Engineering IIB: 4F8 Image Coding - Examples Paper

1. The 2-point one-dimensional Haar transform is given by:

$$\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \mathbf{T} \begin{bmatrix} x(1) \\ x(2) \end{bmatrix} \qquad \text{where} \quad \mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Derive an 8-point transform matrix that is equivalent to 3 levels of the Haar transform, in which only the lowpass output of each pair of transformed samples is subjected to further transformation. (Hint: equation 3.2 in the lecture notes shows the 4-point matrix that is equivalent to 2 levels of the Haar transform.) You should order the rows of your matrix to give increasing bandwidth with row number.

Show that the 8-point transform matrix is orthonormal and briefly describe how it might be used in an image compression application.

2. The n rows of an n-point DCT matrix $\mathbf{T}$ are defined by:

$$\begin{aligned} t_{1i} &= \sqrt{\frac{1}{n}} \qquad \text{for } i = 1 \to n, \\ t_{ki} &= \sqrt{\frac{2}{n}} \cos\left(\frac{\pi(2i-1)(k-1)}{2n}\right) \qquad \text{for } i = 1 \to n, k = 2 \to n. \end{aligned}$$

If n is even, prove that the rows of this matrix are orthogonal to each other and hence show that the matrix is orthonormal.

Why is orthonormality important for image compression?

3. A $2n$-point discrete Fourier transform (DFT) is applied to a vector:

$$\mathbf{x}_{2n} = \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_{rev} \end{bmatrix} \qquad \text{where} \quad \mathbf{x} = \begin{bmatrix} x(1) \\ \vdots \\ x(n) \end{bmatrix} \quad \text{and} \quad \mathbf{x}_{rev} = \begin{bmatrix} x(n) \\ \vdots \\ x(1) \end{bmatrix}$$

Show that the first n Fourier coefficients may be simply related to the coefficients of an n-point DCT on the vector $\mathbf{x}$, and describe the nature of this relationship.

4. A given bandpass sub-image, that results from applying a transform to an image, has pixel values which may be modelled by a Laplacian pdf, given by:

$$p(x) = c\, e^{-|x|/x_0}$$

Calculate c for this to be a valid pdf.

If the sub-image pixels are now quantised by an ideal uniform quantiser with step size q and reconstruction levels at kq, where k is an integer, calculate the discrete probability function p_k of the quantised values.

Calculate the variance and entropy of the quantised values and explain the relevance of these measures to image compression.

5. Using the discrete probability function from the previous question calculate the quantiser step size q which will result in $p_0 = 0.8$. Then, assuming this step size, define a set of Huffman code word lengths for the states $k = -3 \ldots 3$ and calculate the efficiency of this code if an ideal entropy code is assumed for all states outside this range.

Explain the main source of inefficiency and suggest a method for improving the efficiency.

6. Using the JPEG default codes given on pages 48 and 50 (section 3.5) of the notes, express the following 8×8 block of DCT coefficients as a bit stream:

$$\begin{array}{cccccccc}
11 & -9 & 3 & 1 & -1 & 0 & 0 & 0 \\
-7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Calculate the mean bit-rate per pixel to code this block.

7. In the transformed binary filter tree of fig 1, show that the z transfer function $H_{k,1}(z)$ to the bandpass output at level k is related to that at level $k-1$ by:

$$H_{k,1}(z) = H_0(z) H_{k-1,1}(z^2)$$

Given the LeGall analysis filter pair:

$$H_0(z) = \tfrac{1}{2}(z + 2 + z^{-1}) \qquad H_1(z) = \tfrac{1}{8} z^{-1}(-z^2 - 2z + 6 - 2z^{-1} - z^{-2})$$

calculate the transfer functions from the input of the filter tree $\mathbf{x}$ to the outputs $\mathbf{y}_{01}$ and $\mathbf{y}_{001}$.

Repeat this for the inverse LeGall analysis filter pair:

$$H_0(z) = \tfrac{1}{8}(-z^2 + 2z + 6 + 2z^{-1} - z^{-2}) \qquad H_1(z) = \tfrac{1}{2} z(-z + 2 - z^{-1})$$

Considering impulse responses, comment on which arrangement of analysis and reconstruction filter pairs would be preferable for image coding, giving reasons.

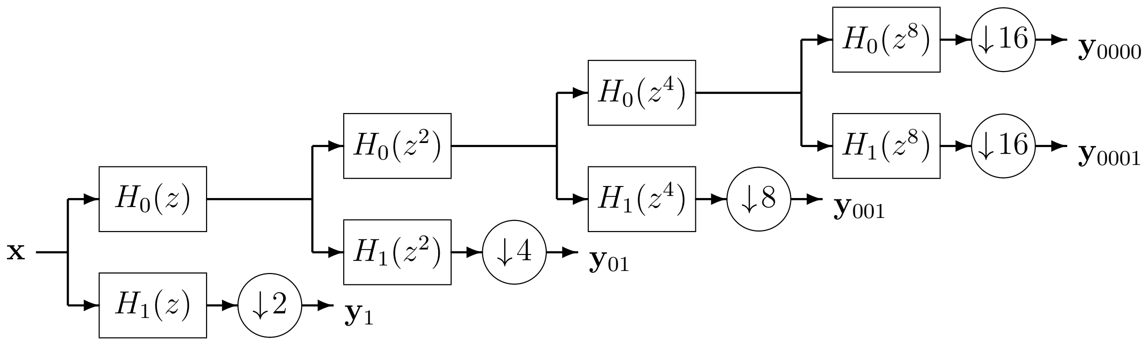

Fig 1: Binary filter tree, transformed so that all downsampling operations occur at the outputs.

8. A 5th-order transformed product filter for a 2-band filter bank is defined in terms of its factors as:

$$P_t(Z) = (1+Z)(1+aZ+bZ^2) \;\; (1+Z)(1+cZ)$$

Derive conditons on a and b in terms of c for $P_t(Z)$ to satisfy the perfect reconstruction (PR) condition.

In addition, it is now required that the left pair of factors of $P_t(Z)$ have the same values as the right pair of factors at $Z = -1, 0$ and $+1$. Show that c may be chosen to meet this requirement and calculate its value. Compare this with the rational value of $c = -\frac{2}{7}$ given in the lecture notes (page 65). Briefly explain why it is desirable for the left and right pair of factors to be similar over this range of Z.

9. A form of the transformation function, used in the design of perfect reconstruction filter banks is:

$$Z = pz^3 + (\tfrac{1}{2} - p)(z + z^{-1}) + pz^{-3}$$

It is required to obtain as many zeros as possible at $z = -1$ for each zero of $P_t(Z)$ at $Z = -1$. Calculate the value of p which achieves this (this is not the value given in the notes!).

With this p, express Z as a function of normalised frequency, ωT_s, and determine how many derivatives of Z are zero at $\omega T_s = \pi$. What does this imply about the magnitude of each $(Z+1)$ factor of $P_t(Z)$ at frequencies near $\omega T_s = \pi$?

If this transformation function is used and the transformed product filter is:

$$P_t(Z) = \tfrac{1}{50}(50 + 41Z - 15Z^2 - 6Z^3) \;\; \tfrac{1}{7}(7 + 5Z - 2Z^2)$$

obtain transfer functions $H_0(z)$ and $G_0(z)$ for the two lowpass filters of the analysis and reconstruction filter banks.

Nick Kingsbury, February 9, 2016.

Relevant Tripos Questions: All past 4F8 papers are relevant to this course.