Review of Feature Extraction and Matching Methods for Drone Image Stitching

M. Dhana Lakshmi, P. Mirunalini, R. Priyadharsini and T. T. Mirnalinee

Abstract Image stitching is the process of combining multiple overlapping images of different views to produce a high-resolution image. The aerial perspective or top view of the terrestial scenes will not be available in the generic 2D images captured by optical cameras. Thus, stitching using 2D images will result in lack of information in top view. UAV (Unmanned Aerial Vehicle) captured drone images tend to have the high aerial perspective, 50–80% of overlapping of information between the images with full information about the scene. This work comprises of discussion about methods such as feature extraction and feature matching used for drone image stitching. In this paper, we compare the performance of three different feature extraction techniques such as SIFT (Scale-Invariant Feature Transform), SURF (Speeded-Up Robust Features), and ORB (ORiented FAST and rotated BRIEF) for detecting the key features. Then the detected features are matched using feature matching algorithms such as FLANN (Fast Library for Approximate Nearest Neighbors) and BF (Brute Force). All the matched key points may not be useful for creating panoramic image. Further, RANSAC (Random sample consensus) algorithm is applied to separate the inliers from the outlier set and interesting points are obtained to create a high-resolution image.

1 Introduction

Many technologies have been developed to produce high-resolution images with a wide view of the scene. However, they have limitation to capture the whole scene at an instance. Panoramic stitching is a technique used widely to overcome this problem. Image or panoramic stitching is the process of combining multiple overlapping images of different views to produce a high-resolution image. The process can be achieved in two ways: the direct pixel to pixel approach and feature-based approach [1]. The direct technique performs matches on each of the pixel to other pixel in

M. Dhana Lakshmi (✉) · P. Mirunalini · R. Priyadharsini · T. T. Mirnalinee
Department of Computer Science and Engineering, SSN College of Engineering, Chennai, India
e-mail: dhanalakshmi1605@cse.ssn.edu.in

D. Pandian et al. (eds.), *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*, Lecture Notes in Computational Vision and Biomechanics 30,
https://doi.org/10.1007/978-3-030-00665-5_59

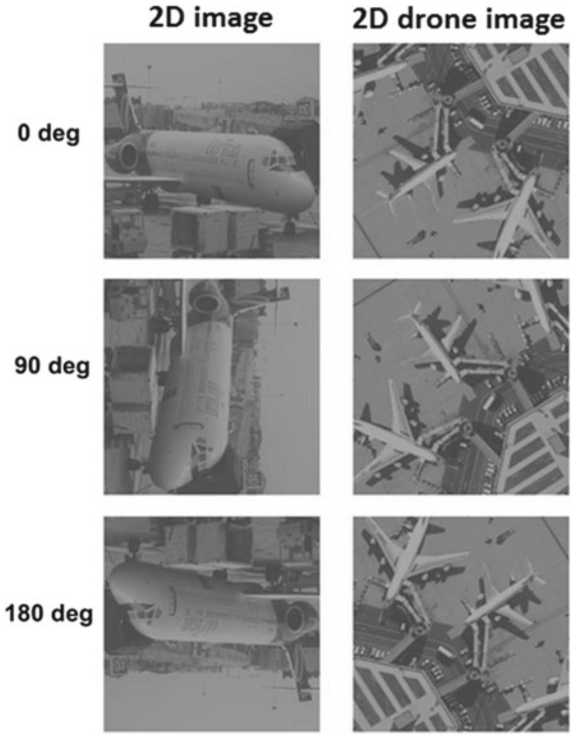

Fig. 1 2D images and 2D drone images with respect to different degree of rotation

order to reduce the mismatches. The feature-based technique, extract the features and perform the matches on the extracted features. Terrestrial scenes (2D images) captured by the digital handheld cameras and smart phones are significantly affected by rotation and become much more difficult to recognize the scene of an image. As the degree of rotation for a terrestrial scene increases, the harder it becomes to identify the scene information as shown in Fig. 1. In stitching of the terrestial scenes, the aerial perspective or top view cannot be retained or captured as such with respect to 2D images. Thus, stitching using 2D images will result in lack of information in top view. When 2D images are used to generate an image stitching of a building, then the roof parts and other structures that are visible from an aerial perspective will not be captured.

But image stitching can also be done using drone images captured by Unmanned Aerial Vehicle (UAV). Drone images contain the high aerial perspective, overlapping of an image and complete information of a scene of an image than the 2D terrestial images. These information helps the image stitching technique to build a high-quality image. The panoramic view of drone images obtained after stitching can be applied in the several fields such as movie industry, civil and mechanical industry.

2 Related Work

The image stitching has been implemented by many researchers in different ways. In [2] the authors reviewed the different feature-based image stitching methods such as SIFT, SURF, and ORB and also proposed a new method A-KAZE. It has been found among three methods that SURF-based visual odometry shows best accuracy for KITTI benchmark dataset. The proposed A-KAZE features demonstrated variation of motion estimation accuracy and computation efficiency. Parallel architecture for image fusion based on ORB feature identifier on a multicore DSP platform has been proposed in [3]. The methodology uses a position weighted image fusion algorithm to stitch the images. A panoramic image stitching technique for rotational images was proposed in [4], which used SIFT and SURF feature detector algorithms and blends the two images using DWT (Discrete Wavelet Transform) after obtaining the matches between the images. A comparison between different feature detector algorithm for the image stitching such as SIFT, SURF, ORB, FAST, Harris corner detector, FAST, MSER detector was done in [5].

3 Feature Extraction Techniques

Keypoints are the dominant features of an image. Features are contributed by the structures and the properties of an image such as color, texture, points, edges, objects, etc. Various feature extraction algorithm such SIFT, SURF and ORB were used to extract key features from drone images.

3.1 SIFT

The SIFT algorithm [6], is invariant to scaling and rotation of an image. It can also handle significant changes in illumination and efficient to run in real time. The SIFT algorithm can be achieved in four steps. It detects a scale space extrema, by generating the several octaves of the original image. Within an octave, images are progressively blurred using the Gaussian Blur operator. Two consecutive images in an octave are taken and one is subtracted from the other. Then the next consecutive pair is taken, and the process repeats. This is done for all octaves. The resultants are the approximation of Gaussian. Localization of keypoint is done by comparing neighboring pixels in the present scale, the successor scale and the predecessor scale. Keypoints can be rejected if they had a low contrast or if they were located on an edge. The assignment of orientation can be done by gradient directions and magnitudes around each keypoint. An image descriptor at each keypoint has been computed using a descriptor generator and stored as a descriptor [2, 7]. SIFT which is scale-invariant, extract the key features of image by resizing the image at different scales. So, all the important features have been extracted due to its invariant property. This property greatly enhances the degree of orientation and performs better in close range and aerial photography.

3.2 SURF

SURF algorithm is fast, robust feature detection and extraction algorithm. It approximates the Laplacian of Gaussian with Box filter and computes local extrema using second-order derivative. Implementation of Haar-like operators over an integral image can fasten the SURF in an efficient manner when compared to SIFT. For orientation assignment, it uses Haar wavelet by applying Gaussian weights for feature description. A keypoint may have connected neighbors which can be chosen and splitted into subparts. The wavelet responses are applied on each of the part to obtain feature descriptor. The features having same type of contrast can perform matching in faster rate [7, 8].

3.3 ORB

ORB is the combination of oriented FAST (Features from Accelerated Segment Test) and rotated BRIEF (Binary Robust Independent Elementary Fast) with some modification in order to enhance the performance of keypoint identification. FAST method is repeatedly applied to each layer of the pyramid in order to achieve the scale-invariant feature. The N keypoints which are computed based on the Harris corner measure are retained and uninteresting keypoints are eliminated. ORB adopts a rotation-aware variant of BRIEF. For detected keypoints, it finds patch centroids by image moments. The moments of a vector, links the keypoint's center to patch's centroid. The binary test pattern is rotated by the moments of a patch which allows feature to be in rotation-invariant form [8, 9].

4 Feature Matching Techniques

Keypoint matching is the process of finding correspondences between two images of the same scene or object. Drone images contain the 50–80% of overlapping between the images. Among the feature points extracted, the points that can be used for stitching are identified by the feature matching methods and the similar points from one image is mapped to points in the other image.

4.1 FLANN

FLANN is a library [10] of optimized algorithms that performs fast nearest neighbor search in high-dimensional features and large datasets. The FLANN uses randomized kd tree algorithm and does the priority search using k-means tree algorithm. Randomized kd-tree algorithm can search multiple trees in parallel by finding a point in the kd-tree which is nearest to a given input point [11]. The search can quickly eliminate the part of the search space by using the tree properties. Priority Search K-Means Tree Algorithm splits the data into M multiple regions and recursively partitioning each zone until the each of the leaf node has no more than M items. Then, picks up the initial centers in random manner [12].

4.2 Brute Force Matcher

BF matcher tries all possibilities and finds the best matches [10]. It takes the descriptor of a feature in an image and compared it using distance measure with all other features in the second image. The nearest point is represented as matched keypoints between the images. The steps involved in Brute force algorithm are as follows:

1. The distance between reference points (first image) and query points (second image) are calculated.
2. The calculated distances are sorted.
3. The *k*-smallest distances are selected as reference points.
4. The steps 1–3 is repeated for all query points.

FLANN considered to be faster since it compares between the nearest points of two images. But BF matcher compares every point in one image with all other point in another image.

5 Image Stitching Technique

An interest point (TP) is a specific location that is recognizable visually in the overlap area between two or more images. Interest points are also considered as inliers which is separated from the outliers to create an image matches.

RANSAC

RANSAC is used to separate matching keypoints (inliers) from non-matching keypoints (outliers) and create an image match. A part of data items is selected aimlessly from the input. An applicable model and the corresponding model parameters are computed using the items of this sample part. The algorithm checks the elements of the whole dataset whether it is consistent with the already represented model. A data element which does not fit the model is considered as outliers. The set of inliers obtained from the applicable model is called concord set. The RANSAC algorithm will run repeatedly until the obtained concord set in certain iteration has enough inliers (valid keypoints). This helps in generation of the interest points used to stitch the images. It also calculates the homography between the images in 3×3 matrix form. This homography matrix establishes the relationship between the two images in order to obtain the stitched image [1, 13–15].

6 Experimental Results

The sample overlapping drone images shown in Fig. 2 are taken from DJI building dataset. These images are captured by the UAV named AscTec Falcon 8 (Ascending Technologies) using Sony NEX-5 (RGB) camera.

The sample images consist of 50–80% of overlapping region. In this paper, feature-based image stitching have been done using Opencv. The methods used for the discussion are as follows:

i. Feature extraction—SIFT, SURF and ORB
ii. Feature matching—FLANN and BF
iii. Interest point generation—RANSAC.

A possible combination of feature extraction and feature matching methods with RANSAC has been performed and the results have been visually compared. The following combination of methods has been carried out and the results are shown below in Fig. 3.

(a) SIFT + FLANN + RANSAC
(b) SIFT + BRUTE FORCE + RANSAC
(c) SURF + FLANN + RANSAC
(d) SURF + BRUTE FORCE +RANSAC
(e) ORB + FLANN + RANSAC
(f) ORB + BF + RANSAC.

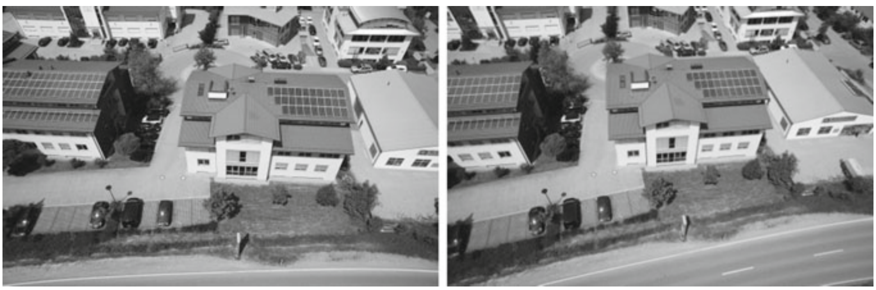

Fig. 2 Sample drone images

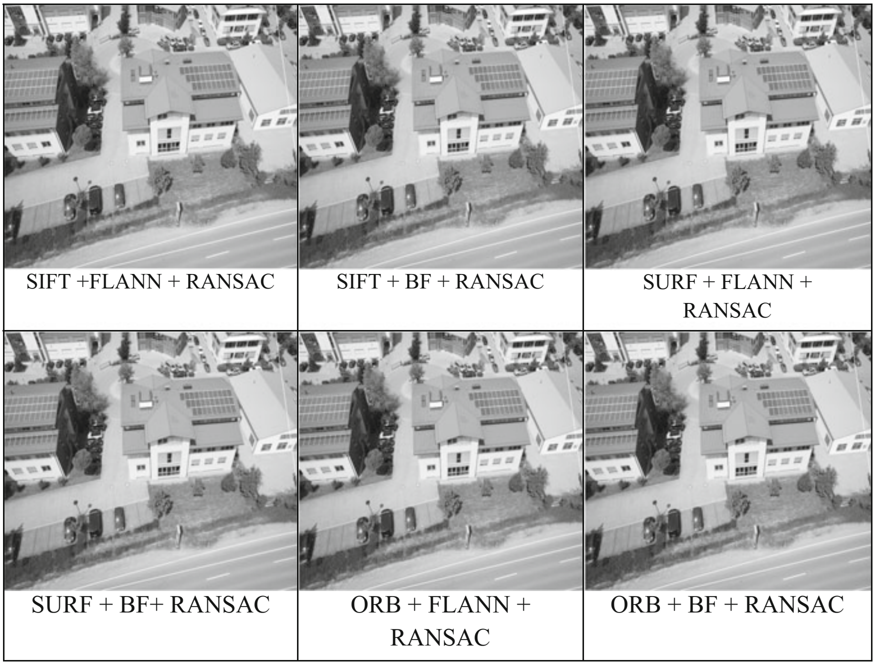

Fig. 3 Experimental results of various feature extraction and matching methods with RANSAC

7 Conclusions

In this paper, SIFT, SURF, and ORB algorithms are used for feature extraction as they efficiently detect the features in distinctive descriptor vector form. On the other hand, ORB extracts the features in binary string. Then feature matching algorithm FLANN and BF are performed on the extracted features. Though the features are matched between the two images of a scene, all matched features are not reliable that is they may not be an interest points used for stitching. RANSAC separates the valid inliers from outliers of matched features. It also calculates the homography between the images in 3×3 matrix form. This homography matrix establishes the relationship between the two images in order to obtain the stitched image. This work has justified that feature-based approach can be used for aerial drone images in image stitching which helps in full view of the scene.

References

1. Szeliski R (2006) Image alignment and stitching: a tutorial. Found Trends Comput Graph Vis 2(1):1–104
2. Chien HJ, Chuang CC, Chen CY, Klette R (2016) When to use what feature? SIFT, SURF, ORB, or A-KAZE features for monocular visual odometry. In: International conference on image and vision computing New Zealand
3. Wang G, Zhai Z, Xu B, Cheng Y (2017) A parallel method for aerial image stitching using ORB feature points. In: 2017 IEEE/ACIS 16th international conference on computer and information science
4. Bind VS, Muduli PR, Pati UC (2013) Robust technique for feature-based image mosaicing using image fusion
5. Adel E, Elmogy M, Elbakry M (2014) Real time image mosaicing system based on feature extraction techniques. In: 9th International conference computer engineering and systems (ICCES)
6. Lowe DG (2004) Distinctive image features from scale-invariant keypoints
7. Karami E, Prasad S, Shehata M (2017) Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images
8. Rao T, Ikenaga T (2017) Quadrant segmentation and ring-like searching based FPGA implementation of ORB matching system for Full-HD video. In: 2017 Fifteenth IAPR international conference on machine vision applications, Nagoya University, Nagoya, Japan
9. Rublee E, Rabaud V, Konolige K, Bradski G (2011) ORB: an efficient alternative to SIFT or SURF. In: 2011 IEEE international conference on computer vision
10. Opencv.org. (2018) OpenCV | OpenCV. [online] Available at: http://opencv.org
11. Suju DA, Jose H (2017) FLANN: fast approximate nearest neighbour search algorithm for elucidating human-wildlife conflicts in forest areas. In: 4th International conference on signal processing, communications and networking
12. Ha Y-J, Kang H-D (2017) Evaluation of feature based image stitching algorithm using OpenCV. In: 10th International conference on (HSI) human system interactions
13. Steedly D (2005) Efficiently registering video into panoramic mosaics. Computer Science Department Faculty Publication Series 84
14. Tidke KS, Banarase SJ (2014) Review on image Mosaicing based on phase correlation and Harris Algorithm. Int J Adv Res Comput Sci Manage Stud 2(2)
15. Agarwal A, Jawahar CV, Narayanan PJ (2005) A survey of planar homography estimation techniques