1 - INTRODUCTORY CONCEPTS

1.1 Give an intuitive explanation for the fact that a pinhole camera has an infinite depth of field.

1.2 What is the effect of adding the same value to the three components of a colour image, represented in the RGB space? (Note: consider that the resulting values are below the maximum possible values for each component)

2 - IMAGE ENHANCEMENT AND FEATURE DETECTION (EDGES, CORNERS, LINES, ...)

2.1
a) The Gaussian filters are, in general, preferable when compared to simple mean filters. Why?

b) One important feature of the Gaussian filters is that they are "separable filters". Explain the concept of "separable filters" and justify the advantage of using it.

2.2 Consider the original image in Fig. 1 and the images in Fig. 2 and Fig. 3, which have resulted from transformation operations to the former image's histogram. It is known that one of those operations was histogram equalization.

a) Which of the two images is the result of that operation? Justify.

b) Indicate a possible operation that the other image is a result of.

Fig. 1 – Original image   Fig. 2 – Transformed image 1   Fig. 3 – Transformed image 2

2.3 A SIFT descriptor has typically 128 values. What is the meaning of these values?

3 SEGMENTATION AND POST-PROCESSING

Consider the image in Fig. 1. Show the result of applying an opening morphological operation in that image using the structuring element in Fig. 2. Describe the general effect of an opening operation.
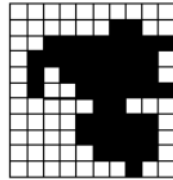

Fig. 1 – Binary image   Fig.2 – Structuring element

4 - GEOMETRICAL MODELLING OF A CAMERA / STEREOSCOPY

4.1 The perspective projection matrix of a camera is the following:

$$\begin{bmatrix} 0 & 1000 & -100 & 100000 \\ -1000 & 0 & -100 & 100000 \\ 0 & 0 & -1 & 1000 \end{bmatrix}$$

What are the pixel-coordinates of the point with coordinates (X,Y,Z)=(100,100,0) in an image captured with this camera?

4.2 Justify the following sentence: "The accuracy of a depth estimate based on stereoscopy varies with the depth itself."

PART2

5 – OBJECT RECOGNITION AND MACHINE LEARNING

5.1 Explain briefly the concept of visual words. In particular, explain 1) how can they be used to compare images, and 2) why is that comparison faster than an approach where the extracted local features are directly compared.

5.2 Suppose that we want to know if a student will pass or fail an exam in the Computer Vision course, given the grades he obtained in the assignments and the grades of the previous courses he attended. Can this be considered a classification problem? Explain why.

6 – CONVOLUTIONAL NEURAL NETWORKS

6.1 Comment the following sentence: "It is possible to replace the fully connected layers of a convolutional neural network (CNN) by an SVM classifier".

6.2 One of the most popular object detection models based on deep learning is YOLO, which stands for "You Only Look Once". Explain why this name is adequate and, in particular, why this is a differentiating factor from other models like Faster R-CNN.

6.3 Explain why the cycle consistency loss is important in the training process of Cycle-Consistent Adversarial Networks (Cycle GANs).

7 – MOTION

7.1 Comment the following sentence: "The optical flow is not always a correct estimate of the motion". Present an example that illustrates this.

8 – CASE STUDY

A company intends to develop a computer program to detect automatically the corners of the penalty area in images captured from football games, such as the one in Fig. 3. Briefly describe the main steps of the processing that you would follow to determine the coordinates of the corners.

Fig. 3 – Example of an image to process