FINAL PROJECT

Jingwen Yang

I. SECTION 1

Obtain the intrinsic parameters of a camera

A. BIG CHECKERBOARD

Use the code developed during the assignments of VMMC-Part I to calibrate your camera, using the large checkerboard image provided. Include in your exam report the following data:

- **a.Size big checkerboard**: 168[mm] (21mm*8);
- **b.Set of images of the screen checkerboard**

Images shown in figure1. The calibrated results are shown in figure2.

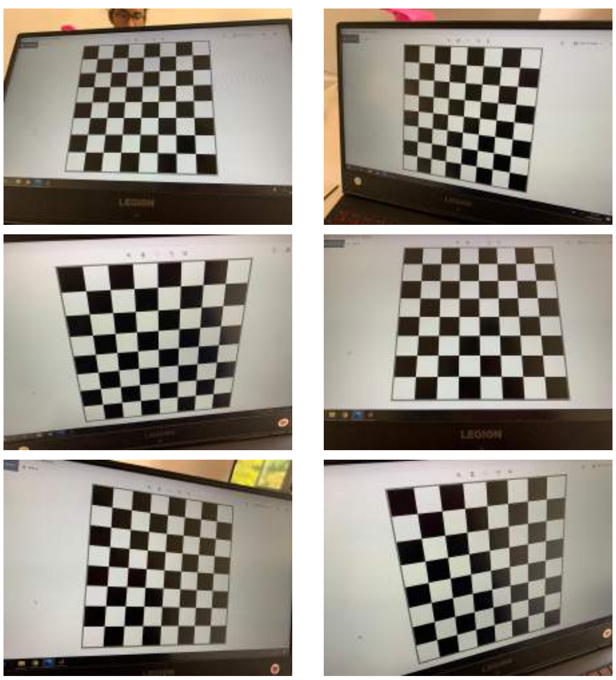


Fig. 1. 1080 x 1080 pixel checkerboard taken with a 1600*1200 resolution

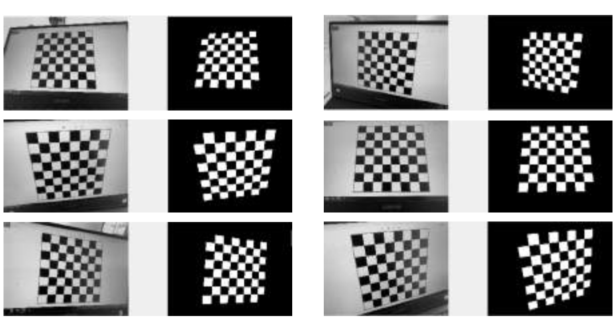


Fig. 2. 1080 x 1080 checkerboard calibrated results

- **c.Resolution of captured images**: 1600*1200[pixel].
- **d.The obtained internal matrix A**

$$A_{1080x1080} = \begin{pmatrix} 1.3081e+03 & 0.7043 & 800.3997 \\ 0 & 1.3076e+03 & 589.4936 \\ 0 & 0 & 1 \end{pmatrix}$$

- COMMENTS

To make the interpretation clear here we set the name for each number in the matrix A asI-A, where each variable are listed in the tableI-A.

$$A = \begin{pmatrix} a & c & u_0 \\ 0 & b & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

TABLE I
VARIABLES IN INTERNAL MATRIX A

Variable	Fact
a	Scale factor of the U axis
b	Scale factor of the V axis
u_0	u_0 coordinate of the principal point
v_0	v_0 coordinate of the principal point

We further interpret what is inside the internal matrixI-A, where each variable are explained in tableI-A

$$A = \begin{pmatrix} f * k_u & -f * k_u * cos(\theta) & u_0 \\ 0 & f * k_v / sin(\theta) & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

TABLE II
VARIABLES IN INTERNAL MATRIX A

Variable	Fact
f	focal length
k_u,k_v	size of pixel
θ	angle of axis of image plane
u_0,v_0	coordinates of the principal point

- **e.Are the pixels of your camera square?**

The pixels of my camera are square. Because from the matrix AI-A, we can calculate the $\theta = 90.0308°$ by equation1, therefore, $sin(\theta) = 1$, and by calculating $a/c = 1$ in equation2, we can conclude that $k_u = k_v$.

$$\theta = acos(-\frac{b}{a}) = acos(-\frac{f * k_u * \cos\theta}{f * k_u}) = 90.0308°. \quad (1)$$

$$\frac{k_v}{k_u} = \frac{a}{c} = \frac{k_v}{k_u * sin(90)} = 1 \quad (2)$$

- **f.Which is the degree of coincidence between the principal point and the center of the image plane?**

The principal point coordinates (u_0,v_0) are (800.4,589.5) according to the internal matrixI-A. Since the resolution of the image is 1600*1200, the center of the image is (800,600), therefore we can calculate the coincidence between the principal point and center of image plane as in equation345.

$$C_{row} = 1 - \frac{\mid 800.4 - 800 \mid}{800} = 99.95\% \tag{3}$$

$$C_{col} = 1 - \frac{\mid 589.5 - 600 \mid}{600} = 98.25\% \tag{4}$$

$$C_{avg} = \frac{C_{row} + C_{col}}{2} = 99.1\% \tag{5}$$

- **g.Are the axes of the image plane orthogonal?**

We have already calculated the angle between image axes θ in equation1 and got $\theta = 90.0308°$. Therefore the axes of image plane are orthogonal.

B. SMALL CHECKERBOARD

Here we do the same process on the small checkerboard.

- **Size small checkerboard:** 160[mm] (20mm*8);
- **Set of images of the screen checkerboard**

Images shown in figureI-B. The calibrated results are shown in figureI-B.

Fig. 3. 720 x 720 pixel checkerboard taken with a 4160x3120 resolution

- **The obtained internal matrix A'**

$$A_{720x720} = \begin{pmatrix} 1.2660e+03 & -8.0708 & 812.8101 \\ 0 & 1.2676e+03 & 588.7873 \\ 0 & 0 & 1 \end{pmatrix}$$

- **Comment on the theoretical and practical relationship between A and A'**:

Theoretically the two A matrices we got should be the same because they are the internal parameters of the same camera. However, practically, they are a little different which possibly because the clicking process brings noise with the participation of human labor.

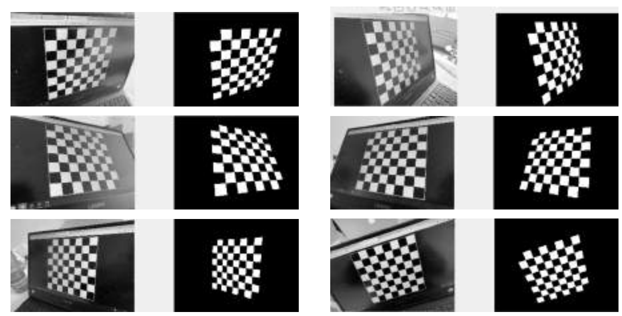


Fig. 4. 720 x 720 pixel checkerboard taken with a 1600*1200 resolution

II. SECTION 2

Finding local matches between several views of an object. For pairs of views, detect, describe, and match feature points using several of the methods explained in class. Select the detector-descriptor couple and a pair of views according to qualitative and quantitative indicators.

A. SCENE CAPTURE

Get different views with a calibrated camera. Varying the angle and distance of the camera, altering light conditions.

- **a. Provide a mosaic representation of your captured views;**

The test sequence includes four images taken in a scene of bedroom with part of the bed and half of a painting on the wall to provide more background information. The first image and the last image have a angle of around 45°. The scene is considered as having a reasonable amount of distance change and angle change.

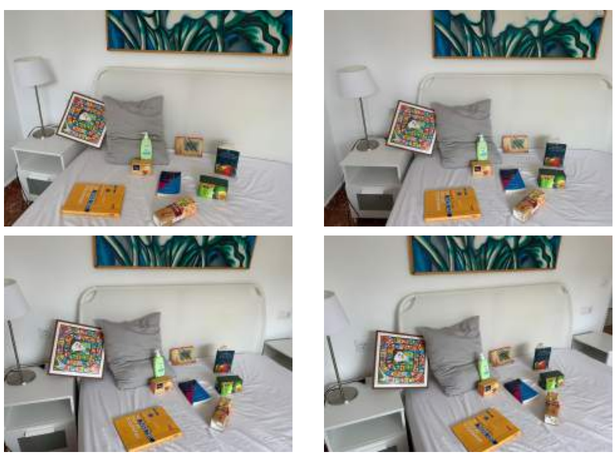


Fig. 5. Captured views

B. DETECTION, DESCRIPTION AND MATCHING OF FEATURE POINTS

Select a representative set from some pairs of views, extract and describe feature points for each view of the pair. And match points between the views. At least test DoH+SIFT, SURF+SURF, KAZE+KAZE, SIFT+DSP-SIFT. For each combination and each selected pair of views:
Here for the test sequence with four images, we are going to choose three pairs of images which are shown in the tableII-B. The choice of image combination satisfies the requirement of representativeness.

TABLE III
VARIABLES IN INTERNAL MATRIX A

Pair Number	Images	Category
PAIR1	image3&iamge4	Consecutive
PAIR2	image2&iamge4	Gap 1 Image
PAIR3	image1&iamge4	Gap 2 Images

- **a.Provide the pair of images with the correspondences overlaid on them.**
- **b.Estimate the homography transformation between the views and include the warped image.**

C. QUALITATIVE& QUANTITATIVE EVALUATION

Estimate the fundamental matrix between pair of views for every selected pair of views and every detector+description combination. For each combination and each selected pair of views:

- **a.Include the estimated fundamental matrix;**
- **b.Qualitative evaluation of the estimated fundamental matrix;**
- **c.Quantitative evaluationby accounting for the number of inliers matching.**

1) Pair 1: Since it is suggested to do the detections, descriptions, matching of feature points together with the qualitative and quantitative evaluations. Here we organize the report by image pair, with the same image pair it is easier to compare the functionality of different detector and descriptor combinations.
First we are going to use the consecutive pair of image3&image4. We will pay more attention to the four mandatory combinations, on which we will do detailed evaluation and analysis. Then we will briefly go through all the possible combinations of detectors and descriptors, just to have a basic idea of how do they work. The images with correspondences and their homography transformations are shown in table6. The four combinations of detectors and descriptors are compared together. From the transformed images we can see that all of them have almost the same performance, which are very good. Maybe it is because of the slightness of the rotation and distance change. But from the feature points distribution we can easily tell the difference between the combinations. The KAZE&KAZE created way too many feature points and thus have way more matchings. But it is computationally expensive. Also most of them focus on the game board with rich information, the other part of the images are poorly extracted. The combination of DOH&SIFT is almost the same as SURF&SURF, which is decent. However, regarding the points matching, we believe the combination of SIFT&DSP-SIFT has the best performance because all the matches are evenly distributed in every part of the images.

The fundamental matrices of each combinaiton is shown in tableIV. The last value is always 1, while the values at (1,1),(1,2),(2,1),(2,2) are always zeros.

DOH&SIFT	SURF&SURF
$\begin{bmatrix} 0.0000 & -0.0000 & 0.0024 \\ 0.0000 & 0.0000 & -0.0177 \\ -0.0033 & 0.0185 & 1.1997 \end{bmatrix}$	$\begin{bmatrix} -0.0000 & -0.0000 & -0.0005 \\ -0.0000 & 0.0000 & -0.0202 \\ -0.0000 & 0.0215 & 1.1996 \end{bmatrix}$
SIFT&DSPSIFT	**KAZE&KAZE**
$\begin{bmatrix} 0.0000 & -0.0000 & -0.0002 \\ -0.0000 & 0.0000 & -0.0210 \\ -0.0003 & 0.0224 & 1.1995 \end{bmatrix}$	$\begin{bmatrix} 0.0000 & -0.0000 & 0.0021 \\ 0.0000 & 0.0000 & -0.0181 \\ -0.0029 & 0.0193 & 1.1996 \end{bmatrix}$

TABLE IV
FUNDAMENTAL MATRICES FOR 4 MANDATORY COMBOS

(a) IMAGE3 (b) IMAGE4

(c) DOH+SIFT

(d) SIFT+DSP-SIFT

(e) SURF+SURF

(f) KAZE+KAZE

Fig. 6. Four mandatory combinations to test on pair 1

The qualitative evaluation is executed by clicking one point on one of the image and see if the epipolar line cross the same point on the other image. We can challenge the fundamental matrix by choosing those points on the region of images where they do not have many matching points. Therefore, in this case, we always tend to click on the corner of the yellow book where is lack of matching points most of the time. Due to the rotation and translation is not significant for this pair, the fundamental matrix performs well. We will further analyse on later image pairs.

The quantitative evaluation is conducted by calculating the ratio of inliers over all the matched points. Apart from that we also output the number of feature points successfully detected, and number of points matched successfully, which are be a mathematical and quantitative way to compare the performance of each combination. As shown in tableII-C1, we can see that they have almsot the same ratio of inliers, which indicates that their performance are similar, as we saw

(a) DOH+SIFT

(b) SIFT+DSP-SIFT

(c) SURF+SURF

(d) KAZE+KAZE

Fig. 7. Qualitative Evaluation for 4 Mandatory Combos,Pair1

in the warped images before. However, the SIFT detector detected way more feature points than the other detectors, the matched points are 30% more than the other combinations, which makes it have a slight better result.

Combination	points detected	pointsmatched	InliersRatio
DOH+SIFT	1809	683	0.5000
SIFT+DSPSIFT	7091	832	0.5012
SURF+SURF	1809	683	0.5007
KAZE+KAZE	1809	683	0.5007

TABLE V
QUANTITATIVE EVALUATION FOR PAIR1

The tableII-C1 lists 10 combinations we want to evaluate besides the four mandatory ones. The figure8 and figure9 are their transformation results. All the LOG, DOG, DOH, K detector are doing a great amount of computation, therefore due to technical reasons here we cannot display their best performance. With the limitation of a low number of points, these detectors still have decent performance in transformation. From the matching points we cannot justify the LOG, DOG, DOG, and K detectors. As for the combination of SIFT&SIFT, SIFT&SURF apparently have some noise, while the SURF&SIFT, SURF&DSP-SIFT are perfectly distributed and matched. These two pairs looks very promising, which will be paid more attention in the evaluation part.

Combination
LOG+SIFT
LOG+SURF
DOG+SIFT
DOG+SURF
K+SIFT
K+SURF
SIFT+SIFT
SIFT+SURF
SURF+SIFT
SURF+DSPSIFT

TABLE VI
10 OPTIONAL COMBINATIONS

For the qualitative evaluation part, as we can see in the table10 and table11, the LOG, DOG, DOH, K detectors all miss the correspondent point in the other image, which can be predicted from the poor matching points, as well as the evidence of a poor fundamental matrix.

As we can see in the tableII-C1, the matching points from LOG,DOG detectors are way less than the others. Although the K detector has as a great amount of matching points as the SIFT,SURF family, its performance is definitely not as good as the SURF&SIFT or SURF&DSP-SIFT. Due to the number of feature points, matching points is not compatible between the combinations, the InlierRatio is also not compatible. Inside the SIFT&SURF family they have the same InlierRatio, so no more information can be extarcted from InlierRatio here.

2) PAIR 2: The second pair is the second image and the fourth image in the sequence. This pair has more rotation then the first pair. Therefore it is more challenging for the detection, description and matching process. Due to the great

LOG+SIFT
$$\begin{bmatrix} 0.0000 & 0.0000 & 0.0269 \\ -0.0000 & -0.0000 & 0.0210 \\ -0.0276 & -0.0228 & 0.9988 \end{bmatrix}$$

LOG+SURF
$$\begin{bmatrix} 0.0000 & 0.0000 & 0.0262 \\ -0.0000 & -0.0000 & 0.0247 \\ -0.0264 & -0.0276 & 0.9986 \end{bmatrix}$$

DOG+SIFT
$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0114 \\ 0.0000 & 0.0000 & -0.0064 \\ -0.0130 & 0.0048 & 0.9998 \end{bmatrix}$$

DOG+SURF
$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0132 \\ 0.0000 & 0.0000 & -0.0060 \\ -0.0146 & 0.0044 & 0.9998 \end{bmatrix}$$

K+SIFT
$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0003 \\ -0.0000 & 0.0000 & -0.0224 \\ -0.0009 & 0.0242 & 0.9995 \end{bmatrix}$$

K+SURF
$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0001 \\ -0.0000 & 0.0000 & -0.0189 \\ -0.0008 & 0.0200 & 0.9996 \end{bmatrix}$$

SIFT+SIFT
$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0006 \\ 0.0000 & 0.0000 & -0.0190 \\ -0.0013 & 0.0201 & 0.9996 \end{bmatrix}$$

SIFT+SURF
$$\begin{bmatrix} 0.0000 & -0.0000 & -0.0001 \\ -0.0000 & 0.0000 & -0.0199 \\ -0.0005 & 0.0213 & 0.9996 \end{bmatrix}$$

SURF+SIFT
$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0050 \\ 0.0000 & 0.0000 & 0.0383 \\ 0.0073 & -0.0574 & 0.9976 \end{bmatrix}$$

SURF+DSPSIFT
$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0050 \\ 0.0000 & 0.0000 & 0.0383 \\ 0.0073 & -0.0574 & 0.9976 \end{bmatrix}$$

TABLE VII
FUNDAMENTAL MATRICES FOR 10 OPTIONAL COMBOS

(a) LOG+SIFT

(b) LOG+SURF

(c) DOG+SIFT

(d) DOG+SURF

(e) K+SIFT

(f) K+SURF

Fig. 8. **first** 6 OPTIONAL COMBINATIONS on PAIR1

Combo	featurePoints	matchingPoints	InliersRatio
LOG+SIFT	200	58	0.5128
LOG+SURF	200	57	0.5088
DOG+SIFT	200	39	0.5128
DOG+SURF	200	37	0.5135
K+SIFT	1809	704	0.5000
K+SURF	1809	683	0.5007
SIFT+SIFT	1809	720	0.5000
SIFT+SURF	7091	714	0.5000
SURF+SIFT	1809	704	0.5000
SURF+DSP-SIFT	1738	185	0.5027

TABLE VIII
QUANTITATIVE EVALUATION FOR PAIR1,OPTIONAL

performance of SURF&SIFT and SURF&DSP-SIFT in the

(a) SIFT+SIFT

(b) SIFT+SURF

(c) SURF+SIFT

(d) SURF+DSPSIFT

Fig. 9. **last** 4 OPTIONAL COMBINATIONS on PAIR1

evaluation of the first pair, here we also going to try these two combination on pair2, so that we can find the best combination among all the pairs.
As shown in the figure12, the homographies are obviously worse comparing to the first pair. From the matching points, we can see there are obvious outliers in DOH&SIFT and SIFT&DSP-SIFT, AND SURF&SIFT. Between SURF&SURF, KAZE&KAZE, and SURF& DSP-SIFT, we can see the matching points of SURF&SURF is more evenly distributed in the image while the matching points of KAZE&KAZE is more gathered at the right top corner where is the painting with rich information. The SURF&DSP-SIFT has less matching comparing those two. Thus the first-step conclusion is the SURF&SURF has the best performance among these six combinations,while the SURF& DSP-SIFT ranks the second, and KAZE&KAZE ranks the third.

The fundamental matrices of each combinaiton is shown in tableIV. The last value is always 1, while the values at (1,1),(1,2),(2,1),(2,2) are always zeros.

In the qualitative evaluation in figureII-C2, we can see the DOH&SIFT and KAZE&KAZE is definitely not doing their job. SURF&DSP-SIFT is slightly worse than SURF&SURF, SURF&SIFT and SIFT&DSPSIFT. We are going to give it one more chance for the third pair of images.

As shown in tableII-C2, we can see that the ratio of inliers are similar. The DSP-SIFT is an excellent descroptor since

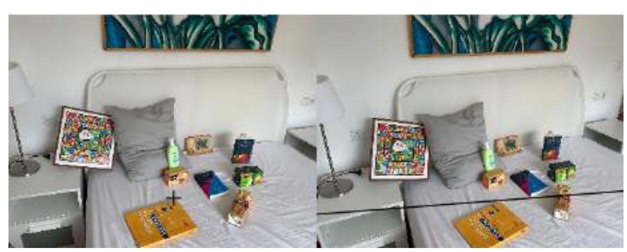
(a) LOG+SIFT

(b) LOG+SURF

(c) DOG+SIFT

(d) DOG+SURF

(e) K+SIFT



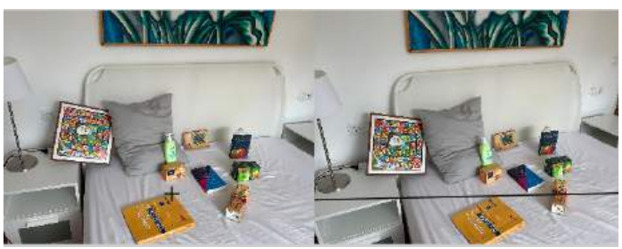
(f) K+SURF

Fig. 10. Qualitative Evaluation for **first** 6 optional combos, Pair1

(a) SIFT+SIFT

(b) SIFT+SURF

(c) SURF+SIFT



(d) SURF+DSP-SIFT

Fig. 11. Qualitative Evaluation for **last** 4 optional combos, Pair1

(a) IMAGE2

(b) IMAGE4

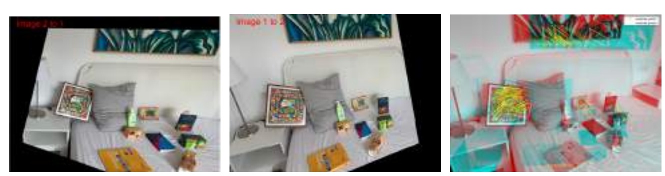
(c) DOH+SIFT

(d) SIFT+DSP-SIFT

(e) SURF+SURF

(f) KAZE+KAZE

(g) SURF+SIFT



(h) SURF+DSPSIFT

Fig. 12. 6 combinations to test on pair 2

it produces more than 200 points in this case, which is outstanding among all the other descriptors. While among all the detectors, SIFT and SURF are both doing a good job. Kaze is computationally expensive and it doesn't ensure a good result. So far after this round, SIFT&DSPSIFT, SURF&DSPSIFT, and SURF&SURF are ranking the top 3.

3) PAIR 3: The third pair is the first image and the fourth image in the sequence. This pair has the greatest difference and therefore would be the most challenging one. Therefore we are going to play around the parameters, for example, increasing the MaxRatio for Matching process, in which case, making sure we are only choosing the matching points we are confident with.

As shown in the figure14, the homographies are obviously worse comparing to the first pair. From the matching points, we can see there are obvious outliers in DOH&SIFT and SIFT&DSP-SIFT, AND SURF&SIFT. Between SURF&SURF, KAZE&KAZE, and SURF& DSP-SIFT, we can see the matching points of SURF&SURF is more evenly distributed in the image while the matching points of KAZE&KAZE is more gathered at the right top corner where is the painting with rich information. The SURF&DSP-

DOH+SIFT

$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0004 \\ 0.0000 & 0.0000 & 0.0048 \\ -0.0009 & -0.0080 & 1.0000 \end{bmatrix}$$

SIFT+DSPSIFT

$$\begin{bmatrix} 0.0000 & 0.0000 & -0.0002 \\ 0.0000 & 0.0000 & 0.0072 \\ -0.0002 & -0.0104 & 0.9999 \end{bmatrix}$$

SURF+SURF

$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0002 \\ 0.0000 & 0.0000 & 0.0075 \\ -0.0001 & -0.0112 & 0.9999 \end{bmatrix}$$

KAZE+KAZE

$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0035 \\ 0.0000 & 0.0000 & -0.0044 \\ -0.0054 & 0.0043 & 1.0000 \end{bmatrix}$$

SURF+SIFT

$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0006 \\ 0.0000 & 0.0000 & 0.0044 \\ -0.0013 & -0.0073 & 1.0000 \end{bmatrix}$$

SURF+DSPSIFT

$$\begin{bmatrix} -0.0000 & -0.0000 & 0.0001 \\ 0.0000 & 0.0000 & 0.0058 \\ 0.0000 & -0.0107 & 0.9999 \end{bmatrix}$$

TABLE IX
FUNDAMENTAL MATRICES FOR 6 COMBOS,PAIR2

Combination	points detected	pointsmatched	InliersRatio
DOH+SIFT	1720	132	0.5000
SIFT+DSPSIFT	8980	279	0.5018
SURF+SURF	1720	192	0.5000
KAZE+KAZE	10656	508	0.5000
SURF+SIFT	1720	132	0.5000
SURF+DSPSIFT	1720	217	0.5023

TABLE X
QUANTITATIVE EVALUATION FOR PAIR1

SIFT has less matching comparing those two. Thus the first-step conclusion is the SURF&SURF has the best performance among these six combinations,while the SURF& DSP-SIFT ranks the second, and KAZE&KAZE ranks the third.
Here attach the fundamental matricesXI

In the qualitative evaluation in figureII-C3, the SIFT&SURF family are still competitive while the KAZE&KAZE is definitely out of this game.

As shown in tableII-C3, we can see that SURF, as a descriptor, lost its game by not producing as many points as DSPSIFT. KAZE produces and matches a lot of points but not doing a good performance.

DOH+SIFT

$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0232 \\ 0.0000 & 0.0000 & 0.1520 \\ 0.0383 & -0.2288 & 0.9611 \end{bmatrix}$$

SIFT+DSPSIFT

$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0018 \\ 0.0000 & 0.0000 & 0.0198 \\ 0.0019 & -0.0300 & 0.9994 \end{bmatrix}$$

SURF+SURF

$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0008 \\ 0.0000 & 0.0000 & 0.0144 \\ 0.0004 & -0.0227 & 0.9996 \end{bmatrix}$$

KAZE+KAZE

$$\begin{bmatrix} 0.0000 & 0.0000 & 0.0013 \\ -0.0000 & -0.0000 & 0.0009 \\ -0.0036 & -0.0008 & 1.0000 \end{bmatrix}$$

SURF+SIFT

$$\begin{bmatrix} -0.0000 & 0.0000 & -0.0067 \\ 0.0000 & 0.0000 & 0.0491 \\ 0.0105 & -0.0757 & 0.9958 \end{bmatrix}$$

SURF+DSPSIFT

$$\begin{bmatrix} 0.0000 & -0.0000 & 0.0065 \\ -0.0000 & -0.0000 & -0.0276 \\ -0.0117 & 0.0412 & 0.9987 \end{bmatrix}$$

TABLE XI
FUNDAMENTAL MATRICES FOR 6 COMBOS,PAIR3

(a) DOH+SIFT

(b) SIFT+DSP-SIFT

(c) SURF+SURF

(d) KAZE+KAZE

(e) SURF+SIFT

(f) SURF+DSPSIFT

Fig. 13. Qualitative Evaluation for 6 Combos,Pair2

D. SELECTION

choose the best pair of views and the best detector+descriptor combination according to results obtained in previous stage.

- **a.Indicate and illustrate your selection.**

(a) IMAGE2

(b) IMAGE4

(c) DOH+SIFT

(d) SIFT+DSP-SIFT

(e) SURF+SURF,MaxRatio0.4

(f) KAZE+KAZE

(g) SURF+SIFT

(h) SURF+DSPSIFT

Fig. 14. 6 combinations to test on pair 3

(a) DOH+SIFT

(b) SIFT+DSP-SIFT

(c) SURF+SURF

(d) KAZE+KAZE

(e) SURF+SIFT

(f) SURF+DSPSIFT

Fig. 15. Qualitative Evaluation for 6 Combos,Pair3

The final choice will be pair 1 with SURF and DSP-SIFT. First of all, quantitatively speaking, KAZE is the best because it creates the most feature points and matching points. But its performance in qualitative evaluation is not good at all. Besides, the DSP-SIFT descriptor is always doing good as well. So the SIFT&DSP-SIFT and SURF&DSP-SIFT are always extrordinary among their peers. However, there are several times the SIFT&DSP-SIFT has bad results in qualitative evaluation. Then qualitatively speaking, both SURD&SIFT

Combination	points detected	pointsmatched	InliersRatio
DOH+SIFT	1738	107	0.5047
SIFT+DSPSIFT	8946	203	0.5025
SURF+SURF	1738	149	0.5034
KAZE+KAZE	10806	377	0.5013
SURF+SIFT	138	107	0.5047
SURF+DSPSIFT	1720	185	0.5027

TABLE XII
QUANTITATIVE EVALUATION FOR PAIR1

and SURF&SURF are good. We suppose SURF is a very good detector. In conclusion, the combination of SURF and DSPSIFT is a pleasant choice.
Since we are just making a decision based on previous qualitative and quantitative evaluation, here attach a easy tableII-D to visualize our judgement simply commenting with 'good' and 'bad' as grades.

Combination	QualitativeEvaluation	QuantitativeEvaluation
DOH+SIFT	BAD	BAD
SIFT+DSPSIFT	BAD	GOOD
SURF+SURF	GOOD	BAD
KAZE+KAZE	BAD	GOOD
SURF+SIFT	GOOD	BAD
SURF+DSPSIFT	GOOD	GOOD

TABLE XIII
QUANTITATIVE EVALUATION FOR PAIR1

III. SECTION 3

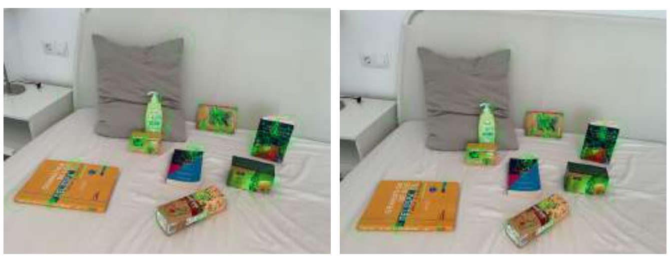
3D reconstruction and calibration First of all, since the scene in section 2 is with a game board which attracts too much feature points and the other part of image is with less feature points. Therefore, to have a sense of a good 3D reconstruction and calibration let's do it on another set of images slightly different from what we are using in section 2. In this sequence, the complicated game board is removed.

A. N VIEW MATCHING

In this section we compute consistent point matches among N views. Get different views with a calibrated camera. Varying the angle and distance of the camera, altering light conditions.

- **a.Provide the images for N-view point matching, indicating detected interest points in each of them.**

As shown in the figureIII-C2, the set of images we are going to use is taken in front of a bed, with some objects distributed on it. The matched points are evenly distributed in the images.



(a) IMAGE1&2

(b) IMAGE3&4

Fig. 16. Bed Sequence

B. INITIAL PROJECTIVE RECONSTRUCTION.

In this section we compute the Fundamental matrix and an initial projective reconstruction from 2 of the cameras.

- **a.Provide the images for the estimation of the Fundamental matrix, indicating the detected interest points and point matches**

As shown in figureIII-B, the first and the last image in the sequence are selected to initialize the projective reconstruction considering the balance between point matching and projection ambiguity.

- **b. Provide the mean re-projection error and the reprojection error histogram.**

As shown in figure18 and tableIII-B, the reprojection error after initialization by the first and the last image in the sequence is evenly distributed along the x axis with a range of [-2,2] while a little bit biased along the y axis with a range of [-6,4]. In the histogram The errors mostly aggregated around 0.

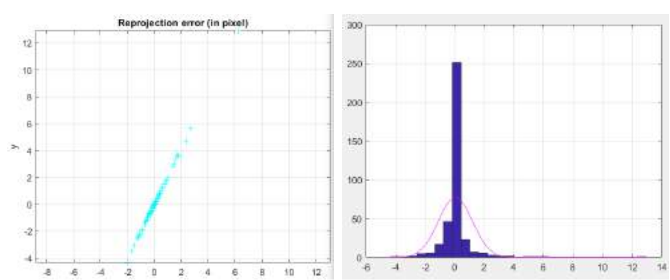
(a) Extracted Feature Points

(b) Matching

Fig. 17. Images for initial projective reconstruction

Residual reprojection error		1.3259
	Axis X	**Axis Y**
Mean	0.03453	0.07101
Std	0.70945	1.46848

TABLE XIV
ERROR STATISTICS AFTER INITIALIZATION



(a) Initialization Reprojection Error

Fig. 18. Reprojection error in pixel and histogram after initialization

C. RESECTION and BUNDLE ADJUSTMENT

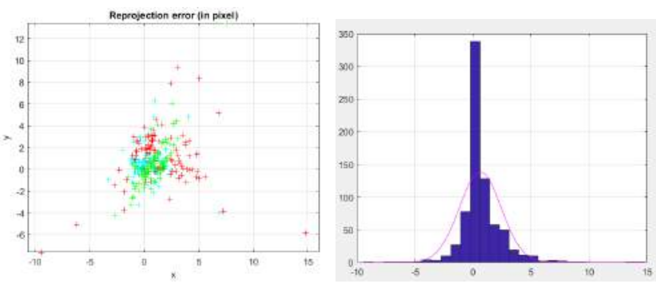
Improve this initial reconstruction by means of a Projective Bundle Adjustment, using a higher number of your images.

- a. **Provide the mean re-projection error and the reprojection error histogram at two points:(i) after the resectioning step, and (ii) after the Projective Bundle Adjustment step.**

1) Resectioning: The resectioning process is an initialization for getting the projection matrices from all the cameras and the homogeneous coordinates of 3D points. As shown in tableIII-C1 and figure19, the errors distributed in the range of [-8.8] pixels for y axis and [-15,15] for x axis. In the histogram it also confirms the errors locate within a range of ± 15 pixels, and most of them gather around 0.

Residual reprojection err resectioning		3.6079
	Axis X	**Axis Y**
MEAN	0.67638	0.54113
STD	1.77829	1.82220

TABLE XV
ERROR STATISTICS AFTER RESECTIONING



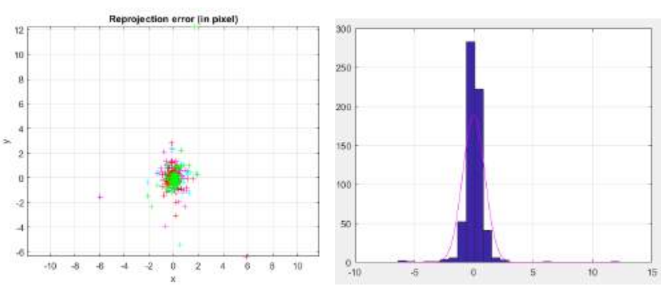
(a) resectioning Reprojection Error

Fig. 19. Reprojection error in pixel and histogram after resectioning

2) Bundle Adjustment: The Bundle Adjustment is for getting the updated projection matrices from all the cameras and the updated homogeneous coordinates of 3D points. As shown in tableIII-C2 and figure20, the errors distributed in the range of 10 pixels for both x axis and y axis. In the histogram it also confirms the errors locate within a range of 10 pixels, and most of them gather around 0.

Residual reprojection err Bundle Adjustment		1.1565
	Axis X	**Axis Y**
MEAN	-0.00000	-0.00000
STD	0.89033	1.23549

TABLE XVI
ERROR STATISTICS AFTER BUNDLE ADJUSTMENT



(a) Bundle Adjustment Error

Fig. 20. Reprojection error in pixel and histogram after Bundle Adjustment

- **b.Comment on the justification of the different reprojection error values in 2.b and the two steps of 3.a**

F	**Recomputed_F**
$\begin{bmatrix} -0.0000 & 0.0000 & -0.1566 \\ -0.0000 & 0.0000 & 0.0065 \\ 0.1540 & -0.0137 & 0.9755 \end{bmatrix}$	$\begin{bmatrix} -0.0000 & 0.0000 & -0.0144 \\ -0.0000 & 0.0000 & 0.0005 \\ 0.0142 & -0.0012 & 0.0997 \end{bmatrix}$

TABLE XVII
FUNDAMENTAL MATRICES FOR 4 MANDATORY COMBOS

First of all, after using the 8 points algorithm to initial the projective reconstruction by two images, the reprojection error is calculated from the 2D coordinates on the two image planes and the reprojected 2D coordinates by multiplying the projection matrices of two cameras and the 3D points we got after initial reconstruction. Then in the resectioning step we are initializing the projection matrices for the rest of cameras, which we took two more cameras into account. Therefore the reprojection error increases a little bit. Then in the bundle adjustment process we update the projection matrices of all the cameras as well as the 3D points. In this way the reprojection error get lower.

D. RECOMPUTE FUNDAMENTAL MATRIX

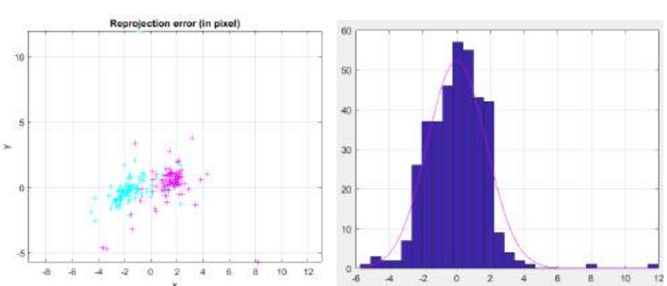
Re-compute the Fundamental matrix between two of the cameras, using the projection matrices obtained after the Projective Bundle Adjustment step.
The fundamental matrices are shown in tableXVII.

E. EUCLIDEAN RECONSTRUCTION WITH ESSENTIAL MATRIX

Use the properties of the Essential matrix (between two cameras) to obtain a Euclidean reconstruction of the scene (use the re-projected points obtained after the Projective Bundle Adjustment step).

- **a.Provide the mean re-projection error and the reprojection error histogram (for these two cameras)**

The reprojection errors after euclidean reconstruction are shown in figure22 and tableIII-E, the errors have a range of [-5,5] for both two axes.



(a) Euclidean Reconstruction Error

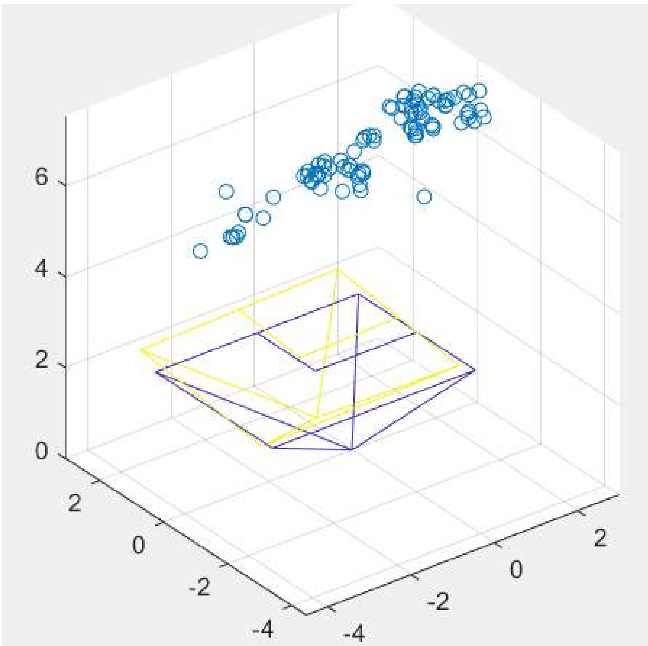
Fig. 21. Reprojection error in pixel and histogram after euclidean reconstruction

- **b.Provide illustrative results (several viewpoints and the 3D Matlab figure) of your 3D point cloud reconstruction.**

The figure22 shows the euclidean reconstructed point cloud.

Residual reprojection err Euclidean Reconstruction		3.039
	Axis X	**Axis Y**
MEAN	-0.04625	0.04990
STD	2.00382	1.44584

TABLE XVIII
ERROR STATISTICS AFTER EUCLIDEAN RECONSTRUCTION

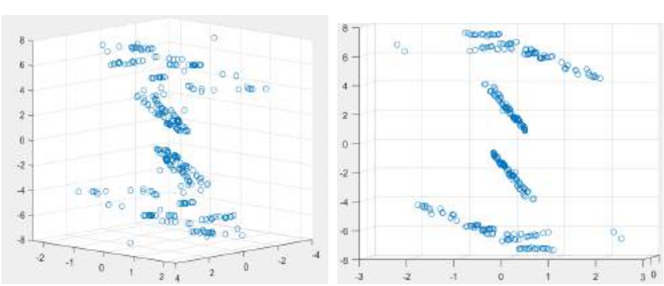


(a) Euclidean Reconstruction

Fig. 22. 3D euclidean reconstruction

- **c.Extra: Provide illustrative results (several viewpoints and the 3D Matlab figure) of an "improved" 3D point cloud.**

You can improve your point cloud with strategies such as: line segments that connect points that are joined by straight lines in your object/scene, "painting" your point cloud with RGB values of the pixels, cluster points from different objects and "paint" them with different colors, etc.
In the figureIII-E we displays the 3D point cloud without cameras.
There are several planes can be easily recognised, especially from the side the points can be seen located as lines.



(a) 3D Point without Cameras

Fig. 23. 3D Point without Cameras