Master in Computer Vision Barcelona

UAB UOC UPC upf.

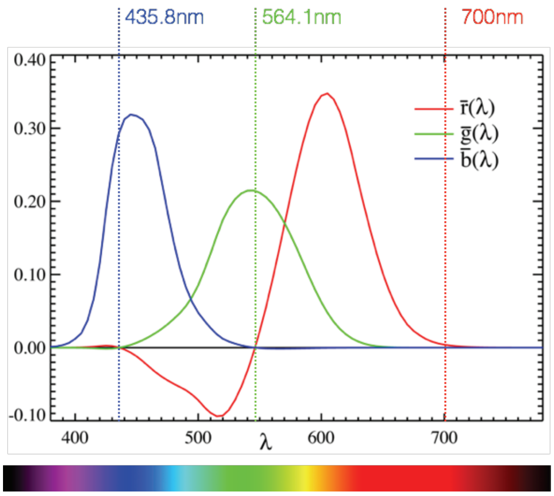
Module: **M1. Introduction to human and computer vision** **Final exam**
Date: November 28th, 2016 **Time: 2h30**
Teachers: Marcelo Bertalmío, David Kane, Ramon Morros, Javier Ruiz, Philippe Salembier, Verónica Vilaplana

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

Problem I David Kane (1 point)

1. Most camera sensors can record a high bit-depth signal (>10-bit). When saving to a low bit-depth image format (typically 8-bit), the signal is first passed through an encoding nonlinearity, before the bit-depth is reduced. The image is then passed through a decoding nonlinearity when it is viewed on a monitor. Explain why the reduction in bit-depth occurs after the encoding nonlinearity and before the decoding nonlinearity.

2. An additive color space describes how a limited set of base colors can be combined to produce a wider range of colors. Explain with reference to the below diagram, how the CIE color space does this and what perceptual experiments were conducted to define this space.



Problem III Marcelo Bertalmío (1 point)

1. Which visual perception property does the "automatic white balance" of cameras try to emulate? Explain the two most popular techniques for in-camera automatic white balance.

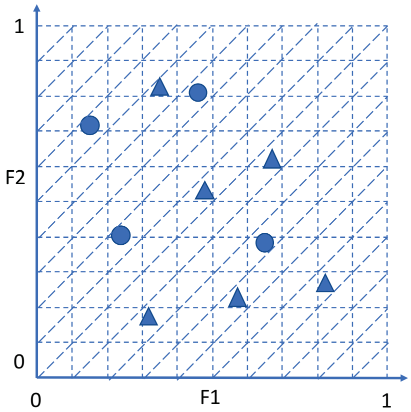
2. Why was the XYZ colorspace introduced alongside RGB? How is the function that transforms an RGB triplet into XYZ, linear or non-linear? Explain why, for any trichromatic display device, there are always colors that we can perceive but that the display is not able to reproduce.

Problem II	Philippe Salembier	(2 points)

1. In order to detect indoor images, we want to analyze the performance of two different features F1 and F2. F1 is a color-based feature and F2 a texture-based feature.
   We have an annotated database of 10 images that forms our ground truth. In involves 4 indoor and 6 outdoor images. When computing the two features on this database, we obtain the following values that are represented as points in the (F1, F2) space; circles represent indoor images and triangles outdoor images. We decide to use the following classification rule:

$$\begin{cases} if\ F1 - F2 \leq Th, & the\ image\ corresponds\ to\ an\ indoor\ \ scene \\ if\ F1 - F2 > Th, & the\ image\ corresponds\ to\ an\ outdoor\ scene \end{cases}$$

   Compute the Precision & Recall curve and define the optimum threshold value.



2. Consider the following gray level image:
   ima =

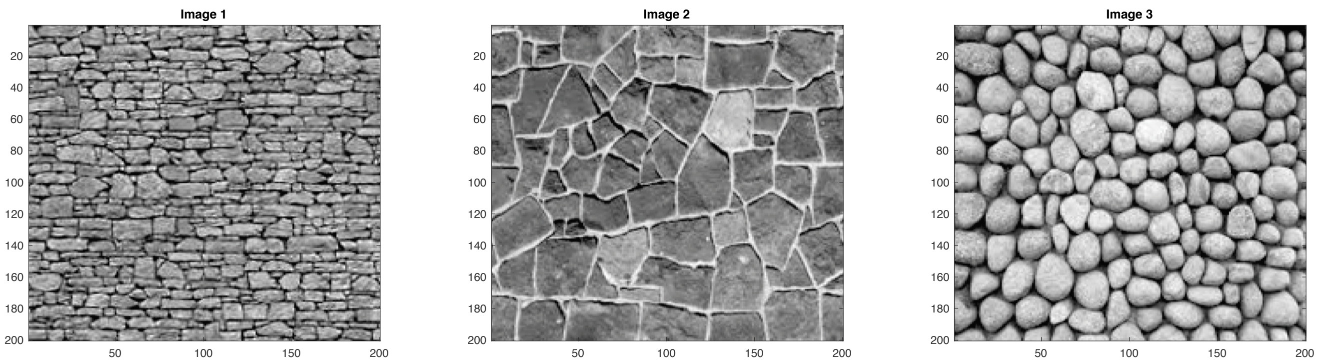
1	1	1	1	1
1	2	2	1	3
2	2	2	2	2
2	2	2	1	2
1	2	3	1	2

   Compute its erosion with the following flat structuring element $se[m,n] = \begin{bmatrix} -\infty & 0 & -\infty \\ 0 & \underline{-\infty} & -\infty \\ -\infty & -\infty & -\infty \end{bmatrix}$.

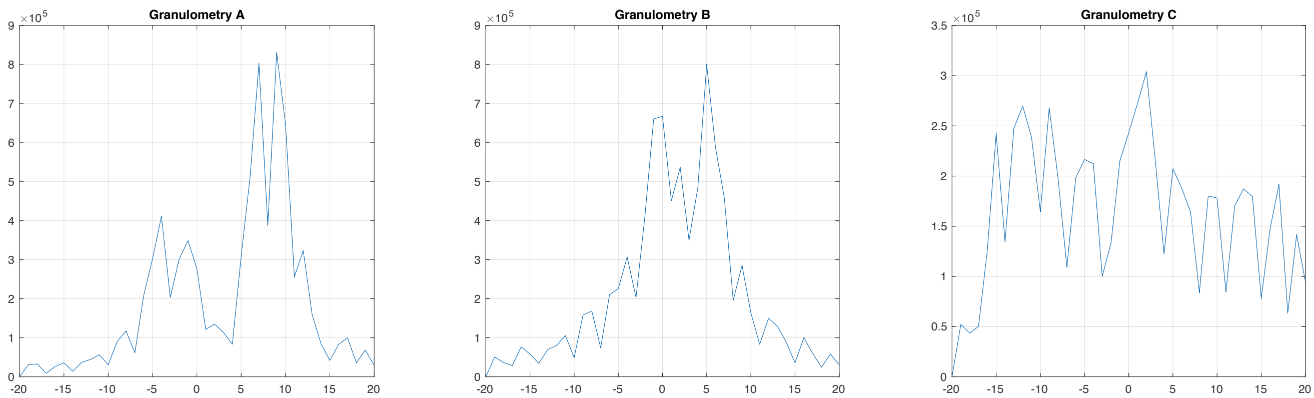
   Note: The underlined element represents the space origin. If necessary assume zero padding.

3. Define the three classical morphological gradients and describe their main difference when used for contour detection.

4. Considering the following three images: image 1, 2 and 3.



   We have computed their granulometry with circular structuring element. The three pattern spectrums: granulometry A, B and C are shown below.



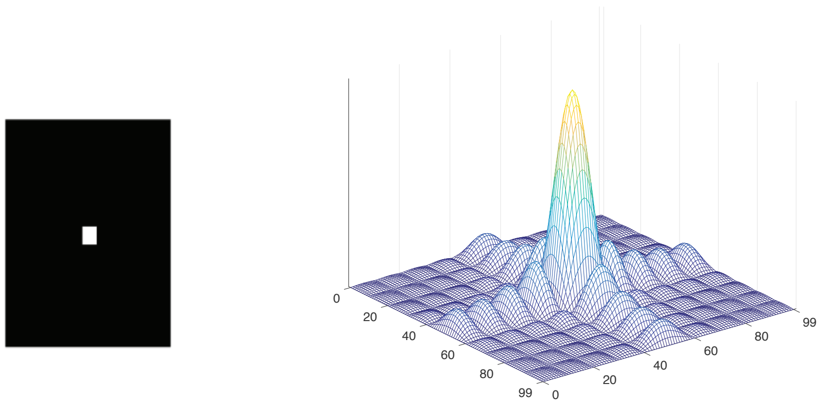
   Define the correspondence between the granulometric curves (A,B,C) and the images (1,2,3).

Problem IV	Javier Ruiz	(3 points)

1. Using the modulation property of the DFT, express the DFT of MxN samples (M even) of the image $x[m,n] \cdot (-1)^m$ in terms of $\tilde{X}[k,l]$, the periodic version of the DFT of MxN samples of x[m,n]

2. Compute the Fourier Transform, $X(F_x, F_y)$, of the image of MxM pixels defined by:
$x[m,n] = \delta[m] + \delta[n]$ with $\delta[k] = \begin{cases} 1 \; if \; k = 0 \\ 0 \; otherwise \end{cases}$

3. Consider the image x[m,n] of 100x100 pixels composed by a black background of level 0 and a white square of level 1 of 10x10 pixels (Figure a). Figure b represents the magnitude of the Discrete Fourier Transform of 100x100 samples using the centered representation. Obtain the value of the DFT at positions X[50,50] and X[60,50].

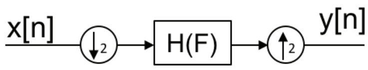


a) Image x[m,n]   b) Magnitude of the DFT with 100x100 samples

4. Detail the discrete impulse response of the Laplacian operator (second derivative). How can it be used to detect contours of an image?

5. Enumerate one advantage and one disadvantage of the Karhunen-Loeve Transform (KLT) versus other linear transformations.

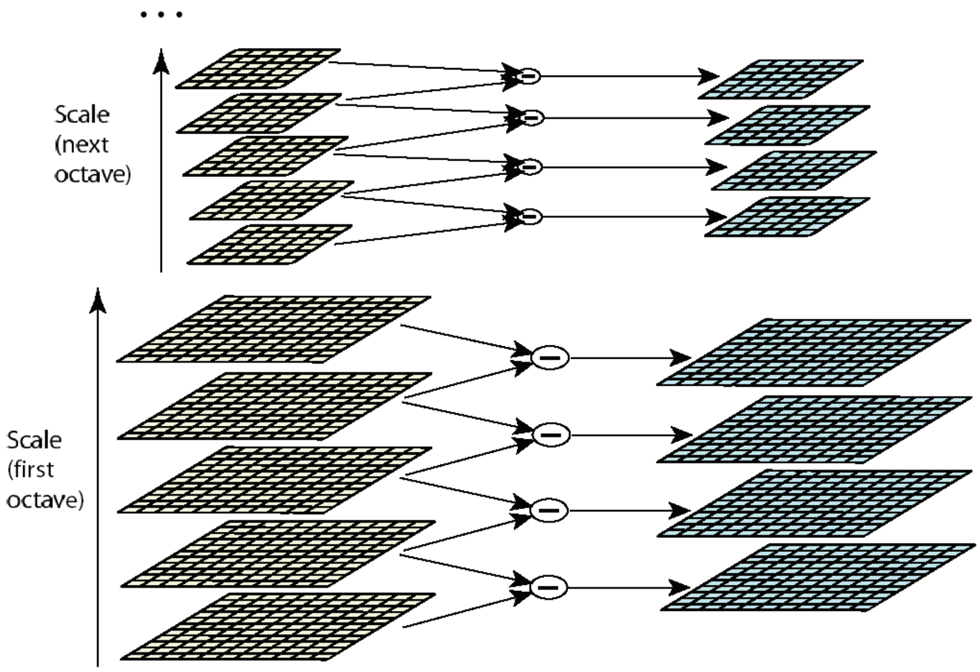
6. Consider the following system using down-sampling and up-sampling processes and a filter with frequency response H(F). Express the Fourier Transform Y(F)=FT{y[n]} as a function of X(F)=FT{x[n]} and H(F).



Problem V	Verónica Vilaplana	(1 point)

1. After running the Canny edge detector on an image, you notice that long edges are broken into short segments separated by gaps. In addition, some spurious edges appear. For each of the two thresholds (low and high) used in hysteresis thresholding, state how you would adjust the threshold (up or down) to address both problems. Assume that a setting exists for the two thresholds that produces the desired result. Explain your answer briefly.

2. The SIFT (DoG) detector finds keypoints using the image pyramid shown in the following figure. Explain how the image pyramid is created and how keypoint candidates are detected using this pyramid.



Problem VI	Ramón Morros	(2 points)

1: Describe the LMedS algorithm and comment its advantages over RANSAC.

2: Describe the basic idea of segmentation on the feature space and explain all steps necessary to obtain an image partition.

3: Formulate the problem of segmentation using GMM. Knowing that the solution can be obtained using Maximum Likelihood Estimation (Equation 1), explain the assumptions made, the parameters to compute and the procedure to obtain the solution.

$$\theta^* = \underset{\theta}{\operatorname{argmax}} \{L(\theta|x)\}$$

4: Describe segmentation using a region-growing approach.