Name: _____

EE435 Spring 2012 Exam 3
Due: Beginning of Class, Tuesday 01 May 2012

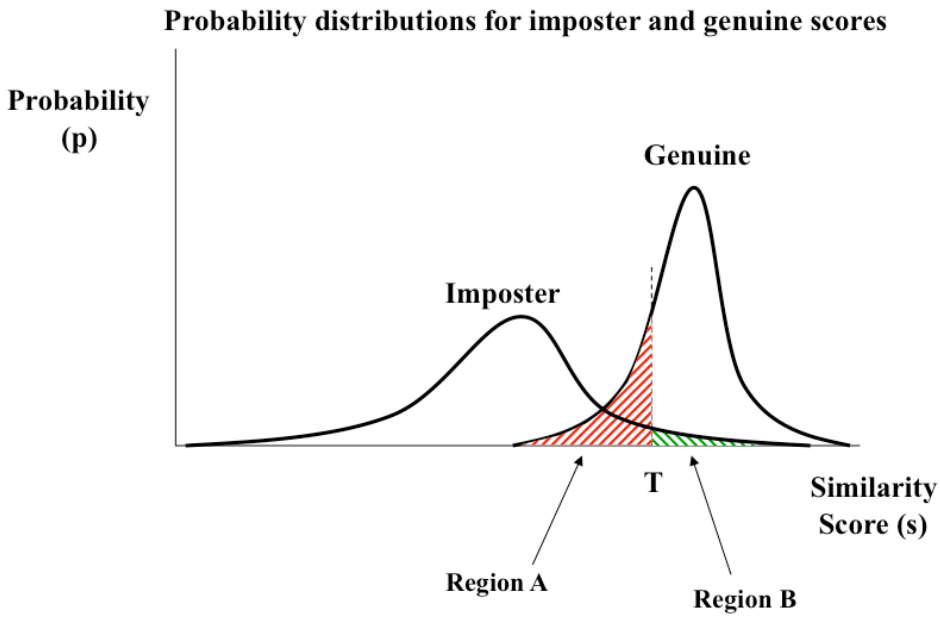
Problem	Points Possible	Points Earned
1	25	
2	25	
3	25	
4	25	
Total	100	

This take-home exam is open book, open notes and open computer. Do not converse with anyone else regarding this test. If you are stuck on a question, you can ask me for some guidance, but the guidance may cost you some points. I will let you know how much it would cost before you commit to accepting my guidance.

For this exam, you may use:
-Class notes/homework/quizzes/exams from the current semester
-MATLAB
-the book "Digital Image Processing using MATLAB"
-the book "Introduction to Biometrics"

Do NOT use any other reference material.

Turn in these sheets along with any scratch sheets you might have used. For problem 4, turn in all code you used in hardcopy.



1. (25 pts total) The figure above represents the results of some testing for the design of a biometric recognition system. The probability distributions for genuine and imposter scores are displayed. Using this figure, answer the following questions:

   a. (5 pts) In terms of expected system performance, what does the shaded Region A represent?

   b. (5 pts) In terms of expected system performance, what does the shaded Region B represent?

   c. (5 pts) What does the vertical line at $s = T$ signify?

   d. (5 pts) If the area of Region A is 0.05, and the area of Region B is 0.015, would you say that the system is more geared towards security or convenience of users? Explain.

   e. (5 pts) What happens to system performance as T is moved to the left, in terms of convenience and security? Explain.

2. The multi-day passes used to enter Disney World use biometrics to ensure that the correct person is using the pass. Each pass is associated with a specific individual using fingerprint recognition. A person purchases the pass, and then enrolls their fingers in the system the first time they try to enter. To enter the park another time using one of these passes for admittance, they must insert their card in a card reader, and then insert their right index and middle finger into a fingerprint sensor. If they match, they get in. If not, they go to customer service to figure out why not.

   Suppose that on March 13, 2007, the following data was collected:
   1607 persons bought multi-day passes, enrolled in the system and attempted to enter the park.
   38,345 persons who already had valid multi-day passes attempted entrance again.
   129 persons who had forged/invalid passes attempted to get in.

   Of the people that attempted to enter the park, 37,417 persons with valid passes were able to get in just fine. The remainder with valid passes who had been successfully enrolled previously, could not enter, and had to go to customer service to fix it. In addition, 7 people managed to get in with invalid passes.

   Using this information, answer the following:

   a. (5 pts) Is using fingerprints in this way *identification* or *verification*? Explain

   b. (5 pts) What is the FAR on March 13, 2007?

   c. (5 pts) What is the FRR on March 13, 2007?

   d. (5 pts) Based on your answers to parts (b) and (c), would you say that Disney World's system was geared more towards security or convenience? Why?

   e. (5 pts) If the FAR is what you computed in part b, how many false rejects would have to have happened such that the system was operating at its EER? Explain.

3. <u>Explain</u> whether the following statements from Midn Durwood are valid or not.

   a. (5 pts) Midn Durwood reports that when he was doing his iris recognition project for the biometrics class using the Daugman algorithm, that when he averaged all of the Hamming distances for the genuine comparisons, the mean value was –0.33, and it was +1.02 for the imposter comparisons.

   b. (5 pts) Midn Durwood says that performing morphological dilation with a circular structuring element can result in objects being reduced in size.

   c. (5 pts) Suppose you have a 2-class pattern recognition problem, male versus female, and height and weight are used as the features. The average male height is 71", average male weight is 185 lbs. The average female height is 67", average female weight is 119 lbs. For all of the collected training data (where ground truth is known), the standard deviation for both male and female height is 1.2, and the standard deviation for both male and female weight is 8.5. A minimum distance classifier will be used for recognition. Each class will be represented by a template located at the mean values of the features. Now you begin testing on unknown pattern vectors. Midn Durwood says it won't matter whether you use the Euclidean distance or the weighted Euclidean distance metric to measure distances…the accuracy of the system will be the same.

   d. (5 pts) Midn Durwood says that the only biometric mentioned in our class that has 100% universality is DNA.

   e. (5 pts) Midn Durwood says that if we turned on all the lights in the biometrics lab, and started collecting video, when no one was in the lab, that all of the video frames would be identical because nothing moves.

4. (25 pts total). In a Daugman iris recognition system, each iris template in the database is composed of two parts: (1) the iris code that represents the features within the iris; and (2), a mask that represents which parts of the iris code are valid to compare. Some parts of the iris code are not valid, because eyelids, eyelashes or glare obscured the actual iris.

The fractional Hamming distance is the numerical value that determines how different two iris templates are. The fractional Hamming distance between two irises (iris A and iris B) is computed as:

$$HD = \frac{\|(\text{code A} \otimes \text{code B}) \cap \text{mask A} \cap \text{mask B}\|}{\|\text{mask A} \cap \text{mask B}\|}$$

Here, the $\otimes$ symbol represents binary XOR, and the $\cap$ symbol represents binary AND. After the XOR and the AND operations are performed, the || x || operator counts the number of logic 1 bits (a summation). Basically the equation above computes the number of valid iris bits that do not match in both iris templates divided by the number of valid bits. In other words, overall, the fractional Hamming distance is the fraction of valid bits that do not match between the two iris templates. Typically, if HD is < 0.32, the two irises are said to be a match.

a. (10 pts) Write a well-commented MATLAB function called *HD.m* that will compute the fractional Hamming distance between two iris templates. Your function should NOT use loops, but if you do use loops, you will lose points.

   Usage: [*fhd, nbits*]=*HD*(*codeA, maskA, codeB, maskB*)

   The four inputs are the iris code and mask from one template (here called codeA and maskA), and the iris code and mask from the other template (here called codeB and maskB). The function returns two values: (1) the fractional HD value based on the equation above; and (2) the number of valid bits compared (this is the denominator of the above equation)

   Error Checking: If the dimensions of the four inputs are not the same, or if the inputs are not binary (all values either 0 or 1), it returns *fhd* = [ ] and *nbits* = [ ], along with an error message,

b. (0 pts: for your benefit only) To test your HD function, download the following four binary images: test_codeA.bmp, test_maskA.bmp, test_codeB.bmp, test_maskB.bmp. Read them into MATLAB and call them test_codeA, test_maskA, test_codeB, test_maskB. Running the following commands should result in the HD values indicated:

```
>> [fhd, nbits]=HD(test_codeA, test_maskA, test_codeB, test_maskB);
```

   results in *fhd*=1 (no bits match, because test_codeB is the <u>exact</u> inverse of test_codeA), and *nbits*=30,000 (because both masks are all 1s, so all bits are valid).

```
>> [fhd, nbits]=HD(test_codeA, test_maskA, test_codeA, test_maskA);
```

   results in *fhd*=0 (comparing identical bits, so all bits are the same), and *nbits*=30,000 (because both masks are the same, all 1s, so all bits are valid).

Note: feel free to test your HD function with any other codes/masks you care to come up with on your own.

c. (10 pts) Download the iris codes and templates for the 5 irises (labeled A, B, C, D and E) at the course website. Fill in the following tables with the HD results using your HD function. The row and column headers indicate which templates are being compared, i.e., in the block right below the letter A, your result would be the Hamming Distance when comparing template A to template A. You do not need to fill in the gray blocks.

Hamming Distances					
	A	**B**	**C**	**D**	**E**
A vs:					
B vs:					
C vs:					
D vs:					
E vs:					

# Bits Compared					
	A	**B**	**C**	**D**	**E**
A vs:					
B vs:					
C vs:					
D vs:					
E vs:					

d. (3 pts) With the recognition threshold set at < 0.32, which, if any, irises (A, B, C, D, E) are from the same eye?

e. (2 pts) Why don't you need to fill in the gray blocks?