NATIONAL UNIVERSITY OF SINGAPORE

SCHOOL OF COMPUTING

EXAMINATION FOR

Semester 1 AY2010/2011

CS4243
COMPUTER VISION & PATTERN RECOGNITION

November 2010 Time Allowed: 2 Hours

INSTRUCTIONS TO CANDIDATES

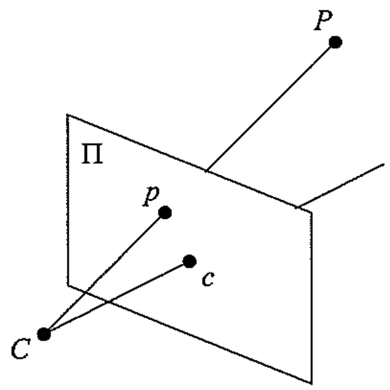
1. This examination paper contains **FIVE (5)** questions and comprises **NINE (9)** printed pages, including this page.
2. Answer **ALL** questions. The maximum mark is 100.
3. Write your answers in the space provided in this booklet. Use the reverse sides if necessary.
4. Write legibly. You may use pen or pencil.
5. This is an OPEN BOOK examination.
6. Please write your Matriculation Number below.

Matriculation No.: ______________________________

This portion is for the examiner's use only.

Question	Marks	Remarks
Q1		
Q2		
Q3		
Q4		
Q5		
Total		

Q1: Perspective Projection (25 marks)



The perspective projection equation that projects a 3D scene point P to an image point p is given by

$$\rho \tilde{\mathbf{x}} = \mathbf{KR}(\mathbf{X} - \mathbf{C}) \tag{1}$$

where $\tilde{\mathbf{x}} = [x, y, 1]^\top$ is the homogeneous coordinate vector of image point p, $\mathbf{X} = [X, Y, Z]^\top$ is the coordinate vector of the 3D scene point P, $\mathbf{K}$ is the camera intrinsic parameter matrix, $\mathbf{R}$ is the rotation matrix that aligns the world frame to the camera frame, $\mathbf{C}$ is the camera center in the world frame, and $\mathbf{c}$ is the image center.

1(a) (5 marks)

From Eq. 1, we can obtain the following equations

$$\rho \mathbf{K}^{-1}\tilde{\mathbf{x}} = \mathbf{R}(\mathbf{X} - \mathbf{C}), \tag{2}$$

$$\rho \mathbf{R}^{-1}\mathbf{K}^{-1}\tilde{\mathbf{x}} = \mathbf{X} - \mathbf{C}. \tag{3}$$

Let $\mathbf{u}$ and $\mathbf{v}$ denote the following: $\mathbf{u} = \mathbf{K}^{-1}\tilde{\mathbf{x}}$ and $\mathbf{v} = \mathbf{R}^{-1}\mathbf{K}^{-1}\tilde{\mathbf{x}}$. Both $\mathbf{u}$ and $\mathbf{v}$ represent the same projection line that projects the 3D scene point P to the image point p. However, their formulae differ by a multiplication of $\mathbf{R}^{-1}$, which usually mean a difference of rotation. But, $\mathbf{u}$ and $\mathbf{v}$ represent the **same** line without rotational difference. So, what does this difference in formulae actually mean?

1(b) (10 marks)

Let d denote the distance from the 3D scene point P to the camera center C. What is the relationship between d and ρ? (Note: Write d in terms of ρ, $\tilde{\mathbf{x}}$, etc.)

1(c) (10 marks)

Consider the case of a special pin-hole camera with unit focal length, i.e., $\mathbf{K} = \mathbf{I}$, $\mathbf{R} = \mathbf{I}$, $\mathbf{C} = \mathbf{0}$. In this case, what is the distance d in terms of $\mathbf{X}$? (Note: Write d in terms of the X, Y, Z.) What is the value of ρ? (Note: Write ρ in terms of X, Y, Z and x, y.)

Q2: Lens Radial Distortion (25 marks)

You join a surveillance company as a specialist in computer vision. In surveillance applications, the images are often distorted by lens radial distortion. You look at a computer vision book and realized that the radial distortion equation is given as follows:

$$\begin{aligned} x' &= x\,(1 + \kappa_1 r^2 + \kappa_2 r^4) && (4) \\ y' &= y\,(1 + \kappa_1 r^2 + \kappa_2 r^4) && (5) \end{aligned}$$

where (x, y) are the coordinates of the undistorted coordinates, (x', y') are the distorted coordinates, $r^2 = x^2 + y^2$, and κ_1 and κ_2 are the radial distortion parameters.

Your task is to estimate κ_1 and κ_2. To solve the problem, you set up an experiment to measure the distorted coordinates (x'_i, y'_i) of a set of image points with known coordinates (x_i, y_i), for $i = 1, \ldots, n$. After the measurement, you set up the following system of linear equations to solve for κ_1 and κ_2:

$$\mathbf{A}\,\mathbf{k} = \mathbf{v} \tag{6}$$

where $\mathbf{A}$ is a matrix, $\mathbf{v}$ is a column vector, and $\mathbf{k}$ is the column vector

$$\mathbf{k} = \begin{bmatrix} \kappa_1 \\ \kappa_2 \end{bmatrix}. \tag{7}$$

2(a) (10 marks)

Write the matrix entries in $\mathbf{v}$.

2(b) (10 marks)

Write the matrix entries in $\mathbf{A}$.

2(c) (5 marks)

What is the value of $\mathbf{k}$? (Note: Write $\mathbf{k}$ in terms of $\mathbf{A}$ and $\mathbf{v}$.)

Q3: Tracking (20 marks)

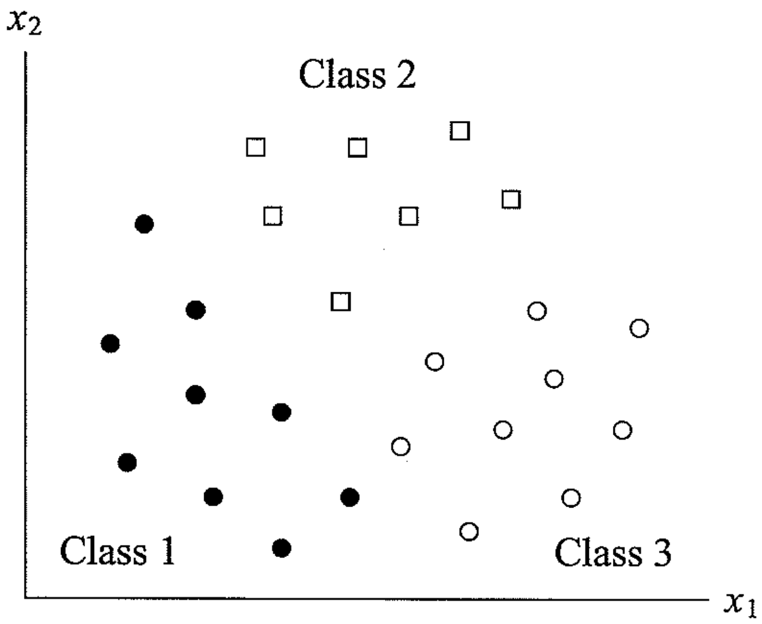
Your surveillance company gives you a second computer vision task to solve. The task is to develop an algorithm that automatically detects people walking in front of a **stationary** surveillance camera, tracks their motion, and plots their exact motion paths in the surveillance video (black curves in the picture below). Each motion path should trace a specific point (e.g., mid-point) of a person's head over time.

Outline the main steps of your algorithm for solving this task. Your algorithm may consist of one or more known algorithms. Your algorithm should differentiate between moving humans and non-human moving objects. It should also handle occlusions. Explain how and why your algorithm works.

(You may use this page for your answer.)

Q4: Pattern Recognition (15 marks)

The following figure illustrates samples of 3 classes in a 2-D space.



Draw the **smallest number** of decision boundaries in the figure above that linearly separate the three classes. Write down the decision rules for classifying the samples into the three classes.

Q5: Multiple View Methods (15 marks)

A 3D scene is captured at n different views. The views are taken such that the coordinates $\mathbf{X}_i$ of a scene point in view i is related to the coordinates $\mathbf{X}_{i-1}$ in view $i-1$ by the rotation matrix $\mathbf{R}_i$ and translation vector $\mathbf{T}_i$ as follows:

$$\mathbf{X}_i = \mathbf{R}_i \mathbf{X}_{i-1} + \mathbf{T}_i.$$

What is the relationship between $\mathbf{X}_n$ and $\mathbf{X}_0$? (Note: Write the mathematical relationship in terms of $\mathbf{R}_i$ and $\mathbf{T}_i$, $i = 1, \ldots, n$.)

— END OF PAPER —