# OLLSCOIL NA hÉIREANN MÁ NUAD

# THE NATIONAL UNIVERSITY OF IRELAND MAYNOOTH

## JANUARY 2017 EXAMINATION

# CS410

# Computer Vision

Dr. D. Charles, Dr. A. Winstanley, Dr. J. McDonald

Time allowed: 2 hours

Answer at least three questions
Your mark will be based on your best ***three*** answers

**All questions** carry equal marks

**[25 marks]**

**1** (a) During the course it was shown how it was possible to model the projection effected by the pin hole camera as a linear transformation from 4-dimensional homogeneous point to a 3-dimensional homogeneous point. This transformation is provided in the expression below: [10 marks]

$$\mathbf{p} = \mathbf{MP} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \mathbf{P} = \begin{pmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \mathbf{m}_3^T \end{pmatrix} \mathbf{P} = \begin{pmatrix} \mathbf{m}_1^T \mathbf{P} \\ \mathbf{m}_2^T \mathbf{P} \\ \mathbf{m}_3^T \mathbf{P} \end{pmatrix}$$

Give details of an approach to calibrating a real-world camera, thereby estimating the values of the individual components of the camera matrix $\mathbf{M}$. In particular, explain how a set of 3D to 2D point correspondences may be used to re-cast the above expression as a homogeneous linear system of the form,

$$\mathbf{Ax} = \mathbf{0}$$

and, explain how it may be solved to compute the camera matrix.

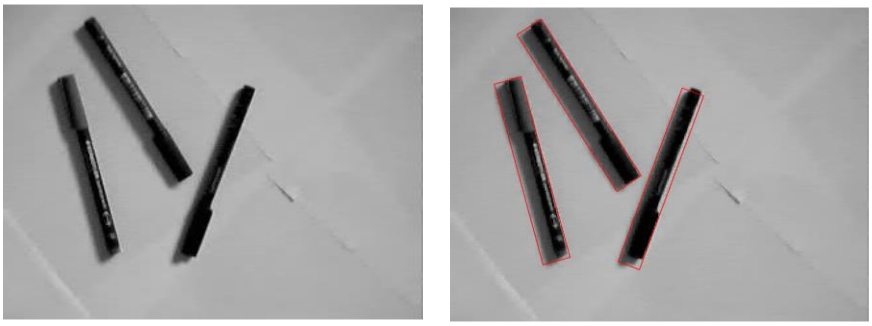
(b) Demonstrate your understanding of the structure of the $\mathbf{A}$ matrix by calculating the two rows resulting from the correspondence between the 3D point $(10,10,15)$ and the image point $(315.2, 282.3)$. [5 marks]

(c) Explain how the above projection model and calibration process can be simplified when computing the projection between a planar target and an image. With reference to the difference between the planar camera model, why is the projection from a planar target to an image is invertible? [10 marks]

How does this simplification affect the structure of the $\mathbf{A}$ matrix from part (a)?

**[25 marks]**

**2** (a) The figure on the left below shows an image of three pens on a contrasting background. The figure on the right shows the same image where three boxes have been overlaid on the pens. [15 marks]



Using the image moments described during the course, detail the complete set of steps required to compute the position and orientation of the boxes in the right image using the image on the left as input.

Your solution should include:

i. an explanation of the set of image processing operations required to generate any intermediate images,

ii. an explanation of the zeroth first moments and second moments of an image, include the terms within each expression,

iii. a “walk-through” of how these moments would be computed for the objects in the above image, and,

iv. how to compute the bounding box for each pen

To assist you in providing a solution the expression below correspond to the second moments of the image.

$$a = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x'_{ij}\right)^2 B[i,j] \qquad c = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y'_{ij}\right)^2 B[i,j]$$

$$b = \sum_{i=1}^{n}\sum_{j=1}^{m} x'_{ij}y'_{ij}B[i,j]$$

where the orientation is given by

$$\theta = tan^{-1}\left(\frac{b}{a-c}\right)$$

(b) Provide a high-level explanation of the set of steps involved in the thinning algorithm explained during the course. [10 marks]

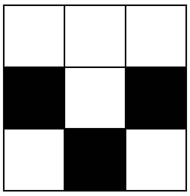
The central step in this algorithm requires identifying whether a pixel is critically connected. We saw that labelling the pixels within a 3x3 region as follows:

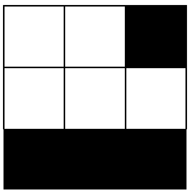
| | | |
|---|---|---|
| 3 | 2 | 1 |
| 4 | 8 | 0 |
| 5 | 6 | 7 |

pixel 8 can be identified as critical connected if the following predicate evaluates to true:

$$
\begin{aligned}
&8.\{[(1+2+3).(5+6+7).\bar{4}.\bar{0}]\\
&+[(1+0+7).(3+4+5).\bar{2}.\bar{6}]\\
&+[3.(5+6+7+0+1).\bar{2}.\bar{4}]\\
&+[1.(3+4+5+6+7).\bar{2}.\bar{0}]\\
&+[7.(1+2+3+4+5).\bar{0}.\bar{6}]\\
&+[5.(7+0+1+2+3).\bar{4}.\bar{6}]
\end{aligned}
$$

Through the use of examples provide an explanation of the structure of this expression. In particular provide an intuitive explanation function of each of the sub-clauses. Using the above expression show how it can be used to determine if the central pixels of the following neighbourhoods are critically connected.





**[25 marks]**

**3** (a) Provide expression for the convolution operation in both the continuous and discrete domains and explain its operation. Why is the operation important in the context of image processing? In your answer you should make reference to the relationship between convolution and linear shift invariant systems. [6 marks]

(b) The two 3x3 images below shows edges oriented at two different angles. Demonstrate how the Sobel operators can be used to compute the <u>magnitude</u> and <u>orientation</u> of each of the edges. You are only required to compute the values for the center pixels in each of the images. [9 marks]

| | | |
|---|---|---|
| 0 | 0 | 200 |
| 0 | 200 | 200 |
| 200 | 200 | 200 |

(a)

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 200 | 200 | 200 |
| 200 | 200 | 20 |

(b)

Note that the two Sobel kernels are given below:

| | | |
|---|---|---|
| 1 | 0 | -1 |
| 2 | 0 | -2 |
| 1 | 0 | -1 |

$S_x$

| | | |
|---|---|---|
| -1 | -2 | -1 |
| 0 | 0 | 0 |
| 1 | 2 | 1 |

$S_y$

(c) Explain how the histogram of oriented gradients (HoG) provides a representation of an image suitable for image and object recognition tasks. How is the HoG representation used within a complete object recognition system? [10 marks]

**[25 marks]**

**4** (a) A driver distraction system uses a camera to detect when a person is fatigued. The system has a detection performance of [10 marks]

$$p(Z = Fatigued|Driver\ fatigued) = 0.95$$
$$p(Z = Fatigued|Driver\ not\ fatigued) = 0.15$$

where $Z$ is the detected state.

The system initialises with a prior probability of

$$P(Driver\ fatigued) = 0.5$$

Provide an expression for Bayes rule and give a brief explanation of each term in the expression.

Starting with the above prior probability the system detects that the driver is in a fatigued state. Using Bayes theorem, compute the probability that the driver is fatigued.

Imagine now that the state of the driver has not changed and the system makes another (independent) detection that the driver is fatigued. Show how Bayes rule can integrate this new information into its probability estimate of the driver being fatigued.

(b) Using Bayes rule in its standard form minimizes the misclassification rate. In scenarios such as the one above this is typically not a good approach to take. Explain why this is the case and give details how minimizing the risk is a more appropriate approach. [7 marks]

In your answer you should provide an expression for the risk and give a high-level explanation of how it can be integrated into the Bayes classifier.

(c) A common requirement of camera calibration systems is to compute correspondences between points on a planar calibration pattern and images of that calibration pattern. This is commonly achieved through the use of a checkerboard, however, in reality any textured planar target could be used with the aid of an interest point detection and matching framework. [8 marks]

Given the image of the book cover below explain how you could use an interest point detector and descriptor to generate a set of correspondences between the image below and an image of the book itself (e.g. appearing at an arbitrary pose in the imaged scene).