CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Computing@CIT

MACHINE VISION

Assignment 2:
Scene reconstruction

DUE DATE

This assignment should be submitted to Canvas before 11:59pm on **Friday 08/05/2020**.

Please submit a single ZIP file with your student number and name in the filename. Your submission should contain:

- A detailed documentation of all code you developed, including the tests and evaluations you carried out. Please make sure that you include a document with every result image you produce referencing the exact subtask and lines of code.

- All Python code you developed in a single .py file that can be executed and that generates the outputs you are referring to in your evaluation. Please make sure that you clearly indicate in your comments the exact subtask every piece of code is referring to.

Please do NOT include the input video and image files.

You can achieve a total of 50 points as indicated in the tasks.

TASK 1 (pre-processing, 8 *points*)

A. Download the file **Assignment_MV_02_calibration.zip** from Canvas and load all calibration images contained in this archive. Extract and display the checkerboard corners to subpixel accuracy in all images using the OpenCV calibration tools [3 points].
B. Determine and output the camera calibration matrix $\boldsymbol{K}$ using the OpenCV calibration tools [1 point].
C. Download the file **Assignment_MV_02_video.mp4** from Canvas and open it for processing. Identify good features to track in the first frame [1 point] using the OpenCV feature extraction and tracking functions. Refine the feature point coordinates to sub-pixel accuracy [1 point].

D. Use the OpenCV implementation of the KLT algorithm to track these features across the whole image sequence [1 point]. Make sure to refine the feature point coordinates to sub-pixel accuracy [1 point] in each step.

TASK 2 (Fundamental matrix, *24 points*)

A. Extract and visualise the feature tracks calculated in task 1 which are visible in both the first and the last frame to establish correspondences $\boldsymbol{x_i} \leftrightarrow \boldsymbol{x}'_{\boldsymbol{i}}$ between the two images [2 points]. Use Euclidean normalised homogeneous vectors.

B. Calculate the mean feature coordinates $\boldsymbol{\mu} = \frac{1}{N}\sum_i \boldsymbol{x_i}$ and $\boldsymbol{\mu}' = \frac{1}{N}\sum_i \boldsymbol{x_i}'$ in the first and the last frame [2 points]. Also calculate the corresponding standard deviations $\sigma = \sqrt{\frac{1}{N}\sum_i(\boldsymbol{x_i} - \boldsymbol{\mu})^2}$ and $\sigma' = \sqrt{\frac{1}{N}\sum_i(\boldsymbol{x}'_{\boldsymbol{i}} - \boldsymbol{\mu}')^2}$ (where $(\,)^2$ denotes the element-wise square) [2 points]. Normalise all feature coordinates and work with $\boldsymbol{y_i} = \boldsymbol{T}\boldsymbol{x_i}$ and $\boldsymbol{y}'_{\boldsymbol{i}} = \boldsymbol{T}'\boldsymbol{x}'_{\boldsymbol{i}}$ which are translated and scaled using the homographies [2 points]

$$T = \begin{pmatrix} \frac{1}{\sigma_x} & 0 & -\mu_x/\sigma_x \\ 0 & \frac{1}{\sigma_y} & -\mu_y/\sigma_y \\ 0 & 0 & 1 \end{pmatrix} \qquad T' = \begin{pmatrix} \frac{1}{\sigma'_x} & 0 & -\mu'_x/\sigma'_x \\ 0 & \frac{1}{\sigma'_y} & -\mu'_y/\sigma'_y \\ 0 & 0 & 1 \end{pmatrix}$$

C. Select eight feature correspondences at random [1 point] and build a matrix comprising the eight corresponding rows $\boldsymbol{a}_{\boldsymbol{i}}^{\boldsymbol{T}} = \boldsymbol{y}_{\boldsymbol{i}}^{\boldsymbol{T}} \otimes \boldsymbol{y_i}'$ to calculate the fundamental matrix using the 8-point DLT algorithm [1 point].

D. Use the 8-point DLT algorithm to calculate the fundamental matrix $\widehat{\boldsymbol{F}}$ for the eight selected normalised correspondences $\boldsymbol{y_i} \leftrightarrow \boldsymbol{y}'_{\boldsymbol{i}}$ [1 point]**.** Make sure that $\widehat{\boldsymbol{F}}$ is singular [1 point]. Apply the normalisation homographies to $\widehat{\boldsymbol{F}}$ to obtain the fundamental matrix $\boldsymbol{F} = \boldsymbol{T}'^{\boldsymbol{T}}\widehat{\boldsymbol{F}}\,\boldsymbol{T}$ [1 point]**.**

E. For the remaining feature correspondences $\boldsymbol{x_i} \leftrightarrow \boldsymbol{x}'_{\boldsymbol{i}}$ not used in the 8-point algorithm calculate the value of the model equation [1 point]

$$g_i = \boldsymbol{x}'^{\boldsymbol{T}}_{\boldsymbol{i}}\boldsymbol{F}\boldsymbol{x_i}$$

Also calculate the variance of the model equation [1 point]

$$\sigma_i^2 = \boldsymbol{x}'^{\boldsymbol{T}}_{\boldsymbol{i}}\boldsymbol{F}\boldsymbol{C_{xx}}\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{x}'_{\boldsymbol{i}} + \boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{T}}\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{C_{xx}}\boldsymbol{F}\boldsymbol{x_i}$$

Use the following point observation covariance matrix of the homogeneous features

$$\boldsymbol{C_{xx}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

F. Determine for each of these correspondences if they are an outlier with respect to the selection of the eight points or not by calculating the test statistic [1 point]

$$T_i = \frac{g_i^2}{\sigma_i^2}$$

Use an outlier threshold of $T_i > 6.635$ [1 point]. Sum up the test statistics over all inliers [1 point].

G. Repeat the above procedure 10000 times for different random selections of correspondences [1 point]. Select the fundamental matrix and remove all outliers for the selection of eight points which yielded the least number of outliers [1 point]. Break ties by looking at the sum of the test statistic over the inliers [1 point].
H. Adapt the display of feature tracks implemented in subtask A to indicate which of these tracks are inliers and which tracks are outliers [1 point]. Also calculate and output the coordinates of the two epipoles? [2 points]

TASK 3 (Essential matrix and 3d points, *18 points*)

A. Use the fundamental matrix $\boldsymbol{F}$ determined in task 2 and the calibration matrix $\boldsymbol{K}$ determined in task 1 to calculate the essential matrix $\boldsymbol{E}$ [1 point]. Make sure that the non-zero singular values of $\boldsymbol{E}$ are identical [1 point]. Also make sure that the rotation matrices of the singular value decomposition have positive determinants [1 point].
B. Determine the four potential combinations of rotation matrices $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$ between the first and the last frame [3 points]. Assume the camera was moving at 50km/h and that the video was taken at 30fps to determine the scale of the baseline $\boldsymbol{t}$ in meters [1 point].
C. Calculate for each inlier feature correspondence determined in task 2 and each potential solution calculated in the previous subtask the directions $\boldsymbol{m}$ and $\boldsymbol{m'}$ of the 3d lines originating from the centre of projection towards the 3d points [1 point]

$$\boldsymbol{X}[\lambda] = \lambda \boldsymbol{m}$$

and

$$\boldsymbol{X}[\mu] = \boldsymbol{t} + \mu \boldsymbol{R}\boldsymbol{m'}$$

Then calculate the unknown distances λ and μ by solving the linear equation system [1 point]

$$\begin{pmatrix} \boldsymbol{m}^T\boldsymbol{m} & -\boldsymbol{m}^T\boldsymbol{R}\boldsymbol{m'} \\ \boldsymbol{m}^T\boldsymbol{R}\boldsymbol{m'} & -\boldsymbol{m'}^T\boldsymbol{m'} \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \boldsymbol{t}^T\boldsymbol{m} \\ \boldsymbol{t}^T\boldsymbol{R}\boldsymbol{m'} \end{pmatrix}$$

to obtain the 3d coordinates of the scene points [1 point]. Determine which of the four solutions calculated in the previous subtask is correct by selecting the one where most of the scene points are in front of both frames, i.e. where both distances $\lambda > 0$ and $\mu > 0$ [1 point]. Discard all points, which are behind either of the frames for this solution as outliers [1 point].
D. Create a 3d plot to show the two camera centres and all 3d points [2 points].
E. Project the 3d points into the first and the last frame [2 points] and display their position in relation to the corresponding features to visualise the reprojection error [2 points].