CS 223b Homework 3 (Due Monday, February 11th)

J. Košecká

e-mail your answers to cs223b@gmail.com

1. **Estimating Planar Essential Matrix.** Suppose that you have two views of a scene captured by a camera which undergoes a planar motion. Assume for simplicity that the optical axis of the camera is z, y-axis is pointing down and x-axis is pointing to the right. In this setting the camera moves in x-z plane and can rotate around y-axis.

   (a) Write down the essential matrix corresponding to the planar motion , i.e. write down the individual entries of the matrix as a function of motion parameters.

   (b) What is the minimal number of corresponding points needed to estimate the planar essential matrix ?

   (c) Given the planar essential matrix come up with a simplified way to decompose it into rotation and translation (just using basic trigonometry).

   Solution: General epipolar constraint has the following form

   $$\mathbf{x}_2^T \widehat{T} R \mathbf{x}_1 = \mathbf{x}_2^T E \mathbf{x}_1 = 0, \tag{1}$$

   where $E = \widehat{T}R$ is the essential matrix[1] In case of planar motion, assuming translation in $x - z$ plane and rotation around $y-$axis by an angle θ, the essential matrix has the follow ing sparse form

   $$E = \begin{bmatrix} 0 & -t_z & 0 \\ t_z c\theta + t_1 s\theta & 0 & t_z s\theta - t_1 c\theta \\ 0 & t_x & 0 \end{bmatrix} \tag{2}$$

   where $s\theta (c\theta)$ denote $\sin\theta (\cos\theta)$ respectively. Given at least four point correspondences, the elements of the essential matrix $[e_1, e_2, e_3, e_4]^T$ can be obtained a s a least squares solution of a system of homogeneous equations of the form (1). Once the essential matrix E has been recovered, the four different solutions for θ and $T = \pm[t_x, 0, t_z]$ can be obtained (using basic trigonometry) directly from the parametrization of the essential matrix. The physically correct solution is then obtained using the pos itive depth constraint.

2. **Camera calibration with a 3D known object.** Download the file hw3-dataset-1.mat from the course web-page. The file contains 3-D coordinates $\mathbf{X}$ of n points in the world coordinate frame and their image (pixel) coordinates in $\mathbf{x}$. Write a Matlab function, which computes the relative displacement $(R, T) \in SE(3)$ between the world coordinate frame and the camera frame and the intrinsic camera parameter matrix $K \in SL(3)$, used to generate this data set. Hand in the result for this particular data set, as well as printout of the Matlab function used to compute it. The function can be called as follows:
   `[R,T,K] = calibration2Dto3D(X,x)`

   Solutions:

   $$R = \begin{bmatrix} 0.3588 & -0.4113 & 0.8379 \\ 0.7972 & 0.6019 & -0.0459 \\ -0.4855 & 0.6845 & 0.5438 \end{bmatrix} \quad T = \begin{bmatrix} 0.1000 \\ 0.1000 \\ 15.0000 \end{bmatrix}$$

   and K is identity.

[1] $\widehat{T}$ denotes a 3×3 skew symmetric matrix associated with vector T.