PART 1

1 - INTRODUCTORY CONCEPTS

1.1 Explain the "depth of field" concept of a camera and tell how it varies with the aperture of the lens.

1.2 What is, in general, the expected result of an operation to equalize the histogram of an image represented in gray levels? Is it possible to predict the general effect of equalizing the histogram of each of the components, R, G and B, of a color image? Justify the answer.

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

2.1 The image of Fig. 1 is corrupted by impulse noise. Explain how to attenuate this noise in order to obtain an image similar to that of Fig. 2. Justify the answer.

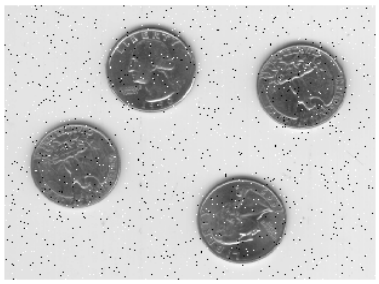

Fig. 1 – Image with noise

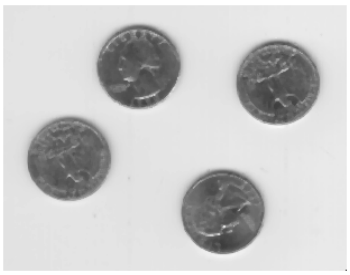

Fig. 2 – Filtered image

2.2 In relation to Canny edges detector, tell what is the purpose of the *non-maximum suppression* and *thresholding hysteresis* steps. If an image contains a long vertical edge, what are the neighbors of a pixel in that edge that are analyzed when performing each of these steps?

2.3 Is it possible to change the sensitivity of the Harris corner detector in order to detect more or fewer corners? If yes, please indicate how.

3 SEGMENTATION AND POST-PROCESSING

3.1 Thresholding techniques are widely used in image segmentation. Indicate two factors that may make it unfeasible to use a single threshold for segmenting an image.

3.2 Consider the pulmonary X-ray images of Fig. 3. The segmentation of pulmonary regions, shown in the figure, is frequently performed using Active Shape Models. How do you justify this option? What is the possibility of applying this type of segmentation technique?

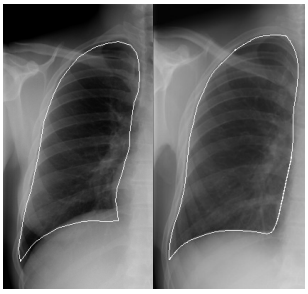

Fig.3 – Segmentation of pulmonary regions

3.3 Consider the images in Figs. 4 and 5. Describe a processing sequence based on morphological operators, starting from the binary image of Fig. 4, to obtain the image of Fig. 5, where the edges edges) of the objects of Fig. 4 are detected.

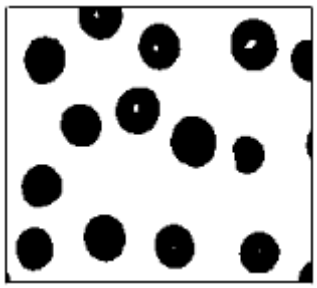

Fig. 4 – Binary image

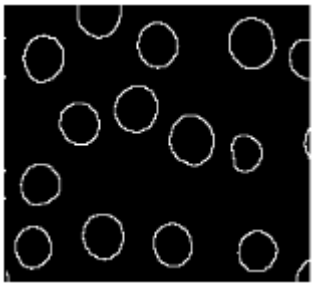

Fig. 5 – Detected edges

4 - GEOMETRICAL MODELLING OF A CAMERA / STEREOSCOPY

4.1 A camera can be modeled, geometrically, by a matrix of 3x4 elements, commonly referred to as a "perspective projection matrix". Give a physical, intuitive explanation for the fact that this matrix is not invertible.

4.2 Briefly describe the rectification of a pair of stereo images and the reason for their use.

PART 2 (not for the 1st mini-test)

5 – LOCAL FEATURES AND BAG OF WORDS

5.1 Analyze the two pairs of images in Figs. 6 and 7. What techniques would you use to detect correspondence in the two pairs of images? Justify your choice, highlighting any differences.

Fig.6 - Pair A

Fig.7 - Pair B

5.2 Consider a bag-of-words descriptor with a dimension of 100. How can you interpret the values that comprise the descriptor, and how can these be obtained?

6 – MACHINE LEARNING AND DEEP LEARNING

6.1 Compare the KNN and SVM classifiers in terms of training and testing time. Justify your answer.

6.2 Consider a Convolutional Neural Network similar to the AlexNet architecture. Give an interpretation of what is being trained in the first layers (convolution) and in the last layers (fully connected).

6.3 Consider the following statement “The Faster R-CNN model can be 10x faster than the R-CNN model.”. Why is this difference possible?

7 – MOTION

7.1 What is the aperture problem and how does it affect motion estimation?

8 – CASE STUDY

8.1 You are involved in a project that aims to create a system that can retrieve an ordered list of images that are similar to an input image. For example, if the user provides the input image (Fig. 1), the system returns some output images (Fig. 2-4) that are visually similar to the input image. Briefly describe the main steps (related to computer vision) that you would follow to accomplish this goal.

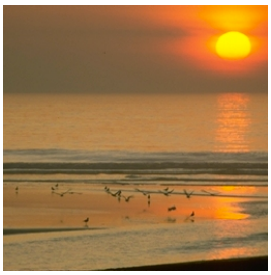

Fig. 1 - Input image

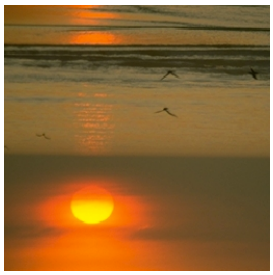

Fig. 2 - Output image 1

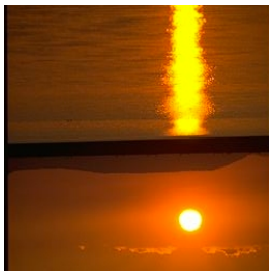

Fig. 3 - Output image 2

Fig.4 - Output image 3