FUNDAMENTALS OF COMPUTER VISION – 2nd PARTIAL EXAM JANUARY 9TH 2024 **RESOLUTION**

Name and Surname:

NIA

1.- What is a feature detector? What is a feature descriptor? Which types of feature descriptors have we studied? Can you put examples (name of the algorithm) of each of the types? (1 point)

(0.25) Feature detector: algorithm that localizes features in a given image. Features are pieces of information that are useful to solve a given task and represent an interesting part of the image.

(0.25) Feature descriptor: algorithm/method that describes a previously detected feature. The result is a measure that, once taken from the detected feature, allows us to distinguish it or compare it with others.

(0.25) Shape, Color, Texture, Motion

(0.25) Shape: Shape Signature, Chain Code, Skeleton, Shape Matrix …

Color: Dominant Color, Mean Grey Value, MPEG7 Color Descriptor, Normalized RGB Histogram, Color Contrast Color Indexing, Color Based Object Recognition

Texture: SIFT, SURF, ORB, GLOH, Local Binary Patterns, Co-ocurrence Matrix, Haar

Motion: Optical Flow, Block Matching, Angular Circular Motion

2.- What is an edge? What is the main difference between edges and corners? How do we detect them? (1 point)

(0.25) A line (sequence of points) that represent a change in depth, orientation, reflectance or illumination. Edge detection is particularly useful to detect object boundaries.

(0.3) Edges represent changes in an specific measure in only one direction (horizontal or vertical), corners represent changes in both directions. Associated eigenvalues are both high and of similar order for corners; for the case of edges, one of them is way bigger than the other

(0.45) Explanation of:

- Harris

- Prewitt, Sobel, Roberts

- Laplacian of Gaussian

- Canny

(see slides for the complete explanation)

3.- Please define what regions represent in an image. Please explain MSER algorithm (1 point)

(0.3) Regions are a set of pixels that differ in a given property (brigthness, color) with respect to neighboring regions. All the points in a blob can be considered in some sense to be similar to each other.

(0.7)

- MSER regions are connected areas characterized by almost uniform intensity, surrounded by contrasting background.
- They are constructed through a process of trying multiple thresholds.
- The selected regions are those that maintain unchanged shapes over a large set of thresholds.

How to create MSER regions

- For each threshold, compute the connected binary regions.
- Calculate Area at each threshold value.
- Analyze this function to determine those regions that have a similar value over multiple thresholds.
- Regions detected at different thresholds have different areas.

4.- Explain what Shape Signature is. Please include in the explanation the type of shape descriptor it is. What are the main drawbacks of this method? Is it rotation and scale invariant? In case it is not, how can we modify it? (1.5 points)

(0.3) This method (one of the Contour-based Global Shape Descriptors) represents the shape of an object by means of an uni-dimensional function which is extracted from the points belonging to the contour of the shape.

There are several possible Shape Signatures such as the centroidal profile, shape radii, complex coordinates, distance to the centroid, tangent or accumulative angle, curvature, arc length, etc.

(0.3) Drawbacks:

- High dependence of the starting point

- Very sensitive to noise

(0.4) Not rotation or scale invariant

How to make it rotation invariant? Correct orientation setting a specific keypoint as origin to perform shift matching

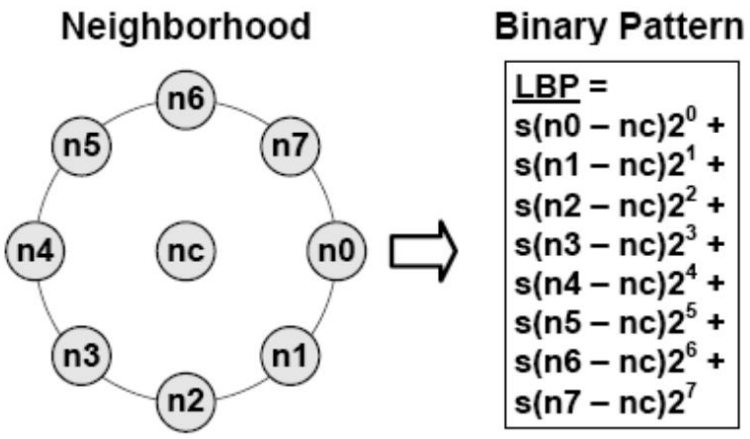
How to make it scale invariant? Normalization

5.- Explain Local Binary Patterns. Which type of descriptor are? How can we calculate LBP value for a given pixel? (1 point)

(0.5)

- The histogram of the binary patterns computed over a region is used for texture description. The operator describes each pixel by the relative grey levels of its neighboring pixels.
- If the grey level of the neighboring pixel is higher or equal, the value is set to one, otherwise to zero.
- Since correlation between pixels decreases with distance, a lot of the texture information can be obtained from local neighbourhoods. Thus, the radius R is usually kept small.

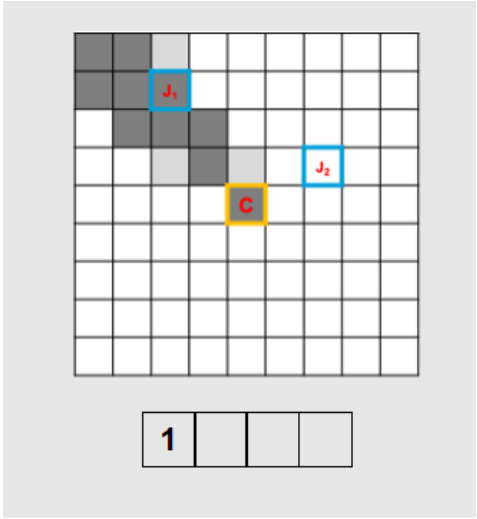
(0.5)



6.- ORB descriptor. Which is the main difference with respect to SIFT? How do we calculate the descriptor? Please use a graphical explanation along with text description (1.5 points)

(1) BRIEF (Binary Robust Independent Elementary Features)

1. We define a neighborhood around each keypoint (smoothing works)
2. Randomly select n pair of points inside the patch P1=(x1,y1) and P2=(x2,y2)
3. Let J1 and J2 the intensity at P1 and P2
4. if (J1 < J2)
    a. output = 1
5. else
    a. output = 0



$n = 4$

(0.5) Main difference: Similar to SIFT but it uses circular regions with 17 spatial bins and 16 orientation bins (272 dimension)

7.- How do we classify a new image using Bag of Words method? Which parameter determines the final number of words in a dictionary? Which is the most suitable metric to assess the performance of an image classification method? (1.5 points)

(0.5) We extract the features (feature detector + descriptor) and we use the previously created dictionary to create the histogram associated to the new image. Comparison (distance between histograms) between the new histogram and those related to the different classes in the dataset will tell us the final class for the new example.

(0.3) If we use k-means to group the different words into clusters, the number of desired clusters will set the final number of words in the dictionary. Other parameters to take into account are the number of classes or the number of examples per class.

(0.7) Confusion matrix, which represents the desired output and the actual output. From this matrix we can extract other per-class metrics such as precision, recall, specificity or accuracy

8.- Practical problem. Given this input image, what would you do to be able to generate the best possible method to achieve pedestrian detection? You can combine multiple feature extractors and descriptors if you want. (1.5 points)

No closed solution