

CS 510 Midterm – Spring 2012

Name ____ **ANSWER KEY** ____

EID _____

Question	Max Points	Points
1	15	
2	15	
3	20	
4	20	
5	10	
6	10	
7	10	
TOTAL	100	

<This Page Left Blank on Purpose>

Q1 Applications of Computer Vision (15 Points)

We devoted two lectures to student presentations of commercial applications of computer vision. Please succinctly summarize three distinct applications, one paragraph for each.

Application 1:

There is no single correct answer. Indeed, any reasonably substantive discourse on any three of the topics covered by the term papers would constitute a good answer. For the record, those topics are:

- Film and Video Production
- Security: Monitoring and Surveillance
- 3D Object Modeling
- Medical Imaging
- People Tracking for Sales and Marketing
- Driver Assistance
- Biometrics
- Gesture Recognition
- Autonomous Vehicles
- Security: Threat Detection
- Sports Video - Referee
- Sports Video - Broadcast

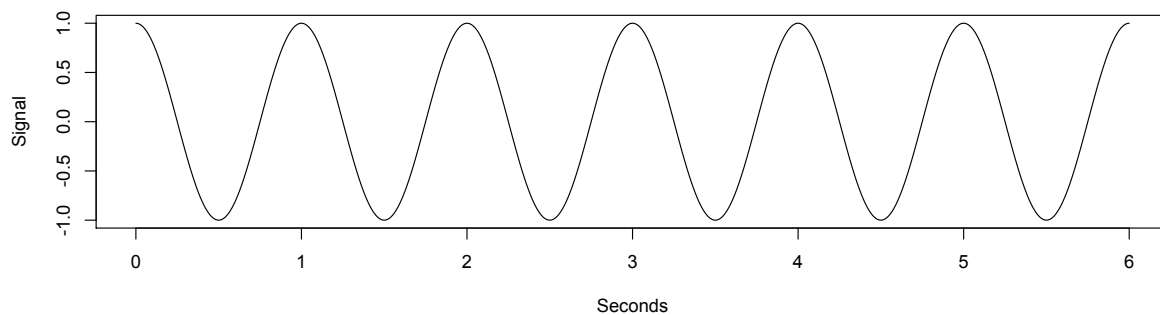
Application 2:

Application 3:

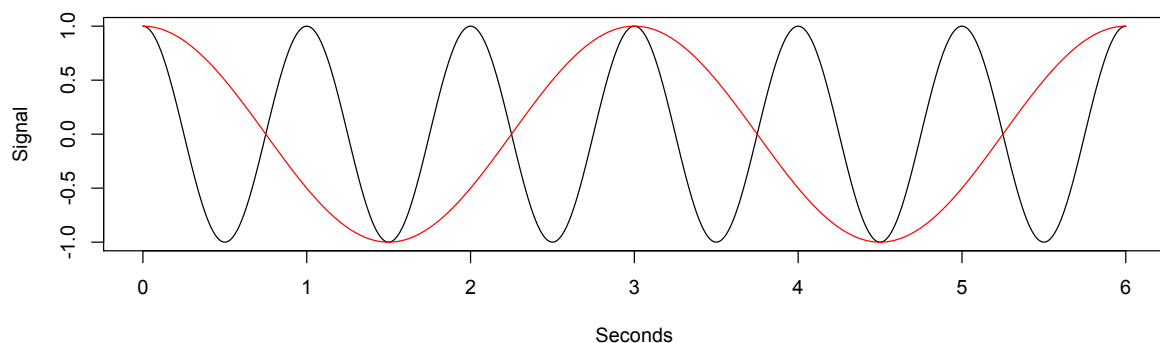
Q2 Limitations of Discrete Sampling (15 Points)

The two related concepts of Nyquist Rate and aliasing play a critical role in understanding one of the most troublesome risks associated with discrete sampling of an analogue signal or even down sampling a discrete signal.

Below is a picture of an analogue signal, a pure cosine wave with a periodicity of one cycle per second. Use this signal to illustrate aliasing and precisely define the Nyquist rate under the presumption that reliable encoding of signals up to one cycle per second is desired.



Here is what happens if samples are spaced at $\frac{3}{4}$ of a cycle (second) intervals.



This is sampling below the Nyquist rate of 2 samples per second (no more than 0.5 seconds between samples). Note on the curve, over 6 seconds there are only 9 samples as opposed to the required 12. Further, 12 is a special case because it can end up sampling straight down the zero crossings, missing the sine wave entirely. Thus, one says sampling must be carried out above (not equal to) the Nyquist rate.

Q3 Comparing Vectors / Images (20 Points)

There is a circumstance fully discussed in lecture in which maximizing correlation is mathematically equivalent to minimizing L2 distance.

First, please precisely define L2 distance between two vectors denoted as:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$L_2 = \sqrt{\sum (v_i - w_i)^2}$$

Now precisely define Pearson's Correlation

$$C = \frac{\sum (v_i - \bar{v}) \sum (w_i - \bar{w})}{\sqrt{\sum (v_i - \bar{v})^2} \sqrt{\sum (w_i - \bar{w})^2}}$$

Finally, state the circumstances under which maximizing correlation is equivalent to minimizing L2 distance and use the formal statement of these circumstances to derive (prove) the equivalence.

First condition, subtract off means

$$\hat{v} = \begin{bmatrix} v_1 - \bar{v} \\ v_2 - \bar{v} \\ v_3 - \bar{v} \end{bmatrix} \quad \hat{w} = \begin{bmatrix} w_1 - \bar{w} \\ w_2 - \bar{w} \\ w_3 - \bar{w} \end{bmatrix}$$

Second condition, make each vector unit length

$$\hat{v} = \frac{\hat{v}}{\|\hat{v}\|} \quad \hat{w} = \frac{\hat{w}}{\|\hat{w}\|}$$

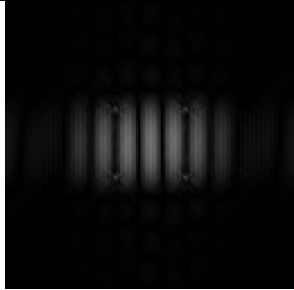
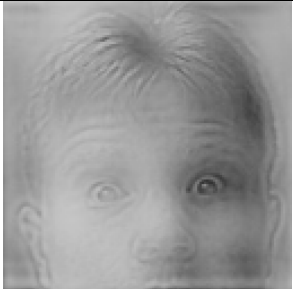
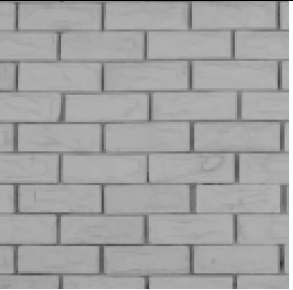
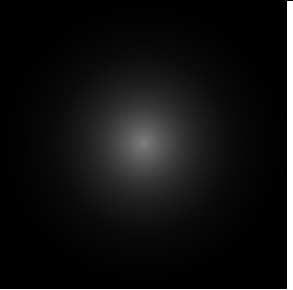
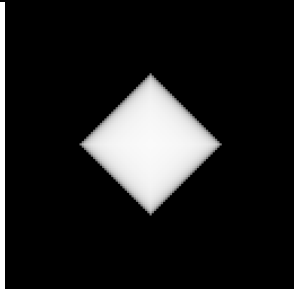
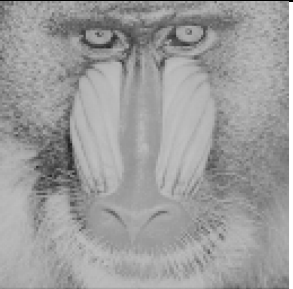
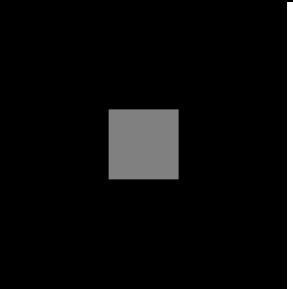
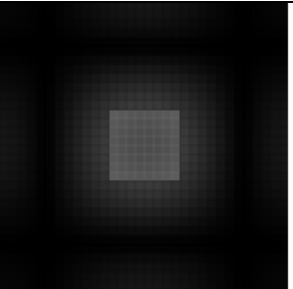
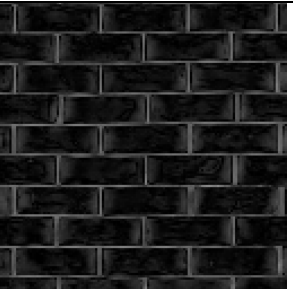
Now, start with L2 and manipulate

$$L_2 = \sqrt{\sum (\hat{v}_i - \hat{w}_i)^2} = \sqrt{\sum (\hat{v}_i)^2 + \sum (\hat{w}_i)^2 - 2 \sum \hat{v}_i \hat{w}_i} = \sqrt{1 + 1 - 2 \sum \hat{v}_i \hat{w}_i}$$

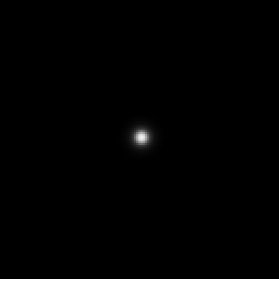
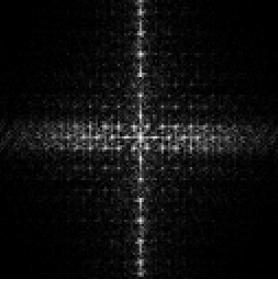
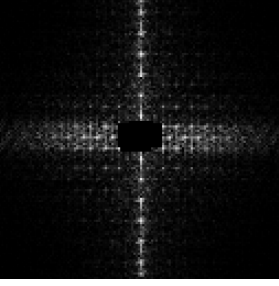
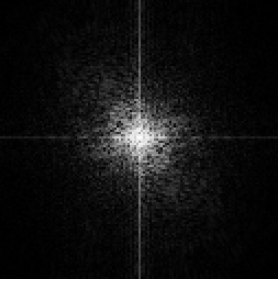
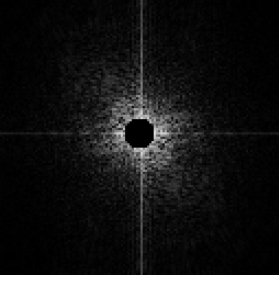
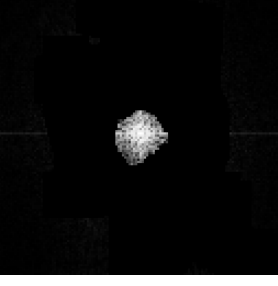
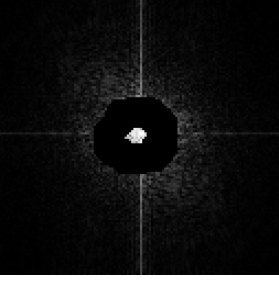
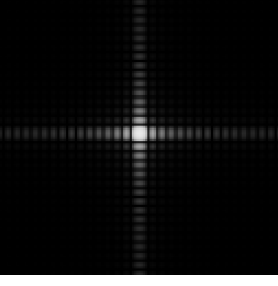
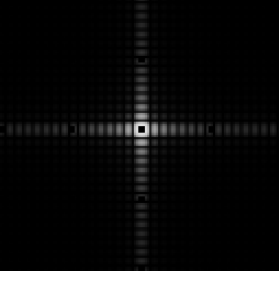
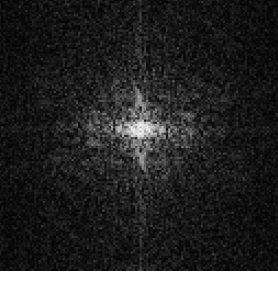
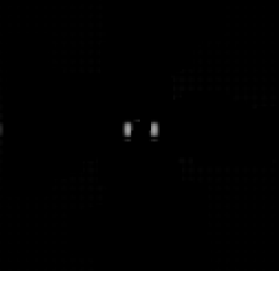
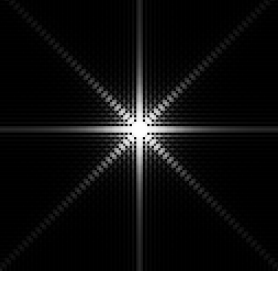
The sum of the squares is just one because of step two and correlation is the dot product because of both steps, so maximizing L2 above is equivalent to minimizing the dot product/correlation.

Q4 The 2-D Fourier Transform (20 Points)

Recall the Fourier Transform Applet on the BrainFlux website by Jason Gallicchio. I have used this tool to create 12 pairs – an image at its corresponding Fourier transform. Images are grayscale and the magnitude of the Fourier transform is shown as a grayscale image. The images are lettered and the transforms are numbered. At the end of this question is a table where you must write the number of the corresponding transform under the letter of the corresponding image.

			
A	B	C	D
			
E	F	G	H
			
I	J	K	L

Q5 The 2-D Fourier Transform (continued)

			
1	2	3	4
			
5	6	7	8
			
9	10	11	12

Write your answer in the table below.

A	B	C	D	E	F	G	H	I	J	K	L
11	7	2	1	12	4	10	8	5	6	9	3

Q5 Geometric Transformations (10 Points)

The following linear algebraic equation is taken directly from the CS 510 course lecture notes:

$$\begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -y_1v_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1v_1 & -y_1v_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2u_2 & -y_2v_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_2 & -y_2v_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3u_3 & -y_3v_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3v_3 & -y_3v_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4u_4 & -y_4v_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4v_4 & -y_4v_4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix}$$

What is the vector on the left hand side of the equation?

These are coordinates of four points after application of a 2-D projective transform.

What is the vector on the right hand side of the equation?

They are the 8 parameters in an 8-DOF projective transform defined as

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$u = u'/w, \quad v = v'/w$$

What do the terms in the matrix signify?

They signify the coordinates of four points prior to application of the 2-D projective transformation as well as the u, v coordinates after the transformation in such a way that it will be possible to solve for the parameters of the transformation (see next question).

How is this equation, which simplifying is in the form $A = MZ$ ultimately put to good use? In other words, with it, what can be accomplished?

By solving for, $Z = M^{-1}A$ we can directly determine the transformation mapping the four points in x,y space to their corresponding locations in u,v space.

Q6 Correlation Space (10 Points)

Consider the following image

6	2	0	6
6	2	4	2
2	6	8	4

Using the image to demonstrate, show the three steps needed to normalize this image so that it may be considered to reside in correlation space.

Step 1	Step 2	Step 3
Unroll the image 6 2 0 6 6 2 4 2 2 6 8 4	Subtract the mean, Which is $48 / 12 = 4$ 2 -2 -4 2 2 -2 0 -2 -2 2 4 0	Divide through by length, yielding a unit length vector. Length is square root of $8*2*2+2*4*4=32+32=64$, or 8 1 8 2 -2 -4 2 2 -2 0 -2 -2 2 4 0

Q7 Basics of Principal Component Analysis (10 Points)

Consider for the sake of illustration an extremely simple problem involving three data points defined in 2-D.

$$p_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad p_2 = \begin{bmatrix} 9 \\ 3 \end{bmatrix} \quad p_3 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

Illustrate the process of constructing the covariance matrix from which the first and second principal components may then be determined. To guide your answer, follow these steps.

Create the data matrix. Since this problem is so simple, avoid normalizing the individual data points (vectors) in any way. However, do show the data matrix before and after it is 'centered'.

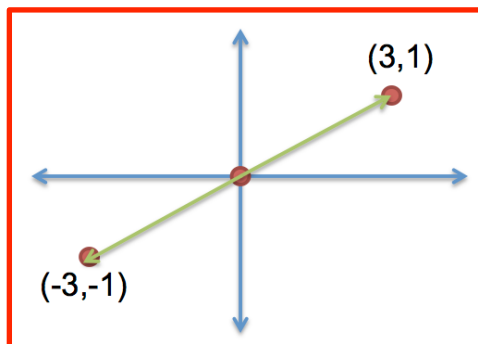
$$X = \begin{bmatrix} 3 & 9 & 6 \\ 1 & 3 & 2 \end{bmatrix}$$

$$\hat{X} = \begin{bmatrix} -3 & 3 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

Compute the covariance matrix from the centered data matrix

$$\Omega = \begin{bmatrix} -3 & 3 & 0 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -3 & -1 \\ 3 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 18 & 6 \\ 6 & 2 \end{bmatrix}$$

Draw a simple sketch of the three data points. See if you can also sketch in the first principal component simply by inspection.



This can be done by inspection because it is such a simple case. The three points lie on a line, with the middle point at the origin after centering. The principal component is an axis defined by the vector from the origin to the point (3,1) [or alternatively to the point (-3,-1), they are the same axis.]