3D RECOVERING OF URBAN SCENES LAB 5

Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

ABSTRACT

This project exposes how to achieve 3D reconstruction from uncalibrated images with a stratified method: projective reconstruction, affine reconstruction and matrix reconstruction. This project will be achieved on both on synthetic data and real data.

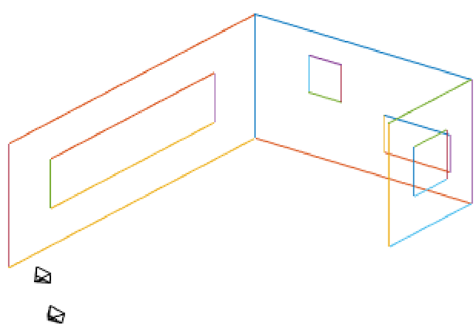
A base code and functions are provided. All the code is attached parallel to this document.

Code split in 3 files: **Lab5.m, factorization.m, vanishing_point.m**

Projective Reconstruction (synthetic data)

The synthetic data which is required to be reconstructed was created in advance:

3D points



Visualize lines

All the information we had are the homogeneous coordinates of 2D points x1, x2.

By rescaling the image coordinates, we can obtain a measurement matrix (the combined image coordinates of all the points in all the images), which is of rank 4. Projective structure and motion can then be obtained by a singular value factorization of this matrix.

The steps of the factorization algorithm for projective shape from motion are below:

- Determine a subset of scene points and cameras so that the measurement matrix is completely determined.
- Normalize the set of points in each image (similarity transform).
- Initialize the projective depth of image1.
- Alternate rescaling the rows of the depth matrix A (formed by projective depths) to have unit norm and the columns of A to have unit norm until A stops changing significantly.
- Build the measurement matrix W.
- Determine the SVD of W = UDV’.
- Get the projective camera matrix P and the Projective Points X.
- Compute the Euclidean distance (d) between data points and projected points in both images and stop when a convergence criterion was obtained. Otherwise, modify the projective depth and go to regenerate the depth matrix A.

3D RECOVERING OF URBAN SCENES LAB 5

Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

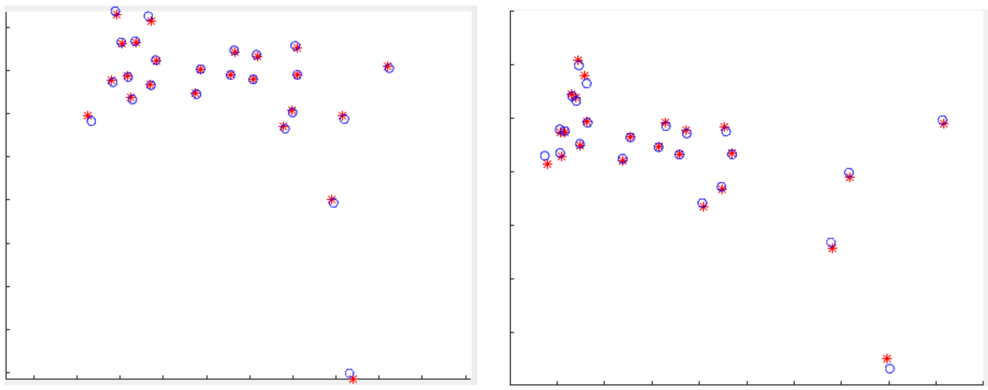
The illustrate of the function is below:

- function '**factorization**'
  - ✓ **Input**: - x : the cell that contains the homogeneous coordinates of 2D points in two images x1 and x2. - Ncam: the number of the cameras.
  - ✓ **Output**: an estimate of Pproj: 3*Ncam x 4 matrix containing the camera matrices, an estimate of Xproj: 4 x Npoints matrix of homogeneous coordinates of 3D points.

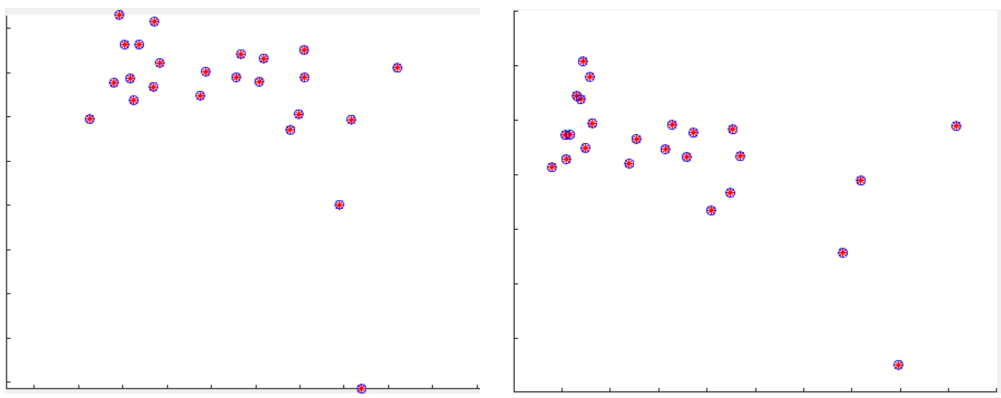
In this function, the output should be the estimates of the camera matrix P (3*N x 4) and the homogeneous coordinates of 3D points X, which achieves the projective reconstruction.

We compared the projected points with the data points and visualize projective reconstruction to evaluate the performance of the function. The result in this parts (synthetic data) are shown below:

- with initialization $\lambda_j^i = 1$ for all i and j



- with initialization $\lambda_j^i = 1$ for all i=1 and $\lambda_j^i = \frac{x_j^{1^T} F_{i1}(e \times x_j^i)}{||e \times x_j^i||^2} \lambda_j^1$:



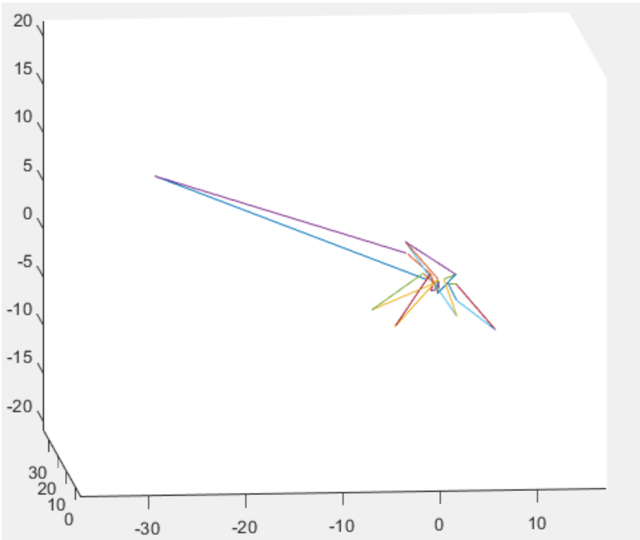
3D RECOVERING OF URBAN SCENES LAB 5

Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

Comments on the results:

- From these results, it can be said that with the second case of initialization, the better results were obtained.
- Since this is a stratified method, after the projective reconstruction, we just finished the first step of the whole method. The lines in 3D space does not make any sense.

To visualize the projective reconstruction, below 3D reconstruction was obtained:



Affine Reconstruction (synthetic data)

Although the first step reconstruction was obtained, there is a projective ambiguity:

$$\mathbf{x} = P\mathbf{X}, \text{ but also } \mathbf{x} = PH^{-1}H\mathbf{X} = \widehat{P}\widehat{\mathbf{X}}$$

Taking into account this problem, a solution to the projective ambiguity problem is the stratified reconstruction. After estimating a projective reconstruction, upgrading the previous reconstruction to an affine reconstruction is a good way to deal with the problem in first step.

$$H_{e\leftarrow p}^{-1} = \begin{pmatrix} K & \mathbf{0} \\ -\mathbf{p}^T K & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} I & \mathbf{0} \\ -\mathbf{p}^T & 1 \end{pmatrix}}_{H_{a\leftarrow p}^{-1}} \underbrace{\begin{pmatrix} K & \mathbf{0} \\ -\mathbf{0}^T & 1 \end{pmatrix}}_{H_{e\leftarrow a}^{-1}}$$

The homography Ha-p need to be found out. To identify the p, we pick three vanishing points in two images. A function named vanishing_point was made for calculate the vanishing points

3D RECOVERING OF URBAN SCENES LAB 5

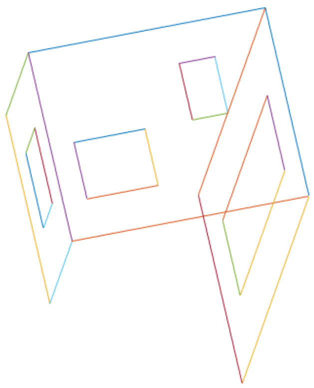
Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

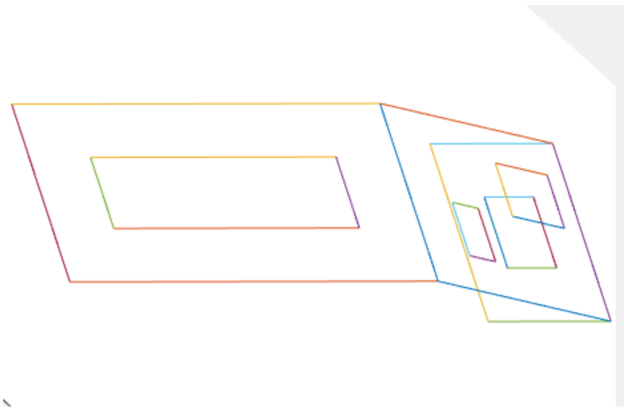
- function ‘**vanishing_point**
  - ✓ **Input**: x1, x2, x3, x4
  - ✓ **Output**: The vanishing point

Three pairs of vanishing points are used to find out their corresponding 3D points V1, V2 and V3 by triangulation.

The p can be obtained by calculate the svd of the matrix A, where A is the matrix contains three 3D vanishing points. When the Ha-p was got, it permits to upgrade the projective reconstruction into an affine one.

The result in this part was got as below:





Comments on the results:

- The essence of the affine reconstruction is to locate the plane at infinity in the projective reconstruction frame.
- The shape of the synthetic data is forming.
- However, although the characteristic of parallel was recovered, the angles of the corners are still not obtained.

Matrix Reconstruction (synthetic data)

To finish the final step of the reconstruction, the only homography He-a need to be calculated. The key to metric reconstruction is to find the image of the absolute conic in one of the images: w= $K^{-T}K^{-1}$.

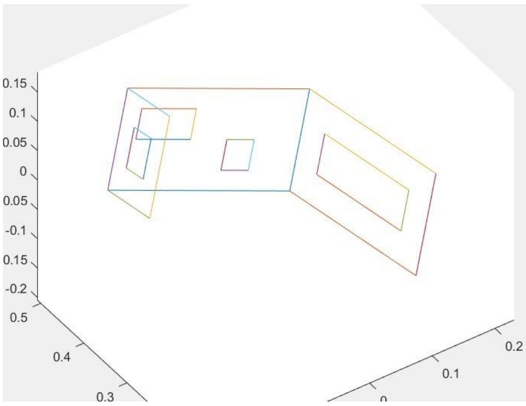
$$\omega = K^{-T}K^{-1} = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{12} & \omega_{22} & \omega_{23} \\ \omega_{13} & \omega_{23} & \omega_{33} \end{pmatrix}$$

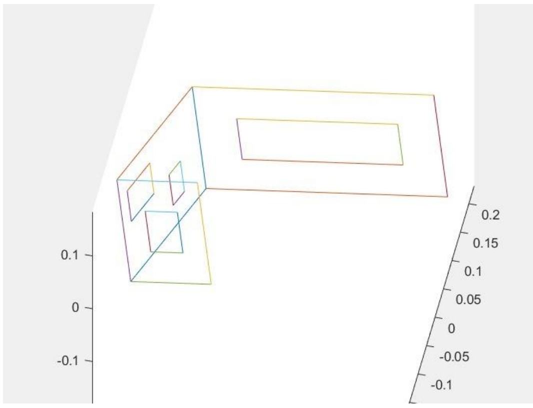
Suppose that the image of the absolute conic is known in some image to be w, and one has an affine reconstruction in which the corresponding camera matrix is given by Pa = [M|m]. Then, the affine reconstruction may be transformed to a metric reconstruction, by applying a 3D transformation of the form:

$$H_{e \leftarrow a} = \begin{pmatrix} A^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

where A is obtained by Cholesky factorization: AA^T=$(M^T \mathrm{wM})^{-1}$. In this way, the He-a can be got and it permits to upgrade the affine reconstruction into a metric reconstruction.

The results are shown as below:





Comments on the results:

- The approach relies on identiying w are various.
- The reconstruction is completed with projective, affine and matrix step by step.

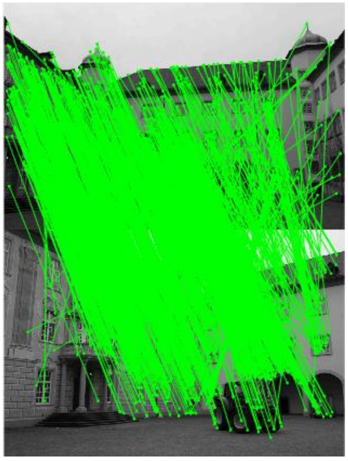
Projective Reconstruction (real data)

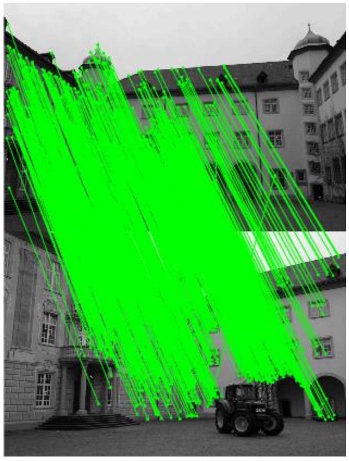
In this part, a pair of real dataset will be used to achieve projective reconstruction.

To get the matching points between two images, we used sift to extract the features. Then we used ransac to get the matching points.

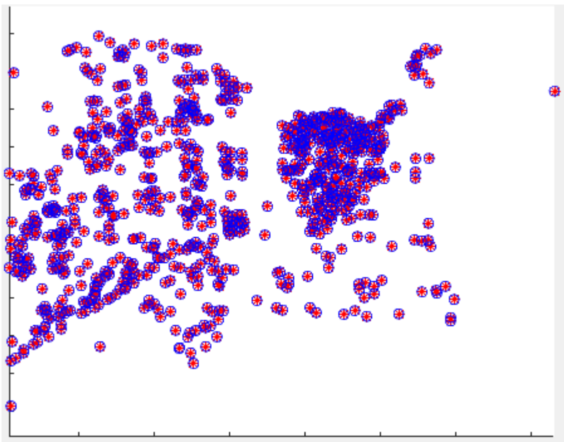
3D RECOVERING OF URBAN SCENES LAB 5

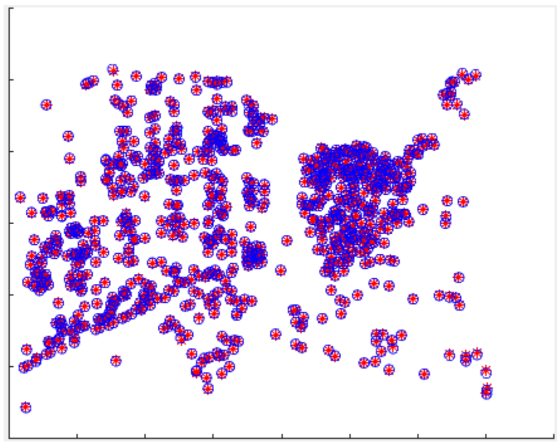
Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer





After that, a set of matching points x1, x2 were obtained, which can be used on the projective reconstruction function. In the same way, the projected points were checked with the data points:





From the result, it can be said that the projected points are perfectly matched with the data points.

Affine Reconstruction (real data)

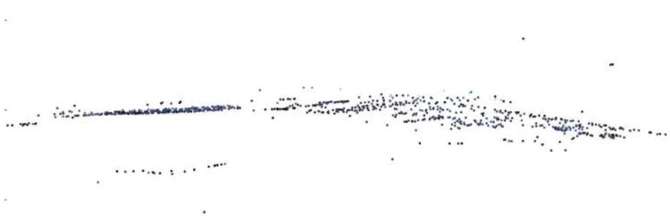
The same way of the affine reconstruction was applied on the real data. The main steps of the affine reconstruction can be integrated as below:

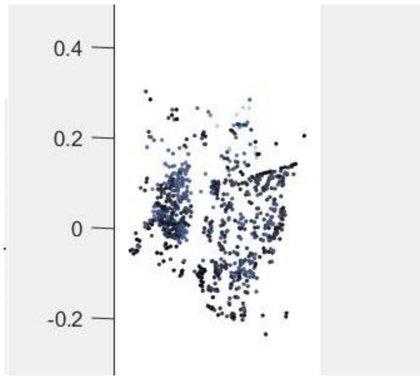
- Pick three vanishing points in two images and find their corresponding 3D points by triangulation.
- Build the A matrix with three 3D points.
- Apply the SVD on A and pick the right null vector of A.
- Get the Ha-p homography and apply it with the camera matrix of the projective reconstruction.

3D RECOVERING OF URBAN SCENES LAB 5

Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

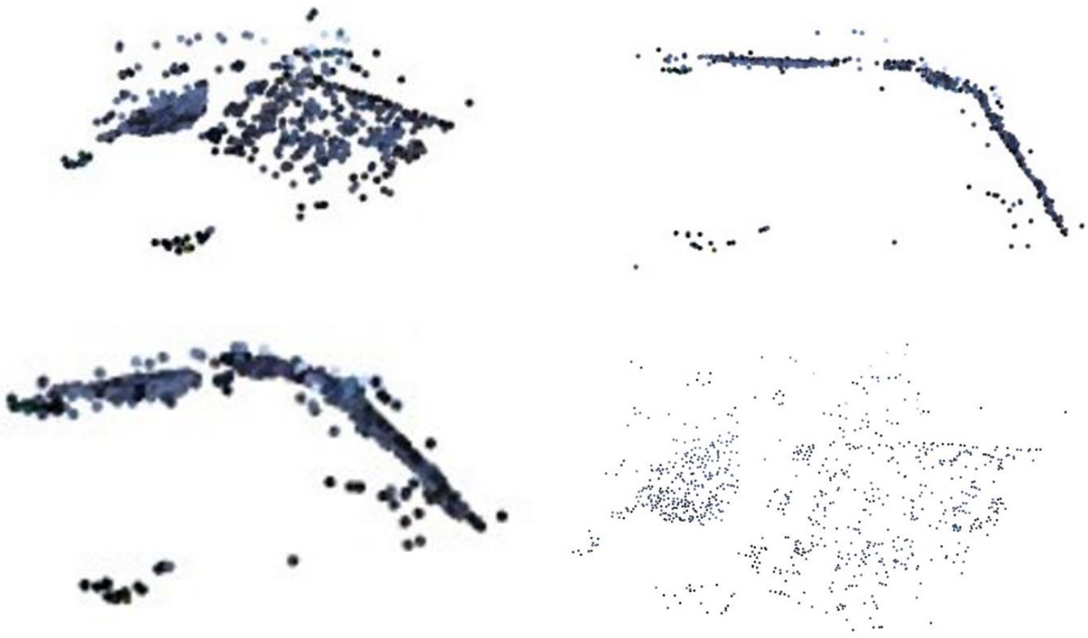
The below results were obtained with the real data:





From the results, the shape of the reconstruction has been slightly formed.

Matrix Reconstruction (real data)



Comments on the results:

- The shape of the building is built on the 3D space after using the metric rectification.
- The goodness of the result depends strongly on the SIFT features and the matching points for two images.
- The more views (images) are used, the better result can be obtained since there is more information of the senses are provided.

3D RECOVERING OF URBAN SCENES LAB 5

Juan Felipe Montesinos, Yi Xiao, Ferran Carrasquer

PROJECT CONCLUSIONS

Once this project done, we can extract some points to consider:

- The 3D construction can be obtained by a stratified method following projective, affine and matrix reconstruction.
- The effect of the 3D construction strongly depends on the matching points and the number of the views. The more views (images) are used, the better result can be obtained since there is more information of the senses is provided.
- For the projective reconstruction, the result does not make so much sense since there is a projective ambiguity.
- From the affine reconstruction, the parallel relationships between some lines can be recovered.
- For the matrix reconstruction, the vanishing points to be used should be coming from pairs of perpendicular lines.