The Australian National University
College of Engineering and Computer Science
Final Exam, First Semester 2022

COMP/ENGN6528 Computer Vision

Question Booklet

Reading time: 15 minutes
Writing time: 2:15 hours
Uploading time: 15 minutes

Instructions on next page

Allotted Time

You will have 2:15 hours to complete the exam plus 15 minutes of reading time (you are allowed to write during this time. An additional 15 minutes has also been allowed to accommodate the additional task of uploading your completed exam to the final exam Turnitin submission portal on the COMP/ENGN4528 Wattle site. Thus you have 2 hour and 45 minutes to complete the exam. NO late exams and submissions will be accepted. You may begin the exam as soon as you download it.

Minimal requirements:

You may attempt all questions

You SHOULD NOT include an assignment cover sheet

You must type your ANU student identification number at the top of the first page of your submission

You must monitor your own time (i.e. there is no invigilator to tell you how many minutes are left).

Your answers must be clear enough that another person can read, understand and mark your answer. 11 or 12 point font with 1.5 spacing is preferred. Scanned images of handwritten equations or diagrams must be legible and of a suitable size. Please be aware that your submitted document should have at least 20 words based on the requirement of Turnitin. If it is not satisfied, it may lead to unsuccessful submission.

Numbering questions

- You must specify the question you are answering by typing the relevant question number at the top the page
- Each question should begin on a new page
- Multi-part questions (e.g. question 1 parts a and b) may be addressed on the same page but should be clearly labelled (e.g. 1a, 1b)
- Questions should be answered in order

You must upload your completed answers in a single document file within the allotted time using a compatible file type for Turnitin (Preference: MS Word's .doc or .docx or .pdf format) **It is the student's responsibility to check that the file has uploaded correctly within Turnitin. No late submission will be accepted (other than the case of prior special arrangement).** Access to the Turnitin practise site can be found here: https://www.anu.edu.au/students/academic-skills/academic-integrity/turnitin

Academic integrity

Students are reminded of the declaration that they agree to when submitting this exam paper via Turnitin:
I declare that this work:

- upholds the principles of academic integrity as defined in the University Academic Misconduct Rules;
- is original, except where collaboration (for example group work) has been authorised in writing by the course convener in the course outline and/or Wattle site;
- is produced for the purposes of this assessment task and has not been submitted for assessment in any other context, except where authorised in writing by the course convener;
- gives appropriate acknowledgement of the ideas, scholarship and intellectual property of others insofar as these have been used;
- in no part involves copying, cheating, collusion, fabrication, plagiarism or recycling.

Mid-Semester Exam: How to

Do's and Don't's:

- You are recommended to record your exam process and keep the video to yourself (for at least 4 weeks). That includes a video from your camera of yourself completing the exam and of your screen. At the start of the video you should clearly show your student card to the camera. (You could use zoom and record to your machine, or to the cloud).
- **Do not submit your video recording.** This file stays private to you, unless we have a specific reason to request this file from you. For this exam we will also randomly select some students and request to view their videos.
- Feel free to **ask question directly posting to "Instructors"** on Piazza (public posting will be disabled for you during the exam) or send email to course convenors. Our class link on Piazza is: https://piazza.com/anu.edu.au/spring2022/engn6528/
- **Do not use any communication system** (other than direct posts on Piazza to Instructors or sending emails to your course convenor) during your exam - if you do, it will be counted instantly as collusion and will have serious academic honesty consequences.
- **Do not upload any material** anywhere (other than to the Wattle upload link at the end of the exam). If you do so, you will also become part of an academic collusion case which will stay on your permanent record at the ANU.
- Be very careful **searching for any material on-line**. If you find yourself seeing references to material which might be the result of collusion (which hopefully will not exist), you are one click away from becoming part of a serious academic honesty case yourself. Remember that all of your activity must appear in your recording, and according to basic academic standards, we also expect you to reference in your pdf file anything which you might have included in your working. You will likely waste valuable time for your exam and expose yourself to serious risks, so we recommend that you do not do so. To be specific, you can access general online material, but any means of communication with other students, including sharing any type of file via social media or web-sites may constitute collusion.

Step-by-step guide on how to sit your actual on-line mid-semester exam:

1. Find yourself a cozy spot and **power down all communication channels**, besides this forum here.
2. Start your full screen **recording now**.
3. **Download** the exam paper from wattle (link will become active at the time of the exam).
4. **Open** the exam in the pdf reader, which you tested before.
5. Take a moment to **read** the whole document. You don't need to spend exactly 15 minutes for reading and you can start working on the exam after you finish reading it at your pace. You can make notes at any time.
6. Fill in **your answers**. You can initially write your answer using your favourite editor. You are strongly recommended to submit a word document or Pdf document. Whichever way you chose, all your activities **must appear on your recording**.
7. Don't forget to **save regularly,** or use a system which does that for you.
8. When you are complete, upload a single file (word or PDF are OK), and upload via the Wattle Link. You can make multiple uploads, they will overwrite the last. Ensure you upload before the end of the time. We have allowed 15 minutes upload time so submit with some time to spare.
9. **Stop your screen recording** and make sure the video file is saved (keep this file for at least 4 weeks). **Do not submit** your video recording unless requested to do so.

There are 6 questions, 70 Marks in total. (Q1-Q6)

Please name your submission as
ENGN6528_exam_u1234567.docx

Questions on the next page

Q1: (13 marks, Easy Questions) [basic concepts]

Answer the following questions concisely. Each of the questions must be answered in no more than 5 lines of text. Longer answers will be penalized.

low

1) In a camera image, what is the principal point? Focal length is also a key factor for imaging. How does having a longer focal length change the appearance of an image? [2 marks]

2) For high level vision problems, what is the difference between a detection and a classification task? [2 marks]

3) For a computer vision problem of model fitting, such as fitting a line to a set of edge pixels in an image, what is the key advantage of using RANSAC vs using least squares? [2 marks]

4) Suppose that you have two cameras that are setup viewing a scene from different viewing angles. You know the intrinsic calibration parameters for each camera, and you have a homography matrix for the pair of cameras. What information does this allow you to recover from the scene? Please describe any restrictions. [3 marks]

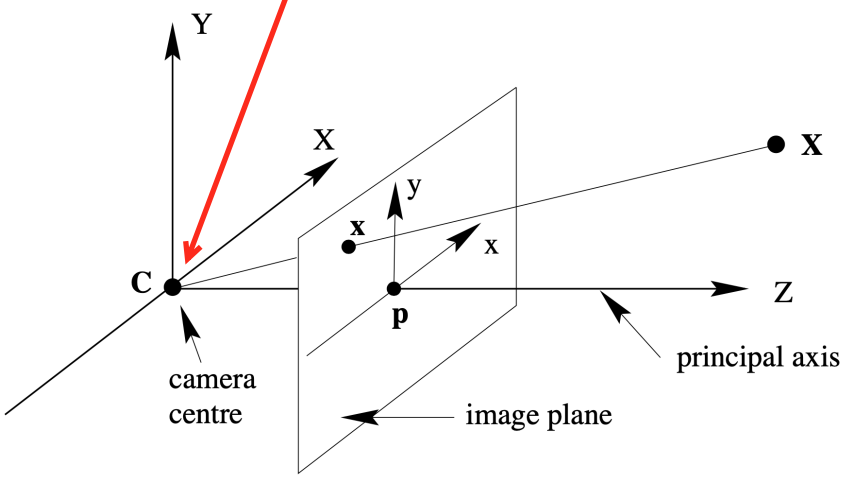
5) In a rectified stereo setup we search along the row of the second image to find a point that matches a point in the first image. [4 marks]
   - What is the advantage of using a sum of squares difference match over a window rather than matching based on single pixel values?
   - Why would you normalize the window value, such as subtracting the mean and divided by the standard deviation, before calculating a sum of squares difference?

Q2: (21 marks, Easy and Moderately Difficult Questions) [3D]

Answer the following questions concisely. Write down working, and if you are unsure about some part along the way, state your best assumption and use it for the remaining parts. Similarly, if you think some aspect is ambiguous, state your assumption and write the answer as clearly as you can.

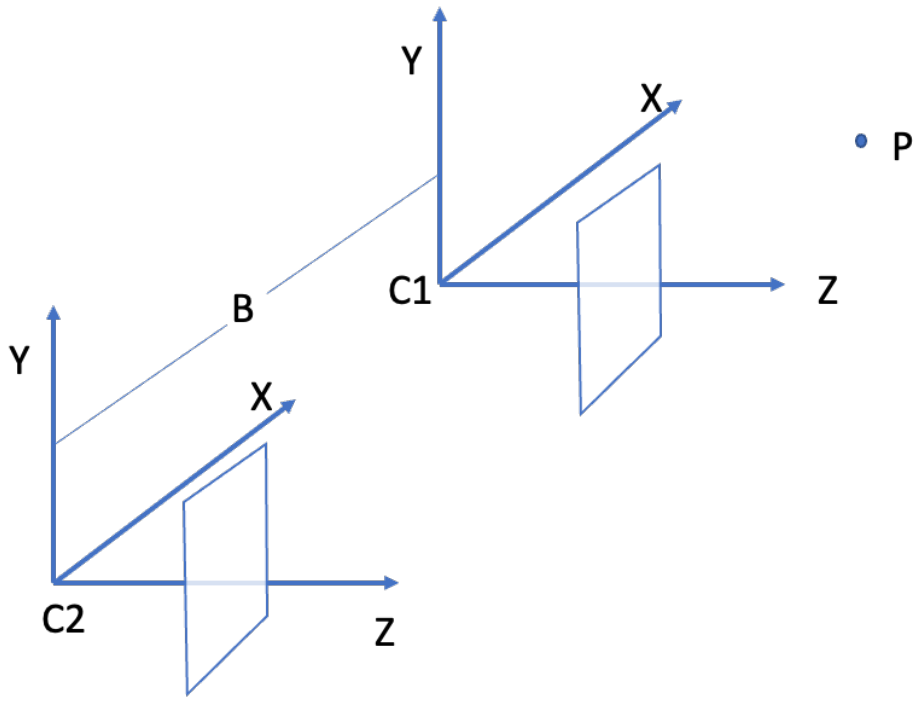
(a) Given two calibrated cameras, C1 and C2, C1 has the focal length (in pixel unit) of 500 in x and 375 in y, the camera has a resolution 512x512, and the camera centre projected to image is at (249, 250), with no skew. Suppose C2 has the same resolution and focal length as C1, but the camera centre projected to image is at (252, 252). Write down the calibration matrix K1 and K2 for C1 and C2 respectively. (Hint: please only write down the final two 3x3 matrices.) [3 marks]

(b) Suppose that a 3D world coordinate system ((X,Y,Z) coordinates as in the below diagram from the lecture notes) is defined as aligned with the camera coordinate system of C1. More specifically, the world origin is at the camera centre of C1, the Z axis is aligned with the optical axis and the X and Y world coordinate systems aligned parallel with the x and y axes of the image of C1. Write down the matrices K[R|t] which define the projection of a point in world coordinate system to the image of C1. (Hint: please only write down the final 3x4 matrix.) [3 marks]



(c) Suppose that the scene has a 3D point, P, that in the world coordinate system defined above that lies at (24, 23, 100) in cm. Note that the points in world coordinate system are measured in cm. What location (to the nearest pixel) will that world point (P1) map to in the image of C1? [2 marks]

(d) Suppose that with respect to the world coordinate system that is aligned with camera C1, camera C2 begins being aligned to C1, and subsequently the centre of C2 is translated by B = 0.2 m to the left of C1 (along the X axis of C1). The two camera centres both remain on the same (X, Z) plane



Write down the matrices K[R|t], which define the projection of points in the world system (i.e, the same coordinate system of C1) to the image of C2. (Hint: please only write down the final 3x4 matrix.) [3 marks]

(e) What is the location (to the nearest pixel) that P maps to in the image of Camera C2?(Hint: Please write down only the final result.) [2 marks]

(f) Define the term epipole. [2 marks]

(g) For camera C1, there is an epipole (or epipolar point) that relates to Camera C2. For the two-camera setup for predicting structure from motion, what is the position of the epipole in camera C1 of camera C2? (Hint: It is a point in the image plane of Camera C1). [2 marks]

(h) Assume we measure the corresponding pixel of a 3D point P2 that is projected in camera C1 at image location (x1, y1), and in camera C2 at image location (x2, y2). How would you find the world coordinates of point P2 given the noisy measurements of (x1, y1) and (x2, y2)? Please provide a robust solution. [4 marks]

Q3(10 marks, Moderately Difficult Questions) [Camera models and SFM]

(a) Given multiple cameras that view the same scene. Suppose that you know the intrinsic parameters of all cameras, and they all view a common object, for which many accurately matched points are available in all images. Suppose that you do not know extrinsic calibration. *What information can you recover about the cameras and the scene from this configuration?* Describe what the parameters are to be recovered. [4 marks]

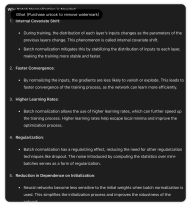
(b) For the above camera setup, suppose that you have many matched points available in both images of a common object, but a small number of the matches may not be correct. Describe a method for recovering camera and scene information given a set of point matches where a small number of the points maybe mismatched. [6 marks]

Q4: (6 Marks, Moderately Difficult Questions) [Shape-from-X]

(a) Shape-from-Shading approaches predict the brightness of an image pixel. Given a point light source at infinity (distant light source), write down the equation that defines the brightness at an image pixel assuming that the camera views a Lambertian surface, please also define the terms of the equation. [2 marks]

(b) Suppose that we have used some other method to know the brightness of the lighting and location of three lights, and the relative location of a single camera, and that we know that the object has a Lambertian surface. We then take three images of an object with each of the lights turned on in turn. The lights, object and camera are kept in precisely the same position. If we just consider the brightness at a single pixel in all three images (the same pixel) what can we deduce about the surface orientation of the object at that pixel from the three brightness measurements? [4 marks]

Q5: (10 Marks, Easy Questions) [Deep Neural Networks, RANSAC, Vanishing Lines and Points]



(a) [Deep Neural Networks]. Describe in two sentences the problem that Batch Normalization is designed to address in Deep Neural Network Learning? [2 marks]

(b) [RANSAC]. Consider the RANSAC algorithm for fitting a plane for a set of points. Let us denote e, the probability that a point being an inlier, we have e = 0.6 for the current case and the number of points in a sample be s = 3. Determine the required number of sampling iterations (N) to ensure that at least one of the samples is free from outliers with probability p = 0.97. [2 marks]

(c) [Vanishing Lines and Points]. Assume there are 4 points in the image denoted as a, b, c, and d shown in Fig.1. In particular, line ab is parallel to the line cd; line ac is parallel to line bd in 3D. The pixel coordinates are a: (78, 78), b: (98, 128), c: (98, 68), d: (118, 93). Please calculate the vanishing points v1 and v2 and the vanishing line. [6 marks]

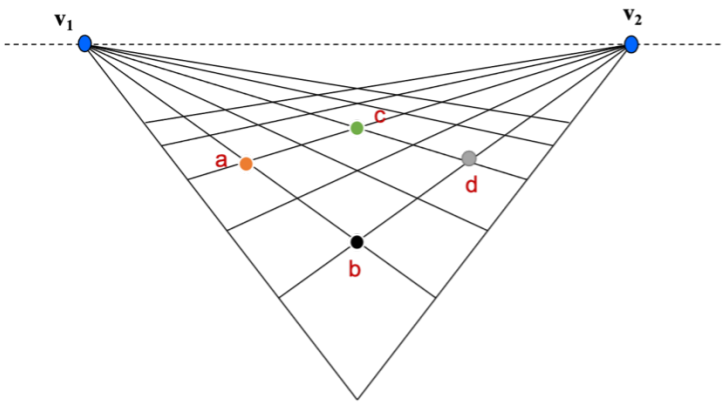


Fig. 1: Vanishing Points and Vanishing Lines

Q6 (10 marks, Challenging Questions). Algorithm Design.

Turn your mobile phone into a 3D scanner for featureless objects (sample image in Fig. 2 below), such as a small statue. Specifically, please design an algorithm to obtain the 3D model of featureless objects. Please note that no other type of sensors, such as lidar or infrared cameras, could be involved in your algorithm.

1) Please briefly describe the key steps of your algorithm. [7 marks]
2) Please also discuss the potential degenerate cases for your setup (namely a set up leading to wrong shape estimation). [3 marks]

Note that you should make your assumptions clear.

Fig 2. sample object image

======= END of ALL QUESTIONS in the EXAM ==========