

Panoramic Image Stitching Using MATLAB

Milind Rane
Dept. of Electronics Engineering
Vishwakarma Institute of Engineering
Pune, India
milind.rane@vit.edu

Deboshrree Roy
Dept. of Electronics Engineering
Vishwakarma Institute of Engineering
Pune, India
roy.deboshrree16@vit.edu

Abstract—Image processing a form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or, a set of features or limits related to the image. Most image processing methods involve treating the image as a two-dimensional signal and using standard signal processing techniques to it. Specifically, image stitching presents different stages to render two or more overlapping images into a unified stitched image, from the detection of features to merger in a final image

Keywords— Image stitching, MATLAB, panoramic image

I. INTRODUCTION

We observe that a effective image stitching algorithm should not only create a even transition within the overlapped region but also preserve the following properties, which are in general agreement with our visual perception: Structure preservation. Unbalanced contrast beyond the overlapped area of a stitched image can be perceptually overblown. Although the structure is well aligned and color transition is smooth within the overlapped area, the abnormal color transition from left to right reveals the unmatched intensities inherent in the input images. The context data of objects in the input images should be taken into account during the stitching process.

II. LITERATURE REVIEW

Panoramic image stitching has an extensive research literature (Szeliski, 2004; Milgram, 1975; Brown and Lowe, 2003) and several commercial applications (Chen, 1995; Realviz. The fundamental geometry of the problem is well understood, and consists of estimating a 3×3 camera matrix or homograph for each image (Hartley and Zisserman, 2004; Szeliski and Shum, 1997). This estimation procedure needs an initialization, which is typically provided by user input to approximately align the images, or a fixed image ordering. For instance, the PhotoStitch software bundled with Canon digital cameras requires a horizontal or vertical sweep, or a square matrix of images. REALVIZ Stitcher version 4 has a user interface to roughly position the images with a mouse, before routine registration proceeds. In newer techniques, we require no such initialization to be provided. In the research methods for automatic image alignment and stitching, Direct methods have the advantage that they use all of the available image data and hence can provide very accurate registration, but they require a close initialization. Feature based registration does not require initialization, but traditional feature matching methods (e.g., correlation of image patches around Harris corners (Harris, 1992; Shi and Tomasi, 1994)) lack the invariance property needed to enable reliable matching of random panoramic image sequences.

II. METHODOLOGY

There are several methods available for panoramic image stitching using MATLAB. A few of these methods are outlined below.

A. Scale Invariant Feature Detection

The SIFT algorithm (Scale Invariant Feature Transform) proposed by Lowe is a method for extracting characteristic invariant features from images. It has been successfully applied to a variety of computer vision problems based on feature matching including object recognition, pose estimation, image recovery and many others. However, in real-world applications there is still a need for improvement of the algorithm's strength with respect to the correct matching of SIFT features. The main objective is to divide the features extracted from both the test and the model object image into numerous sub-collections before they are matched. The features are separated into several sub collections considering the features arising from different octaves, that is from diverse frequency domains. SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of keypoints that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches. The determination of consistent clusters is performed rapidly by using an efficient hash table implementation of the generalised Hough transform. Each cluster of 3 or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally the probability that a particular set of features indicates the presence of an object is computed, given the accuracy of fit and number of probable false matches. Object matches that pass all these tests can be identified as correct with high confidence.

B. RANDOM SAMPLE Consensus algorithm

The Random sample Consensus algorithm (RANSAC) proposed by Fischler and Bolles is a general parameter approximation approach designed to cope with a large proportion of outliers in the input data. Unlike many of the common effective estimation techniques such as M-estimators and least median squares that have been recognized by the computer vision community from the statistics literature, RANSAC was advanced from within the computer vision community. RANSAC is a resampling practice that produces candidate results by using the least number observations (data points) required to estimate the underlying model parameters. As pointed out by Fischler

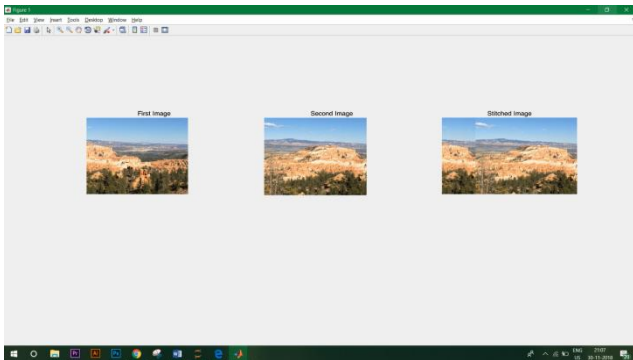
and Bolles, unlike other sampling techniques that use as much of data as possible to obtain an initial solution and then proceed to clip outliers, RANSAC uses the smallest set possible and then enlarges this set with steady data points.

C. Sum of Squares

In this method, the two images we load have one common portion that needs to be used only once. The identification of the mutual portion comes under image registration while identification of optimum point (column) is used as the stitching point. In the code the primary image is read and stored as a new image. A small area (a matrix of twenty rows) is used as a mask. This mask is placed on the first image. The corresponding difference of the pixels is found out, squared and added. The values are stored in a blank matrix. Once the mask covers the pixels of the first image completely, the process is ended. Now the minimum value from the $1 \times$ (columns - 20) matrix is calculated. The matching optimal point (column value) is used as the orientation point for stitching. Now the other image is attached to the original image from the optimal point. This gives us the resulting panoramic image.

III. PRELIMINARY RESULTS

Using the Sum of Squares method, we observe that the final output is a coloured merged image of the two coloured input images.



IV. ALGORITHM

In this code, we first load the images in MATLAB. Next, using 'im2double' we convert the true color image RGB to double precision. We then create a new, empty matrix to store the optimal point that is to be used as a pointer for merging the two images together. A small area (a matrix of 20 rows) is used as a 'mask'. This mask is placed on the first image. The corresponding difference of the pixels is found, squared and added. The values are stored in a blank matrix. Once the mask covers the pixels of the first image completely, the process is ended. Now the minimum value from the $1 \times$ (columns - 20) matrix is calculated. The matching optimal point (column value) is used as the orientation point for stitching. Now the other image is attached to the original image from the optimal point. This gives us the resulting panoramic image.

Steps for implementing SIFT:

- **Constructing a scale space**
Ensure scale invariance. Each octave's image size is half the previous one. In each octave, progressively blur out images. By constructing a scale space, we express the image in different scale. In the following steps, all scales will be taken into consideration and the keypoint representation should be dependent of scales.
- **LoG Approximations**
Fast approximation of Laplacian of Gaussian used to find keypoints.
The results of previous step are used to generate DOG (Difference of Gaussian). That is, the difference between two consecutive scales of each octave. The Laplacian of Gaussian(LoG) operation is effective for finding edges and corners (usually keypoints) on the image. It goes like blurring and calculate second order derivatives. However, LoG is computationally expensive and is not scale invariant. DoG images are approximately equivalent to LoG images but faster.
- **Finding keypoints**
Locate maxima and minima in the Difference of Gaussian image. Iterate through each pixel and check its 26 neighbours (if any). In the following figure, X marks the current pixel. The green circles mark the neighbours. They distribute in three consecutive scales (one octave). Mark maxima and minima. In practice, 5 level of blurs in each octave generates 5 Gaussian blurred images in each octave and will result in 4 DoG images. They can generate extreme images.
- **Getting rid of unnecessary key points**
Edges and low contrast regions are bad keypoints.
- **Removing low contrast features**
- Select a threshold and reject those smaller value pixel. The value is the magnitude of the intensity at the current pixel in the DoG image (minima/maxima in previous step).
Removing edges: Edges are detected by two perpendicular gradients. An edge keypoint has one big gradient (perpendicular to the edge) and one small gradient (along the edge). When both gradients are small, the point is in a flat region. This case is also be rejected. We only accept key points in corners, where two gradients are both big.
- **Assigning an orientation to the keypoints**
Ensure rotation invariant. Calculate gradient magnitudes and orientations for all points in the scale space. Then, an orientation histogram is used to store the orientation of keypoints and their neighbour regions. The peaks in the histogram correspond to dominant orientations.
- **Generate SIFT features**
A 128 dimensions vector is used to represent each keypoint. A 16×16 window around the keypoint is used to generate such distinctive feature. Each pixel/cell has a direction. Every sixteen directions are collected by an orientation histogram. Considering the pixels closer to keypoint (in the

center) affect more than further pixels, we use a “gaussian weighting function” to assign different value. In the following figure, the farther away, the lesser the magnitude (darker).

VI. CHALLENGES OF IMAGE STITCHING

There are many challenges in image stitching such as: Noisy image data or data with uncertainties: An image is often corrupted by noise in its acquisition and transmission, the cost of extracting features is minimized by taking a cascade filtering approach. Very larger images collection need for efficient indexing: large amount of images may lead to high processing time, since each image needs some processing. The main challenge on image stitching is the using of handled camera which may lead to presence of parallax (a shift in apparent object position while observed from different angles of view), small scene motions such as waving tree branches, and large-scale scene motions such as people moving in and out of pictures. This problem can be handled by bundle adjustment. Another recurring problem in creating photo-mosaics is the elimination of visible seams, for which a variety of techniques have been developed over the years. Since the illumination in two views cannot be guaranteed to be the same, stitching two images could create a visible seam. Other reasons for the seam appearing could be the background changing between two images for the same continuous foreground.

In general, the major issues to deal with are presence of parallax, lens distortion, scene motion, and exposure differences.

For panoramic stitching, the ideal set of images will have a reasonable amount of overlap (at least 15–30%) to overcome lens distortion and have enough detectable features. The set of images will have consistent exposure between frames to minimize the probability of seams occurring.

In the non-ideal real-life case, the intensity varies across the whole scene and so does the contrast and intensity across the frames. Lens distortion, motion in the scene, and misalignment all cause ghosting. Also, the ratio of width to height of a panorama image needs to be taken into account to create a visually pleasing composite.

VII. APPLICATIONS OF IMAGE STITCHING

Image stitching is widely used in modern applications, such as the following:

- **“Image Stabilization” feature in camcorders that use frame-rate image alignment**

Image stabilization is a family of techniques that reduce blurring associated with the motion of a camera or other imaging device during exposure. ... With still cameras, camera shake is a particular problem at slow shutter speeds or with long focal length (telephoto or zoom) lenses.

- **High-resolution photomosaics in digital maps and satellite photos:**

Satellite has the high resolution imaging system and is able to collect images with a ground resolution of 0.41 meters (16 inches) in the achromatic or black and white mode. It collects multispectral or color imagery at 1.65-meter resolution or about 64 inches

- **Medical imaging:**

Medical imaging is the technique and process of creating visual representations of the interior of a body for clinical analysis and medical intervention, as well as visual representation of the function of some organs or tissues (physiology).

- **Multiple image super-resolution**

Multi-frame image super-resolution (MISR) aims to fuse information in low-resolution (LR) image sequence to compose a high-resolution (HR) one, which is applied extensively in many areas recently. Different with single image super-resolution (SISR), sub-pixel transitions between multiple frames introduce additional information, attaching more significance to fusion operator to alleviate the ill-posedness of MISR.

- **Video stitching**

It is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image.

VIII. CONCLUSIONS

The panoramic images stitched can be used in applications where the camera is unable to obtain a full view of the object of interest. The full view of the object can be constructed using the image using overlapping regional images acquired for the object. Three main steps of image stitching: registration, calibration, blending the image is also discussed here. Direct technique and feature based technique for image stitching are also described. Furthermore, image stitching model is also described which consists of image acquisition, feature detection and matching, image matching, global alignment and, blending and composition.

REFERENCES

- [1] Y. Yu, K. Huang, and T. Tan, “A Harris-like scale invariant feature detector,” in Proc. Asian Conf. Comput. Vis., 2009, pp. 586–595J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] J. Rabin, J. Delon, Y. Gousseau, and L. Moisan, “RANSAC: A robust algorithm for the recognition of multiple objects,” in Proc. 3D’PVT, 2010.
- [3] Xianyong Fang, Bin Luo, Haifeng Zhao, Jin Tang, Biao He, and Hao Wu. A new multi-resolution image stitching with local and global alignment. IET Computer Vision, 2010.
- [4] M. Krulakova, Matrix technique of image processing in Matlab.
- [5] Beis, J. and Lowe, D. 1997. Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR97)*, pp. 1000–1006.
- [6] R. Benosman and S. B. Kang, editors. Panoramic Vision: Sensors, Theory, and Applications, New York, 2001. Springer.