

Image stitching by extracted key frames using absolute difference method

Mr. Sarvajeet Bhosale¹, Mr. K. R. Desai²

Student of ME, Bharati Vidyapeeth College of Engineering, Kolhapur, India¹

Associate Professor, Bharati Vidyapeeth College of engineering, Kolhapur, India²

Abstract: Image stitching presents different stages to render two or more overlapping images into a seamless stitched image, from the detection of features to blending in a final image. In this process, Scale Invariant Feature Transform (SIFT) algorithm [1] can be applied to perform the detection and matching control points step, due to its good properties. The process of create an automatic and effective whole stitching process leads to analyze different methods of the stitching stages. RANSAC is used for fine results.

Keywords: Image stitching, Video frame extraction, SIFT, RANSAC.

I. INTRODUCTION

Feature-based method is one of methods of Image stitching. And we propose a method based on invariant scale feature, which mainly includes two key parts: image matching and image blending.

Image matching is used to find the motion relationship between two images or several images, and it directly relates to the success rate and the speed of the total process. While image blending is used to eliminate the various illumination of the adjacent image or color does not consecutive caused by the geometric correction or dynamic scene illumination. In that way two images can stitch into a seamless image.

Video frame mosaic is a process that integrates two or more frames into a large-size image with a wide field of view [15]. Image mosaic joints the two images directly in the early times. At this time, if two images don't match that splicing effect not to be unstable, therefore, using image mosaic technology based on the Feature matching [16] causes two large displacement difference images to carry on the examination and the match, the splicing effect is good [5].

Video frames can be used for stitching to create one larger image. Given a sequence of images taken from a single point in space, but with varying orientations, it is possible to map the images into a common reference frame and create a perfectly aligned larger photograph with a wider field of view. This is normally referred to as panoramic image stitching.

II. ACTUAL ALGORITHM

The entire algorithm mainly includes: Extracted video frame from an video sequence, extract SIFT features; match features to get potential feature matches; match image sequence; match the image completely and blend the image, which can be described as the figure 1.

Algorithm:

The algorithm for video frame is represented as,

- Collect a set of overlapping images from the video input (video frames)
- Choose a reference image

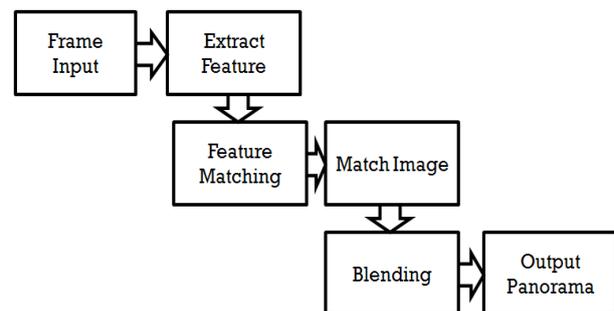


Fig 1: The Entire process of the paper

➤ For each image “x”:

1. Find feature matches between reference and x
 2. Determine transform from x to reference
 3. Transform x and place both on composite surface
 4. Composite becomes new reference.
- Run aesthetic algorithms.

A. Extracted frames from video

The video frame [11] is taken as the input and it is processes using the video processing in Matlab. The key frames are extracted from the video [7, 8]. These frames are given as the input to the SIFT algorithm and are used for creating the panoramic image [12].

For extraction of key frames the sum of absolute difference method is used. Sum of Absolute difference is a simplest method in order to find the relation between two image windows. To represent the difference few of the mathematical formulas are given as:

a) The difference between two points in one dimension is given as:

$$d(A,B) = |x_1 - x_2|$$

b) The difference between two points in two dimensions can be formulated as:

$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

c) The difference between two functions $f(x)$ and $g(x)$ is given as:

$$d(f, g) = \int |f(x) - g(x)| dx$$

d) The differences in two matrix is calculated as:

$$d(A, B) = A - B$$

Sum of difference algorithm is based on calculating the difference between each point in the image window and corresponding point in the template window. Then the difference is added together to measure the similarity between two images. The sum of absolute difference is calculated as:

$$d(A, B) = \sum_i \sum_j |A(i, j) - B(i, j)|$$

B. SIFT

The SIFT algorithm (Scale Invariant Feature Transform) proposed by Lowe is an approach for extracting distinctive invariant features from images [9]. It has been successfully applied to a variety of computer vision problems based on feature matching including object recognition, pose estimation, image retrieval and many others. However, in real world applications there is still a need for improvement of the algorithm's robustness with respect to the correct matching of SIFT features [10]. In this paper, an improvement of the original SIFT algorithm providing more reliable feature matching for the purpose of object recognition is proposed. The main idea is to divide the features extracted from both the test and the model object image into several sub-collections before they are matched. The features are divided into several sub collections considering the features arising from different octaves that are from different frequency domain.

The major stages of computation used to generate the set of image features are Scale-space extrema detection, Orientation assignment/ Key point localization and Key point descriptor.

1. Scale-space extrema detection:

This stage of the filtering attempts to identify those locations and scales which can be identified from different view of the same object. This can be efficiently achieved using a "scale space" function which under certain assumptions can be approximated to a Gaussian function.

2. Key point localization:

This stage attempts to eliminate more points from the list of key points by finding those that have low contrast or are poorly localized on an edge. This is achieved by calculating the Laplacian, value for each key point found in stage 1. Key Point Localization is shown in figure.



Fig 2: two images used for mocaking.

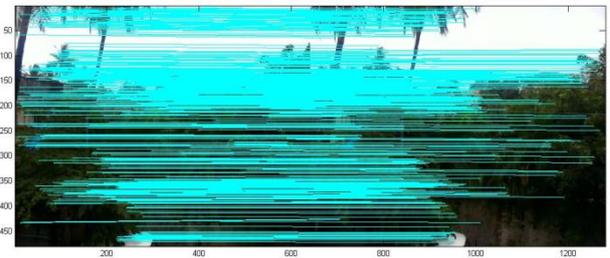


Fig 3: Key point localization.

3. Key point descriptor:

Figure illustrates the computation of the keypoint descriptor. First the image gradient magnitudes and orientations are sampled around the keypoint location, using the scale of the keypoint to select the level of Gaussian blur for the image. In order to achieve orientation invariance, the coordinates of the descriptor and the gradient orientations are rotated relative to the keypoint orientation. For efficiency, the gradients are precomputed for all levels of the pyramid. These are illustrated with small arrows at each sample location on the left side of Fig.

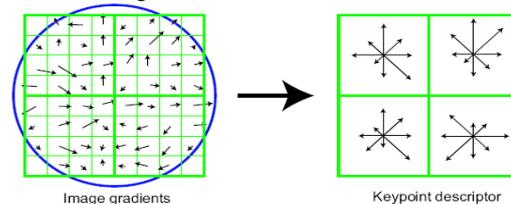


Fig 4: Key point descriptor

C. RANdom Sample Consensus:

The Random sample Consensus algorithm (RANSAC) proposed by Fischler and Bolles is a general parameter estimation approach designed to cope with a large proportion of outliers in the input data [13]. Unlike many of the common robust estimation techniques such as M-estimators and least median squares that have been adopted by the computer vision community from the statistics literature, RANSAC was enveloped from within the computer vision community. RANSAC is a resampling technique that generates candidate solutions by using the minimum number observations (data points) required to estimate the underlying model parameters. As pointed out by Fischler and Bolles, unlike conventional sampling techniques that use as much of the data as possible to obtain an initial solution and then proceed to prune

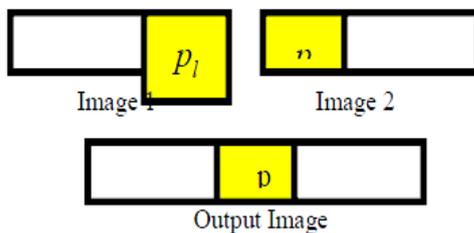
outliers, RANSAC uses the smallest set possible and proceeds to enlarge this set with consistent data points. An advantage of RANSAC is its ability to do robust estimation of the model parameters, i.e., it can estimate the parameters with a high degree of accuracy even when a significant number of outliers are present in the data set. A disadvantage of RANSAC is that there is no upper bound on the time it takes to compute these parameters. When the number of iterations computed is limited the solution obtained may not be optimal, and it may not even be one that fits the data in a good way. In this way RANSAC offers a trade-off; by computing a greater number of iterations the probability of a reasonable model being produced is increased. Another disadvantage of RANSAC is that it requires the setting of problem-specific thresholds.

Algorithm:

1. Selecting feature matches, and using Modified angle filters to test, if satisfying the condition, going to the next step.
2. Using modified length filters to test, if satisfying the condition, going to the next step. Otherwise, returning to the former step.
3. Computing the transformational matrix.
4. Matches satisfying the condition or not.
5. Repeating the upper four steps for N times.
6. Finding the feature matches that mostly satisfy the matrixes, and then computing the final matrix using them.

D. Blending:

Tables must be numbered using uppercase Roman numerals. Table Image obtained by image matching is a overlapping image. Blending minimize intensity difference of overlapping pixels. Pixels near center of image get more weight.



It is given by:

$$P = \frac{d_l}{d_l + d_r} P_l + \frac{d_r}{d_l + d_r} P_r$$

where d_l is the distance between the pixel in overlap region to the border of the left matching image, and d_r is the distance between the pixel in overlap to the border of the right matching image.

E. Results:

Based on extracting invariant scale features, we get potential feature matches then propose a RANSAC algorithm to realize image match precisely. The average weighed method ensures smooth translation between the overlap regions.



REFERENCES

- [1] Tejasha Patil, Shweta Mishra Poorva Chaudhari, Shalaka Khandale, "IMAGE STITCHING USING MATLAB" International Journal of Engineering Trends and Technology- Volume4Issue3- 2013.
- [2] Yanfang Li, Yaming Wang, Wenqing Huang, Zuoli Zhang, "Automatic Image Stitching Using SIFT", IEEE 2008.
- [3] Dianyuan Han, "A standing tree image mosaic algorithm based on SIFT", International conference on Computer Science and Electronics Engineering, 2012.
- [4] Drew Steedly, Charis Pal, Richards Szeliski, "Efficient Registrating Video into Panoramic Mosaic", IEEE, 2005.
- [5] Pengrui Qiu, Ying Liang and Hui Rong. "Image Mosaics Algorithm Based on SIFT Feature Point Matching and Transformation Parameters Automatically Recognizing", proceedings of second International conference on Computer Science and Electronics Engineering, 2013.
- [6] V S Veena Devi, S.M. Prasad and H. Gangadhara Bhat, "An efficient algorithm for image stitching based on scale-invariant feature transform", International Journal of Engineering and Applied Sciences, vol. 2, 2012.
- [7] Chaokui Li, Gang Yang, Jun Wu and Fang Wang, "The UAV Video Image Stitching Based on Improved Moravec Corner Matching Method", International Journal of Remote Sensing Applications, vol. 2, 2012.
- [8] Miguel Lourenco, Joao P. Berreto and Francisco Vasconcelos, "sRD-SIFT: Keypoint detection and matching in images with radial distortion", IEEE transaction on Robotics, vol. 28, 2012.
- [9] Matthew Brown and David Lowe, "Automatic Panorama Image Stitching Using Invariant Features", IJCV, 2011.
- [10] Li Jin, Wang Yanwei, Liang Hong, "Image Mosaic Based on Simplified SIFT", International Conference on Mechanical Engineering and Automation Advances in Biomedical Engineering, Vol.10, 2012.
- [11] Richard Szeliski, "Video mosaic for virtual environment", IEEE computer graphics and application, vol. 16, 1996.
- [12] Yongmin Li, Li-Qunxu, Geoff Morrison, Charles Nightingale & Jason Porphett, "Robust Panorama from MPEG video", IEEE, 2008.
- [13] Stiffen Gauglitz, Tobias Hollerer and Matthew Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking", International journal for computer vision, 2011.
- [14] F. Tarsha-kurdi, T. Landes, P. Grusenmeyer, "Extended RANSAC algorithm for automatic detection of building roof planes from lidar data", The photogrammetric journal of Finland, 2008.
- [15] Michal Irani and P. Anandan, "Video indexing based on Mosaic Representation", ISRAEL, 2000.
- [16] David G. Lowe, "Object Recognition from Local Scale-Invariant Features", proc. of International Conference of Computer Vision, 1999.
- [17] Guozhu Liu, and Junming Zhao, "Key Frame extraction from MPEG video stream", Proceedings of the Second Symposium International Computer Science and Computational Technology, 2009.
- [18] Vladan Rankov, Rosalind J. Locke, Richard J. Edens, Paul R. Barber, "An algorithm for image stitching and blending", proceedings of SPIE, 2005.