Appendix B: Tutorial Problems

Chapter 1: Introduction

Q1.1 What is the primary purpose of standardization in video compression? List two other advantages of standardization.

Q1.2 Using the example of a DVB-T2 terrestrial broadcast system transmitting HDTV video content to the home, explain why digital video compression is needed.

Q1.3 Consider the case of 4 K UHDTV, with the original video in 4:2:2 (a luma signal of 3840×2160 and two chroma signals of 1920×2160) format at 10 bits and a frame rate of 50 fps. Calculate the compression ratio if this video is to be transmitted over a DVB-T2 link with an average bandwidth of 15 Mb/s.

Chapter 2: The human visual system

Q2.1 Assuming that the field of view within the fovea is 2°, compute the number of pixels that fall horizontally within the foveated visual field. Assume a 1 m wide screen with 1920 horizontal pixels viewed at a distance of $3H$.

Q2.2 The following table lists a number of important features of the human visual system (HVS). Complete the table by describing the influence each feature has on the design of a digital video compression system.

HVS characteristic	Implication for compression
HVS is more sensitive to high contrast image regions than low contrast regions	?
HVS is more sensitive to luminance than chrominance information	?
HVS is more sensitive to low spatial frequencies than high spatial frequencies	?
In order to achieve a smooth appearance of motion, the HVS must be presented with image frames above a certain minimum rate (and this rate depends on ambient light levels)	?
HVS responses vary from individual to individual	?

Q2.3 Calculate the normalized contrast sensitivity of the human visual system for a luminance-only stimulus at a spatial frequency of 10 cycles per degree.

Communicating Pictures. http://dx.doi.org/10.1016/B978-0-12-405906-1.00024-6

Chapter 3: Discrete-time analysis for images and video

Q3.1 Plot the sinusoidal signal $x(t) = \cos(20\pi t)$. Compute the frequency spectrum of this sinusoid and plot its magnitude spectrum.

Q3.2 Assume that the sinusoidal signal in Q3.1 is sampled with $T = 0.1$ s. Sketch the magnitude spectrum of the sampled signal and comment on any aliasing issues.

Q3.3 If an HDTV (full HD) screen with aspect ratio 16:9 has a width of 1.5 m and is viewed at distance of $4H$, what is the angle subtended by each pixel at the retina?

Q3.4 The impulse response for the Le Gall high-pass analysis filter is $h_1[n] = \left\{0.25 \quad -0.5 \quad 0.25\right\}$. Compute the output from this filter for the input sequence: $x[n] = \left\{1 \quad 2 \quad 3 \quad 0 \quad 0 \quad \ldots\right\}$.

Q3.5 Compute the z-plane pole zero plots and the frequency responses for the following filter pair: $H_0(z) = 1 + z^{-1}$ and $H_1(z) = 1 - z^{-1}$.

Q3.6 Perform median filtering on the following input sequence using a five-tap median filter and comment on the result:

$$x[n] = \{1, 3, 5, 13, 9, 11, 6, 15, 17, 19, 29\}$$

Q3.7 Compute the basis functions for the two-point DFT.

Q3.8 Consider the two 1-D digital filters:

$$\mathbf{h}_1 = [1 \quad 2 \quad 2 \quad 1]^{\mathrm{T}}$$
$$\mathbf{h}_2 = [1 \quad -3 \quad -3 \quad 1]^{\mathrm{T}}$$

Compute the equivalent 2-D digital filter where $\mathbf{h}_1$ performs horizontal filtering and $\mathbf{h}_2$ performs vertical filtering.

Q3.9 Compute biased and unbiased correlation estimates for the following sequence:

$$x[n] = \{1, 2, 5, -1, 3, 6, -4, -1\}$$

Q3.10 Plot the autocorrelation function, $r_v(k)$, for a white noise sequence $v[n]$ with variance σ_v^2. Form the autocorrelation matrix for this sequence for lags up to ± 3. What is the inverse of this autocorrelation matrix?

Q3.11 A feedback-based linear predictor with quantization uses a predictor $P(z) = (z^{-1} + z^{-2})/2$. Assume an input sequence as given below:

$$x[n] = \{1, 3, 4, 3, 5, 6 \ldots\}$$

and that quantization is performed as follows, with rounding of 0.5 values toward zero:

$$e_Q[n] = \mathrm{rnd}\left(\frac{e[n]}{2}\right); \quad e_R[n] = 2e_Q[n]$$

compute the predictor output sequence $e_Q[n]$ and the reconstructed output signal $y[n]$ from the decoder. Comment on your results in terms of numerical precision.

Q3.12 Compute the entropies of the sequences $x[n]$ and $e_Q[n]$ from Q3.11. Comment on your result.

Chapter 4: Digital picture formats and representations

Q4.1 In video coding schemes it is usual to code the color components in the form Y, C_b, C_r rather than R, G, B. Explain why this approach is justified, its benefits in terms of compression and how it is exploited in image sampling. Explain how pictures can be efficiently stored or transmitted using a 4:2:0 format.

Q4.2 Compute YUV and YC_bC_r vectors for the following RGB vectors:

(a) $[R, G, B] = \begin{bmatrix} 128 & 128 & 128 \end{bmatrix}$
(b) $[R, G, B] = \begin{bmatrix} 255 & 255 & 255 \end{bmatrix}$
(c) $[R, G, B] = \begin{bmatrix} 100 & 0 & 0 \end{bmatrix}$

Q4.3 If a color movie of 100 min duration is represented using ITU-R.601 (720 × 576, 25 fps@8 bits, 4:2:0 format):

(a) What hard disk capacity would be required to store the whole movie?
(b) If the movie is encoded at a compression ratio CR = 50:1 and transmitted over a satellite link with 50% channel coding overhead, what is the total bit rate required for the video signal?

Q4.4 Given a video sequence with a spatial resolution of 1920 × 1080 at 50 fps using 10 bit color sampling, compute the (uncompressed) bit rates of 4:4:4, 4:2:0, and 4:2:2 systems.

Q4.5 If, for a given 1920 × 1080 I-frame in 4:2:0 format, the mean entropy of each coded 16 × 16 luminance block is 1.3 bits/sample and that for each corresponding chrominance block is 0.6 bits/sample, then estimate the total number of bits required to code this frame.

Q4.6 Calculate the MAD between the following original image block $\mathbf{X}$ and its encoded and decoded version $\tilde{\mathbf{X}}$:

$$\mathbf{X} = \begin{bmatrix} 2 & 4 & 4 & 6 \\ 3 & 6 & 6 & 6 \\ 5 & 7 & 7 & 8 \\ 3 & 7 & 7 & 8 \end{bmatrix}; \quad \tilde{\mathbf{X}} = \begin{bmatrix} 2 & 5 & 5 & 5 \\ 4 & 4 & 6 & 6 \\ 5 & 5 & 7 & 8 \\ 4 & 5 & 7 & 7 \end{bmatrix}$$

Q4.7 Assuming a 5 bit digital image block, $\mathbf{X}$, the reconstruction after image compression is given by $\mathbf{Y}$. Calculate the PSNR for the reconstructed signal and provide an interpretation of the results in terms of error visibility.

$$\mathbf{X} = \begin{bmatrix} 3 & 8 & 1 & 8 \\ 7 & 0 & 5 & 0 \\ 2 & 6 & 0 & 5 \\ 4 & 1 & 10 & 2 \end{bmatrix}; \quad \mathbf{Y} = \begin{bmatrix} 4 & 10 & 1 & 9 \\ 7 & 0 & 6 & 1 \\ 3 & 6 & 0 & 4 \\ 5 & 1 & 12 & 3 \end{bmatrix}$$

Q4.8 Assuming an 8 bit digital image, $\mathbf{X}$, the reconstructions due to two alternative coding schemes are given by $\mathbf{Y}_1$ and $\mathbf{Y}_2$ below. Calculate the Peak Signal to Noise Ratio (PSNR) for each of the reconstructions and give an interpretation of the results in terms of error visibility.

$$\mathbf{X} = \begin{bmatrix} 20 & 17 & 18 \\ 15 & 14 & 15 \\ 19 & 13 & 14 \end{bmatrix}; \quad \mathbf{Y}_1 = \begin{bmatrix} 19 & 18 & 17 \\ 16 & 15 & 14 \\ 18 & 14 & 13 \end{bmatrix}; \quad \mathbf{Y}_2 = \begin{bmatrix} 20 & 17 & 18 \\ 15 & 23 & 15 \\ 19 & 13 & 14 \end{bmatrix}$$

Q4.9 Describe a typical GOP structure used in MPEG-2 television broadcasting, explaining the different types of frame coding employed, the predictive relationships between all frames in the GOP, and the transmission order of the frames. If a transmission error affects the second P-frame in the GOP, how many pictures are likely to be affected due to error propagation?

Q4.10 An HDTV satellite operator allocates 10 Mb/s to each program in the DVB-S multiplex. State what color sub-sampling format will be used for this and calculate the required compression ratio for a 1080i25 system.

Q4.11 Compute the gamma corrected version of the following image block for the case where $\gamma = 0.45$. Assume an 8 bit wordlength.

$$\mathbf{X} = \begin{bmatrix} 20 & 17 & 18 \\ 15 & 14 & 15 \\ 19 & 13 & 14 \end{bmatrix}$$

Chapter 5: Transforms for image and video coding

Q5.1 Derive the 1-D two-point KLT for a stationary real-valued process, x, that has the following autocorrelation matrix:

$$\mathbf{R}_x = \begin{bmatrix} r_{xx}(0) & r_{xx}(1) \\ r_{xx}(1) & r_{xx}(0) \end{bmatrix}$$

Q5.2 Prove that the following vectors are an orthonormal pair:

$$\mathbf{a}_0 = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^{\mathrm{T}}; \quad \mathbf{a}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}^{\mathrm{T}}$$

Q5.3 Prove that the four-point DWHT is a unitary transform.

Q5.4 Compute the basis functions for the eight-point 1-D DWHT. Compute the first four basis functions for the four-point 2-D DWHT.

Q5.5 Use the DWHT to transform the following 2-D image block, **S**. Assuming that all data and coefficients are represented as 6 bit numbers and that compression is achieved in the transform domain by selecting the four most dominant coefficients for transmission, compute the decoded data matrix and its PSNR.

$$\mathbf{S} = \begin{bmatrix} 5 & 6 & 8 & 10 \\ 6 & 6 & 5 & 7 \\ 4 & 5 & 3 & 6 \\ 8 & 7 & 5 & 5 \end{bmatrix}$$

Q5.6 Given the four-point DWHT basis functions, if the coefficients after transformation are: $c(0) = 1;\ c(1) = \frac{1}{4};\ c(2) = -\frac{1}{2};\ c(3) = 3$, plot the weighted basis functions and hence reconstruct the original signal waveform.

Q5.7 The 1-D discrete cosine transform is given by:

$$C(k) = \sqrt{\frac{2}{N}}\varepsilon_k \sum_{n=0}^{N-1} x[n] \cos\left(\frac{\pi(n+0.5)k}{N}\right) \quad 0 \le n, k \le N-1$$

where:

$$\varepsilon_k = \begin{cases} 1/\sqrt{2} & k = 0 \\ 1 & \text{otherwise} \end{cases}$$

Calculate the numerical values of the basis functions for a 1-D 4.4 DCT.

Q5.8 The DCT is an orthonormal transform. Explain the term orthonormal and describe what it means in practice for the transform. Write down the formula and the basis function matrix for the 1-D four-point DCT. Using this, show that, for the 1-D DCT:

$$\varepsilon_k = \begin{cases} 1/\sqrt{2} & k = 0 \\ 1 & \text{otherwise} \end{cases}$$

Q5.9 Compute the DCT transform coefficient vector for an input sequence, $\mathbf{x} = [1001]^{\mathrm{T}}$.

Q5.10 Calculate the DCT of the following 2 × 2 image block:

$$\mathbf{X} = \begin{bmatrix} 21 & 19 \\ 15 & 20 \end{bmatrix}$$

Q5.11 Quantize the result from Q5.9 using the following quantization matrix:

$$\mathbf{Q} = \begin{bmatrix} 4 & 8 \\ 8 & 8 \end{bmatrix}$$

Q5.12 Perform inverse quantization and an inverse DCT on the output from Q5.10.

Q5.13 Compute the 2-D DCT of the following (4 × 4) image block:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Q5.14 Given the following block of DCT coefficients and the associated quantization matrix, compute the block of quantized coefficients. Perform zig-zag scanning to form a string of {run/value} symbols (where "run" is the number of zeros preceding a non-zero value) appropriate for entropy coding.

$$\mathbf{C} = \begin{bmatrix} 128 & 50 & -20 & 22 & 12 & 27 & -5 & 7 \\ 40 & -25 & 26 & 20 & -34 & -2 & 13 & -5 \\ -10 & 22 & 12 & 12 & 26 & 12 & 3 & 8 \\ 12 & -2 & 16 & -7 & 9 & 3 & 17 & 17 \\ -32 & 6 & 21 & 9 & 18 & 5 & 4 & 7 \\ -10 & -7 & -14 & 3 & -2 & 13 & 18 & 18 \\ 11 & -9 & -9 & 4 & 8 & 13 & 6 & 9 \\ -7 & 19 & 15 & 8 & 6 & -6 & 18 & 33 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix}$$

Q5.15 Calculate the number of multiply and accumulate (MAC) operations required to compute a conventional (4 × 4) point 2-D DCT. Assume that the separability property of the 2-D DCT is exploited.

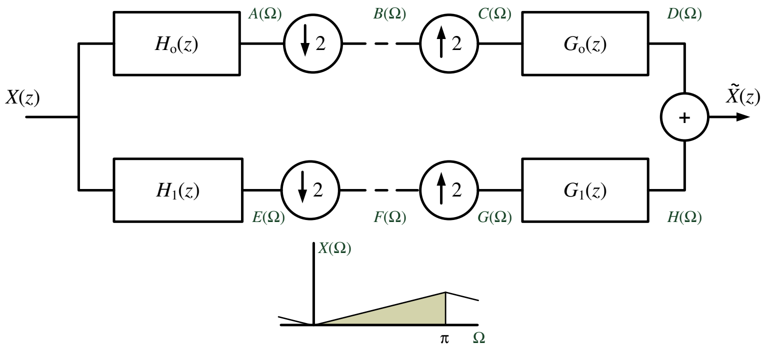
Q5.16 The complexity of the DCT can be reduced using "Fast" methods such as McGovern's Algorithm. Derive McGovern's algorithm for a four-point 1-D DCT. Compare its complexity (again assuming exploitation of separability) with that of the conventional approach for the case of a (4 × 4) point 2-D DCT.

Chapter 6: Filter banks and wavelet compression

Q6.1 For a filterbank downsampler, show that the frequency domain behavior of the output signal, $x_d[n]$, is related to that of the input $x[n]$ by:

$$X_d(\Omega) = 0.5\left[X\left(e^{j\Omega/2}\right) + X\left(e^{-j\Omega/2}\right)\right]$$

Q6.2 The figure below shows a simple two-band analysis–synthesis filter bank and a representative input spectrum.



It can be shown that the output of this system is given by:

$$\tilde{X}(z) = \frac{1}{2}X(z)\left[G_0(z)H_0(z) + G_1(z)H_1(z)\right] + \frac{1}{2}X(-z)\left[G_0(z)H_0(-z) + G_1(z)H_1(-z)\right]$$

where the upsampler and downsampler relationships are given by:

$$C\left(\mathrm{e}^{j\Omega}\right) = B\left(\mathrm{e}^{j2\Omega}\right)$$
$$B\left(\mathrm{e}^{j\Omega}\right) = 0.5\left[A\left(\mathrm{e}^{j\Omega/2}\right) + A\left(-\mathrm{e}^{j\Omega/2}\right)\right]$$

Using the upsampler and downsampler relationships given above, compute and sketch the spectra at points A to H. Assume that all the filters have ideal brickwall responses. Hence demonstrate graphically that the system is capable of perfect reconstruction.

Q6.3 Given the following subband filters:

$$H_0(z) = \frac{1}{\sqrt{2}}\left(1 + z^{-1}\right); \qquad H_1(z) = \frac{1}{\sqrt{2}}\left(1 - z^{-1}\right)$$
$$G_0(z) = \frac{1}{\sqrt{2}}\left(1 + z^{-1}\right); \qquad G_1(z) = \frac{1}{\sqrt{2}}\left(-1 + z^{-1}\right)$$

show that the corresponding two-band filter bank exhibits perfect reconstruction.

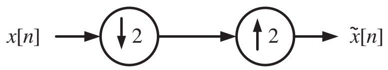
Q6.4 Demonstrate that the following filter relationships produce a filter bank that is alias-free and offers perfect reconstruction:

$$H_1(z) = zG_0(-z); \qquad G_1(z) = z^{-1}H_0(-z)$$
$$P(z) + P(-z) = 2z^{-2}; \qquad P(z) = H_0(z)G_0(z)$$

Assuming that $G_0(z) = H_0(z) = z^{-1}$, what is the output of this filter bank given an input sequence {1,1,0}?

Q6.5 Referring to the figure below, a signal, $x[n]$, is downsampled by 2 and then upsampled by 2. Show that, in the z-domain, the input–output relationship is given by:

$$\tilde{X}(z) = 0.5[X(z) + X(-z)]$$



Q6.6 Demonstrate that the following quadrature mirror filter relationships produce a filter bank that is alias-free and offers perfect reconstruction:

$$H_0(z) = z^{-2} + z^{-3}; \quad G_0(z) = H_1(-z);$$
$$H_1(z) = H_0(-z); \quad G_1(z) = -H_0(-z)$$

What are the limitations of the above filter bank? How, in practice, are more useful QMF filter banks designed?

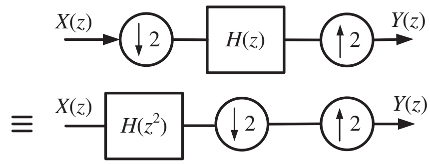
Q6.7 A two-channel single stage 1-D QMF filter bank is constructed using the low-pass prototype filter $H_0(z) = (1 + z^{-1})$. Derive the other filters needed for this system and compute the output signal for an input: $x[n] = \{2, 3, 6, 4\}$.

Q6.8 A 1-D wavelet filter bank comprises two stages of decomposition and uses the same filters as defined in Q6.7 (but factored by $\frac{1}{\sqrt{2}}$ to ensure exact reconstruction). The input sequence is $x[n] = \{1, 1, 1, 1, 1, 1, 1, 1\}$. Using boundary extension, and assuming critical sampling between analysis and synthesis banks, compute the signal values at all internal nodes in this filter bank, and hence demonstrate that the output sequence is identical to the input sequence.

Q6.9 Repeat Q6.8, but this time assume that the bit-allocation strategy employed preserves the lower two subbands but completely discards the high-pass subband. Assuming a wordlength of 4 bits, what is the PSNR of the reconstructed signal after decoding?

Q6.10 Draw the diagram for a 2-D, three-stage wavelet filter bank. Show how this decomposition tiles the 2-D spatial frequency plane and compute the frequency range of each subband. Assuming that the input is of dimensions 256×256 pixels, how many samples are contained in each subband?

Q6.11 Prove that the following two diagrams are equivalent.



Chapter 7: Lossless compression methods

Q7.1 Consider the following codewords for the given set of symbols:

Symbol	C1	C2	C3
a_1	0	0	01
a_2	10	01	10
a_3	110	110	11
a_4	1110	1000	001
a_5	1111	1111	010

Identify which are prefix codes.

Q7.2 Derive the set of Huffman codewords for the symbol set with the following probabilities:

$$P(s_0) = 0.06;\ \ P(s_1) = 0.23;\ \ P(s_2) = 0.30;\ \ P(s_3) = 0.15;$$
$$P(s_4) = 0.08;\ \ P(s_5) = 0.06;\ \ P(s_6) = 0.06;\ \ P(s_7) = 0.06;$$

What is the transmitted binary sequence corresponding to the symbol pattern $s_0,\ s_4,\ s_6$? What symbol sequence corresponds to the code sequence: 000101110? Calculate the first order entropy of this symbol set, the average codeword length and the coding redundancy.

Q7.3 A quantized image is to be encoded using the symbols $I \in \{I_0 \cdots I_{11}\}$. From simulation studies it has been estimated that the relative frequencies for these symbols are as follows: I_0: 0.2; $I_1 \cdots I_3$: 0.1; $I_4 \cdots I_7$: 0.05; $I_8 \cdots I_{11}$: 0.075. Construct the Huffman tree for these symbols and list the resultant codewords in each case.

Q7.4 If the minimum variance set of Huffman codes for an alphabet, *A,* is as shown in the table below, determine the efficiency of the corresponding Huffman encoder.

Symbol	Probability	Huffman Code
a_1	0.5	0
a_2	0.25	10
a_3	0.125	110
a_4	0.0625	1110
a_5	0.0625	1111

Q7.5 Consider the following 4×4 matrix, $\mathbf{C}$, of DCT coefficients, produced from a block-based image coder. Using zig-zag scanning and run length coding (assume a $\{run,\ value\}$ model where *value* is the integer value of a non-zero coefficient and *run* is the number of zeros preceding it), determine the transmitted bitstream after entropy coding.

$$\mathbf{C} = \begin{bmatrix} 2 & 0 & 3 & 0 \\ 0 & 3 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Use the following symbol to codeword mappings:

Symbol	**Code**	**Symbol**	**Code**
{0,1}	00	{1,3}	0111
{0,2}	10	{2,1}	0110
{0,3}	110	{2,2}	0101
{1,1}	1111	{2,3}	01001
{1,2}	1110	EOB	01000

Q7.6 Assuming that the DCT coefficient matrix and quantization matrix in Q5.14 form part of a JPEG baseline codec, derive the Huffman coded sequence that would be produced by the codec for the AC coefficients.

Q7.7 What is the Exp-Golomb code for the symbol index 132_{10}?

Q7.8 Show how the Exp-Golomb codeword 000010101 would be decoded and compute the value of the corresponding symbol index. What is the corresponding Golomb–Rice code for this index (assume $m = 4$)?

Q7.9 Given the symbols from an alphabet, $A = \{a_1, a_2, a_3, a_4, a_5\}$, and their associated probabilities of occurrence in the table below, determine the shortest arithmetic code which represents the sequence: $\{a_1, a_1, a_2\}$.

Symbol	**Probability**
a_1	0.5
a_2	0.25
a_3	0.125
a_4	0.0625
a_5	0.0625

Q7.10 Given the following symbols and their associated probabilities of occurrence, determine the binary arithmetic code which corresponds to the sequence: $\{a_1, a_2, a_3, a_4\}$. Demonstrate how an arithmetic decoder, matched to the encoder, would decode the bitstream: 010110111.

Symbol	**Probability**
a_1	0.5
a_2	0.25
a_3	0.125
a_4 (EOB)	0.125

Q7.11 Given the following symbols and their associated probabilities of occurrence, determine the arithmetic code which corresponds to the sequence: a_1, a_2, a_3 (where a_3 represents the EOB symbol).

Symbol	Probability
a_1	0.375
a_2	0.375
a_3	0.125
a_4	0.125

Show how the bitstream produced above would be decoded to produce the original input symbols.

Repeat this question using Huffman encoding rather than arithmetic coding. Compare your results in terms of coding efficiency.

Q7.12 Derive the arithmetic codeword for the sequence $s_1, s_1, s_2, s_2, s_5, s_4, s_6$, given a symbol set with the following probabilities:

$$P(s_0) = 0.065; \quad P(s_1) = 0.20; \quad P(s_2) = 0.10; \quad P(s_3) = 0.05;$$
$$P(s_4) = 0.30; \quad P(s_5) = 0.20; \quad P(s_6) = 0.10 = \text{EOB}$$

Q7.13 Consider an alphabet, $A = \{a_1, a_2, a_3, a_4\}$, where $P(a_1) = 0.6$; $P(a_2) = 0.2$; $P(a_3) = 0.1$; $P(a_4) = 0.1$. Compute the sequence, S, of three symbols which corresponds to the arithmetic code 0.58310.

Q7.14 Given an alphabet of two symbols $A = \{a, b\}$, where $P(a) = 0.25$; $P(b) = 0.75$, draw a diagram showing coding and probability intervals for the associated arithmetic coder. Derive a binary arithmetic code for the sequence $S = baa$.

Chapter 8: Coding moving pictures: motion prediction

Q8.1 Implement the full search BBME algorithm on the 6×6 search window, **S**, using the current-frame template, **M**, given below.

$$\mathbf{S} = \begin{bmatrix} 1 & 5 & 4 & 9 & 6 & 1 \\ 6 & 1 & 3 & 8 & 5 & 1 \\ 5 & 7 & 1 & 3 & 4 & 1 \\ 2 & 4 & 1 & 7 & 6 & 1 \\ 2 & 4 & 1 & 7 & 8 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}; \quad \mathbf{M} = \begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix}$$

Q8.2 Given the following reference window and current block, show how an N-step search algorithm would locate the best match (assume an SAD optimization criterion). Determine the motion vector for this block. Does this produce the same result as an exhaustive search?

$$\begin{bmatrix} 1 & 4 & 3 & 2 & 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & 2 & 0 & 2 & 3 & 0 & 2 & 1 \\ 0 & 2 & 0 & 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 & 4 & 4 & 0 & 0 & 1 \\ 0 & 3 & 2 & 3 & 0 & 1 & 2 & 2 & 3 \\ 0 & 3 & 3 & 1 & 0 & 1 & 1 & 2 & 2 \\ 0 & 4 & 2 & 3 & 0 & 2 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 2 & 3 & 2 & 2 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 0 \\ 2 & 2 & 0 \end{bmatrix}$$

Q8.3 Use bidirectional exhaustive search motion estimation, such as that employed in MPEG-2, to produce the best match for the following current block and two reference frames:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 2 & 3 \\ 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 3 & 4 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

Q8.4 Explain how the Two-Dimensional Logarithmic (TDL) search method improves the search speed of block-based motion estimation. Illustrate this method for the case of a $\pm 6 \times \pm 6$ search window where the resultant motion vector is [2,5]. Quantify the savings for this particular example over a full search. What is the main advantage and disadvantage of this method?

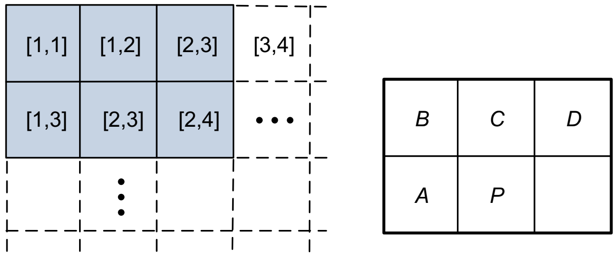
Q8.5 Implement a hexagonal search block matching algorithm on the search window, **S**, using the current-frame template, **M**, as given below. Determine the motion vector for this search. Assume that any search points where the hexagon goes outside of the reference frame are invalid.

$$\mathbf{S} = \begin{bmatrix} 0 & 1 & 2 & 3 & 22 & 19 & 18 & 23 \\ 7 & 3 & 6 & 5 & 33 & 31 & 13 & 22 \\ 4 & 6 & 3 & 7 & 23 & 23 & 15 & 26 \\ 8 & 4 & 1 & 3 & 11 & 22 & 29 & 19 \\ 2 & 8 & 9 & 7 & 8 & 14 & 16 & 18 \\ 5 & 0 & 7 & 3 & 7 & 15 & 12 & 13 \\ 7 & 4 & 6 & 6 & 9 & 8 & 8 & 12 \\ 1 & 2 & 3 & 9 & 10 & 9 & 8 & 12 \end{bmatrix}; \quad \mathbf{M} = \begin{bmatrix} 1 & 4 \\ 10 & 7 \end{bmatrix}$$

Q8.6 Given the following current block P and its set of three adjacent neighbors A, B, C and D, with motion vectors as indicated, use motion vector prediction to initialize the search for the best motion vector for this block.

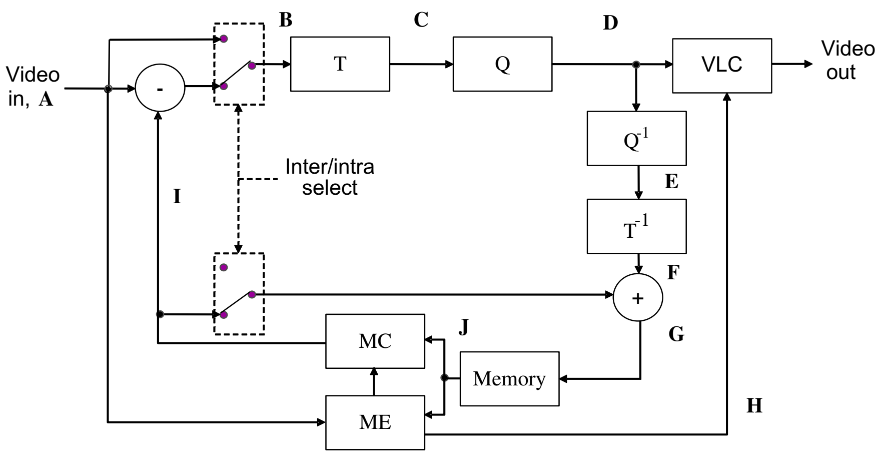
$\mathbf{d}_B$=[1,1]	$\mathbf{d}_C$=[1,0]	$\mathbf{d}_D$=[1,0]	
$\mathbf{d}_A$=[1,2]	P		

Q8.7 Motion vectors for six adjacent blocks in two partial rows of a frame are shown below (shaded). Using the prediction scheme $\hat{\mathbf{d}}_P = \text{med}(\mathbf{d}_A, \mathbf{d}_C, \mathbf{d}_D)$, compute the predicted motion vectors for each of these blocks together with their coding residuals. State any assumptions that you make in computing predictions for blocks located at picture boundaries (see Chapter 12).



Chapter 9: The block-based hybrid video codec

Q9.1 The figure below shows a block diagram of a basic hybrid (block transform, motion compensated) video coding system.



Using the following assumptions and data at time n:

Image frame size: 12×12 pixels
Macroblock size: 4×4 pixels
Operating mode: inter-frame
Motion estimation: Linear translational model (4×4 blocks)

Transform matrix: $$\mathbf{T} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

Quantization matrix: $$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 4 & 4 \\ 1 & 2 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix}$$

Current input block: $$\mathbf{A} = \begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$; assume exact center of frame.

Reference memory: $$\mathbf{J} = \begin{bmatrix} 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 10 & 4 & 7 & 4 & 7 & 4 & 4 & 4 & 3 & 2 & 1 \\ 8 & 4 & 7 & 6 & 7 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 6 & 4 & 6 & 5 & 4 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 10 & 6 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 14 & 4 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 6 & 6 & 6 & 6 & 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 & 0 & 0 & 3 & 4 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 & 5 & 8 & 0 \\ 7 & 7 & 8 & 9 & 0 & 0 & 0 & 0 & 6 & 7 & 0 & 0 \\ 0 & 0 & 0 & 10 & 3 & 0 & 0 & 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Compute:

1. The current motion vector, **H**, at time n.
2. The motion-compensated output, **I**, at time n.
3. The DFD for the current input frame, **B**.
4. The transformed and quantized DFD output, **D**.

Q9.2 If the encoded bitstream from Q9.1 is decoded by a compliant decoder, compute the decoder output at time n that corresponds to the input block, **A**.

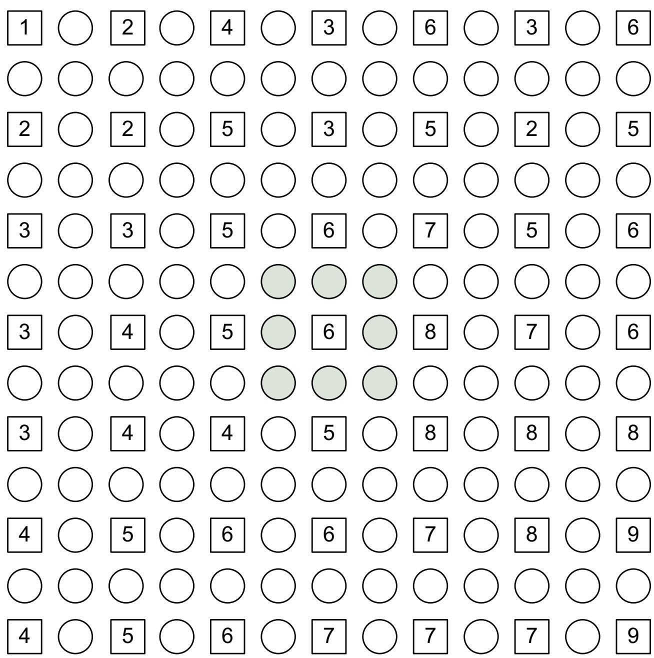
Q9.3 Compute the transformed and quantized output for the current input block given in Q9.1, but this time for the case of intra-frame coding.

Q9.4 Using horizontal, vertical, and DC modes only, produce the H.264/AVC intra-prediction for the highlighted 4 × 4 luminance block below:

1	2	3	4	2	3	5	4	7	8	9	5
2	4	4	6	7	8	4	6	4	4	3	2
4	3	5	6	3	3	3	4	4	4	3	3
4	4	5	6	3	3	3	4	4	4	3	3
1	1	4	6	**2**	**3**	**6**	**3**	×	×	×	×
4	3	3	5	**3**	**2**	**5**	**3**	×	×	×	×
2	2	3	5	**3**	**3**	**4**	**4**	×	×	×	×
1	3	5	3	**1**	**2**	**4**	**4**	×	×	×	×

Q9.5 Following the full search result from Q8.1, refine the motion vector for this search to half-pixel accuracy using the two-tap interpolation filter $(s_i + s_j)/2$ (where s_i and s_j are the horizontal or vertical whole-pixel locations adjacent to the selected half-pixel location).

Q9.6 Use the H.264/AVC subpixel interpolation filter to generate the half-pixel values for the shaded locations in the following search window:



Q9.7 H.264 employs a 4 × 4 integer transform instead of the 8 × 8 DCT used in previous coding standards. Prove that the four-point integer approximation to the 1-D DCT transform matrix, **A**, is given by:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \otimes \mathbf{E}_f$$

where $\mathbf{E}_f$ is a 4×4 scaling matrix. State any assumptions made during your derivation.

Chapter 10: Measuring and managing picture quality

Q10.1 List the primary factors that should be controlled and recorded during subjective video assessment trials.

Q10.2 List the primary attributes that a good subjective database should possess.

Q10.3 Given the following three consecutive frames, compute the temporal activity (TI) for this sequence.

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}; \quad \mathbf{S}_2 = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 3 & 4 & 6 \\ 3 & 4 & 5 & 7 \\ 4 & 5 & 6 & 8 \end{bmatrix}; \quad \mathbf{S}_3 = \begin{bmatrix} 1 & 2 & 4 & 6 \\ 2 & 3 & 5 & 7 \\ 3 & 4 & 6 & 8 \\ 4 & 5 & 7 & 9 \end{bmatrix}$$

Q10.4 Consider the operation of a constant bit rate video transmission system at 25 fps with the following parameters:

$$R_o = 500 \text{ kbps}; \ B = 150 \text{ kb}; \ F_i = 100 \text{ kb}; \ \text{GOP} = 6 \text{ frames}$$

If the pictures transmitted have the following sizes, compute the occupancy of the decoder buffer over time and determine whether underflow or overflow occurs.

Frame no.	Picture size (kbits)
1	20
2–6	10
7	40
8–12	10

Chapter 11: Communicating pictures: delivery across networks

Q11.1 Using the Huffman codes in the following table, decode the following encoded bitstream: 0 0 1 0 1 0 0 1 1 1 1 1 1 1 0 1 0 0.

Symbol	Probability	Huffman Code
a_1	0.5	0
a_2	0.25	10
a_3	0.125	110
a_4	0.0625	1110
a_5	0.0625	1111

Assuming codeword 110 represents an EOB signal, what would be the effect of a single error in bit position 2 of the above sequence?

Q11.2 An image encoder uses VLC and Huffman coding of transform coefficients, based on a set of four symbols with the following mappings:

$$A \leftrightarrow 0; \quad B \leftrightarrow 10; \quad C \leftrightarrow 110; \quad D \leftrightarrow 111$$

Assuming that the sequence transmitted is *ABCDAC*, determine the received symbol sequences for the following scenarios:

1. An error occurring in the 3rd bit position.
2. An error occurring in the 1st bit position.
3. An error occurring in the 4th bit position.
4. An error occurring in the 8th bit position.

For each scenario, assuming that *B* represents the EOB symbol, comment on the impact that the error has on the final state of the decoder and on the reconstructed transform coefficients.

Q11.3 VLC codes for symbols $\{a, b, c\}$ are: $a = 0; b = 11; c = 101$, where: $P(a) = 0.5$; $P(b) = 0.25$; $P(c) = 0.25$. Comment on any specific property these codewords exhibit and on its benefits in a lossy transmission environment. Compare the efficiency of these codewords with that of a conventional set of Huffman codes for the same alphabet.

Q11.4 The sequence of VLC codewords produced in Q7.5 is sent over a channel which introduces a single error in the 15th bit transmitted. Comment on the impact that this error has on the reconstructed transform coefficients and any subsequent blocks of data.

Q11.5 Use EREC to code four blocks of data with lengths: $b_1 = 6$, $b_2 = 5$; $b_3 = 2$; $b_4 = 3$. Assuming that an error occurs in the middle of block 3 and that this causes the EOB code to be missed for this block, state which blocks in the frame are corrupted and to what extent.

Q11.6 A video coder uses VLC and Huffman coding based on a set of four symbols with probabilities 0.4, 0.3, 0.15, and 0.15. Calculate the average bit length of the resulting Huffman codewords. If this video coder is to employ reversible codes (RVLC) to improve error resilience, suggest appropriate codewords and calculate the bit rate overhead compared to conventional Huffman coding.

Q11.7 Assume that $\mathbf{S_1}$ and $\mathbf{S_2}$ below represent corresponding regions in two temporally adjacent video frames. Due to transmission errors, the central 4×4 block (pixels marked as "$\times$") in the received current frame $\mathbf{S}_2'$ is lost. Assuming that the codec operates on the basis of 2×2 macroblocks:

(a) Perform temporal error concealment using frame copying to provide an estimate of the lost block.

(b) Perform motion-compensated temporal error concealment, based on the BME measure. Assume that the candidate motion vectors obtained from the four adjacent blocks are either [0,1], [1,0] or [1,1].

$$\mathbf{S_1} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 & 5 & 5 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 6 & 6 & 7 & 6 & 6 \\ 6 & 6 & 6 & 7 & 6 & 6 & 6 & 6 \\ 7 & 7 & 7 & 7 & 8 & 8 & 8 & 8 \end{bmatrix}; \quad \mathbf{S}_2' = \begin{bmatrix} 2 & 2 & 3 & 3 & 3 & 4 & 4 & 3 \\ 3 & 3 & 4 & 4 & 4 & 4 & 5 & 5 \\ 3 & 3 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 5 & \times & \times & 4 & 4 & 6 \\ 5 & 5 & 6 & \times & \times & 6 & 6 & 6 \\ 7 & 6 & 7 & 6 & 6 & 6 & 6 & 7 \\ 6 & 7 & 7 & 6 & 6 & 8 & 8 & 8 \\ 9 & 9 & 9 & 7 & 8 & 7 & 8 & 8 \end{bmatrix}$$

Q11.8 Assuming again that the central 2×2 block (pixels marked as "$\times$") in the received current frame $\mathbf{S}_2'$ (above) is lost, calculate the missing elements using spatial error concealment.

Q11.9 A block-based image coder generates a slice comprising five DCT blocks of data as follows:

Block Number	Block Data
1	01011011101111
2	10111001011101111
3	1100101111
4	101101101111
5	0001111

where the symbols and entropy codes used to generate this data are as follows (assume E is the End of Block Symbol):

Symbol	Huffman Code
A	0
B	10
C	110
D	1110
E	1111

Code this data using Error-Resilient Entropy Coding (EREC). Choose an appropriate EREC frame size and offset sequence for this slice and show all the encoding stages of the algorithm.

Q11.10 Perform EREC decoding on the bitstream generated from Q11.9.
If the slice data given in the above question is corrupted during transmission such that the values of the last 2 bits in each of block 2 and block 3 are complemented, how many blocks are corrupted after EREC decoding. Compute how many blocks would contain errors if the slice were transmitted as a conventional stream of entropy codewords without EREC.

Chapter 12: Video coding standards

Q12.1 Assuming a ClF format luminance-only H.261 sequence with synchronization codewords at the end of each GOB, calculate the (likely) percentage of corrupted blocks in a frame if bit errors occur in block 1 of macroblock 7 in GOB 2 and block 1 of macroblock 30 of GOB 10.

Q12.2 What parts of a video codec are normally subject to standardization in term of compliance testing? Why is this?

Q12.3 Discuss the concepts of profiles and levels in MPEG-2 indicating how these have enabled a family of compatible algorithms covering a range of applications and bit rates to be defined.

Q12.4 Compare the features of H.264/AVC and MPEG-2 and highlight those features that have enabled H.264 to approximately halve the bit rate of MPEG-2 for equivalent picture quality.

Q12.5 Compare the features of H.265/HEVC and H.264/AVC and highlight those features that have enabled HEVC to approximately halve the bit rate of H.264 for equivalent picture quality.

Q12.6 Why has H.265/HEVC increased its maximum block size to 64×64 pixels?

Chapter 13: Communicating pictures—the future

Q13.1 What are the primary challenges and demands for video compression in the future?

Q13.2 Consider the case of an 8 K resolution video signal in 4:2:2 format with 14 bits dynamic range and a frame rate of 300 fps. Calculate the total bit rate needed for this video.

Q13.3 How might video compression algorithms develop in the future in order to cope with the demands of formats such as that described in Q13.2?