ECSE 683 – Topics in Computer Vision and Robotics
Assignment 3
November 11, 2016

Due: November 25, 2016 at 5:00 pm

In this assignment we will investigate structure from motion. For this assignment you will be using the Pepsi can sequence in pepsi.tar.gz. (Use this for developing your solution – I'll provide a better sequence to test it on ASAP).

Question 1

Implement Algorithm CONSTANT_FLOW (notes, Page 197 Trucco & Verri). Test it on frame pair pepsi04-pepsi05 and plot the resulting flow field. Since the field corresponds to pure translation ($T_z = 0$), recovery of depth up to a scale factor should be possible. Devise a method for recovering depth in this case and plot the resulting depth map. Note, since the height of the Pepsi can is known, determining an overall scale factor should not be too difficult. Hint: consider Equation 8.8 in Section 8.2 of the textbook. Each (v_x,v_y) corresponds to a particular (T_x,T_y,Z).

Question 2

The instantaneous flow field recovered in Question 1 is not the best basis for depth recovery considering the very short baseline. A better approach, given that we have all the frames in advance, would be to track a discrete set of points through the 11-frame sequence, and then use a factorization approach to recover depth.

Implement Algorithm FEATURE_POINT_MATCHING (notes, Page 199 Trucco & Verri), which essentially bootstraps optical flow estimates. Use this to construct the 2N x n measurement matrix, W. Implement and apply the factorization algorithm, MOTSTRUCT_FROM_FEATS (notes, Page 208 Trucco & Verri).

Given that we know the camera is translating parallel to the scene, what is the expected structure of R? How does the value of R estimated by factorization compare to the ideal case? The shape matrix S is defined relative to a coordinate frame with origin at the centroid of P and defined up to a scale factor. Again, with knowledge of any dimension in the image, this scale factor can be recovered. Devise an appropriate method of comparing the dense depth map you recovered in Question 1 with the sparse point set recovered here. Is the shape of the Pepsi can qualitatively correct? Which approach is more "accurate"? Explain any discrepancies in your results.

fpf/Nov. 11, 2016