Real-Time UAV Imagery Stitching Based on Grid-Based Motion Statistics

To cite this article: Cheng Li et al 2018 J. Phys.: Conf. Ser. 1069 012163

View the article online for updates and enhancements.

Related content
- Accuracy of Orthomosaic Generated by Different Methods in Example of UAV Platform MUST Q
  N Liba and J Berg-Jürgens
- An Improved SIFT Algorithm for Unmanned Aerial Vehicle Imagery
  J M Li, D M Yan, G Wang et al.
- Challenges of Integrating Unmanned Aerial Vehicles in Civil Application
  B M Eid, J Chebil, F Albatsh et al.
Real-Time UAV Imagery Stitching Based on Grid-Based Motion Statistics

Cheng Li, Baolong Guo*, Xinxing Guo and Yunpeng Zhi

Institute of Intelligent Control and Image Engineering, Xidian University, School of Aerospace science and technology, No. 2 South Taibai Road, Xi’an, Shaanxi, China.
Email: blguo@xidian.edu.cn

Abstract. In order to receive and stitch sequential images from Unmanned Aerial Vehicle (UAV) synchronously, an improved approach based on feature points matching is proposed. Firstly, global images with overlapping regions are described by Oriented FAST and Rotated BRIEF (ORB) feature. Then Grid-based Motion Statistics (GMS) method is employed to obtain robust feature correspondence from primary matching pairs, these matching pairs are further validated by neighborhood support and used to get transformation matrix. The performance of the proposed algorithm is demonstrated through computer simulated experiments. Experimental results show that the improved method can efficiently solve the problem of smaller-overlapping and less-textured images with real time capability.

1. Introduction

With the rising popularity of cameras mounted on moving platforms, dynamic video image stitching is an important technique to widen the range of application in computer vision. Up to now, it has been widely used in geological prospecting, remote sensing processing, virtual reality and working robot.

Stitching methods exploit geometrically aligning and mosaicking two or more images of the same scene taken into a wide-angle panorama without a seam, in which image registration and fusion are two key techniques. In most cases, the feature points matching methods can deal with translation, rotation and zooming, but it is necessary to raise the efficiency of feature extraction, description and matching in the applications of real-time situation [1].

Image registration techniques based on feature can be generally divided into three categories: function fitting method, statistical model method and graph-based method [2]. LSM (least square method) [3] is a representative function fitting based approach to eliminate false matching points, but the model is built by experimental result, which is not robust for statistical property of image data. RANSAC [4] and its variant [5][6] can rule out outside interference to estimate global optimal parameters by iterative estimation. However, those methods will fail if no sufficient matches are counted. Graph based methods, such as GTM (Graph Transformation Matching) [7], detect false matching via the distribution and neighborhood relation of feature correspondence in a region, which is a KNN based continuously iteration with much time cost.

In this paper, GMS [8] approach is proposed to improve the precision of image stitching. Firstly, ORB [9] features are detected and described in the whole image. Then grid-based motion statistics is detailed to eliminate false matching with low computation cost, which is critical to meet the demands of real time analysis. Finally, several representative match pairs are selected to evaluate the transformation matrix. Experimental results demonstrate that the improved GMS is superior to RANSAC and the proposed algorithm can precisely stitch aerial images at no less than 25fps.
2. Images Description and Registration

The ORB feature makes up the imperfection of FAST and BRIEF with remarkable speed of detection and description. That allows ultra-fast global feature detection with robustness over scale and rotation. The description and registration of two images by ORB is as follows:

2.1. Extraction of Feature Points

The process of feature extraction adopts FAST-9 operator with direction. A corner detector is constructed with given \( n \) (\( n=9 \)), and compare the 16 pixels of the arc of radius \( r \) (\( r=3 \)) to determine the center. Only when the intensity of nine straight pixels are all less than that of center can define it a FAST-9 corner, so that local minima can be avoided.

The direction information of FAST features is produced by intensity centroid, which represents direction with the offset vector between the intensity and the center of a corner. The moments of a patch \( M_{pq} \) is calculated as

\[
M_{pq} = \sum_{(x,y)} x^p y^q I(x,y)
\]  

(1)

where \( I(x,y) \) is the intensity in \((x,y)\), and the centroid of the patch can be represented as

\[
C = \left( \frac{M_{01}}{M_{00}}, \frac{M_{10}}{M_{00}} \right)
\]  

(2)

then, the orientation of the FAST can be defined as

\[
\theta = \arctan \left( \frac{M_{01}}{M_{10}} \right) = \arctan \left( \frac{\sum_{(x,y)} yI(x,y)}{\sum_{(x,y)} xI(x,y)} \right)
\]  

(3)

2.2. Feature Description with Rotation Invariance

The property of rotation invariant is established by rBRIEF descriptor, which overcome the shortcoming of original BRIEF in practicability. Given that a smoothed patch corresponding to one feature point \( p \) can be depicted by intensity contrast, the criterion is defined as

\[
\tau(p;x,y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & p(x) \geq p(y) \end{cases}
\]  

(4)

where \( p(x) \) is the intensity of \( p \), When \( n_d \) pairs of points are selected, the descriptor is defined as a vector of \( n \) binary bit string:

\[
f_{nd}(p) = \sum_{0 \leq d < n} 2^{-d} \tau(p;x_d,y_d)
\]  

(5)

where \( n_d \) generally equals 256 in ORB.

However, the BRIEF acquired above still has no direction, thus an orientation mentioned in (3) is allocated. A steered BRIEF is defined as

\[
g_n(p,\theta) = f_{nd}(p) \big| (x_i, y_i) \in Q_\theta
\]  

(6)

where \( Q_\theta \) is a steered version of \( Q \) with \( R_{\theta} \) being the rotation matrix

\[
Q_\theta = R_{\theta}Q = R_{\theta} \begin{bmatrix} x_1, x_2, \cdots, x_{n-1}, x_n \\ y_1, y_2, \cdots, y_{n-1}, y_n \end{bmatrix}
\]  

(7)
Based on the criterion of highest variance and the lowest correlation, a greedy search aiming to obtain 256 pairs of feature from all correspondence is conducted, which meet the requirement of rotation invariant.

2.3. GMS for Feature Correspondence
Considering that an ORB descriptor is a binary bit string, it is easy to take Hamming distance as matching criterion. However, the correspondence built by ORB inevitably contains mismatches among those good matches, a crucial step is to optimize matching information as fast as possible.

GMS, as well as Grid-based Motion Statistics, is an ideal solution to differentiate true and false matches with high computational efficiency and quality correspondence, which assumes that a true match allows prediction of multiple small region pairs that view the same scenario if motion is smooth over a region. Some notations to be used later is explained as follows:

- Features are detected and matched a feature set \( \{ \mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_N \} \) from an image pair \( \{I_a, I_b\} \);
- \( I_a \) and \( I_b \) are divided into \( G = 20 \times 20 \) non-overlapping cells enumerated as \( i^k \) and \( j^k \) respectively, and \( |X_{i^k, j^k}| \) is defined as the number of matches between cells \( \{i^k, j^k\} \);
- A score for cell-pair \( \{i, j\} \) denoting as \( S_{ij} \) is defined by the equation \( S_{ij} = \sum_{k=1}^{K_{ij}} |X_{i^k, j^k}| \);
- \( s_f \) expresses the standard deviation of the binomial distribution of the false matches;
- The threshold \( \tau \) approximated as \( \alpha s_f \) or \( \alpha \sqrt{n_i} \), which give a per-cell threshold with \( \alpha = 6 \) empirically, where \( n_i \) is the total number of features in the 9-cell neighborhood.

The methodology can be briefed as follows:

- ORB features are detected in both image and then divided into \( G \)-cell grids;
- Cells in an image are looped over for finding the ones which have the highest number of matching feature;
- \( S_{ij} \) and \( \tau \) are calculated for the 9-cell neighborhood of the cell with the highest number of matching, if \( S_{ij} > \tau \), the features matched above are selected as inliers.

3. Algorithm Implementation Process
While GMS builds robust correspondence efficiently, there is many redundancy information to ensure the geometrical relation, which results in large amounts of computation for the projective transformation matrix \( H \). Hence a rough-to-fine mechanism is introduced to sieve the correspondence without representativeness. The following steps explain the complete implementation for stitching:

- Image pair description and registration: A point-to-point correspondence of ORB within the overlapping area is established by GMS for \( \mathcal{X} \);
- Geometrical relation building: For each grid divided into \( G \)-cell, only the match pair with the maximum \( S_{ij} \) value can be preserved in \( \mathcal{X} \) while others are discarded. When \( \mathcal{X} \) completely refines itself, a random sample of \( m \) pairs is extract to calculate \( H \) based the idea of RANSAC, but it requires no iteration;
- Image warping and blending: The target image is warped by \( H \), and then blended together with the reference.

4. Experiment and Analysis
The experiments are performed on a PC with i7 4GHz CPU and 16G RAM, and implemented by C++ with OpenCV 3.1 library. A Matrix 100 UAV by DJ-Innovations is employed for capturing video, which is adjusted to 480p for accommodating slower methods. The scenes are all photographed at
Xidian University, with considered not only the texture complexity but also the overlapping area. The original stitching method based on ORB+RANSAC is used for comparison.

Figure 1, 2 shows the result of ORB+GMS in four conditions. Owing to the ability of GMS to convert multiple matches to high quality correspondence, the proposed method is robust for the overlapped area with low-textured, limited and rotated, in which ORB+RANSAC is prone to failure.
Figure 2. Stitching result in different conditions

| Scenario condition       | Parameters (%, °) | Good Matches | Total Time (ms) |
|-------------------------|------------------|--------------|-----------------|
|                         | Overlapping      | Rotation     | GMS | RANSAC | Proposed | Original |
| Low-textured            | 22.4             | 1            | 596 | 23     | 39.7     | failed   |
| Well-textured           | 31.3             | 0            | 676 | 38     | 40.48    | 55.33    |
| Rotated-overlapping     | 43.8             | 22           | 619 | 37     | 38.8     | failed   |
| Limited-overlapping     | 11.7             | 53           | 320 | 44     | 45.84    | failed   |

Table 1 shows a quantitative comparison of performance of two methods in four conditions mentioned above, in contrast to RANSAC’s lacking sufficient feature matches, GMS efficiently eliminates false matching points and builds smooth consistency among these abundant match pairs, which is convenient for establishing the transformation matrix without large iterative computation and becomes a prerequisite for real-time application.

5. Conclusion
This paper proposes an improved feature points matching approach for UAV image stitching, combining the idea from GMS, which builds a link between feature numbers and match quality, and enables feature correspondence simpler and faster. ORB feature is extracted to describe the candidate video frame, and a rough-to-fine mechanism is introduced to build the spatial relationship. Experiment results show that the improved method can efficiently solve the problems in feature matching based stitch algorithm, such as matching failure result from smaller-overlapping or less-textured and large computation cost by false matching. Moreover, the method satisfies the need for real-time processing.

6. Acknowledgments
This work is supported by the National Natural Science Foundation of China (61571346). The research is also supported by the Fundamental Research Funds for the Central Universities and the Innovation Fund of Xidian University.

7. References
[1] Zhu J, Li C and Xu J. Digital image stabilization for cameras on moving platform [C]. IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2016:255 - 258.
[2] Shan X, Tang P. Review of false matching points detection methods for image matching [J]. Application Research of Computers, 2015.
[3] Yan W D, Tian Z, Wen J H, et al. SIFT mismatching correction based on partial least squares [J]. Journal of Computer Applications, 2012, 32(5):1255-1257.
[4] Tian F, Shi P. Image mosaic using ORB descriptor and improved blending algorithm [C]. IEEE International Congress on Image and Signal Processing, 2015: 693 - 698.
[5] Niedfeldt P C, Ingersoll K, Beard R W. Comparison and analysis of recursive-RANSAC for multiple targets tracking [J]. *IEEE Transactions on Aerospace & Electronic Systems*, 2017, PP (99):1-1.

[6] Vourvoulakis J, Lygouras J, Kalomiros J. Acceleration of RANSAC algorithm for images with affine transformation [C]. *IEEE International Conference on Imaging Systems and Techniques*, 2016: 60 - 65.

[7] Aguilar W, Frauel Y, Escolano F, et al. A robust graph transformation matching for non-rigid registration [J]. *Image & Vision Computing*, 2009, 27(7):897-910.

[8] Bian J, Lin W Y, Matsushita Y, et al. GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2828-2837.

[9] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF [J]. *IEEE International Conference on Computer Vision*, 2011, 58(11): 2564 - 2571.