

**COMP 4900D: Midterm Exam - Solutions**  
Feb. 16, 2006

Student Name: \_\_\_\_\_

Student Number: \_\_\_\_\_

**Instructions:** This is a closed book exam. No calculators are allowed. All questions should be answered on this sheet in the space provided. No electronic devices, phones, etc. are allowed on the desk. There are 17 problems and 25 points in total. You have 80 minutes.

1. (1 point) Let  $A = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix}$ .

- (a) What can you say about the columns of  $A$ ?  
(b) Find the inverse of  $A$ .

(a) They are orthogonal to each other.

(b)  $A^{-1} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$

2. (1 point) What transformation matrix  $A$  takes

$$x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ to } Ax_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ to } Ax_2 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}$$

Answer:  $\begin{bmatrix} 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{bmatrix}$

3. (1 point) The least square solution to  $Ax = b$  minimizes what error function?

Answer:  $\|Ax - b\|^2$

4. (1 point) What is the advantage of having lenses?

Answer: to gather more light.

5. (2 points) For a pin-hole camera with focal length  $f$ , the relationship between a 3D point  $P = [X, Y, Z]^T$  and its image point  $p = [x, y]^T$  is

$$x = f \frac{X}{Z} \quad y = f \frac{Y}{Z}$$

Use homogeneous coordinate to formulate this relationship in a linear form.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x = u/w \quad y = v/w$$

6. (3 points) Suppose a camera has focal length  $f$ , pixel size  $s_x$  and  $s_y$  in  $x$  and  $y$  directions, respectively, and the principal point at  $(o_x, o_y)$ .

- (a) Write down the camera calibration matrix.
- (b) If a point in the world coordinate frame  $P_w = [X_w, Y_w, Z_w]^T$  and camera coordinate frame  $P_c = [X_c, Y_c, Z_c]^T$  can be related by  $P_c = RP_w + T$ , where  $R$  is a rotation matrix and  $T$  is a translation vector, write down the camera projection matrix.
- (c) What are the intrinsic and extrinsic parameters?

$$(a) K = \begin{bmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$(b) M = K[R \ T]$$

(c) intrinsic parameters:  $f$ ,  $s_x$ ,  $s_y$ , and  $(o_x, o_y)$ .

extrinsic parameters:  $R$ ,  $T$

7. (1 point) Describe two characters of salt-and-pepper noise.

- (a) alters random pixels;
- (b) makes their values very different from the true ones.

8. (1 point) Explain the reason that the elements of the convolution kernels for smoothing usually sum to 1.

Answer:

Otherwise, the result image would be either darker or lighter than the original one.

9. (2 point) The discrete Gaussian kernel is defined as  $G(h,k) = e^{-\left(\frac{h^2+k^2}{2\sigma^2}\right)}$  (ignoring the normalization factor) and convolving an  $m \times m$  Gaussian kernel at a pixel is defined by

$$I_G = I * G = \sum_{h=-m/2}^{m/2} \sum_{k=-m/2}^{m/2} G(h,k) I(i-h, j-k)$$

Show that Gaussian filter is separable.

$$\text{Since } e^{-\frac{h^2+k^2}{2\sigma^2}} = e^{-\frac{h^2}{2\sigma^2}} e^{-\frac{k^2}{2\sigma^2}},$$

$$\begin{aligned} I_G &= I * G = \\ &= \sum_{h=-m/2}^{m/2} \sum_{k=-m/2}^{m/2} G(h,k) I(i-h, j-k) = \\ &= \sum_{h=-m/2}^{m/2} \sum_{k=-m/2}^{m/2} e^{-\frac{h^2+k^2}{2\sigma^2}} I(i-h, j-k) = \\ &= \sum_{h=-m/2}^{m/2} e^{-\frac{h^2}{2\sigma^2}} \sum_{k=-m/2}^{m/2} e^{-\frac{k^2}{2\sigma^2}} I(i-h, j-k) \end{aligned}$$

10. (1 point)

- (a) If we try to smooth a  $200 \times 100$  image with a Gaussian kernel of 5 elements, how many multiplications do we need to do?
  - (b) If we want to smooth the same image with a  $5 \times 5$  average kernel, how many multiplications are necessary?
- (a) Since Gaussian kernel is separable, and for each pixel the row and column kernel each need 5 multiplications, the total number of multiplications is  
 $2 \times 5 \times 200 \times 100 = 200,000$
- (b)  $5 \times 5 \times 200 \times 100 = 500,000$

11. (1 point) Suppose we apply a 3x3 median filter to the pixel at the centre. What is the value of the pixel after applying the median filter?

123	125	126	130	140
122	124	126	127	135
118	120	150	125	134
119	115	119	123	133
111	116	110	120	130

Answer: 124

12. (1 point) If we approximate the image derivative with  $I_x = \frac{\partial f(x, y)}{\partial x} \approx f_{i+1,j} - f_{i,j}$ , what is the convolution kernel that computes this derivative?

Answer:  $[-1 \quad 1]$

13. (2 points) The Sobel kernel for computing  $I_x$  is  $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$ .

- (a) What is the result at the centre pixel after applying this kernel?
- (b) What is the advantage of the Sobel kernel over a simple 1x2 kernel?

$a_1$	$a_2$	$a_3$
$a_4$	$a_5$	$a_6$
$a_7$	$a_8$	$a_9$

- (a)  $(a_1 - a_3) + 2(a_4 - a_6) + (a_7 - a_9)$
- (b) Sobel kernel smoothes the noise.

14. (2 points) Write down the key steps for an edge detection algorithm.

- (1) Noise smoothing:  $I_s = I * h$   
(e.g.  $h$  is a Gaussian kernel)
- (2) Compute two gradient images  $I_x$  and  $I_y$  by convolving  $I_s$  with gradient kernels (e.g. Sobel operator).
- (3) Estimate the gradient magnitude at each pixel

$$G(i, j) = \sqrt{I_x^2(i, j) + I_y^2(i, j)}$$

- (4) Mark as edges all pixels  $(i, j)$  such that  $G(i, j) > \tau$ .

15. (2 point) The local intensity variation matrix in a neighbourhood of a pixel is given by  $C = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} = Q^T \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q$ , where  $\lambda_1$  and  $\lambda_2$  are eigenvalues of  $C$  and  $Q$  is an orthogonal matrix. What can you say about  $\lambda_1$  and  $\lambda_2$ ? What is the criterion for a pixel to be classified as a corner?
- (a)  $\lambda_1 > 0$  and  $\lambda_2 > 0$ .  
 (b) if  $\lambda_1 > \lambda_2$  and  $\lambda_2 > \tau$ , where  $\tau$  is a threshold, then the pixel can be classified as a corner.
16. (1 point) Fill in the missing step in the following pseudo code for detecting lines using Hough transform.
1. Quantize the parameter space  
 $\text{int } P[0, \rho_{\max}][0, \theta_{\max}]; // \text{ accumulators}$
  2. For each edge point  $(x, y)$  {  
 For  $(\theta = 0; \theta \leq \theta_{\max}; \theta = \theta + \Delta\theta)$  {  
 $\rho = x \cos \theta + y \sin \theta; // \text{ round off to integer}$   

$P[\rho][\theta]++;$

  
 $}$   
 $}$   
 3. Find the peaks in  $P[\rho][\theta]$ .

17. (2 points) The Euclidean distance from a point  $\mathbf{p}$  to an ellipse  $f(\mathbf{p}, \mathbf{a}) = 0$  is defined by the solution of the following constrained optimization problem:

$$\min_{\hat{\mathbf{p}}} \|\hat{\mathbf{p}} - \mathbf{p}\|^2 \text{ subject to } f(\hat{\mathbf{p}}, \mathbf{a}) = 0$$

- (a) Suppose we want to solve this problem using the technique of Lagrange multiplier, write down the function that we want to minimize.
  - (b) If we know that  $\mathbf{p}$  is close to  $\hat{\mathbf{p}}$ , what is the first-order approximation of  $f(\hat{\mathbf{p}}, \mathbf{a})$  in terms of  $f(\mathbf{p}, \mathbf{a})$ ,  $(\hat{\mathbf{p}} - \mathbf{p})$ , and  $\nabla f(\mathbf{p}, \mathbf{a})$ ?
- 
- (a)  $L(x, y, \lambda) = \|\hat{\mathbf{p}} - \mathbf{p}\|^2 - 2\lambda f(\hat{\mathbf{p}}, \mathbf{a})$  where  $\hat{\mathbf{p}} = [x, y]^T$ .
  - (b)  $f(\hat{\mathbf{p}}, \mathbf{a}) \approx f(\mathbf{p}, \mathbf{a}) + (\hat{\mathbf{p}} - \mathbf{p})^T \nabla f(\mathbf{p}, \mathbf{a})$ .