet transforms**

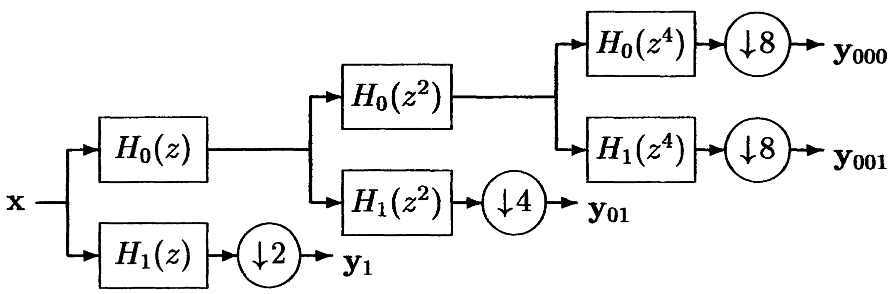
This is the basic 3-level wavelet tree:



3-level wavelet filter analysis tree

To simplify calculation of overall transfer functions, we must move the down-samplers from between the filter stages to the tree outputs, using the equivalence between fig. 3a and fig. 3b, given in the question. Whenever we move a down-sampler from input to output of a filter, we must replace z by z^2 in the filter transfer function.

This is the tree which has been re-arranged so that calculation of overall transfer functions is simplified:



3-level tree, rearranged with all down-samplers at the outputs.

[30%]

(b) **Transfer functions**

In an m-level tree, the signal passes through $m-1$ lowpass filters, followed by either a highpass filter or another lowpass filter, depending on the chosen output.

Hence the transfer function to the lowpass output at level m is

$$H_{0\cdots 00} = \prod_{k=1}^{m} H_0(z^{2^{k-1}})$$

and the transfer function to the highpass output at level m is

$$H_{0\cdots 01} = \left[\prod_{k=1}^{m-1} H_0(z^{2^{k-1}}) \right] H_1(z^{2^{m-1}})$$

[20%]

(c) Transfer functions for a 2-level tree

Using the previous result when $m = 2$ with the given filters:

$$
\begin{aligned}
H_{00} &= H_0(z)\, H_0(z^2) \\
&= \tfrac{1}{4}(z + 2 + z^{-1})\, \tfrac{1}{4}(z^2 + 2 + z^{-2}) \\
&= \tfrac{1}{16}(z^3 + 2z^2 + 3z + 4 + 3z^{-1} + 2z^{-2} + z^{-3})
\end{aligned}
$$

and

$$
\begin{aligned}
H_{01} &= H_0(z)\, H_1(z^2) \\
&= \tfrac{1}{4}(z + 2 + z^{-1})\, \tfrac{1}{8}(-z^2 - 2 + 6z^{-2} - 2z^{-4} - z^{-6}) \\
&= \tfrac{1}{32}(-z^3 - 2z^2 - 3z - 4 + 4z^{-1} + 12z^{-2} + 4z^{-3} - 4z^{-4} - 3z^{-5} - 2z^{-6} - z^{-7})
\end{aligned}
$$

[20%]

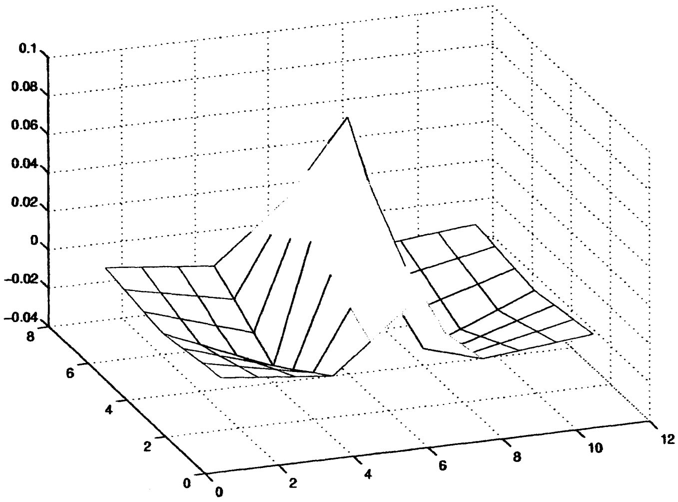
(d) 2-d impulse response

In a separable 2-d filter, when two 1-d filters with impulse responses $\mathbf{h}_a$ (in the column direction) and $\mathbf{h}_b$ (in the row direction) are combined, the matrix representing the 2-d impulse response is given by:

$$\mathbf{H}_{ab} = \mathbf{h}_a\, \mathbf{h}_b^T$$

For a level-2 Hi-Lo filter, the two filters in part (c) are combined with H_{00} applied in the column direction and H_{01} in the row direction. Thus the filter vectors are

$$
\mathbf{h}_a = \frac{1}{16}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 3 \\ 2 \\ 1 \end{bmatrix} \qquad
\mathbf{h}_b = \frac{1}{32}\begin{bmatrix} -1 \\ -2 \\ -3 \\ -4 \\ 4 \\ 12 \\ 4 \\ -4 \\ -3 \\ -2 \\ -1 \end{bmatrix}
$$



and the mesh plot of $\mathbf{H}_{ab}$ is shown on the right, since $\mathbf{h}_a$ is a simple triangular function with peak value $\frac{1}{4}$, while $\mathbf{h}_b$ is a linear interpolation through the points $\frac{1}{32}\{-2, -4, 12, -4, -2\}$. The peak value of $\mathbf{H}_{ab}$ is $\frac{4}{16} \cdot \frac{12}{32} = \frac{3}{32}$.

Since the main feature of $\mathbf{H}_{ab}$ is the vertical ridge down its centre-line, this subband will respond strongly to vertical lines or edges in the image. The mean of $\mathbf{H}_{ab}$ is zero so it will not respond to regions constant intensity. [30%]

J Lasenby and N G Kingsbury, June 2005.