Module: **M1. Introduction to human and computer vision** **Final exam**
Date: December 1st, 2014 **Time: 2h30**
Teachers: Marcelo Bertalmío, Ramon Morros, Nava Rubin, Javier Ruiz, Philippe Salembier, Verónica Vilaplana

- Books, lecture notes, calculators, phones, etc. are not allowed.
- All sheets of paper should have your name.
- **Answer each part in a separate sheet of paper.**
- All results should be demonstrated or justified.

Part I

1: Why was the XYZ colorspace introduced alongside RGB? How is the function that transforms an RGB triplet into XYZ, linear or non-linear? Explain why, for any trichromatic display device, there are always colors that we can perceive but that the display is not able to reproduce.

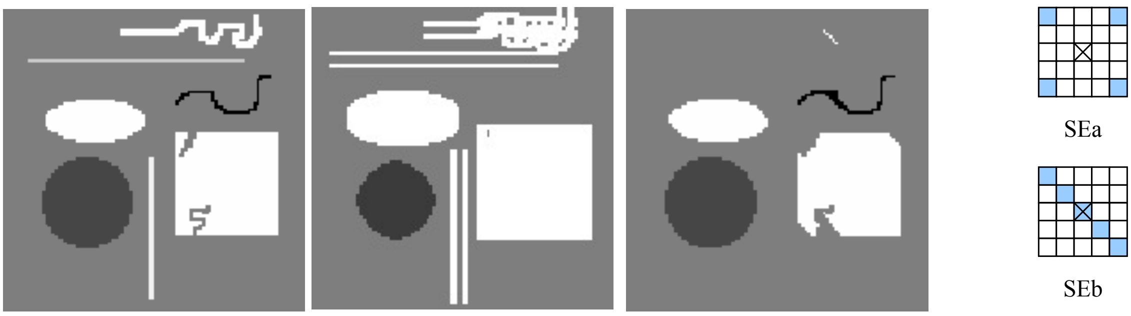
2: What is the color constancy property of the human visual system? How do cameras emulate it?

Part II

3: Histogram equalization redistributes the gray levels of images. Is it possible to increase the entropy of a digital image through histogram equalization? Why?

4: State the three important algebraic properties of an opening and define mathematically each property.

5: An original test image has been processed with two morphological operators T_1 and T_2. The resulting images are shown in the following Figure. The operators can be either an erosion, a dilation, a morphological opening or a morphological closing with one of the structuring elements shown on the right.



Original image | Image 1 (processed with T_1) | Image 2 (processed with T_2) | Structuring element

1. Define the T_1 operator and the structuring element used to create Image 1. Justify your response.
2. Define the T_2 operator and the structuring element used to create Image 2. Justify your response.

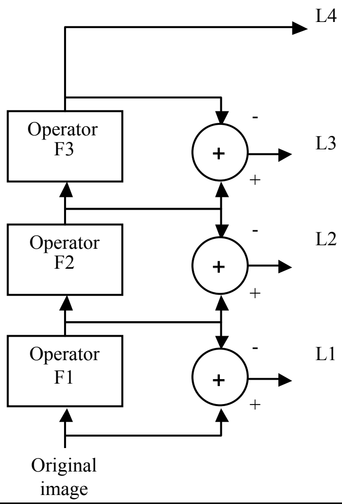
6: Consider the image analysis scheme given on the right side of the page.

In the sequel, various schemes are created with several operators Fi. For each scheme, you are asked to classify images **L1, L2 and L3** in one of the following class:

A: Zero image,
B: Positive image,
C: Negative image,
D: None of the previous cases.

a) System 1:
F1: Dilation with a flat square structuring element of size 5x5.
F2: Erosion with the same structuring element as F1.
F3: Closing with a flat square structuring element of size 3x3.

b) System 2:
F1: Opening with a flat square structuring element of size 7x7.
F2: Closing(Opening(.)) with the same structuring element as F1.
F3: Closing(Opening(.)) with the same structuring element as F1.



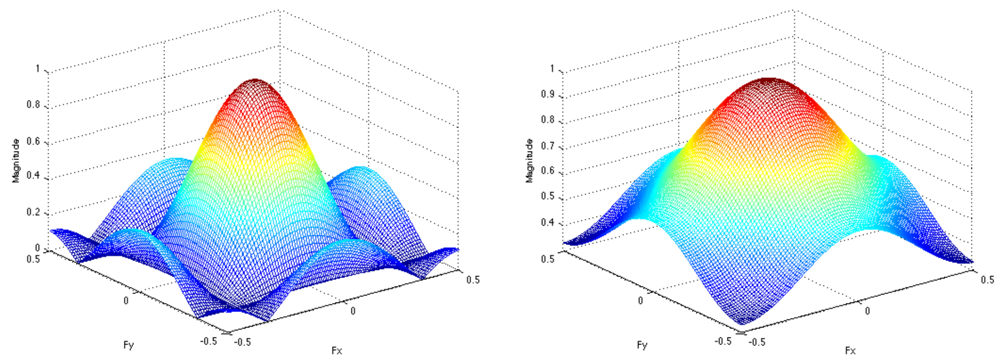
Part III

7: Compute the Discrete Fourier Transform of NxN samples of the image defined by:

$$\boldsymbol{x[m,n] = \delta[m]} \text{ with } \boldsymbol{\delta[m]} = \begin{cases} \boldsymbol{1 \; if \; m = 0} \\ \boldsymbol{0 \; otherwise} \end{cases}$$

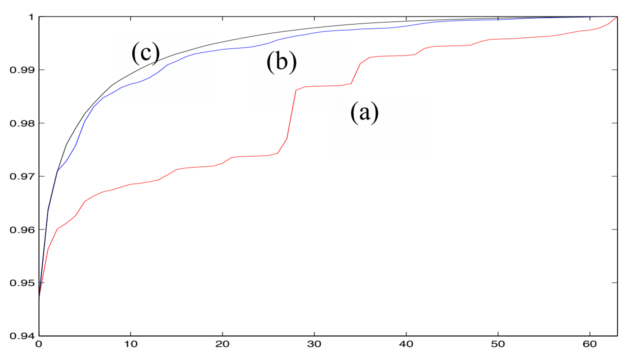
8: Consider the image x[m,n] of NxN pixels with values between [0,1], X[k,l] its Discrete Fourier Transform (DFT) of NxN samples and the image y[m,n] = 1 – x[m,n]. Compute Y[k,l], the DFT of y[m,n], in terms of X[k,l] and the inverse DFT of X[k,l]+Y[k,l].

9: Consider the following frequency responses of two low pass filters (Average and Gaussian) of 3x3 samples. Justify which frequency response corresponds to each filter.

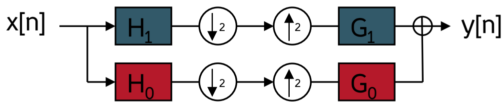


(a) (b)

10: The following figure shows the normalized cumulative energy for the zigzag scanned coefficients of the Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and the Karhunen-Loeve Transform (KLT) using blocks of 8x8 pixels of an image. Justify which curve corresponds to which transformation.



11: Consider the following filter bank decomposition of a 1D signal x[n]. Justify the condition on the filters $h_0[n]$, $h_1[n]$, $g_0[n]$ and $g_1[n]$ to be <u>orthogonal</u> filters.



Why are orthogonal filters usually <u>not</u> used in image processing as filters for the wavelet decomposition?

12: Discuss briefly the differences between Gaussian and Laplacian pyramids.

Part IV

13: List the sequence of steps involved in Canny edge detection, including image preprocessing. Describe the function of each step in one or two sentences.

14: The Harris corner detector finds corners in an image. Conceptually, corner detection can be thought of as an auto-correlation of an image patch. Consider a window which slides over an image patch

a) Describe intensity changes in the window when (i) the image patch is constant or 'flat', (ii) there is an edge in the patch or (iii) there is a corner in the patch.

b) The change in intensity for a shift (u,v) in a neighborhood of a point (x,y) is approximated by a quadratic form:

$$E(u,v) \approx [u \;\; v] \; M \begin{bmatrix} u \\ v \end{bmatrix}$$

Give the expression of M and explain how M is related to the behavior described in (a)

c) Discuss the invariance or covariance of the Harris corner detector with respect to (i) intensity changes and (ii) rotation

15: Provide the pseudo-code for the implementation of the Hough transform to detect circumferences in an image. NOTE: The equation describing a circle with center *(a,b)* and radius *r* can be written as:

$$a = x - r\cos(t)$$
$$b = y - r\sin(t)$$

16: Describe how to combine RANSAC and Least Squares in order to find instances of a given model in an image.

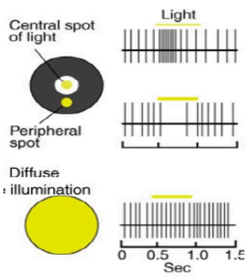
Part V

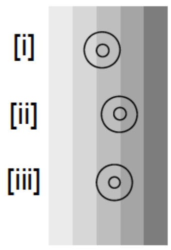
17: Explain briefly the difference between region merging and region growing methods in segmentation.

18: Describe the main problems when using a segmentation technique based on gradient binarization.

Part VI

19:



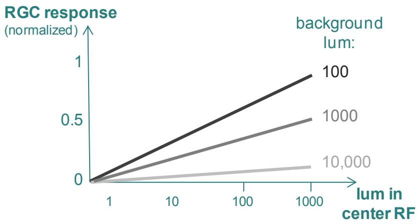


The left figure (taken from the slides shown in class) shows how a typical retinal ganglion cell (RGC), with a center-surround receptive field of type “ON-center”, responds in lab experiments. Three cases are shown: when light is projected only on its center (top panel), only in the surround region (middle) or diffusely all over its receptive field (bottom panel).

For the image shown on the right (vertical stripes of different gray levels), consider three such RGCs with receptive field locations as marked (denoted by [i]-[iii]). For each, indicate whether it will produce <u>more</u>, <u>less</u> or the <u>same</u> number of spikes/sec (on average), compared with its response to a homogeneous gray image.

[i] _______ [ii] _______ [iii] _______

20:



Experiments have shown that RGCs responses can be affected also by changes in the amounts of light falling *outside* their receptive field (RF) center- and surround- regions. Specifically, as shown in the schematic diagram above (taken from the slides shown in class), the average level of luminance in the "background" (a few degrees outside the RF of each RGC) will affect the gain of the cell's response to light within its RF center.
A. What is the purpose of this phenomenon? (ie, what is its computational advantage?)
B. How is it called?
C. What mechanism serves a similar purpose in a film-based (non-digital) camera?
D. What is the main advantage of the RGC mechanism, compared to that of the camera?
E. (Optional – extra points) How is this mechanism achieved (implemented) by RGCs?