Data Provided: None

The University Of Sheffield.

DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING

Autumn Semester 2011-2012 (2 hours)

EEE6082 Computational Vision 4

Answer **THREE** questions. **No marks will be awarded for solutions to a fourth question.** Solutions will be considered in the order that they are presented in the answer book. Trial answers will be ignored if they are clearly crossed out. **The numbers given after each section of a question indicate the relative weighting of that section.**

1. a. Feature extraction is usually the first step in a computer vision task. Explain what properties a good local region detector should have. **4**

b. The salient region detector is based on the unpredictability in the feature space and over scale. Calculate the feature-space saliency of the three image regions in Figure 1.1. The descriptor used here is grayscale intensity. **6**

c. Illustrate how to calculate the feature-space saliency and the inter-scale saliency of the image region with pixel values in bold in Figure 1.2. Here a region is represented as a square and ds=1. Draw and derive how you obtain the results. You are only required to show the formulas. No need to calculate the final results. **5**

d. Repeatability is usually used as a criterion for measuring the robustness of a local region detector. An original image has the size of 100x100 and the transformation is scaling with the factor of 2 both horizontally and vertically, i.e. the transformed image is 200x200. A local region detector detects 4 square regions on the original image. The locations of the top-left corners of the 4 regions are (10, 10), (10, 90), (90, 10), (90, 90) and the size of the regions is all 5x5. The same detector detects 3 square regions on the transformed image. The locations of the top-left corners of the 3 regions are (21, 21), (28, 170), (180, 30) and the size of the regions is all 10x10. Show how the repeatability rate is calculated. Two regions are considered to be corresponding if the overlap error is less than 50%. **5**

5	5	6	6
5	5	6	6
8	8	7	7
8	8	7	7

(a)

5	5	6	6
5	5	6	6
5	5	7	7
5	5	7	7

(b)

2	2	3	3
2	2	4	4
8	7	6	5
10	9	6	5

(c)

Figure 1.1

35	30	30	30	35
30	**25**	**20**	**25**	30
30	**20**	**20**	**20**	30
30	**25**	**20**	**25**	30
35	30	30	30	35

Figure 1.2

2. **a.** What are the advantages and disadvantages of the parts-based and global approaches to image classification? **4**

b. Roughly outline how an object detection via classification algorithm works. **4**

c. In the diagram shown in Figure 2 are the gradient magnitudes and orientations respectively for a grid of pixels. Assuming 6 orientation bins equally spaced from 0-360 degrees, draw a HOG (Histogram of Oriented Gradients) histogram for this grid, marking your measurements on the histogram clearly. Interpolation into orientation bins can be ignored here. **6**

d. Scale Invariant Feature Transform (SIFT) is a very popular method.

i) Which two parts are the SIFT algorithm composed of?

ii) How is orientation normalisation done for a detected SIFT region?

iii) What are the two methods for matching SIFT features in two images? **6**

3 30^0	1 80^0	2 200^0
2 35^0	2 75^0	4 220^0
3 85^0	2 210^0	2 340^0

Figure 2: The gradient magnitudes (above) and orientations (below) of a 3x3 square of pixels.

3. **a.** Roughly describe how the Bag of Features (BoF) model for image classification works. **4**

b. When using the Bag of Features method for image classification, the size of the visual vocabulary has a large effect on the algorithm's accuracy. Explain why both too large and too small a vocabulary is detrimental to accuracy. **4**

c. K-means clustering is usually used for building the visual vocabulary. Roughly outline how the K-means clustering algorithm works. **5**

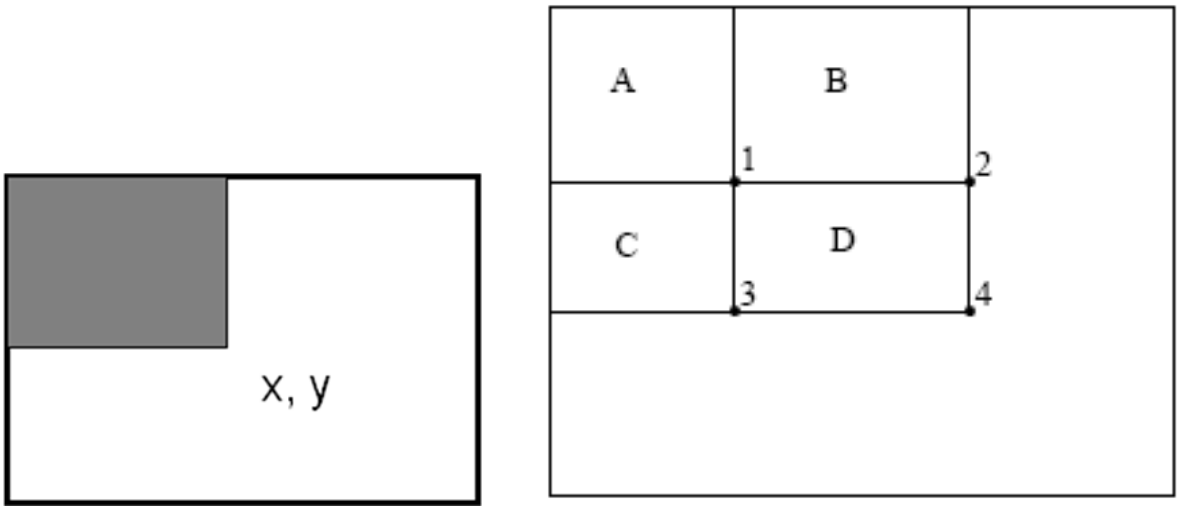
d. What are the differences between actions, activities and events? The BoF method can be also applied to action recognition. What are the advantages of this method in comparison to motion-template based methods? What is the difference between the BoF based action recognition and the BoF based object recognition? **7**

4. **a.** What's the difference between face detection and face recognition? What conditions would make face recognition more difficult? **5**

b. For face recognition, subspace learning techniques are usually used to do dimensionality reduction. What are the two classical subspace learning methods for face recognition and their major differences? **4**

c. The Viola-Jones face detector is very efficient. Integral images are used for calculating Haar-like features. If ii(x,y) = Sum of the values in the grey region in Figure 3(a), show how to compute C-A and B+C-A-D in Figure 3(b). **5**

d. Histograms are often used for representing images or image regions. Name two methods for comparing two histograms. Why colour histogram as a global descriptor would be problematic for image retrieval? Name a simple method to improve the global colour histogram. **6**



(a) **(b)**

Figure 3