Academic integrity is the foundation of an academic community and without it none of the educational or research goals of the university can be achieved. Academic integrity applies to research as well as undergraduate and graduate coursework/exams. Existing policies forbid cheating on examinations, plagiarism and other forms of academic dishonesty. UC Merced students are held to high standards of personal and professional conduct in compliance with the UC Merced Academic Honesty Policy and the UCM Code of Student Conduct.

UCM Code of Student Conduct can be found here: http://studentconduct.ucmerced.edu

By completing this exam, I acknowledge and confirm that I will not give or receive any unauthorized assistance on this examination. It is a close-book exam and I will not look up any lecture notes or textbook for answering questions. I will conduct myself within the guidelines of the university academic integrity guidelines.

You must sign this form before taking the exam. You will not receive any credit if your signature (handwritten or digital) is not on this paper.

Name: ____David Wang________________________________

Signature: David Wang

Note: 20 questions on both sides, maximum 100 points. It is a close book exam. Explain all your answers to earn full credits for each question (including multiple-choice questions). Notice that for each multiple-choice question, you can answer with more than one choices. For example, you can circle d, e for question 5 if you think both d and e are correct (i.e., circle all the correct ones and explain your answers).

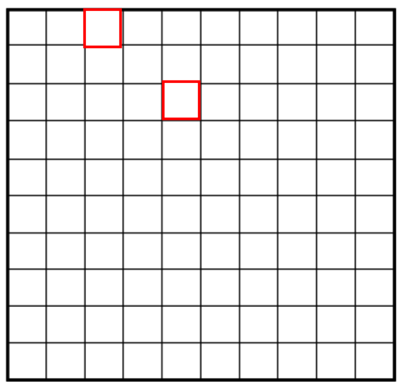
1. [5 points] Given an image $f[\cdot,\cdot]$, calculate the filtered output $h[\cdot,\cdot]$ at the red marked pixels by applying the filter $g[.,.] = \frac{1}{5}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ -------------→ h(1,3) = __, h(3,5) = __;

   Note: assume zero-padding at boundaries

$f[\cdot,\cdot]$

0	0	0	0	0	0	0	0	0	0
0	80	80	80	80	80	80	80	80	0
90	90	90	200	90	90	90	90	90	90
0	70	70	70	70	70	70	70	70	0
60	60	60	60	60	60	60	60	60	60
0	50	50	50	50	50	50	50	50	0
40	40	40	40	40	40	40	40	40	40
0	30	30	30	30	30	30	30	30	0
20	20	20	20	20	20	20	20	20	20
0	10	10	10	10	10	10	10	10	10

$h[\cdot,\cdot]$



Matrix A

$\begin{vmatrix} - & - & - \\ 0 & 0 & 0 \\ 80 & 80 & 80 \end{vmatrix}$

filter $g[]$

$1/5 * \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{vmatrix}$

3x3 Matrix centered at 1,3

Multiplying Matrix A by the filter $g[]$ gives

$\begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -16 & 0 \end{vmatrix}$

Taking the sum of the above matrix gives the result -16. Thus, h(1,3) = -16

Matrix A

$\begin{vmatrix} 80 & 80 & 80 \\ 200 & 90 & 90 \\ 70 & 70 & 70 \end{vmatrix}$

filter $g[]$

$1/5 * \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{vmatrix}$

3x3 Matrix centered at 3,5

Multiplying Matrix A by the filter $g[]$ gives

$\begin{vmatrix} 0 & 16 & 0 \\ 40 & 0 & -18 \\ 0 & -14 & 0 \end{vmatrix}$

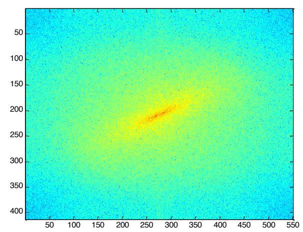
Taking the sum of the above matrix gives the result 24. Thus, h(3,5) = 24

2. [5 points] Match the spatial domain image to the Fourier magnitude image. Explain your answers.

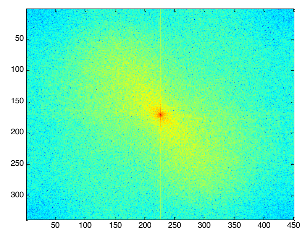
A

B

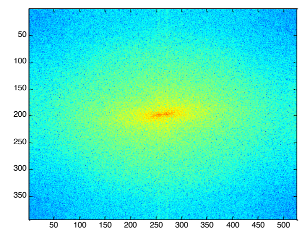
C



/



0



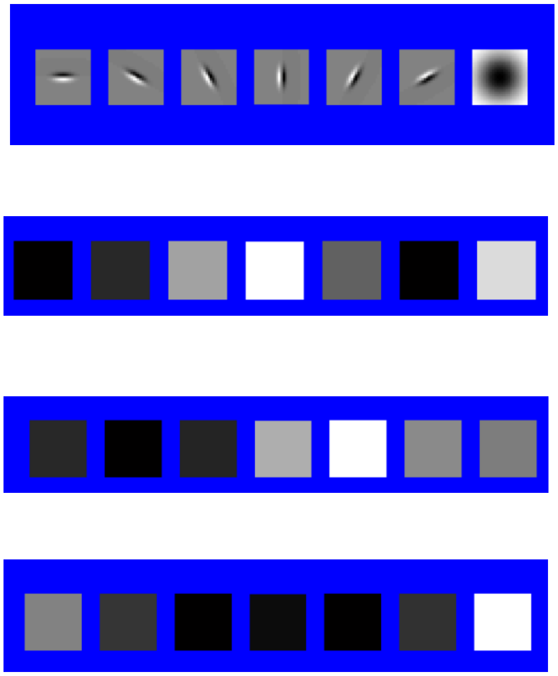
1

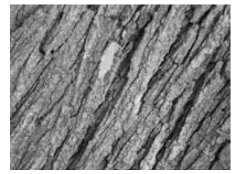
/, Image A matches the most with the 1st image since A has a pattern of lines going diagonally from left to right, while image 1 has the reflection of that pattern. Thus, A matches with 1

2. Image B matches the most with the pattern in image 3, since image B has a pattern with vertical lines, corresponding most to the horizontal line in 3. Thus, B matches with 3

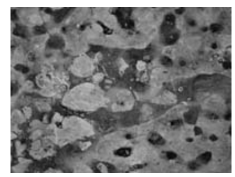
3. Image C matches the most with the pattern in 2, since image C has a pattern with diagonal lines going from bottom left to top right, with mirrors the pattern in 2 going diagonally from top right to bottom left. Thus, C matches with 2

3. [5 points] Match the texture images to the mean absolute responses of filter bank. Explain your answers.



A. 

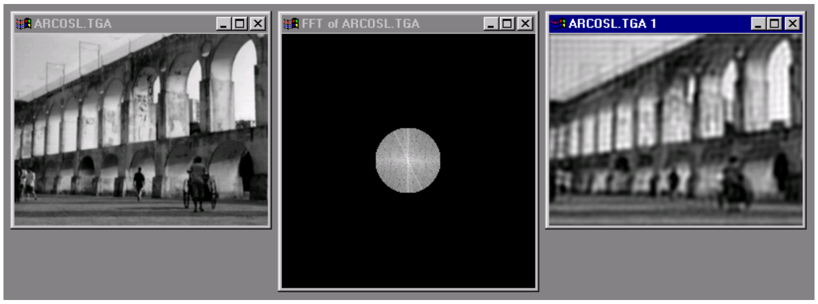
B. 

C. 

1. **For number one, the 4th box is the white box which corresponds to the filter bank's 4th pattern. The 4th box of the filter bank features a vertical pattern. Out of all of the images A-C, this vertical pattern matches most with the second image, B. The vertical black and white patterns match the most with the filter bank's 4th box. Thus, number 1 matches with B.**
2. **For number two, the 5th box is the white box. This corresponds with the filter bank's 5th pattern, the clockwise slanted lines. These lines go in the same direction as the first texture image, A. Thus number 2 matches with image A.**
3. **For number three, the last box is the white box. This corresponds to the last filter bank pattern, the circular pattern. Out of all of the images this matches the most with C, the spotted circular pattern. Thus number 3 matches with image C.**

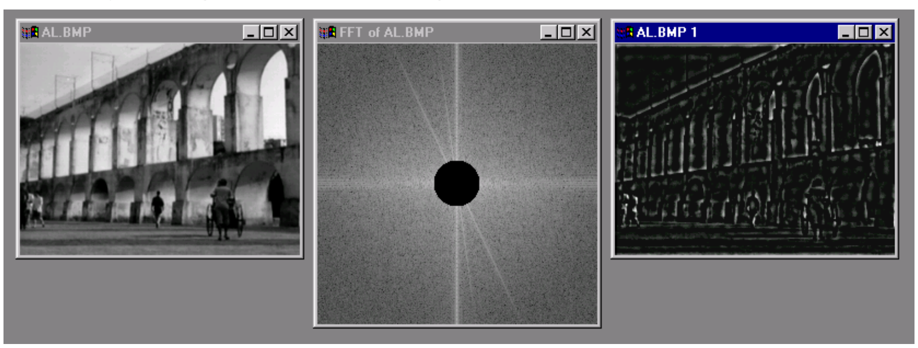
4. [5 points] Explain the type of filter we use on the Fourier domain and why we obtain the resulting images.

   A. Left: Input, Middle: filter. Right: result



The filter used on the Fourier domain is a low-pass filter. The essential structure is kept, resulting in the blurred image.

   B. Left: input image. Middle: filter. Right: Result



The filter used on this Fourier domain is a high-pass filter since the filter took the high-frequency part of the image, resulting in the white lines shown in the resulting image. The lines represent the important edges which defines the image.

5. [5 points] Which of the following statements regarding Canny edge detector are true? Explain your answers.
    a. The non-maximum suppression is used to select pixels that are close to true edges
    b. The edges found by a Canny edge detector are determined by the Gaussian kernel scale
    c. In hysteresis process, we start with low thresholds and then high thresholds
    d. Canny edge detector is a low-pass filter
    e. Hysteresis process is used to connect short edges

Options a and b are true.

Option c is false, since with the hysteresis process we start with *high* thresholds then *low* thresholds; the order is wrong. Option d is false, since edge detectors are high pass filters

Option e is true.

Thus options a, b, and e are true

6. [5 points] Which of the following statements regarding stereo are true? Explain your answers.
   a. Given a pair of images from a calibrated stereo camera, for each pixel in one image, we can use the fundamental matrix to compute the corresponding epipolar line in the other image
   b. Give a pair of images from an uncalibrated stereo camera, for each pixel in one image, we can use the essential matrix to compute the corresponding epipolar line in the other image.
   c. All epipolar lines intersect at the epipole
   d. Epipolar lines are always horizontal lines on an image
   e. When we use larger window for search correspondence, we can capture more details but more noise

A is false since for calibrated stereo cameras, we need to use *essential* matrices, not fundamental matrices, to compute the corresponding epipolar lines.

B is false since essential matrices are used for *calibrated*, not uncalibrated stereo cameras.

D is false, since epipolar lines are not always horizontal lines on an image. It is true only when image planes have been rectified

E is false since *smaller windows*, not larger windows, are used to capture more details

Thus the only correct option is C

7. [5 points] Which of the following statements are true? Explain your answers.
    a. Given one point in one image, the corresponding point in the second image of a stereo par is on a line passing through its epipole
    b. We can use the essential matrix to map a point in one image to a line in the other image.
    c. Depth is inversely proportional to disparity
    d. a and b
    e. a, b and c

Option b is true since essential matrices by definition allow you to map a point in one image to a line in another. Option c is true since in the equation for depth, disparity is in the denominator giving depth and disparity an inversely proportional relationship where if one increases, the other decreases. Option d is false by process of elimination since C is true. Option a is also true by definition.

Thus, the best answer would be e: a, b and c are true.

8. [5 points] Which of the following statements regarding optical flow are correct? Explain your answers.
    a. Assume brightness constancy: the matched pixels between two frames have very similar intensity value
    b. Can handle only small motion: assume points do not move far
    c. Assume spatial coherence: points move like their neighbors
    d. Small window size is robust to noise and occlusion
    e. Use weighted window (e.g., Gaussian) so that center matters more

Option b is false, since multi-scale algorithms can handle large motion. Option d is false, since *large windows* are robust for noise and occlusion. Options A and C are true by definition,

Thus options a, c, and e are true.

9. [5 points] Which of the following statements are true when we use RANSAC to fit data points with an objective function? Explain your answers.
    a. RANSAC is robust to outliers
    b. Applicable for larger number of objective function
    c. Optimization parameters are easier to choose than Hough transform
    d. Computational time grows quickly with fraction of outliers
    e. Not good for getting multiple fits

Running through each of the options, none appear to be false. A through E are true based off of definitions of how RANSAC works.

A is true since Ransac works well with outliers. B is true since Ransac can be used with more objecttive function parameters than Hough transform, and C is true since the parameters are also easier to choose than Hough transform. By definition, Ransac's computational time increases faster with more outliers and parameters so option D is true. E is true since Ransac does not perform as well with multiple fits.

Thus a, b, c, d, and e are all true.

10. [5 points] K-means clustering. Which of the following statements are true? Explain your answers.
    a. K-means algorithm is sensitive to outliers
    b. If we run K-means many times, we may get different clustering results.
    c. The performance of K-means clustering depends heavily on the number of clusters
    d. K-means is useful for segmentation, while agglomerative clustering is useful for summarization
    e. K-means is a supervised learning method

Option d is false, since agglomerative clustering is useful for *segmentation* and K-means is useful for *summarization*. Option e is false. K means is *unsupervised* because points have no external classification. Options A - C are true by definition

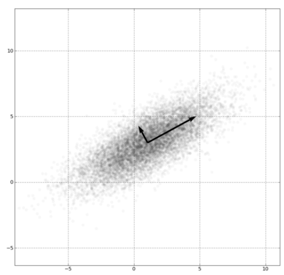
Thus: a, b, and c are true

11. [5 points] Which of the following statements regarding eigenfaces are correct? Explain your answers.
    a. The directions that capture the maximum covariance of the data is the eigenvector corresponding to the smallest eigenvalue of the data covariance matrix
    b. If all the eigenvalues are of the same magnitude, then we do not need to use many eigenvectors.
    c. One of the problems with the eigenface approach is that it is sensitive to alignment errors.
    d. A face image in the dataset can be represented as a linear combination of eigenfaces.
    e. The eigenfaces can be extracted by the principal component analysis (PCA) method.

Option a is false, since the directions that capture the maximum covariance corresponds to the *largest* eigenvalue of the data covariance matrix. Option b is false, since we need to use all eigenvectors when the eigenvalues are of the same magnitude. Options c, d, and e are true by definition

Thus: options c, d, and e are true

12. [5 points] The following figure shows the data destruction and their principal components. Which of the following statements regarding the data points are true? Explain your answers.
    a. The first principal component is along
    b. The first principal component is along
    c. PCA is less useful for data points that are spread out uniformly in a circle.
    d. The direction of the maximal variance is always good for classification.
    e. The principal components are known as the eigenvectors.



Option a is false since the first principle component is along the direction listed in option b. Thus, option b is also true by definition. Options, c and e are true by definition. Option d is false, since the direction of maximal variance is *not* always good for classification in some cases.

Thus: options b, c, and e are true

13. [5 points] Optical flow. Let I(x,y,t) be one pixel value at (x,y) of frame t, and I(x,y,t+1) be a pixel value at (x,y) of frame t+1. With Taylor expansion $(x + \Delta x) = f(x) + \frac{\partial f}{\partial x}\Delta x$, derive the optical flow equation $I_x u + I_y v + I_t = 0$ where I_x, I_y, I_t, are first order derivative of I with respect to x, y, and t coordinate, and u, v are motion vector (displacement).

$I(x, y, t) = I(x + u, y + v, t + 1)$ **Brightness consistency**

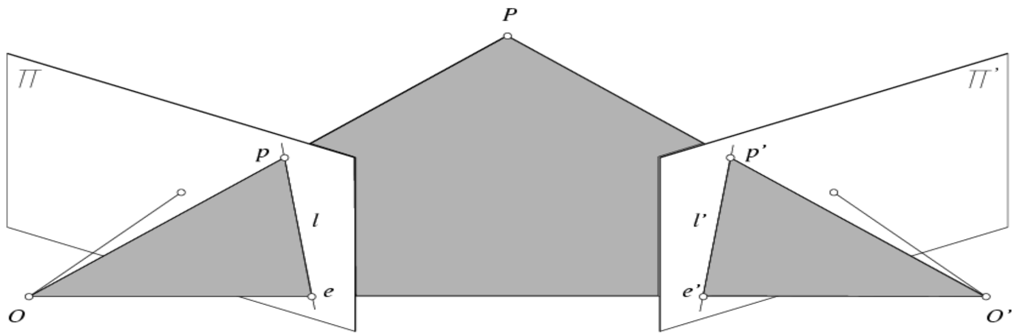
$I(x + u, y + v, t + 1) = I(x, y, t) + I_x u + I_y v + I_y v + I_t$ **Taylor Expansion**

$I(x + u, y + v, t + 1) - I(x, y, t) = I_x u + I_y v + I_y v + I_t$

Since we know that $I(x, y, t) = I(x + u, y + v, t + 1)$, subtracting $I(x, y, t)$ from the right side of the equation gives

$\mathbf{0 = I_x u + I_y v + I_y v + I_t}$

14. [5 points] Epipolar geometry. Given a point P in the world coordinate with two mapped points p and p’ on two image planes with two optical centers O and O’. Derive the following equations.



$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$

Explain all steps (what does the cross product of two vectors do and what does the inner product of two vectors do?) to get full credit.

Points O, O', and P form the plane OO'p.

Op, OO' and O'p' represent different vectors on the surface of the plane.

Taking the cross product of vectors OO' and O'p' gives a surface perpendicular to the plane OO'p, which we can call N.

Since vector Op is perpendicular to N, or OO' x O'p', taking the inner product of vector Op and the cross product N results in the final answer of 0.

15. Extra credit [10 points] Let x=[2 5], y=[1 2] and z=[3 2], compute the covariance matrix.

Given X = | 2 5 |, y = |1 2| and z = |3 2 |, we can define a matrix 2x3 S plugging in the X Y Z values and matrix *U* by adding those values and then dividing by 3

S = | 2 1 3 | *U* = | (2 + 1 + 3) / 3 | = | 3 |
| 5 2 2 | | (5 + 2 + 2) / 3 | | 3 |

S' = S - U = | -1 -2 0 |
| 2 -1 -1 |

Calculating the covariance matrix,

C = S' * S'T = | -1 -2 0 | * | -1 2 |
| 2 -1 -1 | | -2 -1 |
| 0 -1 |

C =
| 5 0 |
| 0 6 |

16. [5 points] Please **debug** the following optical flow code and write down the correct code in each blank line. Suppose ‘frame01.jpg’ and ‘frame02.jpg’ are gray-scale images, and we only compute the optical flow for one time. If there’s nothing wrong, please write down “blank”.

```
1. I1 = imread('frame01.jpg');
2. I2 = imread('frame02.jpg');
3. window_size = 45; r = floor(window_size/2);
5. [h, w] = size(I1);
6. Ix_m = imfilter(I1, [1, -1], 'replicate');
7. Iy_m = imfilter(I2, [1; -1], 'replicate');
8. u = zeros(h, w); v = zeros(h, w);
9. for i = 1 : h-1 -----------------------→   for i = 1+r : h-r
10.     for j = 0 : w ---------------------→   for j = 1+r : w-r
11.         Ix = Ix_m(i-r:i+r,j-1:j+1);
12.         Iy = Iy_m(i-r:i+r,j-1:j+1);
13.         It = I1(i-r:i+r,j-r:j+r)-I2(i-r:i+r,j-r:j+r);
14.         A = [Ix, Iy]; -----------------→   A = [Ix(:),Iy(:)];
15.         b = It; -----------------------→   b = -It(:);
16.         x = A / b; --------------------→   x - A \ b;
17.         u(i, j) = x(1);
18.         v(i, j) = x(2);
19.     end
20. end
```

9. correction:

for i = 1+r : h-r

10. correction:

for j = 1+r : w-r

14. correction:

A = [Ix(:),Iy(:)];

15. correction:

b = -It(:);

16. correction:

x - A \ b;

17. [5 points] Please **explain** the corresponding operations of the following MATLAB code to compute eigenfaces. Assume there is a variable “face_training” with dimension h x w x num_train and k is the number of eigenfaces.

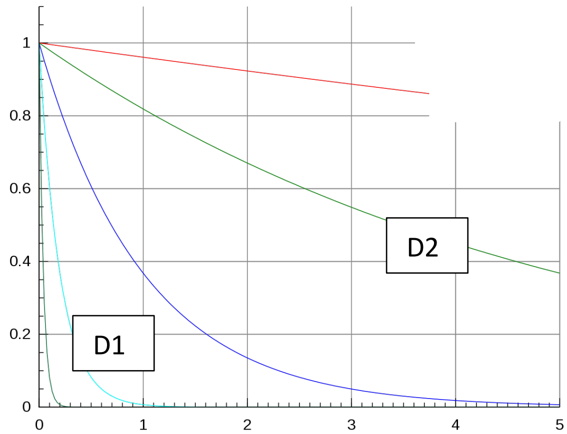
```
1. X = reshape(face_training, h*w, num_train);
2. mu = mean(X, 2); -----------→
3. X = bsxfun(@minus, X, mu); -→
4. C = X * X'; ----------------→
5. [U, S, D] = svd(C); --------→
6. U = U(:, 1:k); -------------→
```

2. Calls the function mean() to compute mean face and mean of the columns in X
3. Calls bsxfun() with inputs @minus, X, mu to calculate
X - mu
4. Computes the covariance matrix by multiplying matrix X and matrix X'
5. Calculates the singular value decomposition by calling svd() with input of C, the covariance matrix
6. Calls U() with 1:k to take in the first column of U as eigenfaces

18. [5 points] Which of the following statements regarding agglomerative clustering are true? Explain your answers.
    a. We will have the same clustering results even when we use different metric functions
    b. Some clusters may have more points than other clusters
    c. We can easily generate a tree based on the hierarchical clustering results
    d. Unlike K-means, we do not need to select the number of clusters
    e. a, b, d are true

Option a is false, since clustering results change when using different metric functions. Option b is true by definition, since in agglomerative clustering pairs of clusters are merged in pairs and may result in some clusters having more points than others. Option c is also true since hierarchical clustering looks for nodes to find clusters closest and merge with similar clusters, forming a tree. Option d is false, since we do need to select the number of clusters for agglomerative clustering. Option e is false by process of elimination, since a is false and d is false.
Thus, options b and c are true.

19. [5 points] Given two image datasets, D1, and D2, where both datasets have the same number of data points and each is represented by a vector of grayscale pixel values, we can compute the corresponding eigenvectors/eigenvalues from two covariance matrices. For each dataset, we plot the eigenvalues in the descending order. The eigenvalues with respect to D1 drop faster than D2 as shown in the following figure (ignore all the other curves).



Which of the following statements are true? Explain your answers.

a. The images in D1 are more similar than those in D2
b. The eigenvectors we need to represent images in D1 are much smaller than those for D2
c. When we project the data points in each dataset onto the subspace spanned by the top two eigenvectors, the scatter plot of the projected points for D2 is much smaller than that for D1
d. It takes more time to compute the eigenvectors for D1 than D2
e. a, b, d are correct

Option a is correct, since eigenvectors represent total covariance. Based on the graph, D1 drops steeper than D2 so it means that D1 has less covariance and therefore more similar images. Option B is correct, since D1 has less covariance it also means that the eigenvectors are smaller in D1 than in D2's higher covariance.

Thus, options a and b are correct

20. [5 points] Which of the following statements regarding eigen-representation (e.g., eigenfaces) are correct? Explain your answers.
    a. Eigenfaces describes holistic (global) appearance variation of faces
    b. Eigenfaces are sensitive to alignment error (e.g., eigenfaces will not perform well when the face images are not aligned well)
    c. Eigenfaces are not sensitive to heavy occlusion (e.g., eigenfaces still perform well when 20% pixels of a face image are occluded)
    d. The first few eigenfaces corresponding to the largest few eigenvalues describe low-frequency components of faces.
    e. a b and d are correct

Option a is correct since eigenfaces compare images of faces in a holistic, parts-based method. Option b is correct since eigenfaces do not work as well when face images are not the same size and are not aligned. Option c is false, since eigenfaces are sensitive and affected by heavy occlusion. Option d is true, since larger images tend to have lower frequency in face images. Option e is true, since options a, b, and d are true.

Thus, options a b and d are correct, and option e is correct.