

CS1674: Essay 1

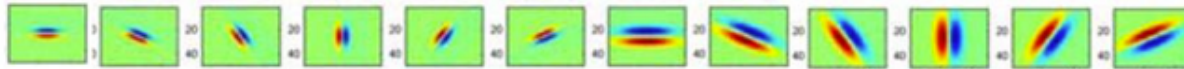
Due: ~~3/25/2021~~ (new due date) 3/26/2021, 11:59pm

This assignment is worth 45 points.

Your task is to provide responses to 9 questions. Most questions are essay-style short responses; a few ask you to complete a task by hand. Each question is worth 5 points. Each response should contain 5-10 sentences. Be as specific as possible.

Any resource, including slides and the web, can be used when drafting your responses. You do not have to use sources beyond the course, but you will need to think beyond just recalling course discussions. If you do use materials beyond those discussed in class, please cite them in your response. Needless to say, you cannot copy or closely paraphrase sentences from papers or blogs on the web (we will check). Also, do not copy-paste or closely paraphrase content from the slides or textbook either. **Do not discuss the questions or answers** with your classmates or with others outside the class. It is fine to discuss the questions with the instructor or TA, but make sure you have carefully considered the question first.

1. What are three applications of computer vision in daily life (broadly defined)? Describe them with a sentence each. Which of these do you consider the most useful, and why? Be specific and detailed.
2. Suppose we form a texture description using a filter bank at two scales and six orientations like the one below. If we rotate Image A by an arbitrary degree (resulting in Image B) and compute the responses to the filters, would the sequence of responses be the same as if we hadn't rotated the image? Why/why not? Then, suppose we compute the mean response of Images A and B to each filter, resulting in a 12x1 feature/descriptor for each image. What can you say about the distance between the two descriptors (for A and B), e.g. would it be 0? If the descriptor is not invariant to rotation, how can we formulate a descriptor that may be invariant to rotation? (Question based on an assignment by Kristen Grauman.)



3. What are the advantages of using responses to a filter bank in order to compute a feature describing an image? What are the disadvantages?
4. Describe what image transformations (e.g. rotation, translation) corner detection is robust to. Then describe what blob detection is robust to. Give reasons for robustness or lack thereof, for each detection method and each transformation.
5. What is an edge? How do we determine where an edge lives in an image? How do we determine how strong an edge is and which way it is oriented? What more complex structures can we form out of a collection of edges, and how?
6. In what ways are (a) a SIFT representation for a keypoint, (b) retrieval or classification based on a bag-of-words representation, and (c) segmentation via clustering, similar? In what ways are they different? Be as detailed as possible.
7. Pick a simple, assymetric image (e.g. house, flag or some other simple shape). Pick two geometric transformations with specific values (e.g. choose the degree of rotation, if using rotation). For the first transformation, write the matrix (with exact values) that describes the transformation, then show what it does to the image (call the output the intermediate result). Now do the same for the second transformation, but apply it on the intermediate result. Finally, starting with the original image, apply the transformations in the opposite order. Describe how the two final outputs compare.

8. Give five examples of techniques that intentionally drop information (e.g. by removing detail or aggregating/summarizing fine information into coarser information). What are the tradeoffs (pros/cons) of dropping information (or making it coarser) in each case? How does this dropping of information relate to examples of the same process in daily life?
9. Imagine that in 10 years you are a computer vision engineer, working for a large company or a small startup. What is the computer vision system that you would be most interested to build? What does it do? How complex is it? What is it good for? What problems can it cause, if any? What knowledge do you think you still need to be able to build such a system? (I'm not looking for technical terms, but rather processes that you don't know how to accomplish.)

Submission:

- essay1.pdf/.docx

Acknowledgement: Adriana Kovashka.