**ECE 538 Homework**

1. a) In Figure 6.1-1 there is a dotted line between *Feature Analysis* and *Application*. Explain. b) In the same figure, there is a feedback loop from *Pattern Classification* and *Application*. Explain.

2. Why might image segmentation be performed before feature extraction and analysis?

3. a) Name the first step in feature extraction. Why is this important? b) Why is it important for a feature to be robust?

4. a) Describe a method to find perimeter of a binary object. b) How can this estimate be improved for objects with curved boundaries and complex shapes?

5. a) What is the thinness ratio of a circle? b) What is the thinness ratio of a rectangle that is 20 pixels wide by 80 pixels high? c) What values do you get in CVIPtools for (a) and (b)? Are these the same as your calculated values? Why or why not? d) Create an ellipse with CVIPtools. Next, use *Utilities->Create Border Mask* to create a version of the ellipse with a wavy border. Extract the thinness ratio with *Analysis->Features* from both. Which object has a smaller thinness measure? Explain.

6. a) What is the aspect ratio of circle with a radius of 25? b) Why rotate an object before finding the aspect ratio?

7. a) For the moment based features defined, why do we need the normalized central moments, instead of the regular moments?, b) Of what use are the RST-invariant moment-based features?

8. Use CVIPtools to explore the RST-invariant features. a) Create binary objects using *Utilities->Create* and *Utilities->Arith/Logic* to OR objects together, b) use *Analysis->Features* to extract the RST-invariant features from the objects, c) add noise to your objects with *Utilities->Noise* and extract the features from the noisy objects. Compare the results for the objects with and without noise, can you still classify the objects? Why or why not?

9. a) What can we say about an image with a narrow histogram? b) What can we say about a histogram skewed toward the left? c) What is a histogram with two major peaks called? What do the peaks typically correspond to?

10. a) What does the standard deviation of the histogram tell us about the image? b) What is maximum value for histogram energy? What image type does this correspond to? c) What does histogram entropy tell us? d) What is the relationship between histogram energy and entropy?

11. a) Given the following binary checkerboard image, where the image is 256×256 pixels and the squares are 32×32, calculate the histogram features, mean, standard deviation, skew, energy and entropy. Verify your results with CVIPtools using *Utilities->Stats->Image Statistics*. Are they the same? Why or why not?



b) Given the following binary circle image, where the image is 256×256 pixels and the radius of the circle is 32, calculate the histogram features, mean, standard deviation, skew, energy and entropy by using the equation of the area of a circle (). Verify your results with CVIPtools using *Utilities->Stats->Image Statistics*. Are they the same? Why or why not?



12. a) Describe the easiest method to obtain color features. b) Why might this method not be what we want? How can we get the information we do want?

13. a) What is the primary metric for spectral features? b) Regarding spectral features explain the statement: “The sector measure will tend to be size invariant, and the ring measure will tend to be rotation invariant”. Sketch images to illustrate this. c) Are the sector or ring spectral features translationally invariant? That is, if an object moves in the image, will the values change?

14. a) As we zoom in on a textured object, how does this affect the spectral features? b) As we zoom out on a textured object, how does this affect the spectral features?

15. Using a pixel distance, *d = 1*, find the gray level co-occurrence matrices for the horizontal, vertical, right diagonal and left diagonal directions, for the following image:



16. a) Find the 5×5 Laws texture energy mask for spots and edges, b) Find the 5×5 Laws texture energy mask for gray level and ripples c) Find the 5×5 Laws texture energy mask for ripples and waves d) What, if any, preprocessing is necessary to use the Laws energy masks?

17. Use CVIPtools to explore feature extraction. a) Select an image(s) of your choice with objects of interest, b) use *Utilities->Create->Border Mask* to create mask images for your objects of interest, c) use *Analysis->Features* to extract features that you think will be of interest for these objects. Examine the feature file. Are they results what you expected? Why or why not?

18. a) Define a feature vector that is useful to classify engineers and non-engineers, b) define a classification rule for these two classes based on your feature vectors.

19. Given the following two feature vectors, find the following distance and similarity metrics:

 

a) Euclidean distance, b) city block distance, c) maximum value, d) Minkowski distance, with *r* = 2, e) normalized vector inner product, f) Tanimoto metric

20. Given the following two feature vectors, find the following distance and similarity metrics:

 

First perform unit vector normalization, using the two vectors to find the component magnitudes.

a) Euclidean distance, b) city block distance, c) maximum value, d) Minkowski distance, with *r* = 2

21. Given the following two feature vectors, find the following distance and similarity metrics:

 

First range normalize the vectors, using the following ranges:

a) Euclidean distance, b) city block distance, c) maximum value, d) Minkowski distance, with *r* = 2, e) normalized vector inner product, f) Tanimoto metric

22. In the following scatter plot we have a two-dimensional feature space with all our sample vectors shown for two classes. Discuss any reasons to remove or add any feature vectors to our data set before we begin developing the pattern classification algorithm.

 **f2**

**f 1**

xxx xx

 x x

o

x x x o o o

 o o o x

 x x o o o

 x x oo o

 x x x o x

 x o oo

 o o o

23. a) When developing a classification algorithm, why do we divide our data into training and test sets? b) Describe two methods for dividing the data into these two sets. Discuss important factors to consider when selecting the training and test sets.

24. Describe the *leave-one-out* and the *leave-K-out* method of developing and testing pattern classification algorithms.

25. Describe an example, other than the ones in the book, which shows why a cost function is important when developing a pattern classification algorithm.

26. Given the following feature vectors, with two classes:

Class 1:  Class 2: 

a) Using the Nearest Neighbor classification method, and the absolute value distance metric, classify the following unknown sample vector as Class 1 or Class 2:



b) Use K Nearest Neighbor, with K = 3

27. Given the following feature vectors, with two classes:

Class 1:  Class 2: 

Using the Nearest Centroid classification method, and the absolute value distance metric, classify the following unknown sample vector as Class 1 or Class 2:



28. Given the following feature vectors, with two classes:

Class 1:  Class 2: 

Using the Nearest Centroid classification method, and the normalized vector inner product similarity measure, classify the following unknown sample vectors as Class 1 or Class 2:

a)  b)  c) 

29. a) What type of preprocessing normalization should we do to apply Bayesian analysis? b) What do we call the n-dimensional form of the linear discriminant function? c) Given the following scatter plot, draw a linear discriminant function to separate the two classes

 **f2**

**f 1**

xxx xx

 x x

 xx xx xx x

x x x o o o

 xxxxxxx o o

x x x o o o

 x x x ooo oo o

 x x x o

 x o oo

 o o o

30. a) What are the main distinguishing characteristics of a neural network? b) What type of preprocessing should be done to apply a neural network? c) Describe three types of activation functions. Which one is used most often? d) In general, how does a learning algorithm work?

1. Given the following feature vectors for a training set, with two classes:

Class 1:  Class 2: 

a) Perform standard normal density normalization (SND) on the feature vectors.

b) Using the Nearest Centroid classification method, and the normalized vector inner product similarity measure, classify the following unknown sample vectors as Class 1 or Class 2:

i)  ii)  iii) 

2. Using a pixel distance, *d = 1*, a) find the gray level co-occurrence matrices for the horizontal, vertical, right diagonal and left diagonal directions, for the following image, b) find the texture features energy, inertia, correlation, inverse difference and entropy.



3. Given the following feature vectors for a training set, with two classes:

Class 1: Class 2: 

a) Perform min-max normalization on the feature vectors, using 0 for the minimum and 1 for the maximum.

b) Using the Nearest Neighbor classification method, and the Minkowski distance measure, with r = 3, classify the following unknown sample vectors as Class 1 or Class 2:

i)  ii)  iii) 

4. Using a pixel distance, *d = 1*, a) find the gray level co-occurrence matrices for the horizontal, vertical, right diagonal and left diagonal directions, for the following 8x8 image, b) find the texture features energy, inertia, correlation, inverse difference and entropy.

$$\left[\begin{matrix}\begin{matrix}\begin{matrix}1&1\\2&3\end{matrix}&\begin{matrix}5&6\\5&7\end{matrix}\\\begin{matrix}4&5\\7&5\end{matrix}&\begin{matrix}7&1\\1&2\end{matrix}\end{matrix}&\begin{matrix}\begin{matrix}7&7\\1&5\end{matrix}&\begin{matrix}2&3\\7&0\end{matrix}\\\begin{matrix}2&2\\5&3\end{matrix}&\begin{matrix}0&2\\5&6\end{matrix}\end{matrix}\\\begin{matrix}\begin{matrix}7&1\\3&5\end{matrix}&\begin{matrix}1&5\\7&0\end{matrix}\\\begin{matrix}5&6\\2&3\end{matrix}&\begin{matrix}7&1\\5&7\end{matrix}\end{matrix}&\begin{matrix}\begin{matrix}6&7\\2&0\end{matrix}&\begin{matrix}7&2\\2&3\end{matrix}\\\begin{matrix}1&5\\1&4\end{matrix}&\begin{matrix}6&7\\5&7\end{matrix}\end{matrix}\end{matrix}\right]$$

5. Using a pixel distance, *d = 2*, a) find the gray level co-occurrence matrices for the horizontal, vertical, right diagonal and left diagonal directions, for the 8x8 image from the previous exercise, b) find the texture features energy, inertia, correlation, inverse difference and entropy.

6. a) Using CVIPtools create images of circles, ellipses, rectangles, squares and each of the object types with holes – this gives eight object classes. Create ten objects for each class of various sizes and orientations. Divide the image set into five of each class for training and five for testing. Use CVIPtools to achieve 100% correct classification. What parameters were used to achieve these results? b) Repeat (a), but blur the image with a 7×7 mean filter. Note that preprocessing steps may improve results. c) Repeat (a), but add zero mean Gaussian noise with a variance of 400. Note that preprocessing steps may improve results. d) Repeat (a), but blur the image with a 7×7 mean filter and add zero mean Gaussian noise with a variance of 400. Note that preprocessing steps may improve results.

7. Consider a system to identify land mines where we are classifying found objects. a) What are the classes? b) What are the relevant cost functions? In other words, what are the risks? c) How can we incorporate the cost functions into our analysis of classification algorithms? How can we apply *specificity* and *sensitivity* metrics?

8. Given the following results from an experiment testing a classification algorithm with four classes, A, B, C, and D, and results shown in the classification matrix, find for each class, the: a) sensitivity, b) specificity, c) precision, d) the F-measure, and, e) Do you think these are good results? Why?

|  |  |
| --- | --- |
| **CLASSIFICATION MATRIX** | **Classifier Results** |
| **A** | **B** | **C** | **D** |
| **Actual Class** | **A** | 8 | 1 | 0 | 1 |
| **B** | 1 | 7 | 2 | 0 |
| **C** | 0 | 0 | 9 | 1 |
| **D** | 0 | 2 | 2 | 6 |

9. Find the correlation coefficient for the following feature vectors:

a) $A=\left[\begin{matrix}2\\4\end{matrix}\right] and B=\left[\begin{matrix}4\\2\end{matrix}\right]$ b) $A=\left[\begin{matrix}2\\4\end{matrix}\right] and B=\left[\begin{matrix}-2\\-4\end{matrix}\right]$

c) Compare your answers for (a) and (b). Explain.

d) Graph the two vectors for (a) and (b). After thinking about the results and examining these graphs, do you think the correlation coefficient is an effective similarity measure? Why or why not?

10. Find the fuzzy similarity between the following feature vectors: a) **A** & **B**, b) **A** & **C**, c) **A** & **D**, d) **B** & **C**, e) **B** & **D**, f) **C** & **D,** g) Explain the results

$$A=\left[\begin{matrix}1\\1\\0\end{matrix}\right] B=\left[\begin{matrix}0.75\\0.75\\0.5\end{matrix}\right] C=\left[\begin{matrix}0.25\\0.5\\0.5\end{matrix}\right]D=\left[\begin{matrix}0.25\\0.25\\0.25\end{matrix}\right]$$

11. a) Given the following samples in a transformed two dimensional feature space, with ‘x’ being one class and ‘O’ being a second class, if we are designing a support vector machine (SVM), which separating line would be selected, the red or green one? b) Explain your answer.

**x x x**

 **x x x xx**

**x x x xx**

 **x x x**

**O O O**

**O O O O**

 **O O O O O**

 **O O O**