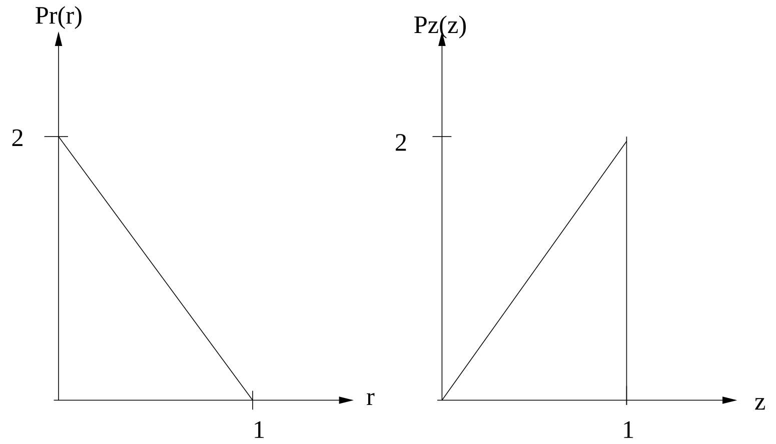
IMAGE PROCESSING (RRY025)

Problems

A IMAGE ENHANCEMENT PROBLEMS

1) An image has the gray level Probability Distribution Function (PDF - or gray level histogram normalised by number of pixels) of $P_r(r)$ shown below left.



a) Find the pixel transformation $y = g(r)$ such that after transformation the image has a **flat** PDF, i.e. which accomplishes histogram equalisation. Assume continuous variables r,y.

b) It is desired to find a transformation $z = f(r)$ such that the transformed image will have the PDF of $P_z(z)$ shown above right. Assume continuous quantities and determine the transformation function $z = f(r)$.

2) a) Two images have the same histogram. Which of the following properties *must* they have in common? (i) Same total power (sum of squares of pixel values) (ii) Same Entropy (sum of I ln I over all pixel values) (iii) Same degree of pixel to pixel correlation?

b) Do histogram equalisation on the following image which has 8 discrete pixel levels (0 - 7), transforming it into a histogram equalised image also with 8 discrete grey levels in the range (0-7).

1	1	1	1	1	1	1	1
0	2	5	5	5	5	2	0
0	3	2	6	7	2	3	0
0	3	3	2	2	3	3	0
0	2	3	2	2	3	3	0
0	3	2	4	4	2	4	0
0	2	6	4	4	4	2	0
1	1	1	1	1	1	1	1

3) a) Match the images (1-6) shown in the figure below with their pixel histograms (A-F).

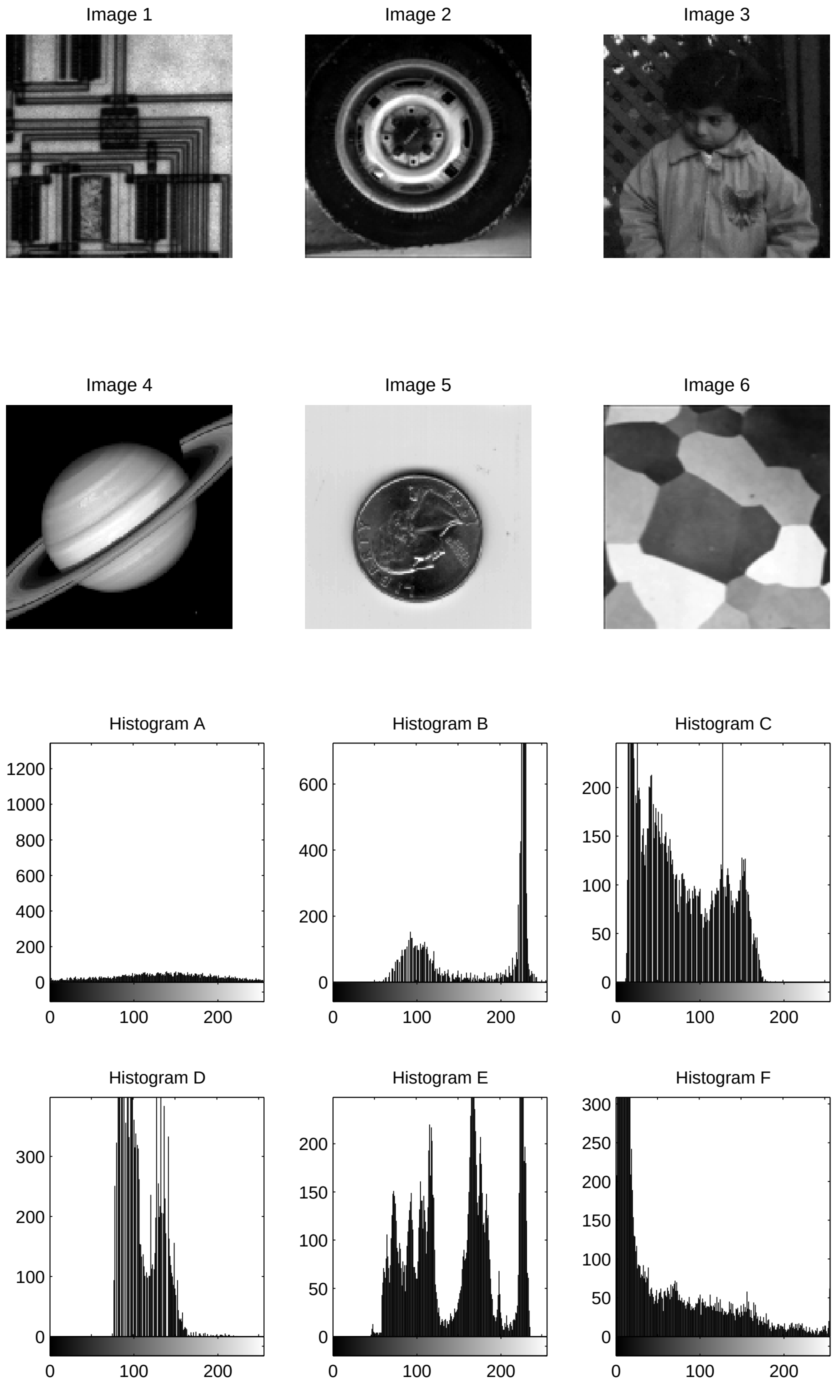


Fig 1. Images and their Histograms **Note** that histogram A has a large spike at x=0.

b) Histogram equalisation is applied to the coin image below (A) and the result is the image (B). Why does the image look so bad, with few details visible on the coin?

c) A different algorithm is applied to (A) and produces the image (C) which enhances both the background and the coin. Describe *in detail* how would write such an algorithm. (Hint: Use the image histogram to select parts of the image on and off the coin).

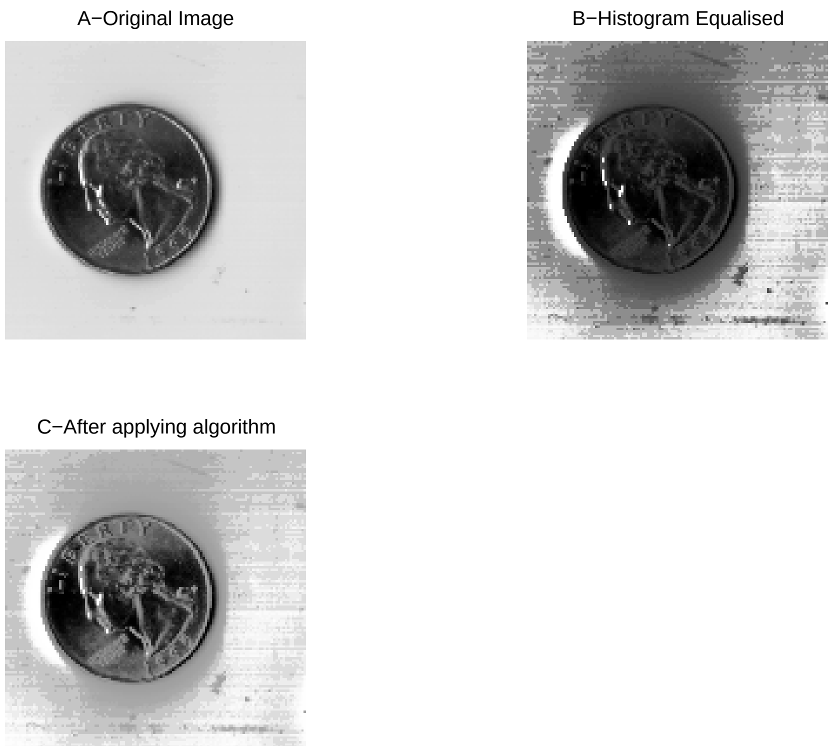


Fig 2. Image enhancement of coin image.

4) Consider the Image 'I' below and the filters 'F' and 'L'

'I'

1	1	1
1	8	1
1	1	1

'F'

	1/8	
1/8	1/2	1/8
	1/8	

'L'

	1	
1	-4	1
	1	

a) Correlate the image 'I' with the filter 'F' above and compute the output image (assume 'same' correlation with zeros outside the input image and *round down* the output pixel values to the nearest integer)

b) Apply a 3 by 3 median filter to the same image 'I' to produce an 3 by 3 output image, again assume zeros outside of the image.

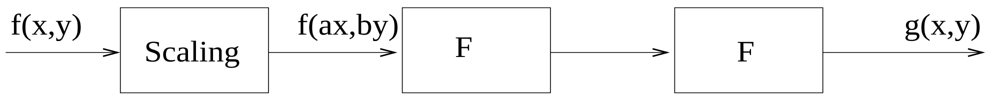
c) The result in part b demonstrates one general advantage and one general disadvantage of the median filter as compared to the mean filter.

d) Show that correlating the Laplace operator 'L' with the image 'I' is equivalent (except for a proportional factor) to locally subtracting a five point local mean from each original value of the image.

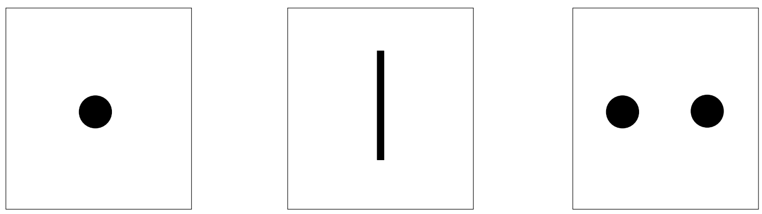
B) FOURIER TRANSFORM PROBLEMS

1) For an image given by the function $f(x, y) = (x + y)^3$ where x,y are continuous variables; evaluate $f(x, y)\delta(x-1, y-2)$ and $f(x, y) * \delta(x-1, y-2)$, where δ is the Dirac Delta function.

2) For the optical imaging system shown below, consisting of an image scaling and two forward Fourier transforms show that the output image is a scaled and inverted replica of the original image f(x,y).



3) Three binary images (with value 1 on black areas and value 0 elsewhere) are shown below. Sketch the continuous 2D FT of these images (don't do this mathematically, try to use instead the convolution theorem and knowledge of FTs of common functions).



4) A real 1D function f(x) can be decomposed as the sum of an even (symmetric through the origin) function and an odd function (antisymmetric through origin).

a) Show that $f_{even} = 0.5[f(x) + f(-x)]$ and $f_{odd} = 0.5[f(x) + f(-x)]$.

b) Show that $FT(f_{even}) = Real(FT(f(x)))$ and $FT(f_{odd}) = iIm(FT(f(x)))$.

5) Consider the images shown below (A to H). Using knowledge of symmetry properties and the convolution theorem (*not exact calculation of the Fourier integrals*), list which images or image which have Fourier transforms, F(u,v) with the following properties.

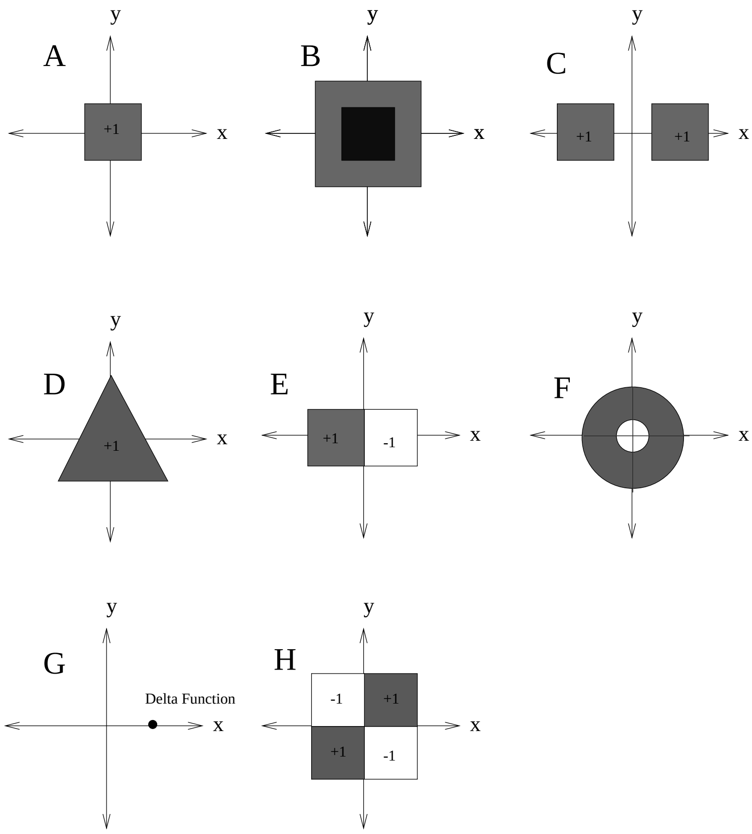
i) The imaginary part of F(u,v) is zero at all u,v

ii) F(u,v) is purely real and positive for all u,v.

iii) F(0,0) = 0

iv) F(u,v) has circular symmetry

v) The real part of F(u,v) is zero at all u,v



Notes Unless otherwise stated the shaded regions have value +1 and outside the shapes the value is 0. Image **B** - is the convolution of A with itself. Images **E** and **H** have regions of negative intensity. Image **G** is a delta function of unit area at $x = x_{xo}$.

6) If a continuous image is Nyquist sampled the discrete version contains all information about the original image, such that the original continuous image can be reconstructed perfectly by low pass imaging. On the other hand if the sampling is not sufficiently dense aliasing occurs and when we attempt to reconstruct we can obtain a completely different continuous image. The following problem illustrated these points.

Consider an input image given by the two dimensional function

$f(x, y) = 2cos[2\pi(4x + 6y)]$

which is sampled on an infinite grid of pixels with separation $\Delta x = 0.1$ and $\Delta y = 0.2$, in the x and y directions respectively. Note that this is *less dense* than required by Nyquist sampling. Consider that we anyway try to reconstruct the original continuous image from the above sampled version. To do this we form the continuous FT of the sampled image; multiply by an ideal low pass spatial filter with cutoff frequencies at half the sampling frequency in each spatial frequency axis and then inverse Fourier transform. Show in your answer what is obtained at each step in the above process, specifically;

a) Draw with labelled axes the Fourier transform of the sampled image (Hint 1; make use of the 2D sampling function ('bed of nails') and the convolution theorem. Hint 2; the FT of $exp(-2\pi i(Ax + By))$ is $\delta(u - A, v - B)$, i.e. a Dirac delta function in the Fourier domain centred on the position $u = A$ and $v = B$.

b) Give the Fourier transform after it has been low-pass filtered.

c) Show that the reconstructed continuous image is given by the mathematical function $2cos[2\pi(4x + y)]$.

d) In order to reconstruct without distortion the original image from sampled data what is the maximum sizes of Δx and Δy that can be used and the size of the low pass filter?

7) a) Determine the *linear discrete convolution* of x(m,n) below with each of the sampled images A, B. For each cases take the origin as the element in the top left corner. Assume that outside each image pixel values are zero and compute the 'full' linear convolution (do the convolutions with pen and paper but check with MATLAB using command 'conv2').

```
 x(m,n)      A(m,n)          B(m,n)

1  4  1       -2             1  2  3

2  5  3        3             1  0  0

               1                                        -
```

b) Form also the cross-correlation of x(m,n) and A(m,n), check your result using the MATLAB command 'filter2'.

c) Form also the *circular convolution* of x(m,n) and B(m,n). Do this with pen and paper but check your result using MATLAB by doing the DFT of the two input functions (with the fft2) commands, multiplying the results element-wise (using .* in MATLAB) and inverse DFT back (using ifft2). The output results should be completely real but in MATLAB you may have to remove small imaginary values from computer rounding errors using the 'real' command.

d) If we wished to *linearly* convolve x and A by means of the DFT to what minimum size must we zero-pad the images? Implement this in MATLAB (type 'help fft2' to find out how to zero pad).

8) Below are shown a series of images (1-6) and the log amplitudes of their Discrete Fourier transforms (A-F). Match the correct transform to each image. Give a one sentence reason for each match. Note the DFT of image 1 is hard to match so concentrate on the other images first.

Image 1

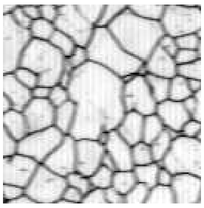


Image 2

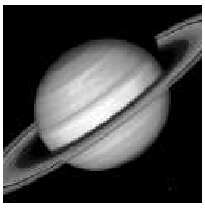


Image 3

Image 4

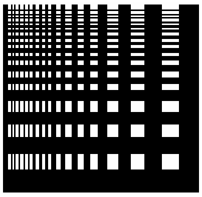


Image 5

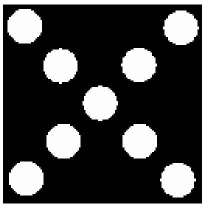
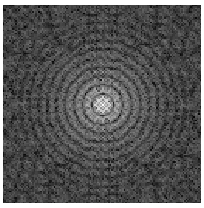
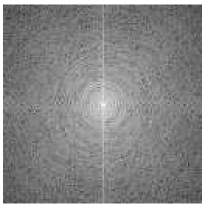


Image 6

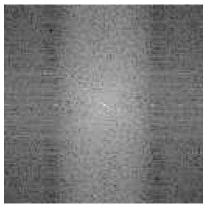
FT Log Amp A



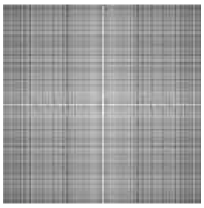
FT Log Amp B



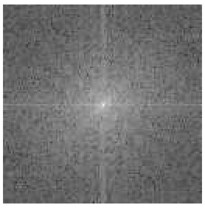
FT Log Amp C



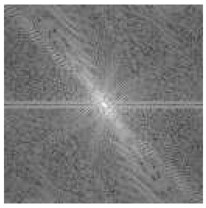
FT Log Amp D



FT Log Amp E



FT Log Amp F



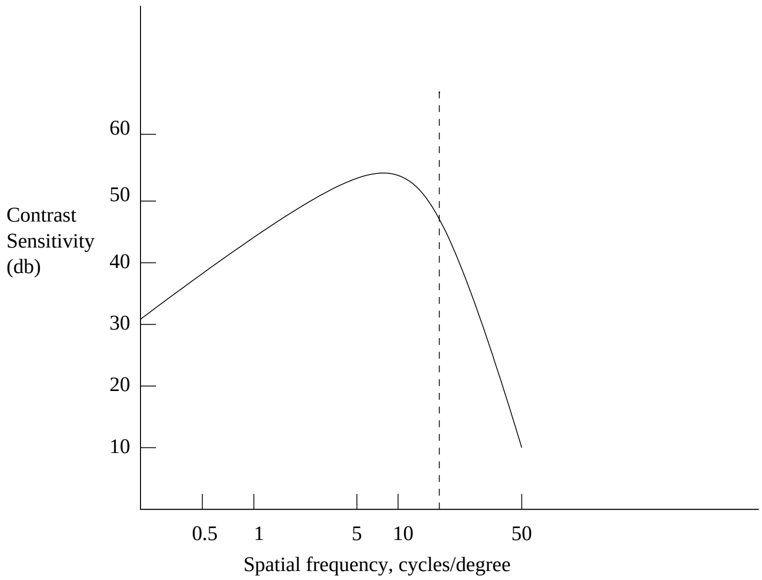
C) HUMAN VISION

1a) Make a sketch of the Modulation Transform Function of the human vision system; that is a plot of the contrast sensitivity versus spatial frequency (spatial frequency filter shape).

b) The figure below illustrates the Mach band effect. Describe what the Mach band effect is and how it is explained in terms of the human MTF.

c) A magazine wishes to publish the above figure so that each gray band appears to the reader to have uniform brightness. Describe how you would process the raw image above to achieve this effect.

2) a) The figure on the next page shows how 'Lateral Inhibition' and other properties of the eye effects the spatial frequencies that are passed onto the brain. Beyond 20 cycles/degree there is a sharp cutoff, so that the information passed to the brain is effectively band-limited. This property is used by some image compression algorithms to allow significant compression without loss of apparent image quality.



Consider a 1024x1024 pixel image on a computer screen with very high resolution of 100 pixels/cm. Assume that a viewer sits 30cm away from the screen. This image has been sent in a compressed form by not transmitting the high spatial frequencies not perceived by the brain. Show that an image can be compressed by this means by a factor of approximately 2 before there is a perceived reduction in quality.

b) In very bright light the spatial resolution of the eye is slightly less and and the amount of compression possible slightly higher, why is this?

3) Draw a sketch of the CIE Chromaticity diagram with chromacity coordinates are $x = X/(X+Y+Z)$ and $y = Y/(X+Y+Z)$. Indicate the following regions;

i) The curved line occupied by the colour perceived when single wavelength light from a prism is shone into the eye.

ii) The 'Line of purples', where purple colours lie.

iii) The place where White lies.

iv) The line of colours from a blackbody radiating from 3000K to 9000K.

v) Where does the colour Brown lie?

D) COSINE AND WAVELET TRANSFORMS

1) a) An image is to be compressed by a transform method and as a first step the image is Cosine transformed. Why does the Cosine transform concentrate more power into fewer components than the DFT?

b) The 1-D Cosine transform of a length N sequence f(x) is *defined* to be;

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) cos[(2x+1)u\pi/2N]$$

where we choose $\alpha(u) = \sqrt{1/N}$ for $u = 0$ and $\alpha(u) = \sqrt{2/N}$ for $u \neq 0$ to give a symmetrical form to the inverse cosine transform..

You wish to do a 1-D Cosine transform in a software package which only has the capability to do Discrete Fourier Transforms, carefully describe each of the steps needed to implement it.

2) Carry out a 2D Haar wavelet transform of the 4x4 letter 'O' image given above, using the functions below for forming Approximation, Horizontal details, Vertical details and Diagonal details. Give the 4x4 image which is the level 1 wavelet transform.

```
1  1  1  1

1  0  0  1

1  0  0  1

1  1  1  1
```

```
Aprox                V-Details       H-Details    D-Details

0.25 0.25            0.25 -0.25      0.25  0.25    0.25 -0.25

0.25 0.25            0.25 -0.25     -0.25 -0.25   -0.25  0.25
```

d) Show that the level 2 wavelet transform equals the image given below.

```
'Haar Level 2 Wavelet Transform'


      3   0   1  -1

      0   0   1  -1

      1   1  -1   1

     -1  -1   1  -1
```

E) IMAGE COMPRESSION PROBLEMS

1) A binary image is to be coded in blocks of M pixels. The successive pixels are independent from each other and 5% of the pixels are 1. What are the Huffman codes for M=1,2,3 ? Also calculate the compression efficiency of each codebook compared to the theoretical maximum given by the Shannon noiseless coding theorem.

2) A set of images is to be compressed by a lossless method. Each pixel of each image has a value in the range (0-3) i.e. 2 bits/pixel. An image from this set is given below; in this image the occurrence of pixels of different values is typical of the set of images as a whole.

3	3	3	2
2	3	3	3
3	2	2	2
2	1	1	0

What is the degree of compression achievable using the following methods.

a) Huffman coding of the pixel values.

b) Forming differences between adjacent pixels (assuming horizontal raster scan) and then Huffman coding these differences.

c) Run length coding, assuming 2 bits to represent the pixel value and 2 bits to represent the run length.

d) In general for case a) how many bits does it take to transmit the Huffman codebook from the coder to the decoder? (assume only that both coder and decoder know that the image has 2 bits/pixel).
Give therefore a estimate of the minimum number of images that must share the same Huffman codebook in order to ensure compression factors greater than 1. (for this calculation assume the same pixel statistics as in part a).

3) a) An image compression algorithm such as JPEG uses uses 'block transforms', what are these? Show that if each block is transformed by a 'fast' algorithm such as the Fast Fourier Transform or Fast Cosine Transform that a 8x8 block transform of a 512x512 pixel image is almost exactly 3 times faster than transforming the whole image directly.

b) As the block size is decreased the transform becomes faster, what therefore sets the limit for the smallest useful block size for image compression? If the correlation coefficient between adjacent pixels is $\rho = 0.9$, why is the choice of 8 for the block size a good one?

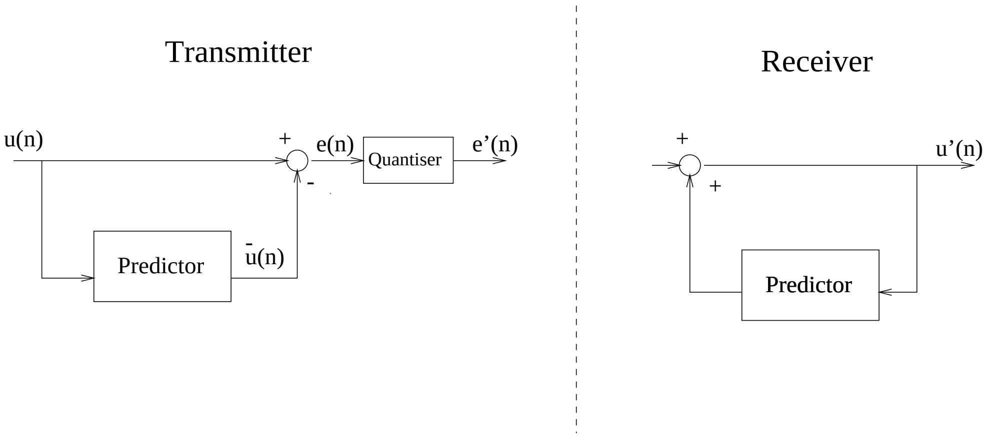
c) Why in general does the Discrete Cosine Transform for each block produce a slightly higher degree of compression than using the Discrete Fourier Transform?

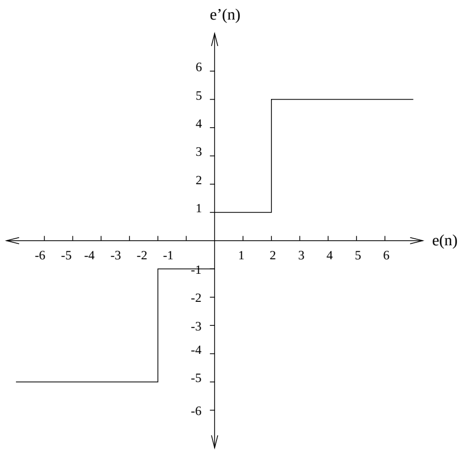
4 a) Describe the difference between lossy and non-lossy image compression. Mention one application of each type of compression

b) In Pulse Code Modulation (PCM) each pixel value of an image with a value between 0 and 127 is transmitted as a fixed length 7 bit binary number. Such an image is compressed by instead transmitting the difference between adjacent pixels on a line, the total range of differences that must be transmitted is between -127 and +127, yet it is possible to transmit a particular image *without loss of information* using an average of only 4.3 bits/pixel. How is this achieved? Why in general can the sequence of pixel differences in an image be compressed more than the raw pixel values?

c) To achieve even higher compression (i.e 2 bit/pixel) a lossy compression scheme is used in which differences between pixels are re-quantised before being transmitted using the schematic and quantiser illustrated below and the predictor rule $\overline{u(n)} = u(n-1)$.

For the following input sequence $u(n) = 101, 110, 107, 108, 105, 102$ what is the output sequence $u'(n)$? Assume the first pixel value is transmitted directly without error so that $u'(1) = u(1)$). Tabulate the values of $u(n), \overline{u(n)}, e(n), e'(n)$ and $u'(n)$, for each n.





d) A DPCM system using the same quantiser and predictor rule can significantly reduce the errors in transmission while maintaining the same transmission rate of 2bit/pixel. What would the output sequence be using such a system for the input sequence given in c)?

5) Image Compression algorithms work by removing redundancy in images. Three types of redundancy exist, coding, inter-pixel and psycho-visual. Furthermore some types of compression are lossless (fully reversible) and others not. Copy the table below and indicate which compression algorithms make use of each type of redundancy and which are lossless.

	Coding Redundancy	Interpixel Redundancy	Psycho-visual Redundancy	Lossless
Huffman coding of pixel Values				
Simple Run Length Coding of a binary image				
Delta Modulation				
Standard JPEG				

6) A FAX machine transmits black and white binary images with each pixel being either a 0 (white) or a 1 (black). On average 90% of of the pixels are expected to be white and there is no correlation between pixels. To speed transmission image compression is used.

a) Based on the Shannon noiseless coding theorem what is the theoretical minimum number of bits per pixel needed to transmit the message without error, and what then is the maximum lossless compression ratio achievable?

b) Entropy coding methods are used. Groups of pixels on each line are encoded via the Huffman coding scheme. What is the Huffman codebook and achievable compression if we code groups of 2 pixels.

c). An alternative method of compression is via Run Length Coding. How does this work?

Assume we code a maximum run length of M pixels where $M = 2^m - 1$, where m is an integer, and the probability of a 0 is p. What is the probability of having a run of n zeros followed by a 1 (where $0 \leq n \leq M - 1$), and what is the probability of having M zeros?

d) It can be shown that the mean number of pixels in a run is given by

$$\mu = \frac{(1 - p^M)}{(1 - p)}$$

What therefore is the achievable compression ratio for M=7 and p = 0.90.

e) At the beginning of the FAX transmission, because of noise on the telephone line a bit is received as a 1 instead of a 0. Describe what happens to the received image in the cases of Huffman and Run length coding respectively.

F) IMAGE RESTORATION PROBLEMS

1) An input image $g(x, y)$ is blurred by convolution by a blurring function $h(x, y)$ and then has random noise $n(x, y)$ added. Give a mathematical expression for the resulting image $f(x, y)$. Describe how to restore the image using the Inverse Fourier filter, i.e. how to estimate $g(x, y)$ given $f(x, y)$ and knowing $h(x, y)$. Give an expression for the noise distribution $n'(x, y)$ in the restored image.

b) An alternative restoration filter is the Wiener Fourier filter. Give the mathematical expression for this filter as a function of spatial frequency, carefully defining all the quantities. Name one advantage and two disadvantages of the Weiner filter over the Inverse filter.

c) A weather satellite is launched into geostationary orbit so it looks down over the same part of the Earth. Unfortunately after launch it is found to be rapidly spinning on its axis so that images taken with even the shortest exposure are rotationally blurred. Describe a system to restore the images and save the mission.

2) State for the following image distortions whether the distortion is linear or non-linear and whether the point spread function is space-invariant or not.

i) A photograph taken through the front window of a car moving at constant velocity on a flat road.
ii) A telephoto image of a distant object taken out of the side window of a jeep moving on a bumpy dirt road.
iii) A image of one of Saturns moons taken from an accelerating space probe, where the velocity and acceleration vectors are parallel to each other and perpendicular to to vector between space probe and moon.
iv) A photograph taken by a digital camera with a maximum readout value so that all pixels greater than I_o in $I(x, y)$ equal I_o in $I'(x, y)$.
v) An image effected by multiplicative noise so that $I'(x, y) = I(x, y)n(x, y)$ where the elements of $n(x, y)$ are drawn from a random distribution.
vi) An image rotated 180°.

3) A sampled image I(m,n) is distorted by convolution with either the space invariant point functions h_1 or the function h_2 where.

$$h_1 = \delta(m, n) + \delta(m-1, n) + \delta(m+1, n) + \delta(m, n-1) + \delta(m, n+1)$$

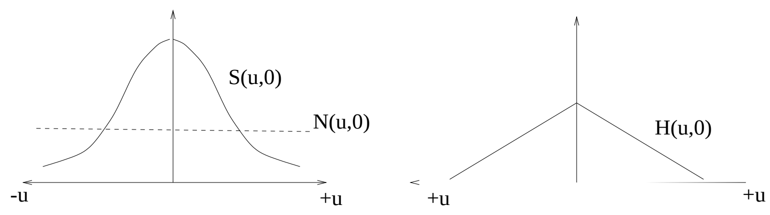
and

$$h_2 = 5\delta(m, n) + \delta(m-1, n) + \delta(m+1, n) + \delta(m, n-1) + \delta(m, n+1)$$

What are the 'Optical Transfer Functions' of these two convolving functions?. Assuming that the distorted images also contain random additive noise then in one image the distortion can be effectively removed using an Inverse Fourier filter, while the other requires the Pseudo-Inverse filter. Which image must use the Pseudo Inverse filter and why?

4) A particular image is drawn from a set of images having spectral power density S(u,v). This image has been distorted by convolution, described by multiplication of the image Fourier transform by the Optical Transfer Function H(u,v). White noise is also added with spectral density N(u,v).

S(u,0), N(u,0) and H(u,0) are shown in the figure below. Sketch on a copy of this figure the inverse and Wiener filter to be applied in order to restore the image.



5) Consider the problem of restoring images blurred by image motion.

a) Consider an image moving at a speed s past a camera whose exposure time is Δt. What is the Point Spread Function, and the Optical Transfer Function? Why is the Inverse Fourier filter unsuitable for restoring this image if there is any additive random noise?

b) Calculate the appropriate Wiener filter to apply to the above problem. Assume that the additive noise in the Fourier domain is white with power spectral density $s\Delta t$ while typical undistorted images have power spectral density $1/u$ where u is the spatial frequency. What is the Wiener filter value at $u = 1/(s\Delta t)$?

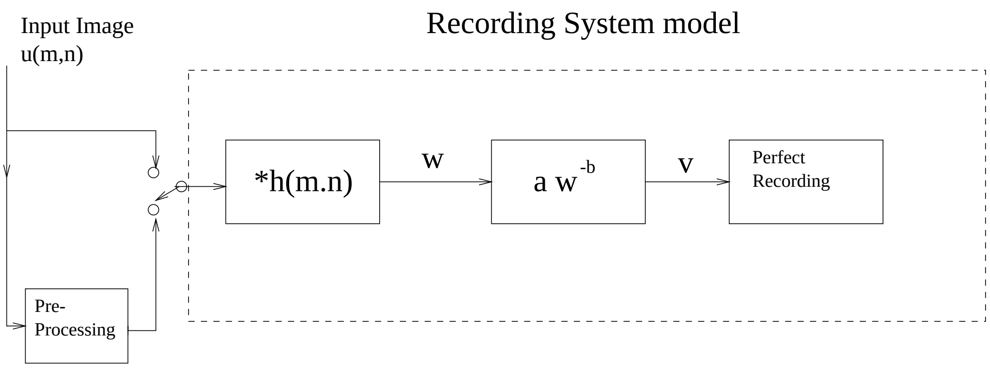
c) Consider a camera taking a picture of the front of a house out of a car side window. If the car is at rest at the start of an exposure of length Δt but accelerates so that $x(t) = at^2/2$, what is the Point Spread Function in the distorted image? The resulting Optical Transfer function differs in two important ways from the one obtained in the case of uniform motion, what are they?

d) A photograph is taken out of a side window of a car driving at a constant 40km/hour down Gibraltargatan towards Chalmers. Why in general is it not possible to use Inverse or Wiener filters to restore the blurring in this image?

6) A digital image stored in a computer is to be recorded on a a photographic film using a a flying spot scanner. The effect of the spot size of the scanner and the film non-linearity is equivalent to first passing the image through the model given below and then recording it on a perfect recording system.

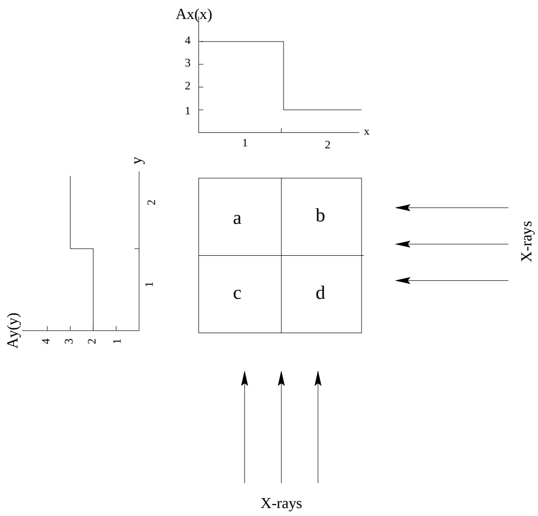
What preprocessing should be applied to the image before recording the image to compensate for (a) scanner spot size, assuming $b = -1$ and (b) film non-linearity, ignoring the spot size. The effect of the scanner spot size is to convolve the image with with $h(m, n)$ defined as

$$h(m, n) = \delta(m, n) + (1/4)[\delta(m - 1, n) + \delta(m + 1, n) + \delta(m, n - 1) + \delta(m, n + 1)]$$



If instead $h(m, n) = sin^2(m)sin^2(n)/m^2n^2$ **and** simultaneously $b = -0.5$ how should we preprocess the image so the photographic recording is as near to the original as possible? Why is it likely to be impossible to get a perfect reconstruction in this case, even in the absence of noise or numerical rounding errors?

7) The contents of a closed box are investigated by taking two X-rays from two directions 90° apart. One X-ray is taken with the X-rays shining through in the x direction and the other in the y-direction. The two pixel projected images of the amount of absorption are $A_y(y)$ and $A_x(x)$. We use image reconstruction to estimate the values of the absorption within the box at a,b,c and d.



i) The first pixel of the projected image $A_x(1) = a + c$, etc. Write the relationships between the four measurements $A_x(1), A_x(2), A_y(1), A_y(2)$ and the four unknowns a,b,c,d. and express as a matrix equation, $\mathbf{y} = \mathbf{B}\mathbf{x}$. Explain why, given only the data, there is no unique solution for a,b,c,d.

ii) Demonstrate that any solution which fits the data can be written as;

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Which values of γ give physically reasonable results?

iii) It is argued that images with lower total energy are more likely. Which image fits the data but also minimises the sum of squares of all the pixel values?