Data Provided: None

The University Of Sheffield.

DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING

Autumn Semester 2012-2013 (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. The salient region detector is based on unpredictability in the feature space and over scale. What is used to define the feature space saliency? A histogram is usually used to approximate the probability density functions. Between a peaked histogram and a flattened/distributed histogram, which is more salient? Why? **4**

b. Calculate the feature-space saliency of the two image regions in Figure 1. The descriptor used here is grayscale intensity. **4**

c. In the salient region detector, central pixels in a region are more reliable than pixels on the edge. What sampling strategy can be used to incorporate this?

What method can be used to make discrete circular regions smoother?

When a region is small, the resultant histogram is usually sparse. What methods can be used to make the histogram less sparse? **6**

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, and rotation with an angle of 90 degrees clockwise. A local region detector detects 4 circular regions on the original image. The coordinates of the centres of the 4 regions are (10, 10), (10, 90), (90, 10), (90, 90) and the scale (radius) of all the regions is 5. The same detector also detects 4 circular regions on the transformed image. The coordinates of the centres of the 4 regions are (21, 21), (28, 170), (180, 18), (90, 90) and the scale (radius) of all the regions is 10. Show how the repeatability rate is calculated. Two regions are considered to be corresponding if the overlap error is less than 50%. **6**

5	5	5	5
5	5	5	5
8	8	8	8
8	8	8	8

(a)

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

(b)

Figure 1

2. **a.** Briefly describe how keypoint matching between two images works. **5**

b. Briefly outline how automatic scale selection in region detection works. **3**

c. Figure 2 shows the gradient magnitudes and orientations, respectively for a grid of pixels. Assuming 4 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. Where appropriate, you should perform linear interpolation on the orientations. **6**

d. **i)** What are the differences between Scale Invariant Feature Transform (SIFT) and HOG?

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	6 60^0	9 200^0
2 45^0	3 120^0	4 225^0
4 135^0	6 210^0	3 330^0

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

3. **a.** Briefly describe how the Bag of Features (BoF) model for action recognition works. How is the 'visual vocabulary' usually constructed? **5**

b. K-means clustering is a popular unsupervised learning method. If the features are pixel values including (50, 60, 80, 100, 120, 150, 170), briefly illustrate how K-means clustering works using these features. K = 3 here and you can assign the initial cluster centres yourself. **6**

c. Once you get the 'visual vocabulary' in the BoF model, how is an action sequence represented using the visual words? What is the dimension of this representation? **4**

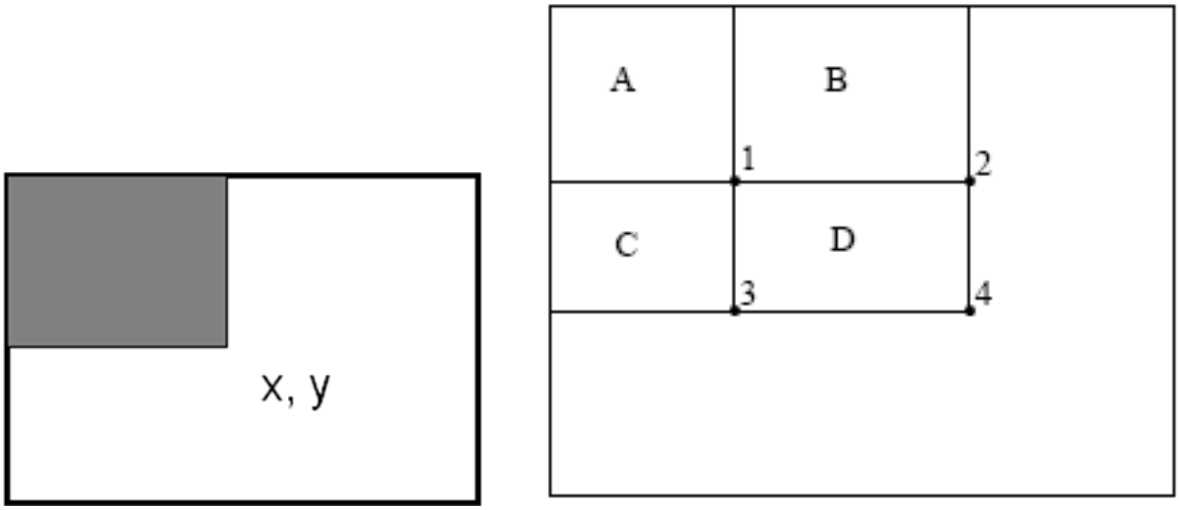
d. Motion template is another representation for actions. What's the difference between a Motion Energy Image (MEI) and a Motion History Image (MHI)? Briefly describe the advantages and disadvantages of BoF and motion template for action representation. **5**

4. **a.** Briefly describe how face detection works? How would face detection work for both frontal faces and profile faces? **6**

b. The Viola-Jones face detector is very popular. What are the three major components of this detector? **3**

c. Integral images are used for fast calculation. If ii(x,y) = Sum of the values in the grey region in Figure 3(a), show how to compute D-A and C+D-A-B in Figure 3(b). **5**

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



(a) **(b)**

Figure 3