Homework #3 for Computer Vision (E1.216) 2021

Notes

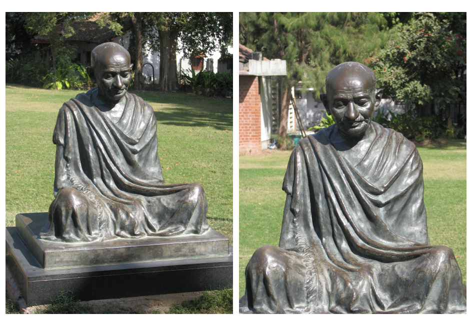
- This homework is worth 12% of your grade, so pay attention and earn it!
- Grading will be based on the quality of your results and the merits of your analysis.
- This homework is due by 5pm on 29th May 2021. Absolutely no extensions possible.
- Link for uploading homework will be provided later.

Q1: Normalised Cuts (30 points)

Implement a method for image segmentation using normalised cuts. Use your code to segment colour images. Your similarity measure should encode both image colour and/or feature information as well as spatial coherence, i.e. for every pixel use a local neighbourhood (4 neighbours, 8 neighbours, patch etc.).

Demonstrate your method on at least 3 images downloaded from the internet and at least 3 images taken by you using a mobile phone camera. Make sure you pick different types of scenes for the images (varied properties such as texture etc). The report should contain the original and the output segmented images along with important details such as speed of your code, largest image size you can handle etc. Also provide key details of your implementation, explain the implications of different design choices and briefly describe key observations. For this question, you need to demonstrate that you have given adequate thought to the design, implementation and testing of your solution.

Q2: Epipolar Geometry (30 points)



The two images shown are of a life-statue of Mahatma Gandhi at Sabarmati Ashram in Ahmedabad. For this problem, you should use the attached zip file for this dataset as indicated. For your convenience, a set of matched points are provided. Use these matches for part 1 and 2. The report should contain the implementation details (like algorithmic details, resolving ambiguity among multiple solutions, parameters used etc) and tables for error comparison.

For the pair of images as shown above, compute the epipolar geometry as described below :

Part 1: Calibrate the two cameras using the calibration information provided. Subsequently compute the essential matrix using the eight point algorithm and find the rotation and translation direction between the two images. The dataset also provides the correct (ground truth) motion between the two cameras. Compare your rotation and translation direction estimates with the ground truth. Repeat this exercise using the Hartley normalisation for the eight point algorithm. Explain the differences you observe. [10 points]

Note: To compare translation directions, measure the angle between estimate and ground truth.

For comparing rotation estimates, use the following. Any 3×3 rotation matrix $\mathbf{R}$ can be defined as a rotation by angle θ about an axis of rotation $\mathbf{n}$. Specifically, given $\mathbf{R}$ we can define the angle of rotation and axis as

$$
\begin{aligned}
\theta &= \cos^{-1}\left(\frac{tr(\mathbf{R}) - 1}{2}\right) \\
\mathbf{n} &= -\frac{1}{2\sin\theta}\left(\mathbf{R}(2,3) - \mathbf{R}(3,2), \mathbf{R}(3,1) - \mathbf{R}(1,3), \mathbf{R}(1,2) - \mathbf{R}(2,1)\right)
\end{aligned}
$$

where $tr(.)$ is the matrix trace and $\mathbf{R}(i,j)$ has the usual meaning. Further define the 'difference rotation' $\delta\mathbf{R} = \mathbf{R}_e\mathbf{R}_g^{-1}$ where $\mathbf{R}_e$ and $\mathbf{R}_g$ are the estimated and ground-truth rotations respectively. Then the angular distance between the two rotations is defined by the angle θ of $\delta\mathbf{R}$ and you can also compare the axes of rotations.

Part 2: The second part of the exercise is to understand the uncertainty involved in the estimation process. For this exercise do not use the calibration of the cameras, i.e. calculate the fundamental matrix and not the essential matrix.

Run a bootstrap experiment as follows: For N correspondences, if we use only $(N - k)$ correspondences at a time (i.e. drop k correspondences) we will get a new estimate of the epipolar geometry. We can repeat this many times by randomly picking $(N - k)$ out of N correspondences. Let us pick a suitable value such as $k = 10$ (you can also chose another appropriate value of k) and repeat this random selection many times. For each estimate of the fundamental matrix extract the left and right epipoles. Thus, if you run T such random trials, you have T estimates for the left and right epipoles. For both cases, compute the mean and covariance and display it. Repeat the same experiment using Hartley's normalisation. Explain the differences you observe. [10 points]

Part 3: Repeat the exercise in Part 1, except that here you will also compute the matching points yourself using SIFT and a matching scheme. You can download the appropriate code from `http:\\www.vlfeat.org`, specifically `vl_sift` and `vl _ubcmatch`. Now you will need to do a robust estimation, for which you can use RANSAC to remove outlier correspondences. How consistent is your estimate here compared with that obtained in Part 1 where you were provided good correspondences. [10 points]