

Video Stitching using Improved RANSAC and SIFT

Aswathy Ashokan ¹, Ligi Achuthan ²

¹Computer Science Department, College of Engineering Munnar

² Asst Prof. Computer Science Department, College of Engineering Munnar

Abstract - The goal is to explore techniques such as image correspondence using interest points, robust matching with improved RANSAC, homography, and background subtraction and blending. The basic idea of stitch several images into a panorama is to map all the images onto a reference plane. In this project, we choose frame as the reference plane and the homography matrices between other frame images and reference frame are computed using SIFT and improved RANSAC algorithms. Identify key points and matches using SIFT. Then the key point correspondences between two frames are filtered out by the default threshold of descriptor matching. First choose correspondences from the matches, and implemented Normalized Direct Linear Transformation (DLT) to estimate the homography matrix. This process is then automated by improved RANSAC that is iterated, randomly choosing 4 correspondences each time. The degree of match is evaluated by calculating the error of other correspondences based on such homography. . The best the homography matrix is then found with most inliers. By using improved RANSAC algorithm . Once the projection transform updated in real time, we still need to blend the frames to compensate for exposure differences and other misalignments

Key Words: stitching, RANSAC, SIFT

1. INTRODUCTION

The goal is to explore techniques such as image correspondence using interest points, robust matching with improved RANSAC, homography, and background subtraction and blending. The basic idea of stitch several images into a panorama is to map all the images onto a reference plane. In this project, we choose frame as the reference plane and the homography matrices between other frame images and reference frame are computed using SIFT and improved RANSAC algorithms. Identify keypoints and matches using SIFT. Then the keypoint correspondences between two frames are filtered out by the default threshold of descriptor matching. First choose correspondences from the matches, and implemented Normalized Direct Linear Transformation (DLT) to estimate the homography matrix. This process is then automated by improved RANSAC that is iterated, randomly choosing 4 correspondences each time. The degree of match is evaluated by calculating the error of other correspondences based on such homography. . The best the homography matrix is then found with most inliers. By using improved RANSAC algorithm. Once the projection transform updated in real time, we still need to blend the frames to compensate for exposure differences and other misalignments

2. FEATURE IDENTIFICATION USING SIFT

The automatic construction of large, high-quality panoramas from regular hand-held photographs is one of the recent success stories of computer vision, with stitching software bundled with many digital cameras and photo editors. The SIFT algorithm is widely used due to various advantages, including its robustness to rotation, scaling, and changes in luminance [4]. This algorithm consists of the follows four steps: scale-space extreme detection, key point localization, orientation assignment, and a key point descriptor. In the first step, images are reproduced with different scales and are defined as the octave [4]. A difference of Gaussian (DoG) image with different sigma values is then calculated for each octave, and key point candidates selected as the local minimum or maximum using a 3X3 mask for the adjacent DoG images [4]. In the second step, two methods are used to extract more stable. Features from the key point candidates, where the first sets a critical coefficient for smooth regions in the DoG images, while the other uses a Hessian matrix for edge regions [4]. After localizing the key points, one or more orientations are assigned to each key point location based on the local image gradient directions. In the third step, the orientation is quantized using 36 bins of ten degrees in a 16x16 sample array window. In the last step, a key point descriptor is computed based on eight directions aligned in a 4x 4 grid [4]. As a result, the descriptor includes a 128-element feature vector for each keypoint. In addition to reduce the effects of changes in the illumination intensity, the feature vector is modified using unit length normalization [4]

The scale-invariant features are efficiently identified by using a staged filtering approach [6]. The first stage identifies key locations in scale space by looking for locations that are maxima or minima of a difference-of-Gaussian function[9].Each point is used to generate a feature vector that describes the local image region sampled relative to its scale-space coordinate frame[9]. The features achieve partial invariance to local variations, such as affine or 3D projections, by blurring image gradient locations. The resulting feature vectors are called SIFT keys. In the current implementation, each image generates on the order of 1000 SIFT keys, a process that requires less than 1 second of computation time. The SIFT keys derived from an image are used in a nearest-neighbor approach to indexing to identify candidate object models. Collections of keys that agree on a potential model

3. IMAGE MOSAICKING FOR PANORAMIC VIDEO

Frame selection from sequential video frames is the first step for creating a panoramic video. At this point, two images should have an overlapping region, which is identified using phase correlation that indicates the overlapping rate of the two images based on an inverse Fourier transform after calculating the cross power spectrum. The SIFT algorithm along with improved RANSAC (Random Sample Consensus) algorithm [5] is used to match the descriptor in the two overlapped images. The improved RANSAC homography algorithm based on the modified media flow filter, to detect wrong matches for improving the stability of the normal RANSAC homography algorithm. The method improved the local registration between neighboring images. Experiments and Statistical Analysis show that this mosaic method is robust.

4. USING IMPROVED RANSAC ALGORITHM

To the normal algorithm, usually only a small number of inliers are returned. But after applying the improved RANSAC homography algorithm, usually there are more number of inliers returned and the homography can be accurately returned[5]

5. VIDEO FRAME BLENDING

Once the projection transform updated in real time, we still need to blend the frames to compensate for exposure differences and other misalignments. In our stitching work, we deal only a few different source videos in current, firstly we align image by epipolar transform and then blend frame by frame. So the algorithm of blending is must be less time exhausted for real time. However, it is difficult in practice to achieve a pleasing balance between smoothing out low-frequency exposure variations and retaining sharp enough transitions to prevent blurring by these method. A fast and effective approach to make the panoramas nature and reduce blurring and ghost error utter mostly. Firstly we define a range T ($0 < T < \text{region width}$), and in the T , the picture will be natural.

6. PANORAMA USING KEY FRAMES

A preliminary panorama is then created from key frames. The goal is to map all the frames onto the plane corresponding to the reference frame. Mapping frame s which share a little area is difficult Therefore we need to perform a two stage mapping . Since our source frames come from a 30fps video, there is a large amount of overlap between the frames. In particular, this means that the values of the background pixels of each frame map to the same pixels on the reference plane. Then in order to get just the background, it suffices to take a mean of all pixels of the image of the reference plane. For each pixel of the reference plane (background) image, compute the mean of every frame that ever has a pixel on this background pixel.

7. CONCLUSION

This paper presents an efficient for stitching video sequences into wide-range and high-quality panoramic video. The algorithm utilized SIFT ALGORITHM along with an improved RAN SAC to estimated initialization projection transform and compensates it frame by frame. A fast blending method can reduce ghost error and blurring effectively

REFERENCES

- [1]Kang, S.B., Szeliski, R., Uyttendaele," Seamless Stitching Using Multi-Perspective Plane Sweep". Microsoft Research, Tech. Rep. MSR-TR-2004-48 (2004)
- [2]Zelnik-Manor, L., Peters, G., Perona, " Squaring the Circle in Panoramas". In: Proc. 10th IEEE Conf. on Computer Vision (ICCV 2005), 2005
- [3] David G. Lowe, " Distinctive Image Features from Scale-Invariant Keypoints" January 5 2004
- [3] Oh-Seol Kwon and Yeong-Ho Ha,"Panoramic Video using Scale-Invariant Feature Transform with Embedded Color-Invariant Values" , Vol. 56, No. 2, May 2010
- [4] David G. Lowe," Object Recognition from Local Scale-Invariant Features"
- [5] Fuli Wu ," An Improved RANSAC homography Algorithm for Feature Based Image Mosaic"
- [6] Bin He1, Gang Zhao, Qifang Liu, YangyangLi ,"Video Auto Stitching in Multi-Camera Surveillance System" 2010 The 3rd International Conference on Machine Vision (ICMV 2010)