Computer Vision 2 Exam

27 May, 2015

The original articles are allowed on the table. NO lecture slides/book. Good luck.

Question 1: Projections 20p

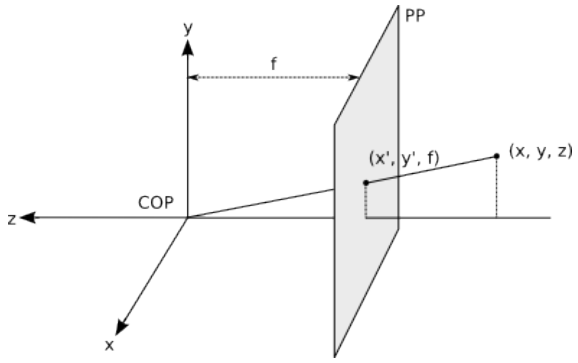


Figure 1: Projection from 3D to 2D.

Consider a Center of Projection (COP), and a 3D point (x, y, z), as illustrated in figure 1.

- a - 5p) Show the relation between the 3D point (x, y, z) and its 2D projection (x', y') on the Projection Plane PP that has a distance f from the COP.
- b - 5p) Is the relation in (a) linear? Why (not)? (hint: you may add a drawing or curve)
- b - 5p) Give the homogeneous matrix that gives the projection in (a).
- c - 5p) Show that this matrix, when multiplied with the 3D point, gives the same inhomogeneous result as in (a).

Question 2: Epipolar geometry 20p

- a - 6p) Draw a schematic representation of epipolar geometry, including epipoles, baseline, and epipolar plane.
- b - 4p) Indicate and explain the epipolar constraint.
- b - 4p) Why is the epipolar constraint useful? Can you give examples?
- c - 3p) What is the difference between the Essential matrix and the Fundamental matrix?
- d - 3p) Can you think of advantages and disadvantages between a large and a small baseline?

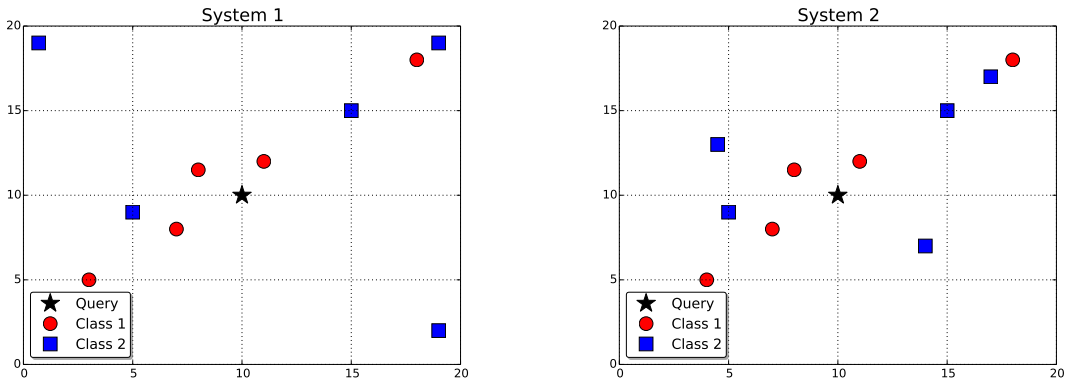
Continue on the next page. Please write your name and student number on all answer pages.

Question 3: Bag of words 20p

- a - 4p) In the first lecture we discussed 9 reasons why computer vision is difficult. The popular Bag of Features (bag of visual words) model deals with several of these difficulties. Choose 4 of these difficulties, and explain WHY and HOW Bag of Features solves them.
- b - 4p) Choose 4 of these difficulties, and explain WHY and HOW Bag of Features does **not** solve them. They can be 4 different, or the same difficulties as in (a).
- a - 4p) Describe the Bag of Features model in less than 10 lines.
- b - 4p) Describe the Spatial pyramid in the Bag of Features model in less than 10 lines.
- c - 4p) Describe ambiguity modeling in the Bag of Features model in less than 10 lines.

Question 4: Evaluation 20p

In the following we compare the image retrieval results of two different systems. Results for system 1 (left) and for system 2 (right) are:



The star denotes the query image, the red dots (class 1) and blue squares (class 2) represent images that belong to two classes. The axis represent a feature space, and we rank images based on how close they are to the query image. Assume the query is supposed to be from class 1 (i.e. class 1 is the target class).

- a - 4p) Draw the Precision-Recall curves for both systems based on the query ranking (class 1 is the target class).
- b - 4p) Draw the ROC-curves for both systems from the ranking for the query. Assume the query is from class 1 (i.e. class 1 is the target class).
- c - 4p) Compute Average Precision for both systems. Assume the query is from class 1 (i.e. class 1 is the target class). Show the calculation, keep it in fractions if you need to.
- d - 4p) Compute AUC for both systems. Assume the query is from class 1 (i.e. class 1 is the target class). Show the calculation, keep it in fractions if you need to.
- e - 4p) Compare and analyze the scores for both systems, system 1 and 2, and for both measures, AP and AUC.

Continue on the next page. Please write your name and student number on all answer pages.

Question 9: Papers 20p

- a - 2p) In a pinspec camera, i) what happens when you move the occluder? ii) What happens if you make the occluder bigger. Be as precise as you can.
- b - 2p) In the future, new cameras may be invented that can record a holographic representation of the scene. Explain WHY and HOW the Plenoptic function should be adapted to deal with future hologram-cameras.
- c - 2p) Consider (Matzen and Snavely, 2014, "Scene Chronology"). Imagine, we want to apply this method on all images taken in the working room of the American President in the White house (the so-called "oval office"). What is the main thing you would you change to the algorithm to work in the room of the President?
- d - 2p) Consider (Achanta et al, 2011, "SLIC superpixels"). What is their single most important parameter? Which results are influenced by it? Give two examples when you want to change the setting for this parameter.
- e - 2p) Consider (J. Hosang,et al. 2015: "What makes for effective detection proposals?"). i) Why not use AUC to evaluate object detection?
- f - 2p) Consider (Lampert et al.: Attribute based classification for zero-shot Visual Object categorization, 2013). Below you find three examples of animals as drawn in the 18th century. They are based on field-notes that were passed to a sketch artist. What is the relation of the sketch artist with Lampert et al.?

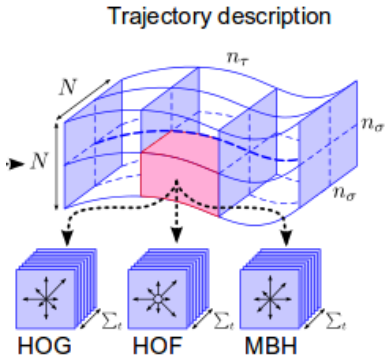
Walrus

Hippo

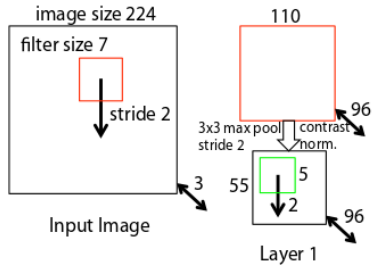
Giraffe

Continue on the next page. Please write your name and student number on all answer pages.

- g - 2p) Below you see descriptors as used by Wang et al, in "Dense Trajectories and Motion boundary histograms for Action Recognition". Compute the dimensionality of the HOG, HOF, and MBH descriptors, for the exact image as you see it here (Explain how you got to the answer):



- h - 2p) Below is a part of a figure by Zeiler et al, 2014, "Visualizing and understanding convolutional networks". Please explain all numbers in the figure.



- i - 2p) In (Rothganger et al., 2006, 3D Object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints) they state that "objects are always planar in the small". They use a scale invariant feature detector (the Laplacian). How is scale invariance related to the planarity of a patch?

Good luck! Please write your name and student number on all answer pages.