Computer Vision 1: Image Alignment and Stitching

Bogdan Teleaga, 11413565
Mircea Mironenco, 11164182

March 13, 2017

1 Image alignment

The first part of the assignment was about computing the best affine transformation that maps between two images. As a prerequisite, we need to find some interesting similar points that are common for both images. This was done using David Lowe's *SIFT* algorithm provided by *VLFeat*.

1.1 SIFT

SIFT [1] provides a feature description of an image, by detecting points of interest. The description can then be used to match or find specific objects from the image, in other images. When a new image is compared to one from which we have extracted SIFT features, candidate matching features are found based on the Euclidian distance of their feature vector. SIFT features are invariant to rotation, shifting and scaling, and partially invariant to illumination changes.

The algorithm initially applies a Gaussian blur continuously. The smoothed and resampled images are then transformed using a difference of Gaussians, and key points are minima and maxima of the resulting image transformations. Low contrast candidate points and edge response points along an edge are discarded.

1.2 RANSAC

RANdom SAmple Consensus or RANSAC is then used for computing the affine transformation. Being given a number of common points of interest shared between the two images we can compute a linear transformation that maps between these points. Since we're essentially transforming a plane, we need three points to find a well defined transformation. We can also see this algebraically since our system of equations has 6 unknowns and each point gives us two equations. Having computed a suitable transform for those three points we use it to map all of the features found earlier. We then compare the distance between the features transformed by our candidate transformation and their position in the target image.

Figure 1: Features mapped between *boat1* $\rightarrow$ *boat2* and *boat2* $\rightarrow$ *boat1* respectively. We chose to only plot a small number for clarity. We used N = 5 and P = 3

The points that fall within a certain threshold are called *inliers*. Since we are sampling, we could get unlucky and sample three points on the same line which means our linear system is ill defined. Moreover, the SIFT feature pairs are not always perfect. To prevent these problems we can either sample more points or re-sample a couple of times, always keeping the transformation that gave the best results with respect to the number of inliers.

1.3 RANSAC Transformation

Given 6 linear independent equations, that correspond to 3 matches of 3 keypoints, we have chosen the transformation to produce as little outliers as possible (given a threshold of 10). We have aligned boat1.pgm and boat2.pgm, the results for our implementation with linear interpolation can be seen in Figure 2 and for the MATLAB function 'maketforms' and 'imtransform' in Figure 3.

From our experiments we have found that around 3-4 iterations are necessary to find a good proportion of inliers, given the data points. We can observe that the algorithm is very

Figure 2: RANSAC best transformation

Figure 3: RANSAC MATLAB transformation using 'maketforms' and 'imtransform'

Figure 4: Final stitching between the bus images.

robust even with very few iterations.

2 Image stitching

We apply the computed transformation to the first image and use linear interpolation to fill gaps where pixels are unknown. This is done using essentially with a mean filter, which takes into account the neighbouring pixels at a given position, and averages them. We construct a larger bounding image to contain both of our original images in the transformed space, the stitched image will have the dimension of the bounding box image. We also shift corners of the image that are negative given the transformed coordinate space.

The result of the image stitching of left.jpg and right.jpg can be seen in Figure 4.

References

[1] David Lowe. Distinctive image features from scale-invariant keypoints. 2004.