Image Stitching

Alperen Gezer

November 25, 2021

1 Introduction

In image stitching, we aim to connect images that is taken from different angles, and create a single panoramic view.

For this purpose, we use homography, which can be represented as below.

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \tag{1}$$

Note that H is normalized by making h_{33} equal to 1.

2 Selecting Points

As the first step, we put two images that we want to stitch side by side, and select 12 sets of correspondence points. We extract matched points to a variable in order to utilize in the following computations.

3 Estimating Homography

We can rearrange the Equation 1 as follows.

$$\begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' & -y_i' \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = 0 \tag{2}$$

In order to estimate the homography matrix between 2 images, we need to pick at least 4 matching points. Nevertheless, for more precise mapping, we can use more points and then calculate least squares solution.

For exact solution for 4 points, 8 elements of matrix H is used and h_{33} is set to be $\sqrt{1 - h_{11}^2 - h_{12}^2 - h_{13}^2 - h_{21}^2 - h_{22}^2 - h_{23}^2 - h_{31}^2 - h_{32}^2}$

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 x_1' & -y_1 x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 y_1' & -y_1 y_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2 x_2' & -y_2 x_2' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2 y_2' & -y_2 y_2' \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3 x_3' & -y_3 x_3' \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3 y_3' & -y_3 y_3' \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4 x_4' & -y_4 x_4' \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4 y_4' & -y_4 y_4' \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = 0 \tag{3}$$

For solutions for n, more than 4, points, least-squares solution is found by using all elements and scaling to make $h_{33} = 1$.

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 & -y'_1 \\ & & & & \vdots & & & & \\ x_n & y_n & 1 & 0 & 0 & 0 & -x_nx'_n & -y_nx'_n & -x'_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -x_ny'_n & -y_ny'_n & -y'_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = 0 \tag{4}$$

In order to solve the linear equation, we use singular value decomposition, and then take the last vector of V as the solution, and multiply the whole matrix with $\frac{1}{h_{33}}$.

3.1 Normalization

In order to achieve a equally distributed matrices, we can normalize the selected points. To normalize a set of coordinates;

$$\begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

Where;

$$T = \begin{bmatrix} \frac{1}{\sigma_x} & 0 & -\frac{\mu_x}{\sigma_x} \\ 0 & \frac{1}{\sigma_y} & -\frac{\mu_y}{\sigma_y} \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Let's define two normalization transforms for both images T and T'. Then Equation 1 becomes;

$$sT'x' = H_nTx \tag{7}$$

In order to go back from H_n to H, we can rearrange the Equation 1 and 7.

$$H = T'^{-1}H_nT \tag{8}$$

4 Image Warping

From Equation 1, we know how to transform a pixel to its matching position on another frame. As we calculate H matrix from the previous steps, we can apply the transformation.

However, when we use H matrix to calculate new positions of the pixels, we will get floating points, and that will cause empty areas in the resulting image. Instead, we inverse the H matrix, and we scan the result image to calculate which pixels on the source image is closer, and interpolate to improve the result.

5 Merging

After warping image and saving the edge coordinates, we apply a bending mask and merge two images together.

6 Experiments

6.1 5-Point Fitting

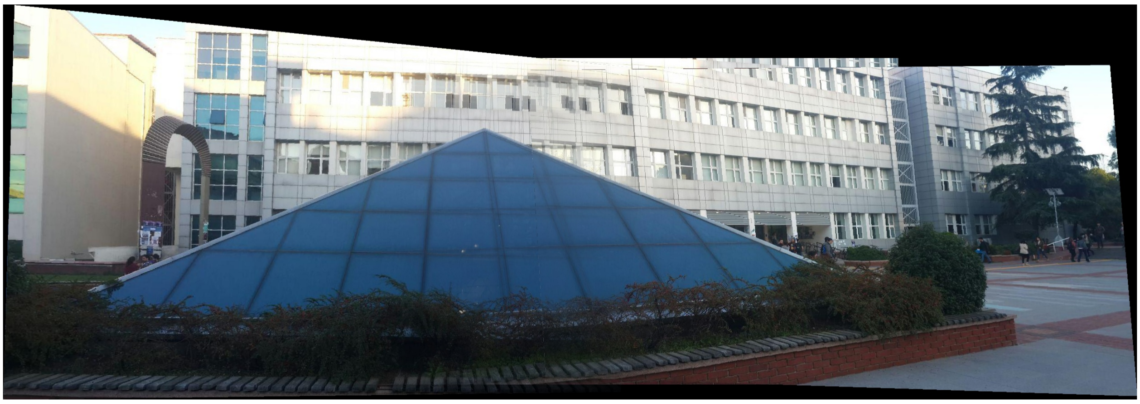

Figure 1: Experiment result for 5-point fitting.

In this experiment, we used limited number of points to calculate the homography. A misalignment in seats and building windows can easily be noticed.

6.2 12-Point Fitting

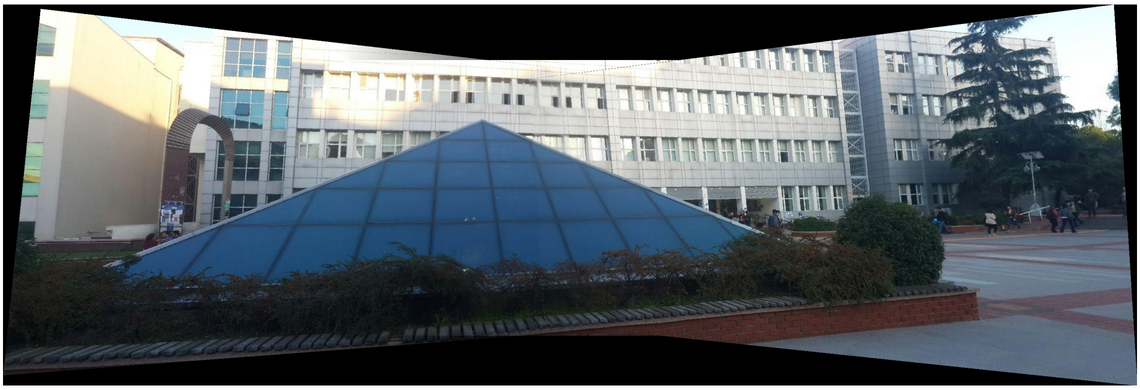

Figure 2: Experiment result for 12-point fitting.

Increasing the number of matching points resulted in better fitting. However, there are still misalignments, especially on the building

6.3 3 Wrong Points without Normalization

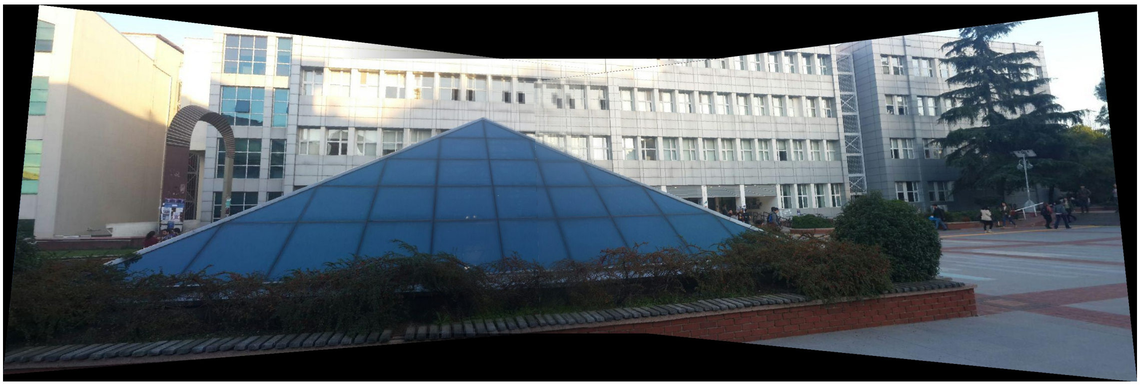

Figure 3: Experiment result for 3 wrong points without normalization.

When we introduce wrong matches, we observe building windows distorting a lot, creating a mess.

6.4 3 Wrong Points

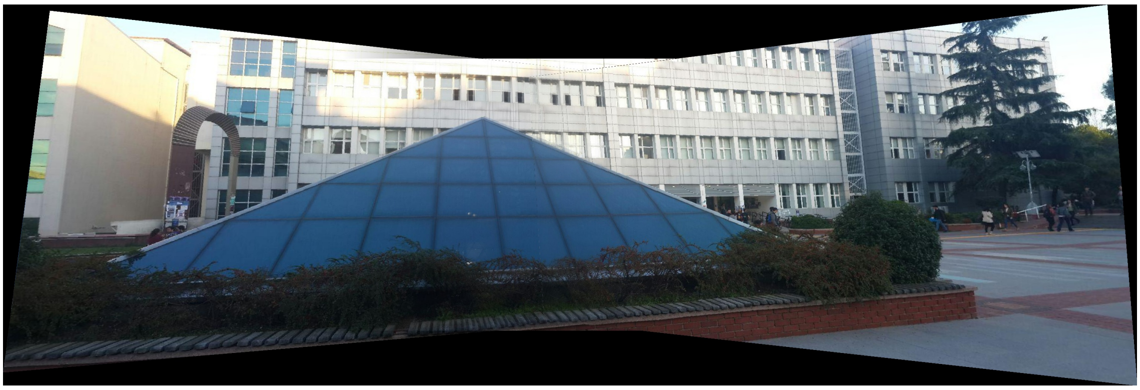

Figure 4: Experiment result for 3 wrong points with normalization.

Introducing normalization to the scene improves results. In this step, choosing wrong points randomly resulted in non-sense (taking infinity to calculate) homography solutions. Thus, an orientation vector based on reference points is extracted, and wrong matches are chosen randomly but by preserving the moment of the image.

6.5 5 Wrong Points

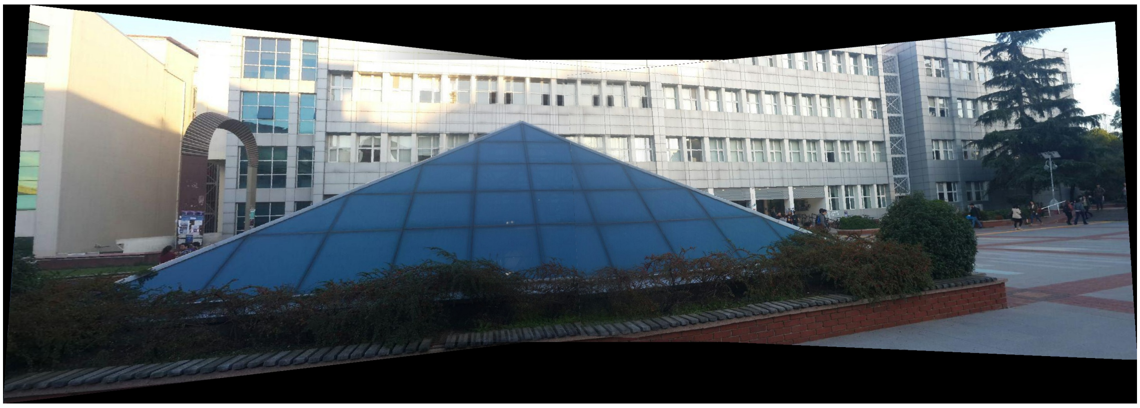

Figure 5: Experiment result for 5 wrong points with normalization.

Adding more wrong matches, again, decreasing the quality of the stitching.

6.6 Noisy Points

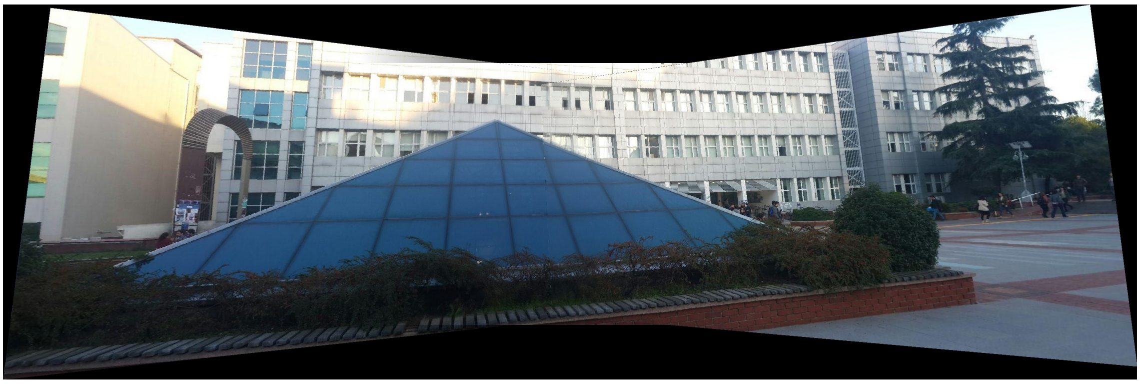

Figure 6: Experiment result for noisy points.

When points have little noise, the alignment of the images does not corrupt dramatically. Nevertheless, little distortion is introduced.

6.7 Panorama

Finally, we add all images subsequently.]

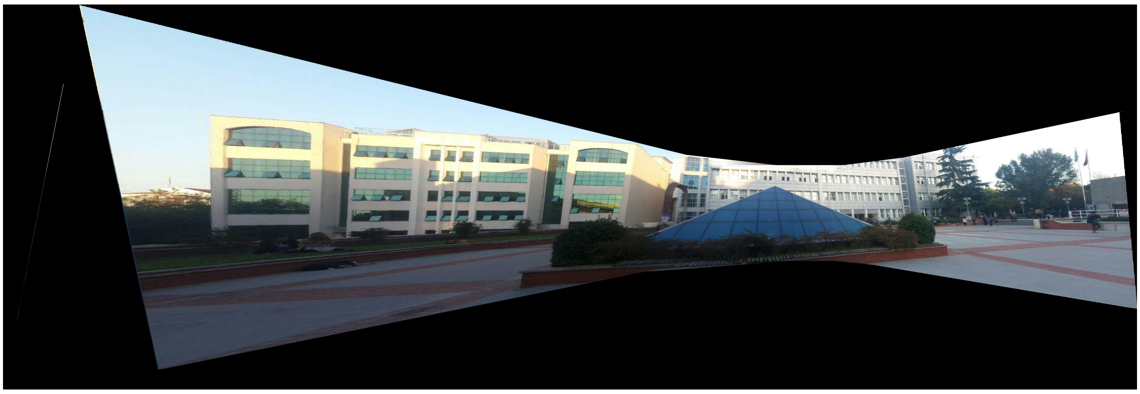

Figure 7: Experiment result for panorama.