Exercise 3 Solution

Qiao Qiao

Robotics, Cognition, Intelligence - Technische Universität München

1 Exercise 2: Essential Matrix

Derive the computation of the essential matrix $x_L^T E x_R = 0$ from the transformation between two cameras.

λ is depth value. X is 3D point in right camera frame. R is rotation matrix from right camera frame to left camera frame. T is the translation vector from right camera frame to left camera frame. x is the projection in homogeneous coordinates.

Left camera: $\lambda_L x_L = RX + T$ Right camera: $\lambda_R x_R = X$

Eliminating X: $\lambda_L x_L = R(\lambda_R x_R) + T$

$\hat{T}$ is the skew-matrix of T ($\hat{T}v \equiv T \times v$)

Multiply $\hat{T}$ on the left side, since $\hat{T}T = 0$

$\lambda_L \hat{T} x_L = \lambda_R \hat{T} R x_R$

Multiply x_L^T on the left side, since $x_L^T \hat{T} x_L = 0$

$x_L^T \hat{T} R x_R = 0$

So, $E = \hat{T}R$

2 Exercise 4: Bag-of-Words for Place Recognition

- What is the main difference between match_all and match_bow functions in src/sfm.cpp?
  In match_all, the algorithm compare all possible image pair combinations except the image pairs at the same time step.
  In match_bow, for every image, at first, the image is converted to bow vector. And then find num_bow_candidate closest matches to the bow_vector in the inverted index. These candidate image matches are later used to calculate if some corner in them are really match and if these corner matches are inliers. And then the bow_vector that corresponds to this frame and this camera was added to the inverted index.

- What does the num_bow_candidates parameter control?
  This parameter controls how many closest candidate matches to the bow_vector in the inverted index.

- In our case we have 2 × 82 images. What would be the number of candidate pairs for 2 × 1000 images for match_all and match_bow functions?

  In our case, when using match_all, there are 13284 candidate pairs.
  It is calculated by: $\sum_{i=1}^{2\times 82} i - 82 = 13366 - 82 = 13284$
  When using 2 × 1000 image pairs: $\sum_{i=1}^{2\times 1000} i - 1000 = 1999000 - 1000 = 1998000$.
  There will be 1998000 candidate pairs.

  In our case, when using match_bow with 25 bow candidates, there are 3649 image pairs.
  The number of image pairs depend on not only the number of images but also the number of bow candidates. The number of candidate pairs is smaller than number of images × number of bow candidates.
  $3649 < 82 \times 2 \times 25 = 4100$

  When using 2 × 1000 image pairs with 25 bow candidates. The number of candidate pairs is $3649 < 1000 \times 2 \times 25 = 50000$, which is much smaller than using match_all (1998000).