IMAV2018-37
http://www.imavs.org/pdf/imav.2018.37

Automatic Combination of Line and Point Descriptors for Thermal Aerial Image Mosaicing

L. González-Guzmán*[1], J. Martinez-Carranza[1,2],
[1]Instituto Nacional de Astrofísica, Óptica y Electrónica, Sta. María Tonantzintla, CP 72480, Puebla
[2]University of Bristol, Bristol, UK, BS8 1UB

Abstract

Thermal aerial image mosaicing is a challenging problem, but it becomes more challenging in low-textured images. Thermal images are a particular case where the image may exhibit large portions of the image with low texture. State of the art methodologies using point-based descriptors often obtain poor results when facing the mosaicing task. Motivated by the latter, we propose a novel approach for thermal aerial image mosaicing that automatically combines line-based and point-based descriptors. Line descriptors are used to obtain a fast and robust estimation of the image transformation model, while point-based descriptors are used to support regions without enough line segments information. The proposed approach was evaluated using thermal aerial image sequences captured at 70 meters and comparing results with three of state of the art algorithms for point-based descriptors.

1 Introduction

Image mosaicing is an important process in Computer Vision and Remote Sensing, it has various applications such as: creation of maps, augmented reality, 3D reconstruction, autonomous vehicle navigation, and tracking. Image mosaicing is an active research area in which it seeks to accelerate its process without losing robustness. However, most of the works reported in the literature have focused on the use of visible spectrum images. Consequently, other bands of the electromagnetic spectrum have been little explored. It is important to mention that a visible spectrum image compared with an infrared spectrum image is very different. While visible spectrum images represent information with light, color and texture; infrared images represent information from the temperature of objects, forming blobs or regions with similar intensities in the image, consequently these images have little texture and absence of colors in the scene [1]. Methods of image mosaicing based on points that use thermal images, generally do not obtain good results due to the lack of texture in this kind of images.

*Department of Computer Science at INAOE. Email addresses: {luis-gonzman, carranza}@inaoep.mx

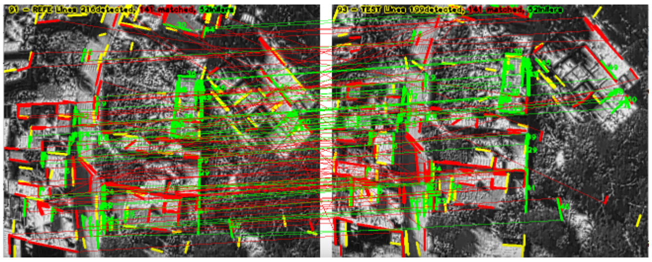


Figure 1: Line segments correspondences between a reference image (left) and a test image (right). Yellow lines represent lines without correspondence. Red lines show wrong matches. Green lines represent inlier matches (good matches). https://youtu.be/TdJ2Xj8nSxo

This research proposes a robust and efficient alternative for the creation of thermal aerial image mosaics using line primitives. Using line segments and their descriptors, it is possible to obtain a reliable description in low texture images, since each line segment is represented by a larger region when it is compared to the region representing by a point. In addition to the extension of a line segment, which represents more information in the discrimination of features, the geometric information, such as angles and intersection points, are used as criteria to obtain an image transformation model namely, a homography. Figure 1 shows a pair of images from which lines were extracted and correspondences were searched to determine a transformation model between two images.

On the other hand, we present an automatic and fast approach based primarily on quadtrees to decide what keypoints should be used to support the line segments. The main idea is to use only those points lying on image areas where no line intersections are located, seeking to make the most of all the available information in the image, this is, extracting points only in certain missing features regions and not in the whole image as would normally be done in a blind strategy, thus reducing the execution time and providing much less points, but quaility points, for the estimation of the transformation model, see Figure 2.

Experiments carried out in this research concentrate on the comparison of point-based techniques with respect to line-based techniques using the proposed approach for thermal image mosaicing. The results obtained show that the proposed methodology is an automatic and efficient alternative

IMAV2018-37
http://www.imavs.org/pdf/imav.2018.37

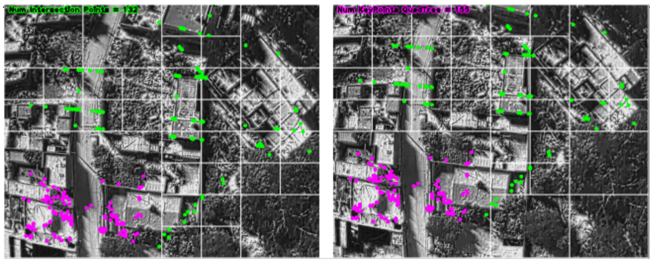


Figure 2: Point correspondences: Green points represent intersection points from inlier line segment correspondences (green lines from Figure 1) and pink points show quadtree-keypoint matches in areas where intersection points were not found.

for thermal aerial image mosaicing. In addition, experiments carried out indicate that our approach is faster than the techniques based on primitive points such as SIFT [2], SURF [3] and ORB [4]. In the case of ORB algorithm that has been commonly used in real-time applications, the proposed approach obtains better results in the final mosaic because due to the adequate selection of features, either line only, points only or both, to be used to generate the homography required to carry out the stitching task.

2 Related work

Image mosaicing in the visible spectrum is a well-studied area where there are numerous works that try to obtain a balance by reducing the execution time and conserving robustness. However, in the literature few works address the problem of image mosaicing in low texture visible or IR images. Some related works presented below that make use of line segments or keypoints in visible or infrared spectrum images.

2.1 Line segments for image mosaicing in visible spectrum images

J. Zhu and M. Ren [5] propose a feature matching method based on SIFT to extract points and form targeted line segments. This method uses the Harris corner detector to extract keypoints in order to build graphs directed from the extracted points. Subsequently, it describes directed line segments using SIFT algorithm and compares them with the line segments determined in another image to achieve the approximate correspondence of points. Finally, the matching points are adjusted and the erroneous pairs are eliminated through the RANSAC method to achieve the mosaic of images. The line segments description of the proposed method continues to preserve the robustness of SIFT algorithm to transformations such as image rotation, distortion and scaling, as well as illumination. However, this method proposed by J. Zhu and M. Ren, has some disadvantages in low texture images such as thermal images, because the techniques they use to extract features (Harris and SIFT) are mainly based on the extraction of information of corners and textures. A similar approach is carried out by Z. Yang et al. [6], where the authors propose to use keypoints detected by SURF to construct directed line segments, the approach proposed by the authors improves the execution time. However, the techniques they use to extract keypoints and form lines depend on regions rich in texture, therefore, their proposal is not robust in images of low texture or thermal images.

2.2 Points for image mosaicing in thermal images

The approach proposed by P. Shah [7] uses thermal images to generate mosaics by implementing a lens aberration correction and a brightness correction method. The developed method uses SIFT for the extraction and description of keypoints due to its robustness features in most transformations. For the estimation of the homography, the author used RANSAC algorithm to eliminate outliers. However, its method is not very accurate in images with little overlap because SIFT does not obtain a sufficiently good set of key points in the thermal images.

On the other hand, Y. Wang et al. [8] performs the creation of mosaics from frames of video captured by a thermal camera. The detection of low level features was performed using SIFT algorithm. The authors comment that although SIFT is one of the most robust methods in the literature, it has problems with the small number of common points between pairs of thermal images and traditional methods such as “Least Square” induce the erroneous registration of correspondences . To deal with this problem Y. Wang et al. propose an easier to implement and highly robust method called “Random M-least square” as a replacement for bundle adjustment. The proposed method is partially related to least squares and also maximizes the number of points of adapted features that fit the transformation model within an accepted registration error. The use of their proposed algorithm increases the correspondence of the points detected by SIFT but does not attack the feature detection problem, so its algorithm is limited to the set of points that the feature extractor detects, this makes that its method work well when SIFT finds a broad set of key points.

2.3 Line segments for image mosaicing in thermal images

In the literature there are no related works where line segments are used directly for the creation of mosaics in thermal images. However, there is a work published by L. Wei et al. [1] that uses line primitives to only estimate the affine matrix as a transformation model between the images. This approach proposes a method of eliminating matches of erroneous line segments, it consists of two steps: The calculation of the affine transformation matrix by intersection points of three randomly selected lines using a modification of RANSAC algorithm, and finding the best transformation by a match score ordering. Once the authors calculate the transformation matrix, they evaluate each line segment by transforming it and measuring the error of the transformation with the center point of the line and a tolerance angle, if this error is less

IMAV2018-37
http://www.imavs.org/pdf/imav.2018.37

than a threshold, the line segment it will be considered as an inlier, otherwise as an outlier. This proposed method has certain drawbacks in not considering perspective transformations between the images and another important aspect is that the central distance of a line is not a reliable metric because the size the line segment may vary between images.

Finally, this research uses line segments to estimate a transformation model and quadtree-keypoints are used as support when no line was detected or when there is not enough line information. Details of our proposal are shown in Section 3.

3 Methodology

This research uses intersection points from line segments to estimate a fast and robust transformation model. Subsequently, the intersection points of the lines labeled as inliers by our RANSAC algorithm are used to generate qudtrees of sub-regions from the image, with the aim of locating regions without sufficient intersection points (green regions, see Figure 4(b)) according to an established threshold. If the percentage of green regions is greater than 60%, keypoints are extracted to obtain a transformation model based on points (see Figure 4(c)) and a comparison is made using the keypoints detected between the line model, points and the combination of both, always looking for the best model that obtains greater number of inlier keypoints. On the other hand, if the number of inliers is relatively close, the model with smaller standard deviation is chosen. Figure 3 shows a diagram of the proposed approach, each stage is described below.

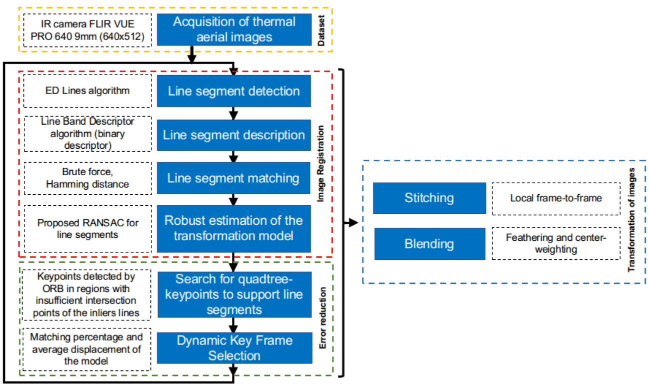


Figure 3: Proposed approach

3.1 System overview

The algorithms developed in this research were implemented in C++ using GNU Compiler Collection (GCC) 5.4.0 and using Open Source Computer Vision Library (OpenCV) 3.3.0 on a personal computer with an Intel Core i7-4720HQ (2.60GHz) processor and 16GB of RAM memory.

3.2 Dataset

Sequences of thermal aerial images were captured using a FLIR VUE PRO 640 infrared camera mounted on a UAV, the images have a spatial resolution of 640x512 pixels. Some of the aerial thermal sequences acquired in this research can be downloaded in the following link: Thermal Aerial INAOE Dataset.

3.3 Line segment detection, description and matching

The detection of line segments was performed by ED-Lines [9] algorithm, which obtains segments of a pixel of thickness and it is considered one of the fastest algorithms to extract this kind of feature. On the other hand, the description of line segments was performed by the Line Band Descriptor [10] algorithm, this descriptor uses bands (sub-regions) near the line segment to extract information by means of local and global convolutional masks, finally it generates a binary vector of mean and standard deviation with the information of the bands. Once line segments were described and detected, brute force and Hamming distance were used to determine correspondences between the pair of images.

3.4 Elimination of erroneous correspondences and robust model estimation

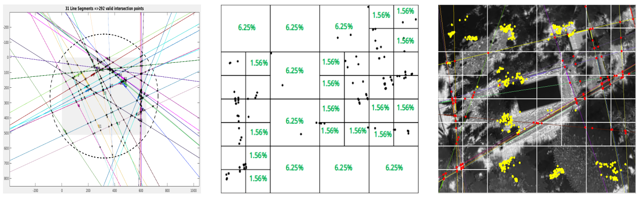
These stages were carried out by a proposed adaptation of RANSAC [11] algorithm for line segments. Similar to the work of L. Wei et al. [1], we use intersection points to estimate the transformation model and eliminate erroneous correspondences. Using minimum distance between line segments and using two transformation error thresholds (distance and angle), it is possible to determine inlier line segments and estimate a robust model that uses geometric information of the structures present in the image. Once the best model was found, our algorithm recalculates the transformation model using all the intersection points of the inlier lines as long as their transformation error is less than or equal to 3 pixels and they are within an established range. Figure 4(a) shows an example of valid intersection points used to recalculate the model.

3.5 Search of support keypoints

In this stage quadtrees were used to identify regions with insufficient intersection points (less than 4 points per quad), see Figure 4(b). If the percentage of quads with insufficiency of intersection points is greater than 60% of the complete image then keypoints of support will be extracted only in these regions and a model based on keypoints will be estimated, otherwise the line model will be used only. The main objective of this stage is to extract points distributed throughout the image and preserve the best model obtained from line segments, keypoints or a mixture between lines and points (the average of line-based and point-based models), which best describes the relationship of the images, see Figure 4(c).

3.6 Transformation of the images

Traditionally in the stitching process, the approaches proposed in the literature consider all the images of the input frames set even if the camera remains immobile or if the change between frames is minimal. It is important to consider



(a) (b) (c)

Figure 4: Search for quadtree-keypoints to support line segments. (a) Valid intersection points. (b) Quadtree generated, green quads represent regions where keypoints will be extracted. (c) Intersection points (red) and quadtree-keypoints (yellow).

that the execution time is high and that errors in the estimation of the homography are propagated for the following ones that fit the canvas because each model depends on the previous ones to form the transition of the sequence of images in the final mosaic. For this reason, this stage was performed frame-to-frame using Dynamic Key Frame Selection (DKFS) proposed by J. Li et al [12] to reduce execution time and accumulated error in the mosaic, see Figure 5. In our DKFS implementation, we consider a relevant frame if its match percentage is less than 40% or if the average displacement of the model is greater than 20% of the image's size. Furthermore, the feathering and centered-weighting techniques were used to fade borders between regions with coincidence in the final mosaic.

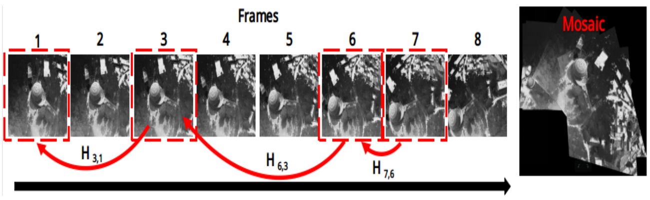


Figure 5: Dynamic Key Frame Selection used to select relevant frames according to sequence transition.

4 Experimental results and analysis

Experiments were performed on a sequence of thermal aerial images composed of 335 frames, this sequence was captured by an UAV on a rectangular route of approximately 660 meters, see Figure 7(a).

In order to evaluate the proposed approach, the comparison of performance, execution time and quality of the final mosaic were made with three of the main algorithms for keypoint detection of state-of-the-art. Some tests were performed using keypoints detected by SURF, SIFT, and ORB algorithms in order to create mosaics of thermal aerial images. On the other hand, the same test was performed using only line segments, and using the proposed methodology combining line segments in conjunction with quadtree-keypoints detected by the three keypoint detection algorithms.

4.1 *Inliers error evaluation*

Table 1 shows the performance obtained in each test, evaluating the average number of features detected, matches, inliers determined by our RANSAC algorithm, average distance error (minimum distance for line segments), average angle error (always 0 if points were evaluated), and finally, the number of frames considered in the mosaic by our DKFS algorithm. In the evaluation, 2000 features were established as limit in the detection of each technique. On the other hand, a feature inlier was defined if it satisfies the following conditions: It has 3 pixels of error or less (points and lines) and it has 3 degrees of error or less (lines only).

Table 1: Performance comparison.

Execution	Average Number			Average Error				Number of frames used in the mosaic
				Distance		Angle		
	Features	Matches	Inliers	Mean	SD	Mean	SD	
SURF	1994	1325	711	1.466	0.727	0.000	0.000	115 / 335
SIFT	1994	1324	670	1.390	**0.724**	0.000	0.000	120 / 335
ORB	1994	1316	726	1.580	0.737	0.000	0.000	93 / 335
LBD	175	124	61	**0.970**	0.809	0.796	0.553	103 / 335
LBD-QT-SURF	1102	668	278	1.305	**0.751**	**0.195**	**0.140**	152 / 335
LBD-QT-SIFT	1011	636	246	1.228	0.758	0.299	0.212	142 / 335
LBD-QT-ORB	499	304	135	**1.215**	0.780	0.421	0.292	116 / 335

The results presented in Table 1 show that the test performed only with line segments (LBD) obtains the lowest average error in distance but the worst standard deviation of distance, also, the average error of angle obtained is the worst of all the tests. On the other hand, the approaches based on the use of line segments and quadtrees-keypoints (LBD-QT-SURF, LBD-QT-SIFT and LBD-QT-ORB) reduce the error of angles compared with the approach of only line segments. Likewise, the test performed with line segments and quadtrees-keypoints detected by ORB (LBD-QT-ORB) obtains the best performance in terms of average distance error, while the lowest average angle error is obtained by the test performed with line segments and quadtrees-keypoints detected by SURF (LBD-QT-SURF). Besides, Table 2 shows the list of models used in the 334 iterations performed by each test.

Table 2: Comparison of the type of model selected in each iteration.

Execution	Type of model selected (Number of iterations)		
	Only Points	Only Lines	Mixed Model
SURF	334	0	0
SIFT	334	0	0
ORB	334	0	0
LBD	0	334	0
LBD-QT-SURF	246	86	2
LBD-QT-SIFT	198	132	4
LBD-QT-ORB	**149**	**180**	**5**

The results show that the test carried out with line segments and quadtrees-keypoints detected by ORB shows a balanced pattern between the use of line and point models, this pattern is the expected one due to the fact that half of the scene lines are abundant (for buildings, roads, etc.), while in the other part of the scene there are very few lines for the vegetation present, see Figure 7(a).

4.2 Run time evaluation

Table 3 shows the average execution times of each stage of image mosaicing according to the previous tests. The results show that the test performed with only line segments is the one with the shortest execution time. Tests performed where points are used without quadtrees (SURF, SIFT and ORB), consume too much execution time for the creation of mosaics. On the other hand, line segments in conjunction with quadtrees-keypoints (LBD-QT-SURF, LBD-QT-SIFT and LBD-QT-ORB) reduce the execution time in all the tests due to the fact that quadtrees locate sub-regions where it is necessary to extract points to support the lines, instead of extracting them throughout the image as originally done in the literature. Figure 6 shows a comparative graph of the average run time per frame of each test.

Table 3: Average run time per frame in each stage of image mosaicing.

Execution	Average Run Time (milliseconds)							
	Reading	Detection and Description	Matching	RANSAC	Quadtrees	Model Adjustment	Stitching and Blending	Total Average Run Time
SURF	4.474	87.199	39.972	349.077	0.000	0.004	0.081	480.806
SIFT	1.203	197.101	77.962	357.441	0.000	0.004	0.052	633.764
ORB	1.045	**21.167**	17.223	369.448	0.000	0.003	0.064	408.949
LBD	0.989	27.727	**0.325**	**147.741**	0.000	0.006	0.085	**176.873**
LBD-QT-SURF	0.974	27.738	0.337	151.710	255.667	0.009	0.053	436.488
LBD-QT-SIFT	0.965	27.985	0.342	149.205	406.651	0.008	0.055	585.212
LBD-QT-ORB	0.979	27.926	0.338	148.598	**102.552**	0.007	0.071	**280.472**

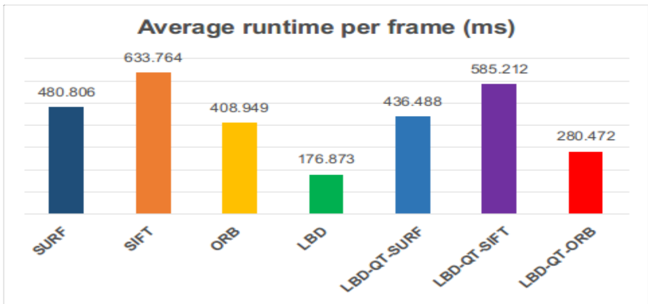


Figure 6: Comparison of average run time per frame in milliseconds.

4.3 Qualitative evaluation of the mosaics

Figure 7 presents the mosaics generated by each test performed. The results show that the use of line segments is not enough for scenarios with a lot of vegetation because the accumulated error affected the aligning of the initial and final parts of the mosaic, see Figure 7(b). On the other hand, the tests where only points are used, like the lines, they present error in certain regions of the mosaic generated by the great ambiguity of intensity in regions with vegetation. Tests done with only keypoints (SURF, SIFT, and ORB) show an approximate reconstruction to the real world scene, however, the runtimes of these tests are very high, see Table 3.

Finally, tests performed with line segments in conjunction with quadtrees-keypoints (LBD-QT-SURF, LBD-QT-SIFT, and LBD-QT-ORB) also show error in some regions of the mosaic. In the case of the test performed with line segments and quadtrees-keypoints detected by ORB (LBD-QT-ORB), see Figure 7(h), it is possible to notice that there are effects of ghosting but this test retains better the shape of the reconstructed region, this is because ORB algorithm determines keypoints near the contours of the objects present in the scene while intersection points of the line segments provide information in the outlines and outside of them, see Figure 2.

5 Conclusions

Image mosaicing in thermal images is an area with great research opportunities, where the majority of state of the art algorithms focus on the use of point-based descriptors. These descriptors are successful in color images, where texture and color patterns are exhibited. In contrast, thermal images lack this information and therefore, we have presented a robust and efficient alternative to solve the problem of image mosaicing in thermal aerial images using line segments in conjunction with a search strategy for support points.

Thus, we have described a novel approach for the generation of thermal aerial image mosaics. Our proposed framework decides in each iteration whether to use a model based on points, a model based on lines or a mixture of both, in order to take advantage of information coming from the structures present in the scene, thus obtaining a model that best describes the relationship between images, without requiring to extract several points in the whole image, hence reducing execution time.

According to our experiments, our proposed approach obtains comparable results with respect to state of the art techniques based on points. We also obtain less execution time and less transformation error.

As a future work, we will push for an optimize version of our approach in order to achieve real-time thermal image mosaicing, also including transferring our methodology to visible images.

6 Acknowledgements

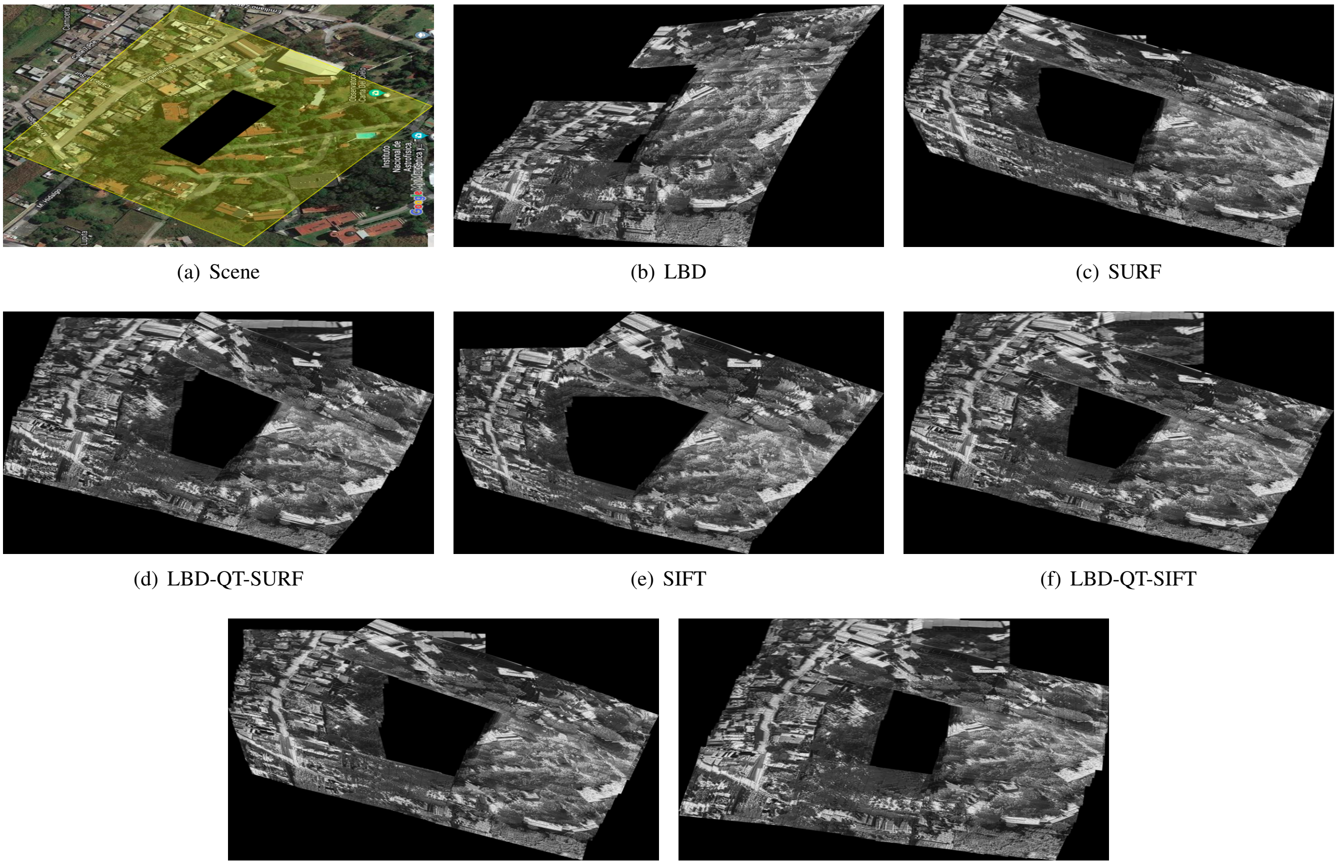
This research has been supported by CONACYT grant 608901 and CONACYT-INEGI project 268528.

References

[1] Lijun Wei, Yong Li, Hang Yu, Liangpeng Xu, and Chunxiao Fan. An improved RANSAC algorithm for line matching on multispectral images. *Proc.SPIE*, 10225:10225 – 10225 – 5, 2017.

[2] David G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

IMAV2018-37
http://www.imavs.org/pdf/imav.2018.37



(a) Scene (b) LBD (c) SURF

(d) LBD-QT-SURF (e) SIFT (f) LBD-QT-SIFT

(g) ORB (h) LBD-QT-ORB

Figure 7: Comparison of mosaics generated using line segments, keypoints and line segments in conjunction with QuadTree-keypoints.

[3] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 2008.

[4] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. ORB: An efficient alternative to SIFT or SURF. In *Computer Vision (ICCV), 2011 IEEE international conference on*, pages 2564–2571. IEEE, 2011.

[5] Jun Zhu and Mingwu Ren. Image mosaic method based on SIFT features of line segment. *Computational and Mathematical Methods in Medicine*, 2014:11, 2014.

[6] Zhanlong Yang, Dinggang Shen, and Pew Thian Yap. Image mosaicking using SURF features of line segments. *PLoS ONE*, 12(3):1–15, 2017.

[7] Preeyanka Shah. Image Processing Aerial Thermal Images to Determine Water Stress on Crops. Technical report, Stanford University, Stanford, CA, 2013.

[8] Yi Wang, Aldo Camargo, Ronald Fevig, Florent Martel, and Richard R. Schultz. Image mosaicking from uncooled thermal IR video captured by a small UAV. *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 161–164, 2008.

[9] Cuneyt Akinlar and Cihan Topal. EDLines: A real-time line segment detector with a false detection control. *Pattern Recognition Letters*, 32(13):1633–1642, 2011.

[10] Lilian Zhang and Reinhard Koch. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. *Journal of Visual Communication and Image Representation*, 24(7):794–805, 2013.

[11] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[12] Jing Li, Tao Yang, Jingyi Yu, Zhaoyang Lu, Ping Lu, Xia Jia, and Wenjie Chen. Fast aerial video stitching. *International Journal of Advanced Robotic Systems*, 11, 2014.