# THE UNIVERSITY OF AUCKLAND

**FIRST SEMESTER, 2012**
**Campus: City**

**COMPUTER SCIENCE**

**COMPSCI 773: Intelligent Vision Systems**

**(Time allowed: TWO hours)**

**NOTE:** Attempt **all** questions!

Write the answers in the boxes below the questions.

Marks for each question are shown just before each answer box.

This is an open book exam. Candidates **may** bring calculators, notes, reference books, or other written material into the examination room.

| *Section:* | **A** | **B** | **C** | Total |
|---|---|---|---|---|
| *Possible marks:* | 30 | 50 | 20 | 100 |
| *Awarded marks:* | | | | |

SURNAME:

FORENAME(S):

STUDENT ID:

QUESTION/ANSWER SHEETS FOLLOW

Student ID: ________________________________________

# Section A: Camera calibration and stereo vision

## Calibration

1. Consider a 3 by 3 calibration matrix with $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ colum vectors. In both the Tsai and Zhang's calibration, $\mathbf{r}_1$ and $\mathbf{r}_2$ are obtained first. Explain why and how $\mathbf{r}_3$ can be inferred from $\mathbf{r}_1$ and $\mathbf{r}_2$? [2 marks]

2. **Application** of Question 1: Compute $\mathbf{r}_3$ for $\mathbf{r}_1 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$ and $\mathbf{r}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. [1 mark]

3. Let $\mathbf{P}_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$ and $\mathbf{P}_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$ denote Cartesian coordinates of a 3D point in the world reference frame and the camera reference frame, respectively. Consider the rigid body transformation from the world reference frame to the 3D camera coordinate specified by the $3 \times 3$ rotation matrix R and the $3 \times 1$ translation vector $\mathbf{T}$. Write the $4 \times 4$ matrix $\widetilde{M}$ which relates $\widetilde{\mathbf{P}}_c$ to $\widetilde{\mathbf{P}}_w$ in homogeneous coordinates, $\widetilde{\mathbf{P}}_c = \widetilde{M}\widetilde{\mathbf{P}}_w$ as a function of the rotation matrix $R$ and the translation vector $\mathbf{T}$. [2 marks]

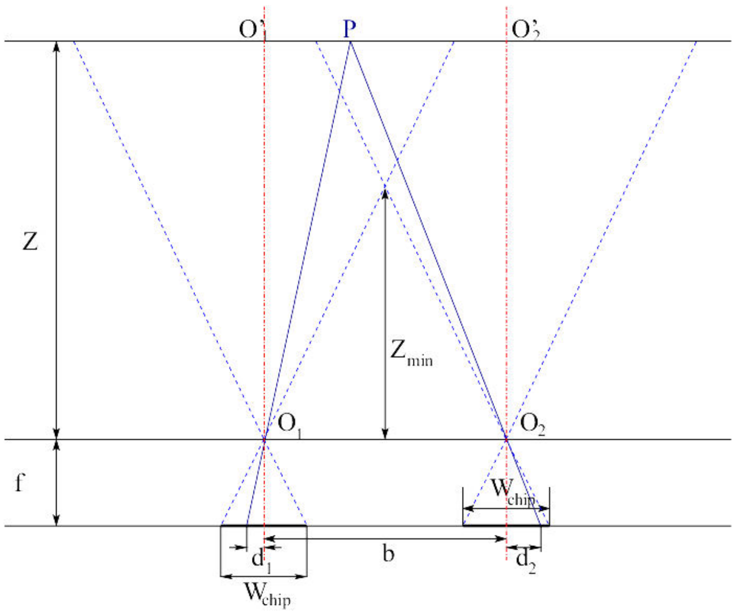
Student ID: ________________________________________

4. Conversely, write the $4 \times 4$ matrix $\widetilde{M'}$ which relates $\widetilde{\mathbf{P}}_w$ to $\widetilde{\mathbf{P}}_c$ in homogeneous coordinates, $\widetilde{\mathbf{P}}_w = \widetilde{M'}\widetilde{\mathbf{P}}_c$, as a function of the rotation matrix $R$ and the translation vector $\mathbf{T}$. [3 marks]

5. **Application** of Questions 3 and 4: Compute $\widetilde{M}$ and $\tilde{M}'$ for rotation $R = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and translation $\mathbf{T} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Show your working. [2 marks]

Student ID: ______________________________

**Stereo vision**




6. Let two identical cameras be placed in canonical geometric position with optical centres $O_1$ and $O_2$ and parallel optical axes ($O_1O_1'$ and $O_2O_2'$) as displayed in the above figure. The distance between the optical centres is the baseline length $b$ and the focal length for both the cameras is $f$. Let $p$, $n$, and $W_{chip}$ denote the physical width of one pixel on the camera's sensor, the number of pixels along one scanline, and the width of the camera's sensor, or the scanline width, respectively. A point $\mathbf{P}$ at depth $Z$ appears at different position along a scanline in each image. The disparity $D_{\mathbf{P}}$ of point $\mathbf{P}$, is given by $D_{\mathbf{P}} = d_2 - d_1$.

   Derive the formula which links the disparity $D_{\mathbf{P}}$ to the depth $Z$ at point $\mathbf{P}$, the focal length $f$ and the baseline $b$ (Show your working). [2 marks]

Student ID: ____________________

7. Consider rotation matrices and translation vectors, $R_1, T_1$ and $R_2, T_2$, which relate camera 1 and camera 2, respectively, with the **same** calibration object. Find the matrix $M$ which relates the optical centers of these cameras, $\widetilde{\mathbf{O}}_2 = M\widetilde{\mathbf{O}}_1$ in homogeneous coordinates, as a function of the matrices $R_1$, $T_1$, $R_2$, and $T_2$. Show your working. [3 marks]

8. **Application** of Question 7: $R_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$, $T_1 = \begin{bmatrix} 100 \\ 0 \\ 500 \end{bmatrix}$, $R_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$, and $T_2 = \begin{bmatrix} -100 \\ 0 \\ 500 \end{bmatrix}$. Compute the baseline length $b$. [2 marks]

Student ID: ______________________

9. Consider the setup from Question 6. As a function of parameters of this stereo system, what minimal, $Z_{\min}$, and maximal, $Z_{\max}$, distances $Z$ are achievable (measurable)? [2 marks]

10. **Application** of Question 9: The $W_3$ camera has a baseline of 76 mm and 3 focal lengths from 6mm to 18 mm. Each $W_3$ sensor is 6 Mpixels with 3000 pixels per line and the pixel width of 5 micrometers. Compute the distances $Z_{\min}$ and $Z_{\max}$. [2 marks]

Student ID: ____________________

11. Consider a well-made calibration setup where a planar target is placed ideally parallel to a camera image plane at distance $Z$ along the optical axis and the target centre, the camera optical centre, and the image sensor centre are aligned. The camera's internal parameters are the pixel width $w$ and the number of pixels per line $N$. The camera focal length $f$ is known from the Tsai calibration. Assume that the radial lens distortion is only of the first order $\kappa_1$. Although in our **ideal** calibration setup, the calibration target line at a vertical distance, $Y_{\text{line}}$, from the centre on the target should be projected into a horizontal image line, the radial distorsion bends it. Let the projected calibration line in the image be at the minimal vertical distance, $r_1$, from the principal image point (the trace of the optical axis) and at the distance $r_2$ from the principal point when this line intersects the right side of the image. Draw the experimental set-up (calibration target, image and projected lines) including all the above-defined parameters. [4 marks]

Student ID: ____________________

12. **Application** of the setup in Question 11: Provide equations linking $r_1$ and $r_2$ with $Z$, $f$, $\kappa_1$, $Y_{\text{line}}$, the pixel width $w$, and number of pixels per line $N$. [2 marks]

13. **Application** of the equations in Question 12: Given the measured $Z = 40$ cm, $f = 5$ mm, $N = 2000$ pixels, $w = 5$ micrometers, $Y_{\text{line}} = 400$ mm, and $r_1 = 980$ pixels, compute $\kappa_1$. How do you call this type of radial distortion and how does this phenomenon relate to your usual camera, e.g. the digital or mobile phone one? [3 marks]

Student ID: ____________________

## Section B: Epipolar Geometry and Stereo Matching: 50 marks

14. Let the baseline of a two-camera system coincide with the $X$-axis of the world $XYZ$-coordinates. How are both the cameras placed one with respect to another if the epipole in the left image coincides with the principal point (trace of the optical axis) and the epipole in the right image is sitting infinitely far along the $X$-axis and has zero $y$-coordinate? [4 marks]

15. How are two cameras placed one with respect to another if epipoles in both images are sitting infinitely far along the $Y$-axis of the world co-ordinate frame and have the same $x$-coordinate? [4 marks]

16. Given a camera with the projection matrix $P_1 = \begin{bmatrix} 0.25 & 0 & 0 & -4 \\ 0 & 0.5 & 0 & -2 \\ 0 & 0 & 0.25 & -1 \end{bmatrix}$, determine the optical centre of this camera? [6 marks]

Student ID: ______________________________

17. Given the same camera as in Question 16 above, and the second camera with the projection matrix $P_2 = \begin{bmatrix} 0.5 & 0 & 0 & 4 \\ 0 & 0.25 & 0 & 2 \\ 0 & 0 & 0.5 & 1 \end{bmatrix}$, determine the point $\widetilde{\mathbf{D}}_1 = \begin{bmatrix} \mathbf{D}_1 \\ 0 \end{bmatrix}$ at the infinity of the ray, projecting the 3D point with homogeneous coordinates $[1, 1, 1, 1]^\mathsf{T}$ to the image plane of the first camera, and project $\widetilde{\mathbf{D}}_1$ to the image plane of the second camera. [6 marks]

18. What relationship does exist between the fundamental matrix $\mathbf{F} = [F_{i,j}]_{i,j=1}^{3}$ of a pair of cameras and the homogeneous coordinates $\widetilde{\mathbf{p}}_1$ and $\widetilde{\mathbf{p}}_2$ of corresponding points with the Cartesian coordinates $\mathbf{p}_1$ and $\mathbf{p}_2$ in the left and right images, respectively, of a stereo pair captured by the cameras. [5 marks]

19. In terms of the relationship in Question 18, specify the epipolar line $\mathbf{a}^\mathsf{T}\widetilde{\mathbf{p}}_1 = 0$ in the left image that corresponds to the point with Cartesian coordinates $\mathbf{p}_2 = (x_2, y_2)$ in the right image. [4 marks]

20. In which point(s) do all the epipolar lines of the right and left image of a stereo pair intersect? [3 marks]

Student ID: ______________________________

21. Describe, in brief, main reasons why stereo matching that searches for corresponding areas in a stereo pair of images is an ill-posed, in the math sense, problem. [4 marks]

22. 3D stereo reconstruction of human heads / faces typically uses stereo pairs captured with cameras having a vertically oriented baseline. Explain in brief why such pairs may be more appropriate than the pairs with the conventional horizontal baseline. [4 marks]

23. Describe in brief which differences between the corresponding image signals are taken into account in the the correlation based matching.

    *Hint*: Consider math models of signals and noise that lead to the matching score. [5 marks]

24. Describe in brief which problem does the dynamic programming stereo (DPS) solve. [5 marks]

    *Hint*: Consider simplifications of a 3D surface model and the choice of the matching score leading to DPS.

Student ID: ____________________

# Section C: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA): 20 marks

25. A face recognition problem is to be solved using the PCA on a database of 1000 images, each of size 2000 (rows) $\times$ 3000 (columns) pixels, that evenly represent 10 different persons (100 images per person). Describe the idea behind the PCA, the size of the covariance matrix $C$, the number of eigenvalues and the size of eigenvectors inferred, and the underlying computational problem. [5 marks]

26. Consider the solution given by Turk and Pentland to the computational problem in Question 25 and provide again the size of the computed matrix $M$, the number of eigenvalues, and the size of eigenvectors inferred, given the database in Question 25. [3 marks]

27. Consider LDA instead of PCA to solve the face recognition problem in Question 25 using the same database separated into 10 different classes (one class per person). Briefly explain the main differences between the PCA and LDA. [2 marks]

Student ID: ________________

28. Consider Question 27 and the database in Question 25. What are the sizes of the between-class scatter matrix $S_{\mathrm{b}}$, of the within-class scatter matrix $S_{\mathrm{w}}$, and of the eigenvectors of the underlying generalised eigenvector equation (as was defined in class). [5 marks]

29. Consider as in Assignment 2, a face database of the class, such that includes the R (red), G (green), B (blue), H (hue), and S (saturation) colour channels, as well as the depth channel. What are and how to tackle the challenges of using the LDA on such a database? *Hint*: Consider redundancy amongst other issues. [5 marks]

Student ID: ______________________________

# Overflow page 1

Student ID: ______________________________

# Overflow page 2

Student ID: ________________

# Overflow page 3