

# Exercise 2 - Theory

Abdelaziz, Ibrahim

Somkiadcharoen, Robroo

Berg, Oliver

December 14, 2017

## 1 Theory

### 1.1 Homography Definition

In case of  $P^2$  (2D Projection plane) we have homogeneous coordinates as  $[x_1, x_2, x_3]$  and the  $H$  transformation matrix is of size  $3 \times 3$  as  $\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$  where 8 of them are independent ratios(DOF) and another one is the gain.

Using the same logic as above, you can get a point from  $P^n$  (n-dimensional Projection Space) as  $[x_1, x_2, \dots, x_n, x_{n+1}]$  and the  $H$  transformation matrix is of size  $(n+1) \times (n+1)$ . Thus,  $(n+1)^2 - 1$  DOF. Motivated From [1][2]

### 1.2 Line preservation

Given that a point  $x = [x_1, x_2, x_3]$  is a point in 2D Projection plane which is also on a line  $l$ , and all the points are on  $l$  which gives  $l^T x_i = 0$ . We can derive

$$l^T x_i = 0 = l^T H^{-1} H x_i \quad (1)$$

From (1) we get that the points  $x' = H x_i$  that is transformed lie on the line  $l' = l^T H^{-1}$ . In other words, we can perceive from the equation that points are transformed by  $x'$  and line is transformed by  $l'$

Highly Motivated by [3] [4]

## 2 Implementation

You may find the implementation code inside the `main.py` file. To run the code, call `python main.py`.

### 2.1 Relative rotation estimation from homography

To 4. : As stated in the Exercise "H2 was computed after manually rotating the camera", since rotating the camera manually would cause some translation which we neglected in solving for  $R2$ , then we will need to do correction to it.

### 2.2 Camera pose estimation from homography

To 2. "meaning of  $t = -RC$ ": Motivated by [5], because the chessboard lies within the xy-plane, it holds

$$[K|0_3] \begin{bmatrix} R & -RC \\ 0_3^T & 1 \end{bmatrix} (XYZ1)^T \quad (2)$$

being subject to  $Z = 0$  corresponds to

$$K[r_1 r_2 t](XY1)^T \quad (3)$$

where  $t$  now responds to the initial  $-RC$ . We translate the virtual camera to the real camera.

To 3. "3rd element of  $t$  being negative": With the translation being negative, this means the camera resides within the positive X-Y-Z quadrant of the world coordinate system, such that we need to do negative translation to align the origin of the camera coordinate system with the origin of the world coordinate system.

## References

- [1] Homogeneous matrix has eight independent ratios of matrix elements? [Online]. Available: <https://stackoverflow.com/questions/9534453/homogeneous-matrix-has-eight-independent-ratios-of-matrix-elements>
- [2] Dynamic pn to pn alignment. [Online]. Available: <https://www.cs.tau.ac.il/~wolf/papers/dyn-alignment.pdf>
- [3] Projective geometry. [Online]. Available: <http://www.umiacs.umd.edu/~ramani/cmsc828d/ProjectiveGeometry.pdf>
- [4] Projective geometry, camera models and calibration. [Online]. Available: <http://www.cse.iitd.ernet.in/~suban/vision/geometry.pdf>
- [5] Camera calibration. [Online]. Available: <http://www.epixea.com/research/multi-view-coding-thesis9.html>