3D Computer Vision Homework 1

Spring Semester, 2015

Bodong Zhang u0949206

3D- Computer Vision

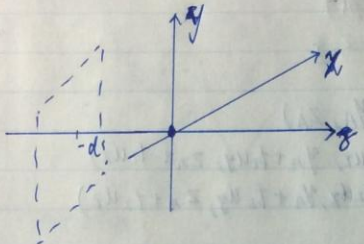
Bodong Zhang
u0949206

Assignment 1

Problem 1 : Pinhole Camera

a Assume pinhole ~~center~~ locates at original point.

Projection plane is $z=-d$.



Assume straight line is : $(x_0+u_x t, y_0+u_y t, z_0+u_z t)$ $-\infty<t<\infty$

point (x,y,z) projected to $(-\frac{d}{z}x, -\frac{d}{z}y, -d)$

So straight line is projected to $\left(-\frac{x_0+u_x t}{z_0+u_z t}d, -\frac{y_0+u_y t}{z_0+u_z t}d, -d\right)$

① if $u_z=0$, Projection is $\left(-\frac{x_0}{z_0}d+\frac{u_x}{z_0}d\,t, -\frac{y_0}{z_0}d+\frac{u_y}{z_0}d\,t, -d\right)$

So it is a straight line that goes across $\left(-\frac{x_0}{z_0}d, -\frac{y_0}{z_0}d, -d\right)$

and its direction is $\left(\frac{u_x}{z_0}d, \frac{u_y}{z_0}d, 0\right)$

② if $u_z\neq 0$, Projection is $\left(\frac{\frac{u_x}{u_z}(z_0+u_z t)-\frac{u_x}{u_z}z_0+x_0}{z_0+u_z t}d, \frac{\frac{u_y}{u_z}(z_0+u_z t)-\frac{u_y}{u_z}z_0+y_0}{z_0+u_z t}d, -d\right)$

$$=\left(\frac{d u_x}{u_z}+\frac{(-\frac{u_x}{u_z}z_0+x_0)d}{z_0+u_z t}, \frac{d u_y}{u_z}+\frac{(-\frac{u_y}{u_z}z_0+y_0)d}{z_0+u_z t}, -d\right)$$

set $\frac{1}{z_0+u_z t}=t'$ because $-\infty<t<\infty$, so $-\infty<t'<\infty$

$$=\left(\frac{u_x}{u_z}d+(-\frac{u_x}{u_z}z_0+x_0)d\,t', \frac{u_y}{u_z}d+(-\frac{u_y}{u_z}z_0+y_0)d\,t', -d\right)$$

So it is also a straight line that goes across $\left(\frac{u_x}{u_z}d, \frac{u_y}{u_z}d, -d\right)$

and its direction is $\left(-\frac{u_x d}{u_z}z_0+x_0 d, -\frac{u_y}{u_z}d z_0+y_0 d, 0\right)$

Problem b

$(x, y, z) \rightarrow \left(f\frac{x}{z}, f\frac{y}{z}, f\right)$

Assume A, B, C are three collinear points

so

assume $A: (x_A, y_A, z_A)$

$B: (x_A + t_1 u_x,\ y_A + t_1 u_y,\ z_A + t_1 u_z)$

$C: (x_A + t_2 u_x,\ y_A + t_2 u_y,\ z_A + t_2 u_z)$

After Projection:

$$A' = \left(f\frac{x_A}{z_A},\ f\frac{y_A}{z_A},\ f\right) = (x_A', y_A', f)$$

$$B' = \left(f\frac{x_A + t_1 u_x}{z_A + t_1 u_z},\ f\frac{y_A + t_1 u_y}{z_A + t_1 u_z},\ f\right) = \left(f\frac{\frac{x_A}{z_A}(z_A + t_1 u_z) - \frac{x_A}{z_A} t_1 u_z + t_1 u_x}{z_A + t_1 u_z} \dots\right)$$

$$C' = \left(f\frac{x_A + t_2 u_x}{z_A + t_2 u_z},\ f\frac{y_A + t_2 u_y}{z_A + t_2 u_z},\ f\right)$$

$$B' = \left(f\frac{x_A}{z_A} + f\frac{t_1}{z_A + t_1 u_z}\left(-\frac{x_A}{z_A}u_z + u_x\right),\ f\frac{y_A}{z_A} + f\frac{t_1}{z_A + t_1 u_z}\left(-\frac{y_A}{z_A}u_z + u_y\right),\ f\right)$$

$$C' = \left(f\frac{y_A}{z_A} + f\frac{t_2}{z_A + t_2 u_z}\left(-\frac{x_A}{z_A}u_z + u_x\right),\ f\frac{y_A}{z_A} + f\frac{t_2}{z_A + t_2 u_z}\left(-\frac{y_A}{z_A}u_z + u_y\right),\ f\right)$$

set $t_1' = \frac{t_1}{z_A + t_1 u_z} = \frac{1}{\frac{z_A}{t_1} + u_z}$ $\quad -\infty < t_1 < \infty \quad \therefore \quad -\infty < t_1' < +\infty$

set $t_2' = \frac{t_2}{z_A + t_2 u_z}$ $\quad$ where $-\infty < t_2' < +\infty$

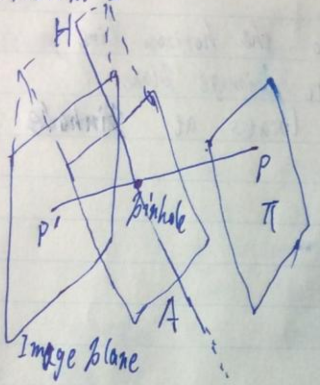
$\therefore A' = (x_A', y_A', f) \quad B' = \left(x_A' + t_1' f\left(-\frac{x_A}{z_A}u_z + u_x\right),\ y_A' + t_1' f\left(-\frac{x_A}{z_A}u_z + u_y\right), f\right)$

$C' = \left(x_A' + t_2' f\left(-\frac{x_A}{z_A}u_z + u_x\right),\ y_A' + t_2' f\left(-\frac{x_A}{z_A}u_z + u_x\right), f\right)$

$\therefore$ The three points are still collinear after projection.

Problem 2

a.



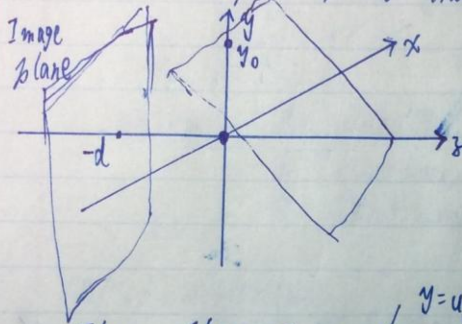
The convergence point of two parallel lines on plane π locates at infinity.
Assume p' is the projection point of P.
So PP' must parallel to plane π and go across pinhole.

So PP' must locates at the plane A that parallel to π and pass through the pin hole, so P' is on plane A. Also p' is on Image plane.

So p' locates at the intersection of plane A and Image plane, that is staight line H.

Problem 2

b Assume x-axis is parallel to the horizon line,
choose y-axis that parallel to image plane
choose original point that locates at pinhole



Then plane π : $\begin{cases} y = u_y t_1 + y_0 \\ z = u_z t_1 \end{cases}$ $-\infty < t_1 < +\infty$

point (x, y, z) on the plane π will be projected to
$(-\frac{d}{t_1 u_z} x, -\frac{d}{u_z t_1}(u_y t_1 + y_0), -d)$

Two parallel lines converge at infinity, so $t_1 = \infty$ or $x = \infty$

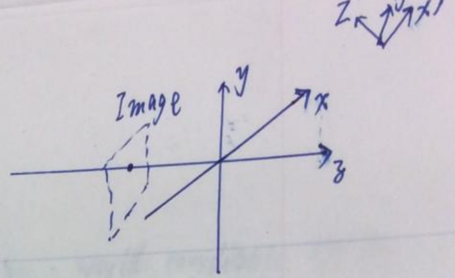
① $t_1 = \infty$, x is finite
then converge at $(0, -\frac{d u_y}{u_z}, -d)$

② t_1 is finite, $x = \infty$,
then converge at ∞

③ $x = k t_1 \to \infty$ k can be any real number
then converge at $(-\frac{d}{u_z} k, -\frac{d u_y}{u_z}, -d)$
if $u_z \neq 0$, then $-\frac{d}{u_z} k$ can be any real number

Besides, the plane parallel to π and passing through pinhole
is $\begin{cases} y = u_y t_1 \\ z = u_z t_1 \end{cases}$ $-\infty < t_1 < +\infty$ the intersection with image plane
is $(m, -\frac{d u_y}{u_z}, -d)$ $-\infty < m < +\infty$ prove: $u_z t_1 = -d$ $t_1 = -\frac{d}{u_z}$
$y = u_y(-\frac{d}{u_z}) = -\frac{d u_y}{u_z}$

It is the same line as ③
So converge on the horizon line H



Problem 3

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z}\begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ Z \\ 1 \end{pmatrix}$$

where
$$\begin{pmatrix} x \\ y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} & t_x \\ R_{21} & R_{22} & R_{23} & t_y \\ R_{31} & R_{32} & R_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ Z' \\ 1 \end{pmatrix}$$

$$M = \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} & R_{13} & t_x \\ R_{21} & R_{22} & R_{23} & t_y \\ R_{31} & R_{32} & R_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad O = \begin{pmatrix} x' \\ y' \\ Z' \\ 1 \end{pmatrix}$$

Because $\begin{pmatrix} x \\ y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ at the original point (optical center)

$$\therefore \; MO = \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

Problem 4

a No matter how far or close the scene is, pinhole camera can project a clear image, so it has infinite depth of focus

b $\Delta Z_0^- = \dfrac{Z_0(Z_0-f)}{Z_0+f\frac{d}{b}-f}$

$$-\delta z = \frac{z(z-f)}{z+f\frac{d}{b}-f}$$

$$f = \frac{zz'}{z+z'}$$

$$\therefore \quad z+f\frac{d}{b}-f = \frac{z(z-f)}{-\delta z}$$

$$\frac{z}{f} = \frac{z+z'}{z'}$$

$$f\frac{d}{b} = \frac{z(z-f)}{-\delta z} + f - z$$

$$b = \frac{fd}{\frac{z(z-f)}{-\delta z}+f-z} = \frac{d}{\frac{z(\frac{z}{f}-1)}{-\delta z}+1-\frac{z}{f}} = \frac{d}{\frac{z(\frac{z}{z'})}{-\delta z}+1-1-\frac{z}{z'}}$$

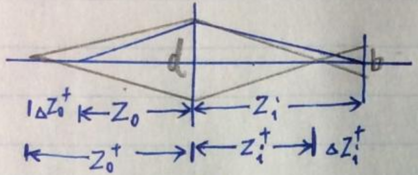
$$b = \frac{d}{-\frac{z}{z'}\left(\frac{z}{\delta z}+1\right)}$$

The diameter of blur circle is $\left|\dfrac{d}{-\frac{z}{z'}\left(\frac{z}{\delta z}+1\right)}\right|$

d is the diameter of lens

Problem 4

c



$$Z_o^+ = f\,\frac{Z_i^+}{Z_i^+ - f} \qquad Z_i = f\,\frac{Z_o}{Z_o - f}$$

$$Z_i^+ = Z_i\,\frac{d}{d+b} = f\,\frac{Z_o}{Z_o - f}\,\frac{d}{d+b}$$

$$Z_o^+ = f\,\frac{f\frac{Z_o}{Z_o-f}\frac{d}{d+b}}{f\frac{Z_o}{Z_o-f}\frac{d}{d+b} - f} = f\,\frac{\frac{Z_o}{Z_o-f}d}{\frac{Z_o}{Z_o-f}d - (d+b)} = f\,\frac{Z_o d}{Z_o d - (d+b)(Z_o - f)}$$

$$= f\,\frac{Z_o d}{Z_o d - Z_o d - bZ_o + fd + fb} = f\,\frac{Z_o d}{f(d+b) - bZ_o}$$

$$\Delta Z_o^+ = Z_o^+ - Z_o$$

$$= \frac{fZ_o d - Z_o f(d+b) + bZ_o^2}{f(d+b) - bZ_o} = \frac{bZ_o^2 - Z_o f b}{f(d+b) - bZ_o}$$

$$= \frac{Z_o(Z_o - f)}{f\left(\frac{d}{b} + 1\right) - Z_o}$$

Practical Problem

Introduction:

In this experiment, the purpose is to calibrate a camera and then access the experiment results. A point has different coordinates in the world coordinate system and in the camera coordinate system. But by doing camera calibration, we can find the connections between them so that we can predict the location of a point in image given its coordinates in the world coordinate system.

Algorithm:

First, the location of a point is relevant to the coordinates in camera frame, described as follows.[1]

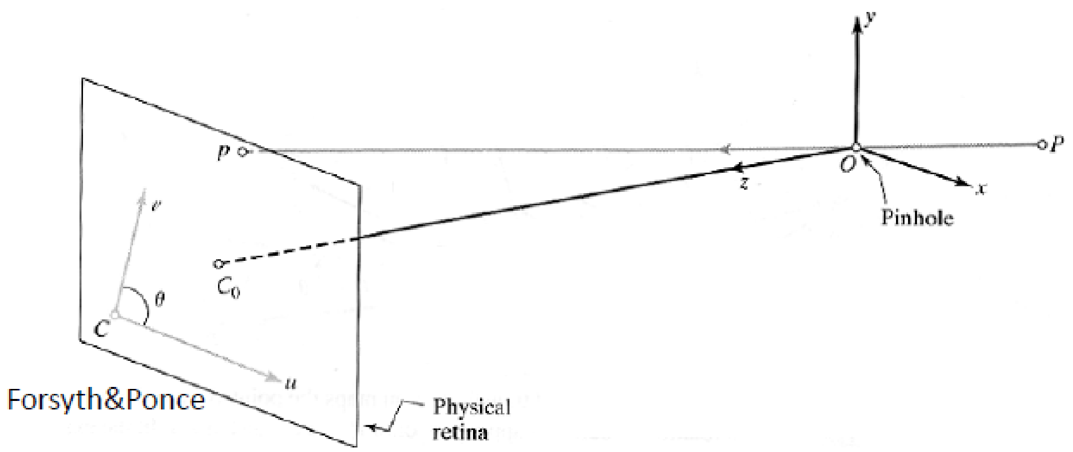


Figure 1: camera frame and image

$$u=f\frac{x}{z}$$

$$v=f\frac{y}{z}$$

Because u, v are measured by cells and x, y, z are measured by mm, so we change f into α and β where α equals to f*k1 and β equals to f*k2. k1 is the number of pixels in one millimeter in horizon direction and k2 is the number of pixels in one millimeter in vertical direction. Also, we don't know the original center in image, so we add u_0 and v_0 to translate the original point. As a result, the formula is described as follows:

$$u=\alpha \frac{x}{z}+u_0$$

$$v=\beta \frac{y}{z}+v_0$$

Moreover, we consider that the u axis and v axis are not exactly perpendicular, so we assume the angle between them is θ. Thus, we get

$$u=\alpha \frac{x}{z}-\alpha \cot(\theta) \frac{y}{z}+u_0$$

$$v=\frac{\beta}{\sin(\theta)} \frac{y}{z}+v_0$$

In sum,

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \alpha & -\alpha\cot(\theta) & u_0 & 0 \\ 0 & \dfrac{\beta}{\sin(\theta)} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

The camera frame is different with world frame, so we need the rotation and translation matrix to transform.

$${}^{C}\vec{p} = {}^{C}_{W}R\ {}^{W}\vec{p} + {}^{C}_{W}\vec{t}$$

$$\begin{pmatrix} {}^{C}\vec{p} \\ \\ \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}R & {}^{C}_{W}\vec{t} \\ 0\ \ 0\ \ 0 & 1 \end{pmatrix} \begin{pmatrix} {}^{W}\vec{p} \\ \\ \end{pmatrix}$$

By combining these two matrixes into M, p=$\frac{1}{z}$MP, M is matrix, p and P are vectors that show the location in image and world frame, separately.

$$\mathcal{M} = \begin{pmatrix} \boldsymbol{m}_1^T \\ \boldsymbol{m}_2^T \\ \boldsymbol{m}_3^T \end{pmatrix} \Longrightarrow z = \boldsymbol{m}_3 \cdot \boldsymbol{P}, \quad \text{or} \quad \begin{cases} u = \dfrac{\boldsymbol{m}_1 \cdot \boldsymbol{P}}{\boldsymbol{m}_3 \cdot \boldsymbol{P}}, \\ v = \dfrac{\boldsymbol{m}_2 \cdot \boldsymbol{P}}{\boldsymbol{m}_3 \cdot \boldsymbol{P}}. \end{cases}$$

In experiment, we have many pairs of p and P. After rewriting the formula, the equations are as follows[1]:

$$\underbrace{\begin{pmatrix} P_{1x} & P_{1y} & P_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1P_{1x} & -u_1P_{1y} & -u_1P_{1z} & -u_1 \\ 0 & 0 & 0 & 0 & P_{1x} & P_{1y} & P_{1z} & 1 & -v_1P_{1x} & -v_1P_{1y} & -v_1P_{1z} & -v_1 \\ & & & & & & \cdots & \cdots & \cdots & & & \\ P_{nx} & P_{ny} & P_{nz} & 1 & 0 & 0 & 0 & 0 & -u_nP_{nx} & -u_nP_{ny} & -u_nP_{nz} & -u_n \\ 0 & 0 & 0 & 0 & P_{nx} & P_{ny} & P_{nz} & 1 & -v_nP_{nx} & -v_nP_{ny} & -v_nP_{nz} & -v_n \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix}}_{m} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

$$Q \qquad m = 0$$

The constraint is that m is unit vector. This is overconstrained problem and we want to make sure ||Qm||is as small as possible.

In the slides, the algorithm is by using SVD to get best m.

```
%Perform SVD of P
[U S V] = svd(P);
[min_val, min_index] = min(diag(S(1:12,1:12)));

%m is given by right singular vector of min. singular value
m = V(1:12, min_index);
```

Figure 2: Program of Performing SVD[1]

After getting matrix m, we need to extract intrinsic and extrinsic parameters from matrix m.

The algorithm[2] is

Estimation of the intrinsic and extrinsic parameters

Write $M = (\mathcal{A}, \boldsymbol{b})$, therefore $\rho(\mathcal{A} \quad \boldsymbol{b}) = \mathcal{K}(\mathcal{R} \quad \boldsymbol{t}) \Longleftrightarrow \rho \begin{pmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \boldsymbol{a}_3^T \end{pmatrix} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T - \alpha \cot\theta \boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T \\ \frac{\beta}{\sin\theta} \boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T \\ \boldsymbol{r}_3^T \end{pmatrix}$

Using the fact that the rows of a rotation matrix have unit length and are perpendicular to each other yields

$$\begin{cases} \rho = \varepsilon / |\boldsymbol{a}_3|, \\ \boldsymbol{r}_3 = \rho \boldsymbol{a}_3, \\ u_0 = \rho^2 (\boldsymbol{a}_1 \cdot \boldsymbol{a}_3), \\ v_0 = \rho^2 (\boldsymbol{a}_2 \cdot \boldsymbol{a}_3), \end{cases} \quad \text{where} \quad \varepsilon = \mp 1.$$

Since θ is always in the neighborhood of $\pi/2$ with a positive sine, we have

$$\begin{cases} \rho^2 (\boldsymbol{a}_1 \times \boldsymbol{a}_3) = -\alpha \boldsymbol{r}_2 - \alpha \cot\theta \boldsymbol{r}_1, \\ \rho^2 (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \dfrac{\beta}{\sin\theta} \boldsymbol{r}_1, \end{cases} \quad \text{and} \quad \begin{cases} \rho^2 |\boldsymbol{a}_1 \times \boldsymbol{a}_3| = \dfrac{|\alpha|}{\sin\theta}, \\ \rho^2 |\boldsymbol{a}_2 \times \boldsymbol{a}_3| = \dfrac{|\beta|}{\sin\theta}. \end{cases}$$

Thus,

$$\begin{cases} \cos\theta = -\dfrac{(\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)}{|\boldsymbol{a}_1 \times \boldsymbol{a}_3| |\boldsymbol{a}_2 \times \boldsymbol{a}_3|}, \\ \alpha = \rho^2 |\boldsymbol{a}_1 \times \boldsymbol{a}_3| \sin\theta, \\ \beta = \rho^2 |\boldsymbol{a}_2 \times \boldsymbol{a}_3| \sin\theta, \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{r}_1 = \dfrac{\rho^2 \sin\theta}{\beta} (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \dfrac{1}{|\boldsymbol{a}_2 \times \boldsymbol{a}_3|} (\boldsymbol{a}_2 \times \boldsymbol{a}_3), \\ \boldsymbol{r}_2 = \boldsymbol{r}_3 \times \boldsymbol{r}_1. \end{cases}$$

Note that there are two possible choices for the matrix $\mathcal{R}$ depending on the value of ε.

Figure 3: Estimation of the intrinsic and extrinsic parameters

After getting matrix m, we should normalize m so that vector(m31, m32, m33) is unit vector and then follow the algorithm above.

Experiment:

In the experiment, the camera used to for calibration is iPad Air Front Camera that has fixed focal length (2.15mm). First, take a picture of two orthogonal checkerboard planes to get sets of two coordinate systems(world and camera frames) which shows the connections between them.

Figure 4: Calibration Pattern

By using “step1” program to click corner points in image, we will get samples that show connection between world frame coordinates and camera frame coordinate system. The length of little square is 28mm. The “coordinates” data has format: (u, v, x, y, z). There are 96 samples used for estimation of parameters.

After getting samples, we rewrite samples as Q matrix and use SVD algorithm stated in “algorithm” part to get best vector m.

-0.0027
8.3352e-04
7.0839e-04
0.5765
-0.0014
-0.0016
-0.0016
0.8171
-1.2289e-06
-1.3405e-06
1.1135e-06
8.6803e-04

Figure 5: vector m(unit vector)

The third step is to estimate the result by ploting ground truth points location and estimated points location deduced from m.

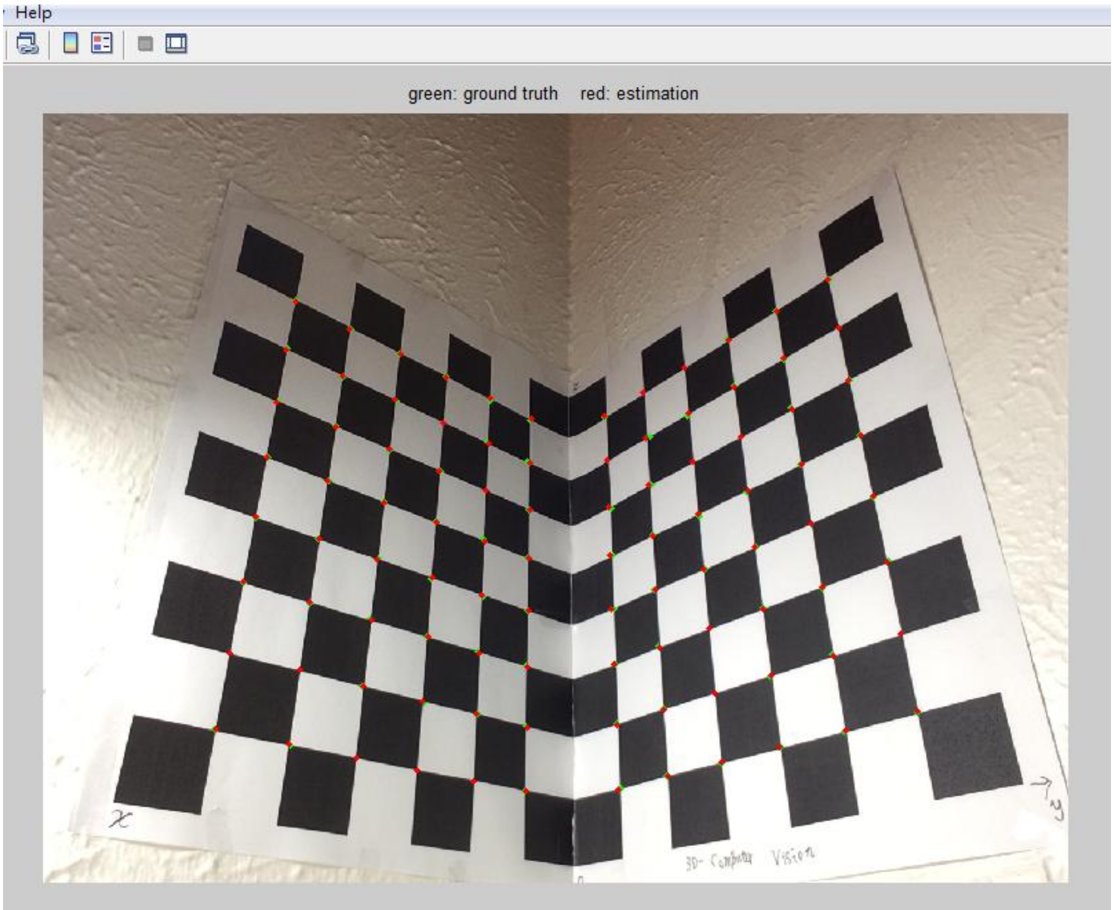


Figure 6: Plot difference between ground truth location and estimated location

Green points are ground truth by labeling, red points are estimated location deduced by world frame coordinates and m-matrix. Please notice that red points and green points are so close so that red points may cover green points. The average distance between points is 1.7241 pixels. So the estimation is very accurate.

The fourth step(also the last step) is to extract extrinsic parameters and intrinsic parameters by using Samer M Abdallah, Beirut[2] algorithm. After normalization so that $||\mathbf{m}_3||=1$, the experiment result is

$\cos(\theta)=0.0071$

$\alpha=1183.9$

$\beta=1172.6$

$u_0=646.8$

$v_0=457.0638$

extrinsic matrix

-0.7415	0.6709	-0.0108	7.1446
-0.3436	-0.3934	-0.8528	168.1127
-0.5763	-0.6286	0.5222	407.0832
0	0	0	1.0000

Intrinsic matrix

1.0e+003 *

(10^3)*

1.1839	-0.0084	0.6468	0
0	1.1726	0.4571	0
0	0	0.0010	0

Analysis:

$\cos(\theta)=0.0071$, $\theta=89.60$ and is very close to 90, so it is a reasonable answer.

The focal length is 2.15mm of front camera in iPad Air. We also assume that the size of pixel in sensor is similar to iphone 6 whose size is 1.5 um. So the α and β should be close to $2.15*\frac{1000}{1.5} \approx 1433$. The estimated α and β is 1183.9 and 1172.6, so the estimated results are plausible.

The distance of the optical center to the world coordinate origin measured by ruler is 433mm. According to the equation, $\mathbf{m}_3 \cdot (x', y', z') + t_z \approx z$. $t_z = 407.08$, set $x'=0$, $y'=0$, $z'=8*28$, $\mathbf{m}_3 \cdot (x', y', z') + t_z = 429$ and is very close to 433mm. So the estimated parameters are plausible.

References:

[1]slides in class

[2]slide Samer M Abdallah, Beirut