DATE: 26. February 2017

University of Stavanger

EXAM IN: ELE510 IMAGE PROCESSING with ROBOT VISION - Solution

DURATION: 4 hours

ALLOWED REMEDIES: Defined, simple calculator permitted.

THE SET OF EXERCISES CONSISTS OF 4 EXERCISES ON XXX PAGES

NOTES: Formulas are found on page YYYY.

Exercise 1

(25%)

a) Give a brief explanation of the following expressions and concepts:

1. point spread function
   Answer: When image transformation is performed using a linear operator, the operator can be defined by the "point spread function". This corresponds to an impulse response in signal processing. Let the input to the operator be a point source. Then the output of the operator corresponds to the point spread function and describes the operator.
2. image resolution
   Answer: the number of pixels used to represent an area ($nppixels/cm^2$) as well as the number of bits pr. pixel (i.e. the number of different gray levels that is possible to use) defines the resolution.
3. contrast in an image
   Answer: If an image has good contrast, all levels from white to black are being used.

b) i) What type of operators or filters can be represented by a filter mask. Explain how the filter mask is used to produce an output image (you can use a sketch).
Answer: When we have a shift invariant filter with finite support, we can use a filter mask (kernel). Linear filters can now use convolution (or correlation).

ii) Show example of filter masks that would blur the image, and sharpen or extract edges of an image.
Answer: Blur: mean filter, gauss filter
Sharpen/extract edges: Prewitt/Sobel

iii) Does the order of which we apply two linear operators on an image make any difference on the result?
Answer: Not if the operators are shift invariant.

c) A linear and separable operator can be written as:

$$g = h_c^T f h_r \tag{1}$$

Consider a small image:

$$f(x,y) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \tag{2}$$

i) Design the matrix h_r that will give the average of the pixel position with the next in the row. Let $g_{temp} = fh_r$, What is the content of h_r and g_{temp}?

ii) If we want to do the similar smoothing over the columns, define h_c, and provide the final g.

iii) Comment on the last row and column of the output image g.

d) The definition of 2D Discret Fourier Transform (DFT) and its invers is found in the Formula section. Define matrix U with element: $U(x,\alpha) = \frac{1}{N} e^{\frac{-j2\pi x\alpha}{N}}$, and let $\mathbf{u}_i$ denote a column in U
i) Describe in words (1-2 sentences) what it means that the 2D DFT is separable.
Answer: When a 2D transform is separable it means that we can do a 1D transform over the rows, and thereafter do another 1D transform over the columns of the result, and we sould get the same answer as if we did the columns first and thereafter the rows, and this is the answer of the 2D transform. Thus a seperable 2D transform can be implemented by a cascade of two 1D transforms.

ii) For the image $f(k,l)$, write the DFT, $\hat{f}(m,n)$, as a function of U and $f(k,l)$.
Answer: $\hat{f}(m,n) = Uf(k,l)U$
remember $U^T = U$. The convolution is taken care of by the matrix formulation. Can also be written:

$$\hat{f}(m,n) = \sum_{k=0}^{M-1} \frac{1}{M} e^{-j2\pi \frac{km}{M}} (\sum_{l=0}^{N-1} f(k,l) \frac{1}{N} e^{-j2\pi \frac{ln}{N}}) \tag{3}$$

iii) Why is the Fourier transform more commonly used than other transforms?
Answer: convolution theorem and detailed basis functions

Exercise 2

(25%)

a) i) Otsus method is much used for image segmentation. Describe the core idea of the method, use figures and equation (5) in the explanation. *Answer: Otsus method binarize an image by finding a threshold. The threshold is found by the use of equation 5 which gives the between class variance. Otsus method starts with a small threshold t, and evaluates $\sigma B(t)$ for increasing values until the variance starts decreasing. Assumes that the variance is well behaved without to many local maximums.*

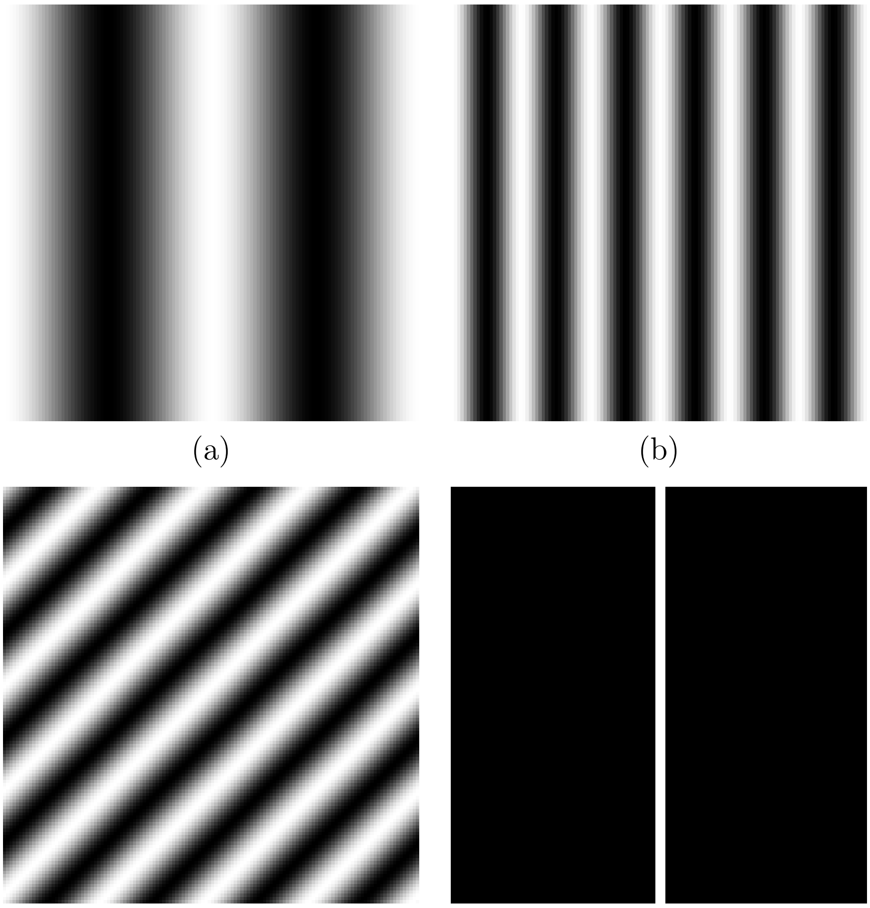
ii) Mention other method(s) for segmentation of images. *Answer: Otsus method is an example of thresholding algorithms. We have looked at other ways of finding the threshold, but Otsus is often a good solution to thresholding. There are many segmentation algorithms that are not based on thresholding, but on for example local regions etc. We have mentioned region growing, split and merge, and Watershed in class.*

b) Real world images are usually exposed to some kind of noise. Explain some different types of noise, and how we can deal with them.
Answer: Gaussian additive noise : Linear filters, Gaussian filtering etc. Salt and pepper noise: median filtering is much better than linear filters speckle (multiplicative) noise: homomorphic filtering (logarithmic function, thereafter linear filtering)

c) i) Why is Gaussian filters much used in image processing?
Answer: Gaussian in space domain is Gaussian in frequency domain. Does not impose ringing etc.

ii)What is the Laplacian of Gaussian (LoG) filter? Please use sketches when explaining.
Answer: Laplacian filter is second order derivative. This gives a double edge image, and the zero crossings can tell the exact location of an edge (as oposite to a first order derivative filter like sobel). Derivation is sensitive to noise, therefore it is important to do a smoothing before a second order derivative. Since the order of linear shift invariant operators does not matter, we can make a filter that is the laplacian (double derivative) of a Gaussian filter. By convolving an image with this filter (LoG) it corresponds to first smooth the filter and thereafter find the laplacian.

iii) The LoG filter is often approximated by the difference of Gaussian (DoG) filter. Explain how we can find the DoG filtered output of an image.
Answer: DoG is difference of Gaussian filter. It is found by taking the Gaussian of an image, and finding the difference between the original image and the Gaussian smoothed image. It gives a high pass variation of the image, and can be regarded as appriximation to LoG of the image.



(a) (b) (c) (d)

Figure 1: Figure to problem 2d)

d) Sktech the 2D Fourier transform of the images in Figure 1

Formulas

Discrete Fourier transform (DFT) and the invers DFT:

$$\hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{1}$$

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{2}$$

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_y \sum_x f(x,y)h(\alpha - x, \beta - y) \tag{3}$$

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y)\cdot r(x,y) \tag{4}$$

Between class variance:

$$\sigma_B^2(t) = \frac{[\mu(t) - \bar{\mu}\theta(t)]^2}{\theta(t)(1-\theta(t))} \tag{5}$$

LoG function:

$$LoG = [\frac{x^2+y^2-2\sigma^2}{\sigma^4}]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{6}$$

$$\mathcal{H} = \sum_{window} \{(\nabla I)(\nabla I)^T\} \tag{7}$$

$$= \sum_{window} \begin{bmatrix} I_x I_x & I_x I_y \\ I_y I_x & I_y I_y \end{bmatrix}. \tag{8}$$

$$R = \det(\mathcal{H}) - k(\frac{\mathrm{trace}(\mathcal{H})}{2})^2. \tag{9}$$

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{10}$$

$$\lambda\mathbf{p} = \mathcal{K}\Pi_0\mathbf{T}\mathbf{R}\,{}^W\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{11}$$

Here $\mathbf{p} = [x\ y\ 1]^T$ is the image coordinates in number of pixels and ${}^W\mathbf{P} = [X\ Y\ Z\ 1]^T$ the world coordinates in meter.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha \cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

where $\alpha = kf = \frac{f}{\Delta x}$ and $\beta = lf = \frac{f}{\Delta y}$.

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{13}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{14}$$

$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{15}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{16}$$