

The University of Nottingham Ningbo China

SCHOOL OF COMPUTER SCIENCE

A LEVEL 3 MODULE, SPRING SEMESTER 2021-2022

COMPUTER VISION

Time allowed: 2 hours

Candidates may complete the front cover of their answer book and sign their desk card but must NOT write anything else until the start of the examination period is announced

Answer ALL THREE Questions

Only silent, self contained calculators with a Single-Line Display or Dual-Line Display are permitted in this examination.

Dictionaries are not allowed with one exception. Those whose first language is not English may use a standard translation dictionary to translate between that language and English provided that neither language is the subject of this examination. Subject specific translation dictionaries are not permitted.

No electronic devices capable of storing and retrieving text, including electronic dictionaries, may be used.

DO NOT turn examination paper over until instructed to do so

1. Feature Detection and Stitching

- (a) Sobel operator, sometimes called Sobel filter, is used in image processing and computer vision, particularly within edge detection algorithms where it creates an image emphasising edges. Compute G_x and G_y , gradients of the image by performing the convolution of Sobel filters with the image below. Note that you only need to compute for the pixels highlighted in grey.

[8 marks]

100	100	50	50
100	100	50	50
100	100	50	50
100	100	50	50

- (b) Histogram of Oriented Gradients (HoG) is a popular feature descriptor used in computer vision for the purpose of object detection. Consider an image patch with the size of 64×128 . If an 8×8 cell (as defined in HoG) and a 16×16 block (as defined in HoG) are used, what is the dimension of the HoG vector for this image? Show your working steps and an answer without the intermediate steps will not receive any marks.

[8 marks]

- (c) SIFT is one of the most commonly used feature descriptors. Briefly outline the main steps used to extract SIFT features from a given image.

[8 marks]

- (d) Random sample consensus (RANSAC) is an iterative method to estimate parameters of a mathematical model from a set of observed data. Briefly describe how to use RANSAC to estimate Homography.

[10 marks]

2. Optical Flow, Camera Calibration and Stereo

- (a) Briefly describe the differences between motion field and optical flow. What is the assumption used when estimating optical flow?

[10 marks]

- (b) You are given a video capturing a pedestrian walking from left to right in front of a wall. The video is captured by a camera statically mounted on a tripod without moving the lens. The wall fills in the whole video frame and the width of the wall captured by the video is about 10 meters. The walking speed of the pedestrian is about 2 km/h. The size of the video frame is 729×486 pixels and the framerate is 60 fps. Based on these facts, should you use Lukas-Kanade method to compute optical flow between two consecutive frames and why?

- (c) Briefly describe what camera calibration is and what camera calibration can be used for. Give an example of how we can calibrate a camera, e.g. how to use a checkboard to do so.
[12 marks]
- (d) What is the correspondence problem in stereo matching as discussed in the lecture? How can camera geometry make the problem easier to solve?
[6 marks]

3. Visual Recognition and Detection

- (a) Bag-of-Features model is widely used in large-scale image search applications. Describe how such model is used, ie. how a large-scale image database is built when using the Bag-of-Features model and how to use it to perform search.
[10 marks]
- (b) Compare and contrast fully supervised object detection and weakly supervised object detection techniques.
[3 marks]
- (c) Name three specific object detection techniques that use Convolutional Neural Networks (CNNs). Choose one of these techniques to describe its main contributions and network architecture. You may draw a diagram to help you illustrate your points.
[11 marks]
- (d) In Convolutional Neural Network (CNN), non-linearity is often used after convolutional layer. Briefly describe the effect of adding non-linearity to CNN and write down two non-linearity functions.
[8 marks]