

Photogrammetric Analysis of Images Acquired by an UAV

Moises Diaz-Cabrera¹, Jorge Cabrera-Gamez¹, Ricardo Aguasca-Colomo¹,
and Kanstantsin Miatliuk²

¹ Instituto de Ingeniera Computacional (SIANI)
University of Las Palmas de Gran Canaria, Spain

moisdca@gmail.com, jcabrera@iusiani.ulpgc.es, raguasca@diea.ulpgc.es

² Automation and Robotics Department, Białystok University of Technology, Poland
k.miatliuk@pb.edu.pl

Abstract. Processing of aerial imagery is a broadly topic discussed nowadays. An Unmanned Aerial Vehicles (UAV) developed in our laboratory was used as experimental platform for the present research. An analysis of the possible application of SURF feature-based algorithm to match outdoor images is introduced. Experimental data comprise selected images taken from different heights (100 and 150 m), different lighting conditions, different pitch, roll and yaw angles, among others effects. The obtained results are validated by using low cost equipment and a low quality video sequence.

Keywords: keypoints detectors, local descriptors, mapping, aerial photography, Unmanned Aerial Vehicles (UAV).

1 Introduction

Unmanned Aerial Vehicles (UAVs) have many applications and they are usually operated by remote control. It saves a human pilot, weight and safety considerations. Since they house sensory devices such as inertial systems or video cameras in particular, it is possible to have an aerial view, augmented by additional physical information. Several missions are often successfully achieved by using this kind of platform. For instance, captured images are determinant in trial issues. Even military missions are usually solved with this kind of vehicles. Thus, people rarely realise if they fly around urban areas. They are ideal to measure devastated areas or interest regions. The range of designed UAV is vast: from micro vehicles, which reach around 500 ft of altitude to heavyweight aerial vehicles, which work in international regions and could weight over 30000 lb.

Improving visual information supported by commercial aerial imagery as Google Maps or Microsoft Virtual Earth, is the motivation of this study. Many areas around the world lack of high quality information, mainly in rural areas. Low cost equipment could provide a new higher resolution cartographic. The Fig. 1 introduces the interest regions which has been analysed in this paper.



Fig. 1. Scene taken from a commercial cartographic (right) and the scene taken from our UAV (left)

That area belongs to a rural area in Gran Canaria Island, which is not correctly mapped yet by commercial aerial imagery.

Our research is oriented towards evaluating the usability of SURF [1] features in solving the matching between cartographic images and images taken by the UAV. We have used an UAV with a single camera, installed in nadir position during the acquisition. Our main contribution in this paper is the study of SURF as tool to detect characteristic points in a set of images from a piece of land.

The outline of the paper is as follows: the section 2 analyses some reference works, the section 3 describes the used vehicle, the main study to detect interesting points by the SURF feature-based algorithm is presented in section 4. The section 5 reveals the results and the 6th section closes the paper with the conclusions.

2 Related Works

In the literature some authors have previously coped with creating orthomosaic by using different techniques. Several organized steps are described to develop orthorectified single images in [4]. This method focuses on the correction of distorted images using the GPS and IMU data. It requires at least three control points on the land and a Digital Elevation Model (DEM) of the surface.

A description for commercial software to create aerial maps is introduced in [6]. The system takes orthorectified and geographically registered imagery. The technique is based on matching feature points, clustering and RANSAC to carry out the correct stitching and develop a map. They have to hand-label a minimum set of control points. In our work, we have analysed a hard area without apparently internal structure. Our goal is focused on an automatic correct matching under the mentioned conditions.

The robustness of an efficient algorithm to detect Maximally Stable Extremal Regions (MSER) is demonstrated in [5]. The robust matching of local features

and the use of multiple scaled measurement regions is exposed as well. This fast detection algorithm could be other possible technique in order to detect extremal points in unclear images, as in our case.

In this work [3], authors have dealt with photogrammetric techniques as well. The achieved results aim at rectification and georeferencing individually each image before generating a composite image of the view.

A hybrid method is committed to get a general mosaic of a region in [7]. The images are received from a set of UAVs flying at low-altitude. They combine inaccurate information on the cameras position and orientation, and the image data to get a high-resolution map. They are able to reduce the GPS errors and the computational time.

In [2], a reflex camera was mounted in an aerial platform to create an aerial map. To do the orthomosaic, two successful ways are proposed: *i*) Taking three control points of a region from other cartographic and *ii*) Using the GPS, IMU for automatic generation and the control points to correct and test the results.

Finally, detector-based algorithms are found in the literature. In spite of the fact single-scale detection approaches offer satisfactory results (i.e.: Harris, Shi-Tomasi or SUSAN detector), multi-scale detectors are often more robust. Harris - Laplace and SIFT algorithms are a suitable option to find singular points on images. However, precise matching study is performed using SURF because of their lower computational cost, beside its robustness both for detectors and descriptors.

3 The Test UAV

Low cost camera: It is a mini camera 1/3" Sony CCD 420 TVL, with a 3.7 mm pinhole lens. It is rather light, just 23 g, without wires. One of the drawbacks is the images quality. Fig 2 depicts the installed camera on the plane.

The Fig. 2 shows the used UAV in the present work. This vehicle is provided with low cost camera. It is a mini camera 1/3" Sony CCD 420 TVL, with a



Fig. 2. Used Unmanned Aerial Vehicle, the camera position and orientation

3.7 mm pinhole lens. It is rather light, just 23 g, without wires. One of its drawbacks is the images quality. Note the camera was in nadir position for the study, although the camera appears in other position in the picture. Firstly, all data are mixed in an on board OSD (On Screen Display). The vehicle incorporates a 2.4 GHz transmitter where GPS, artificial horizon for vehicle orientation and camera data are sent. The video is stored in a multimedia disk. Also it is digitized and transferred to a laptop. The fly route is carried out in a manual way. The UAV sends its position and orientation data by the telemetry channel.

On the one hand, the Zang method was used to correct the distortion of the frames. As measurement of the calibration, the rear projection of the pattern calibration's corner points was used. Corrected images prove that the distortion is more relevant at the images' edge. Important translations and missing data at the images' edges are produced. However, this solution is faster and easier for processing instead of the pincushion distortion processing.

On the other hand, the mode to view the video in the camera is interlaced. We have performed a de-interlaced filter. It is based on a simple horizontal sub-sampling with a bilinear de-interlacing. It allows cleaning each individual frame, improving the sharpness and the clarity. Nevertheless, the size on the images is reduced in 1/4 of proportion, as negative effect.

4 Feature Point Extraction

The study with feature-based SURF algorithm is performed over selected frames from a rural scenario. The vehicle had flown for the same area at different altitude levels, in different illumination and rotation conditions. Fig. 3 shows the two set of images for a same scenario. It is exposed the low quality data and the perceptual aliasing as well: many parts in the photos lack structure.

Original SURF algorithm over these images detects a large number of easily confused keypoints. Confusing keypoints should be filtered during the matching process. For instance, there are repetitive shapes on the images such as



Fig. 3. Set of images captured at different altitude: higher (above) and lower (below) altitude. Red box denotes the reference image.

shrubberies, which are identified by SURF algorithm. To remove keypoints, we establish the next rule: *keypoints with similar descriptors should be filtered*. So, we calculate the interdistances matrix $\mathbf{I}_{k \times k}$ of all descriptors by using the Euclidean norm where k indicates the total number of the more relevant keypoints. In that matrix, the rows and the columns represent the interdistance of one keypoint with the rest of keypoints. After that, the mean from interdistances matrix's rows are computed in a vector $\{\mathbf{m}\}_{i=1}^{i=k}$. Keypoints with the highest values are considered the most remarkable keypoints. Otherwise, the lowest values designate similar keypoints between them. The final selected keypoints will held the highest value in the vector \mathbf{m} .

4.1 Manual Matching

For an in-depth analysis, nine of the most distinctive keypoints, selected using the interdistances matrix, were chosen. Then, SURF algorithm was applied to each image of the set. If one of these keypoints was found by SURF, their descriptors would be stored. The descriptors of the other corresponding keypoints would be stored as well. Finally, nine matrices \mathbf{Kp}_i contained the descriptors of the same keypoints detected in different conditions: illumination and rotation mainly from our low quality set of images. Then, the interdistance matrix was computed on these 9 matrix based on Euclidean norm.

In the ideal case, all values should be zero. It means that the descriptors in all cases were the same. Nevertheless, in real case, the interdistances are not zero, as it is shown in Fig. 4. It manifests the quality measurement of our images. As a reference value for our set of images, $e_m = 0.4$ means the median value for the best keypoints relationship. Since they are the best visual keypoints, $e_m = 0.6$ could be the more conservative threshold to discriminate possible false correspondences.

5 Results

Two experiments were performed. The results are given in term of percentages of the correct and the wrong matching. The first experiment validates the improvement of the designed filter. Then, the second one tries to validate whether our low-cost equipment could create orthomosaics. After a visual inspection of the images, the most significant keypoints were selected.

5.1 First Experiment

This experiment was conducted to expose the effect of the designed filter. An image taken at higher altitude was selected as the reference image. The rest of images were compared with the reference image (see Fig. 3). Firstly, the right and wrong correspondences of a single image with the reference image are shown in table 1 (on the right). Although more than 200 keypoints were identified in all

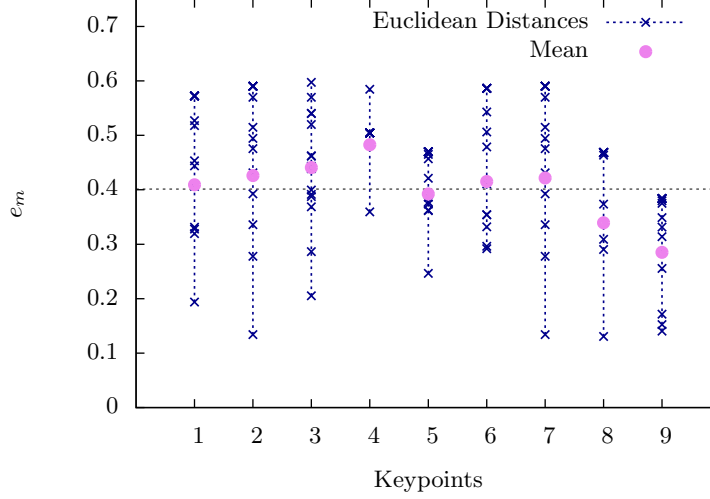


Fig. 4. Quality measurement keypoints based on Euclidean norm

images, only a few good matches were eventually found (see the *Match.* column). The right and wrong correspondences are shown in the last two columns.

Then, we selected only the 25 % of the best keypoints in the reference image, according to our designed filter. The distinguished keypoints were 74 in this case. Now, the procedure is similar: the rest of images were individually matched with the reference image. The matching results are shown in the last two columns in the table 1 (on the left). Here, the right matching are labelled with letter *R* and the wrong matching with letter *W*.

Contrasting both tables, we could deduce that in images captured at the same flight altitude, the right matching rate improve *vs* wrong detections. Comparing images with a wide range of altitude is hard. So, the results were not improved with the images 4, 5 and 6. Although the number of total matching with the filter is obviously lower, the improvement rate is better. Fig. 5 shows an example of matching between two images taken at different flight altitude.

5.2 Second Experiment

The selected nine points in the reference image were matched using the automatically detected points with SURF on the rest of single images.

The column *Finding* in the table 2 represents the found keypoints in the rest of the single images. The rest of the columns have the same meaning as table 1. As curiosity, the 6th image has not got any automatic matching despite having two found keypoints regarding reference image.

According to this table, the next impression is deduced: There were two subset of images: images taken at the same altitude (1,2,3) and images taken at different

Table 1. Correspondences using directly SURF on all images (on the right) and using our filter on the reference image (on the left). Right matching are labelled with letter *R* and wrong matching with letter *W*.

Img.	Kp	Match.	R	W
Ref.	298	-	-	-
1	282	34	23	11
2	287	37	25	12
3	198	34	22	12
4	246	21	9	12
5	207	37	24	13
6	366	21	12	9

Img.	Kp	Match.	R	W
Ref.	74	-	-	-
1	282	14	11	3
2	287	13	10	3
3	198	12	9	3
4	246	7	3	4
5	207	18	10	8
6	366	11	4	7

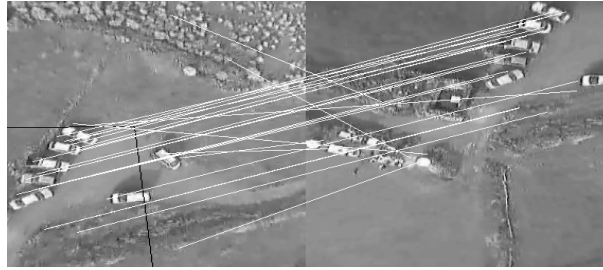


Fig. 5. Example of matching between two images captured at different flight altitude

Table 2. Analysis of matches by using 9 selected points in the reference image

Img.	Kp	Finding	Match.	R	W
Ref.	9	-	-	-	-
1	282	9	3	3	0
2	287	8	3	3	0
3	198	9	6	6	0
4	246	2	1	1	0
5	207	4	3	3	0
6	366	2	0	0	0

altitude (4,5,6) regarding the reference image. So, images taken at the same altitude get more matches than the others. Certainly, the success rate is higher in this last case. Although only 3 keypoints were successfully matched on average, they could be quite enough to deal with the map creation problem.

6 Conclusions and Future Lines

In the present work a complete system to send and receive video and sensor data has been presented. The processing software architecture is organized around two threads that run concurrently: one to receive video data and the other one to tackle the orthomosaic creation.

An experimental study using SURF detector with selected images have been shown. It is focused on the descriptor behaviour in hard experimental outdoor conditions: low quality images, areas with repetitive structure and prone to perceptual aliasing, low descriptor stability and very similar descriptions between several keypoints. In this scenario, descriptors of the same point in two images with the same scene are easily confused. That paper has dealt with searching representative points in outdoor areas with neither structure nor singular points.

A drawback using only SURF to create orthomosaics is the suffered distortion with two images taken at different scales. However, this study is not focusing only in that feature-based algorithm to get the stitching problem, but a simple study to reduce false matching. Obviously, other sensor such as GPS or IMU, among others, could improve the correct matching results. Currently, they could be send trough the telemetry channel and synchronized for each frame for future research. The obtained conclusions encourage improving the results using a higher quality equipment in future research.

References

1. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110, 346–359 (2008)
2. Consentino, F.: Development of an integral system for aero photogrammetric flight planning, data acquisition and imagery ortho-mosaicking. In: *AeroDreams UAV*, Buenos Aires, Argentina (2008)
3. Ladd, G., Magchaudhuri, A., Earl, T., Bland, G., Mitra, M.: Rectification, georeferencing, and mosaicking of images acquired with remotely operated aerial platforms. In: *Proceedings of Annual Conference of American Society for Photogrammetry and Remote Sensing*, Reno, Nevada (2006)
4. Lopez, F.: Formacion de mosaicos ortorectificados utilizando imagenes aereas digitales. Ph.D. thesis, Dissertation thesis in Geography, UNAM, pp. 1-103 (2006)
5. Matas, J., Chum, O., Martin, U., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: *Proceedings of British Machine Vision Conference*, London, pp. 384–393.
6. Pesti, P., Elson, J., Howell, J., Steedly, D., Uyttendaele, M.: Low-cost orthographic imagery. In: *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS 2008*, pp. 24:1–24:8 (2008)
7. Yahyanejad, S., Wischounig Strucl, D., Quaritsch, M., Rinner, B.: Incremental mosaicking of images from autonomous, small-scale uavs. In: *7th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS 2010)*, pp. 329–336 (2010)