

# A METHOD TO ENHANCE HOMOGENEOUS DISTRIBUTION OF MATCHED FEATURES FOR IMAGE MATCHING

B. Gulmez<sup>1,2\*</sup>, F. Demirtas<sup>1,2</sup>, I. Yildirim<sup>1,2</sup>, U.M. Leloglu<sup>3</sup>, M. Yaman<sup>1</sup>, E. T. Guneyi<sup>1</sup>

<sup>1</sup> ESEN Sistem Entegrasyon, Titanyum C Blok Kat 2, ODTU Teknokent, 06800, Çankaya, Ankara, Türkiye

<sup>2</sup> Hacettepe University, Department of Electrical and Electronics Engineering, 06800 Cankaya Ankara, Turkey

<sup>3</sup> Middle East Technical University, Geodetic and Geographic Information Technologies Department, 06800 Çankaya Ankara, Turkey  
( baran.gulmez , fatih.demirtas irem.yildirim , ugur.leloglu , mustafa.yaman , eylem.guneyi )@esensi.com.tr

**KEY WORDS:** Feature, Homogeneity, Feature Matching, Image Matching, Feature Distribution

## ABSTRACT:

Features are distinctive landmarks of an image. There are various feature detection and description algorithms. Many computer vision algorithms require matching of features from two images. Large number of correct matches with homogeneous distribution in the images is needed for robustness of the image matching. The matches are generally obtained using a feature distance threshold and ambiguous matches are rejected using a ratio test. This paper proposes a method that can be added to image matching pipeline for enhancing homogeneous distribution and increasing the number of matched feature points. After successfully matching an image pair, spatially close feature points go through an elimination process which aims to decrease ambiguity at the second matching step. Then, a coarse geometric transformation between two images is calculated, through which the detected feature points in one image (i.e. the moving image) are projected to the other image (i.e. the fixed image). Then, feature points from the moving image are matched to neighboring feature points of the fixed image within a pre-determined spatial distance. This narrows down the possible candidates and enables less correct matches being rejected because of the ratio test. The effectiveness and feasibility of our method is demonstrated with experiments on images acquired from a drone camera during flight.

## 1 INTRODUCTION

Image matching is defined as computing correspondence between a pair of images and is intensely used in computer vision for various purposes like image registration (Doucette, Antonisse, Braun, Lenihan, & Brennan, 2013), stereo vision (Mur-Artal & Tardos, 2017), image stitching/mosaicking (Ghosh & Kaabouch, 2016), image stabilization (for motion detection) (Brox & Malik, 2011), object detection (Cheng & Han, 2016), visual odometry (Sanfourche, et al., 2012), pose estimation (Lu, Xue, Xia, & Zhang, 2018), simultaneous localization and mapping (Taketomi, Uchiyama, & Ikeda, 2017) and many more.

Homogeneous distribution of image features used for image matching is crucial for the quality of the matching since the computed transformation should cover the whole image area as much as possible for gaining better accuracy (Zhu, Wu, & Xu, 2006). If matched points are clustered in certain parts of the image, the model fitted to these points may cause large errors in other parts of the image for either pose estimation or mosaicking applications. However adding even a single distant point in the unmatched area reduces such deviations significantly.

In this study, we present an algorithm that enhances homogeneous distribution and number of matched feature points for image matching. The proposed method is composed of two parts. The first part is standard feature-based image matching pipeline (Zitová & Flusser, 2003) and the second part is our addition to this pipeline. The proposed algorithm is tested on real aerial images.

This paper is organized as follows: In Section 2, the background is introduced, and related works are reviewed. Next, proposed

method is explained in Section 3. In Section 4, the experiments and results of the proposed method is presented. Finally, the paper is concluded in Section 5.

## 2 BACKGROUND AND RELATED WORK

Image matching methods are mainly divided into two categories: area-based and feature-based methods. In this paper, we mainly focus on improving feature matching methods (Zitová & Flusser, 2003). Features are distinctive landmarks in an image. A feature might be an edge, a corner or a blob (Li, Wang, Tian, & Ding, 2015). Feature detection is the process of detecting these distinctive landmarks in an image. Some algorithms not only detect the features (give location of features), but also create a feature vector using the neighboring pixels around location of the detected feature point. Features are then matched according to the distance (dissimilarity) between their descriptors, which can be Euclidean distance for non-binary descriptors and Hamming distance for binary descriptors (Canclini, et al., 2013). Some descriptors are invariant to scale, illumination or orientation differences between the images (Moreno-Noguer, 2011; Salahat & Qasaimeh, 2017; Sedaghat & Mohammadi, 2018).

A simple algorithm for matching the features is to measure the distances of all possible pairs and choose the ones with the distance below a predetermined threshold as correct matches. However, this parameter by itself is not enough for reliable matching. Another step in the matching process is comparing the distance of the closest neighbor to the second closest neighbor. If this ratio is higher than a maximum distance ratio threshold, the pair is discarded. This method is used in (Baumberg, 2000) and (Lowe, 2004) to prevent ambiguous matches. Also a similar method is used by (Deriche, Zhang, Luong, & Faugeras, 1994)

---

<sup>1\*</sup> Corresponding Author

for the same purpose. Ratio test is useful because distance between the feature vectors of a true match pair must be significantly smaller than that of the second closest match pair. Figure 1 shows the probability density function for correct and incorrect matches in terms of the ratio of the second closest match pair for each feature as proposed by (Lowe, 2004). As can be observed in Figure 1, if this ratio is greater than 0.8, the probability of the incorrect matches is higher. Thanks to this method, most of the incorrect matches can be eliminated in return of discarding a small portion of the correct matches.

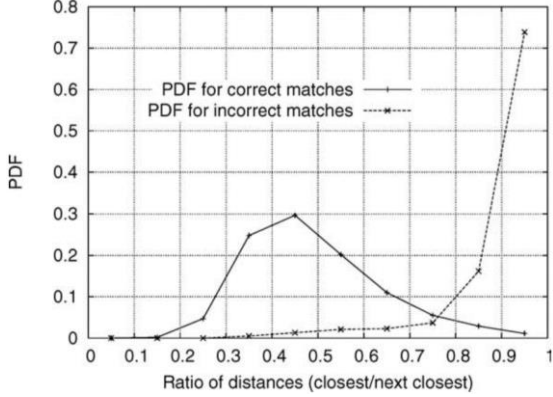


Figure 1. Ratio of distances (Lowe, 2004)

Many methods have been proposed to increase image matching performance in the past several decades. Especially feature-based methods gained more popularity due to their low computational cost and robustness to affine transformations. For example, (Zhang, Wang, & Wang, 2015) propose a two-step method. In the first step a coarse transformation matrix between the moving and fixed images is calculated after matching images using a modified MSER (Matas, Chum, Urban, & Pajdla, 2004). In the second step, the transformation matrix is improved by using the information acquired at the first step. Their method also improves illumination invariance which is a well-known drawback of feature-based methods. (Mishkin, Matas, & Perdoch, 2015) propose MODS which incorporates multiple detectors and utilizes progressive matching. The result is more robust against viewpoint changes than ASIFT (Morel & Yu, 2009) even though its computational cost is lower. Another example of coarse-to-fine method is proposed by (Yu, Yang, Deng, & Yuan, 2018) to register oblique images. Initial matches were acquired with MSER and SIFT (Lowe, 2004) detectors and used to calculate a coarse geometric transform. Then, propagative matches are added using affine invariant normalized cross-correlation. (Yahyanejad & Rinner, 2015) proposed RFAE (Robust Features Across Edges) to match highly affine-distorted thermal and visual images.

On the other hand, it is necessary to have a metric for measuring the homogeneous distribution of the matched features so that the performance of different algorithms could be compared. (Yahyanejad & Rinner, 2015) evaluate their results with a metric proposed for the quality of matching (Eq. 1).

$$Q = \frac{|\cup_i \{x: \|x-f_i\| < \delta\}|}{A} \times \max_{i,j} \frac{\|f_i-f_j\|}{d} \quad (1)$$

The first term in Eq. 1 is the coverage ratio calculated by dividing aggregated area of circles centered at feature points with given radius  $\delta$ . The second term is the maximum distance between two

points normalized by the diagonal of the image. In Eq. 1,  $x$  is a point,  $f$  is an inlier point in the image and  $i, j$  are the indexes of the inliers. Such a metric is needed because other metrics to evaluate the matches on an image such as average nearest neighbor distance ratio (Bansal & Ardell, 1972) decreases with added points. However, even though an added correct match may not significantly increase the matching performance, it also does not decrease it. Their metric perfectly meets these properties while keeping its simplicity. They also indicate that at least nine pairs of matches are needed to reliably match a pair of images. This quality metric is suitable for measuring how well the matches are distributed in an image, hence we use it as a metric of homogeneous distribution of the matches.

### 3 PROPOSED METHOD

The proposed method is composed of adding several steps to the feature-based image matching pipeline. Figure 2 shows the flowchart of the proposed method.

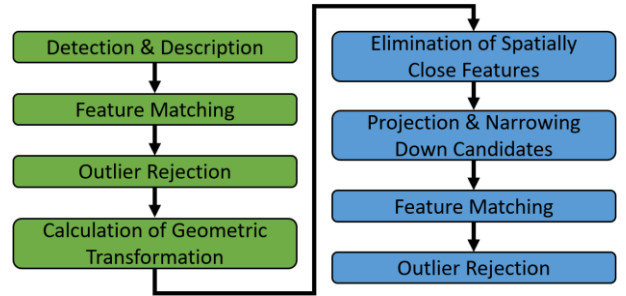


Figure 2. The flowchart of the proposed method

#### Algorithm 1: Elimination of Spatially Close features

**Input:**  $P$  : Feature points of the image  
 $d$  : 2D Spatial distance threshold  
 $m$  : Feature distance threshold  
 $s$  : Scale threshold  
**Output:**  $NC$  : Feature points after the elimination  
**Variable:**  $C$  : Eliminated feature points, initially empty

---

**for all**  $P \notin C$  **do**  
 $P_2 \leftarrow$  All other  $P$  in distance  $d$   
 $C \leftarrow \{ P_2 \mid (P_2 \notin C) \text{ and } (\text{Feature Vector Distance} < m) \text{ and } (\text{Scale Difference} < s) \}$   
**end for**  
 $NC \leftarrow P \setminus C$

---

The algorithm is composed of two iterations. The initially detected features are matched and a coarse geometric transformation between the images is calculated as the first iteration, which is actually running the standard feature-based image matching pipeline (the green part of the flowchart). Then spatially close features are eliminated and using the computed initial geometric transformation, the detected feature points of one image (i.e. the moving image) are projected to the other image (i.e. the fixed or reference image). Each feature point of the moving image is then matched to feature points of the fixed image if it is closer than a predefined spatial distance ( $\lambda$ ) as a threshold. Match parameters at this step may be less tight than the initial matching step, because the candidates for a feature to match are narrowed down so the probability of incorrect matching is decreased. As the number of the candidates for a feature point to match decreases, the ratio test rises in importance such that small changes in the ratio test threshold significantly affects number of matches. That is why spatially close features are eliminated. Spatially close features may also be close in terms

of feature vector distance undesirably. This situation occurs more often for scale invariant feature detectors such as SIFT, SURF (Bay, Ess, Tuytelaars, & Van Gool, 2008) and KAZE (Alcantarilla, Bartoli, & Davison, 2012). Because close locations are detected at different scales. Features that are close in terms of both spatial and feature vector distance prevent potential correct matches due to ratio test. Therefore, at the beginning of second iteration feature points of the fixed image go through an elimination process (see Algorithm 1). For every feature point in the image other features in spatial distance  $d$  are found. Then the candidates that were constrained by spatial distance are also constrained by feature vector distance. Then the remaining candidates constrained by scale as a final similarity measure. This process is applied only to fixed image. Two instances are shown in Figure 3. The final step of matching is the outlier rejection repeated for the newly computed set of matched features.

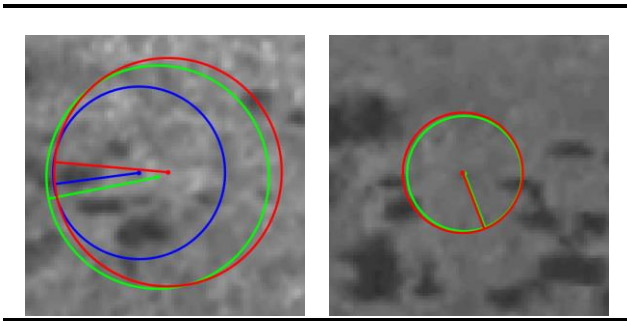


Figure 3 Elimination of spatially close features: green feature is compared to others, red features are eliminated, blue feature is not eliminated due to scale difference.

For computing the initial transformation from the feature points of the moving image to those of the fixed image, the projective transformation is used (Goshtasby, 1986). Projective transformation includes translation, scaling, skew (angle stretch) and rotation. Other geometric transformation methods such as local weighted mean method (Goshtasby, 1988) are not preferred even though they make better approximations. This is because local approximation methods have higher computational cost than projective transformation and does not perform well at regions of image which are distant from matches. Besides, a coarse transformation accuracy is adequate in the first step.

In the proposed method, SURF is used for detection of features and KAZE is used for computing the descriptor vectors of the features. KAZE is selected since it has a highly distinctive descriptor which is scale and rotation invariant.

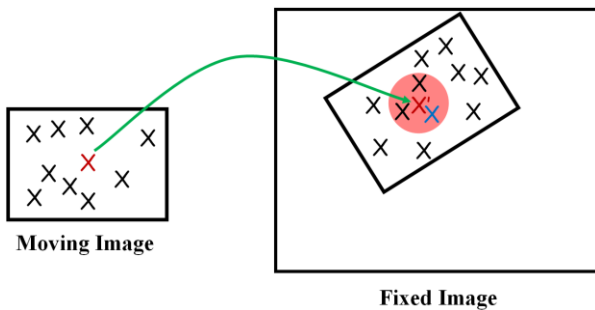


Figure 4. Projection of a point to the fixed image and narrowing down of candidates

The process of projection and narrowing down candidates is depicted for one feature point in Figure 4 where, the feature matching is limited to a spatial distance threshold based on the initial transformation. The process is applied to each detected feature point in the moving image. The feature point indicated with red X is projected to the fixed image and a circular area around the projected feature point (shown with red X) is used for narrowing down candidates with spatial distance.

## 4 EXPERIMENTAL RESULTS

### 4.1 Experimental Setup & Dataset

*MATLAB-2017b* is used to perform the experiments presented in this paper on a desktop PC with *Intel(R) Core (TM) i7-8700 CPU @ 3.20 GHz*.

Tests are realized for 10 pairs of images. 10 aerial photos taken from a drone camera in August 2018 in a rural area at around 500 m AGL altitude are matched to three different orthophotos (Figure 6) with different sizes which are taken in June 2015. The orthophotos are acquired from General Directorate of Mapping which is the national mapping agency of Turkey. The size of all drone images is 2048x4086 with a GSD (ground sampling distance) around  $\sim 0.29$  where the size of the orthophotos are 3219x3281, 2910x3343 and 4231x4595 respectively with a GSD of  $\sim 0.25$  m. Drone images and orthophotos are paired as shown in Table 1.

Drone Image Index	1	2	3	4	5	6	7	8	9	10
Orthophoto Index	a	a	b	b	b	b	b	c	c	c

Table 1. Drone Image - Orthophoto Pairings

The drone images were captured by *DJI Matrice Pro 600* drone with *Basler acA 4096-11ac* camera installed approximately looking at nadir with a 6 mm lens which has 99° of field of view.

A sample image from the drone camera used in the experiments are shown in Figure 5. The contrast enhanced images are presented in Figure 7 for better presentation since the original images has low contrast.

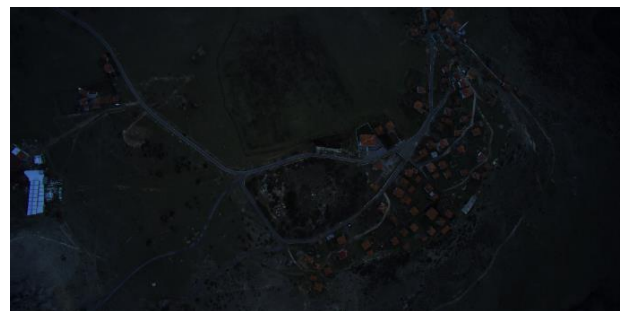


Figure 5 Sample image taken from Drone Camera (Image 6)



Figure 6. The orthophotos used in the experiments as the fixed images.

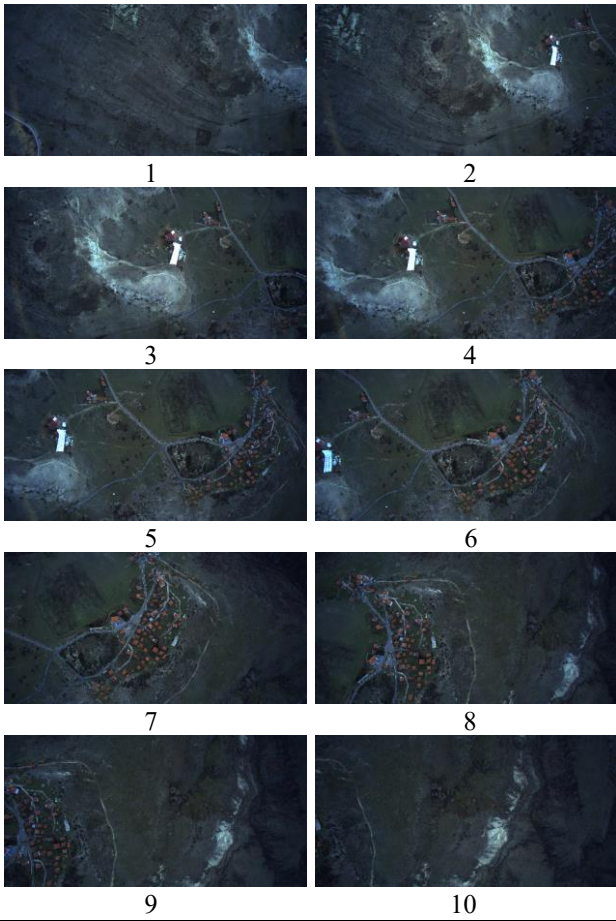


Figure 7 Images taken from the Drone Camera (contrast enhanced versions)

#### 4.2 Experiments & Results

The image matching experiments are conducted using the 10 pair of images in the constructed dataset as explained in the previous subsection.

In the experiments, the distance vectors for feature matching is calculated using sum of squared differences (SSD). Feature distance threshold is 3% and maximum ratio threshold is 60% for elimination of repeated features  $d$ ,  $m$ ,  $s$  are 10 pixels, 1.6 and 25 percent respectively. The  $\lambda$  value for narrowing down candidates after projection is 80 pixels.

The obtained results are provided in Table 2 and Table 3. In Table 2, homogeneous distribution metric for all pairs of images are shown. For all pairs, matching performance is increased with the use of our method. In Table 3, number of matches and inliers for the 1<sup>st</sup> and 2<sup>nd</sup> iterations are shown. An example of inliers which are obtained at the 1<sup>st</sup> and 2<sup>nd</sup> iterations are shown in Figure 8 for the 6<sup>th</sup> drone image.

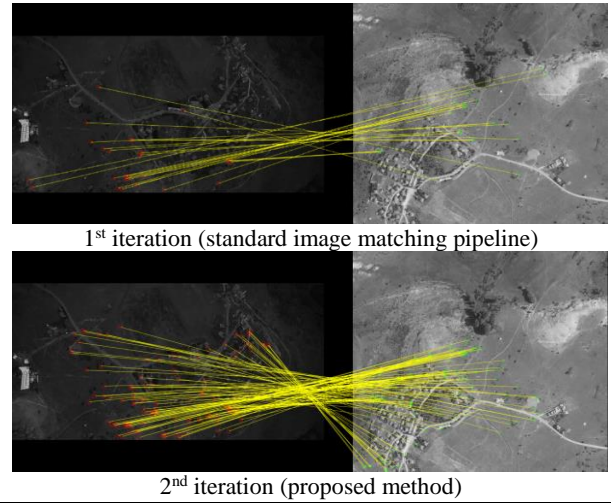


Figure 8 Drone image 6, inliers 1<sup>st</sup> and 2<sup>nd</sup> iterations

3<sup>th</sup> and 4<sup>th</sup> drone images show how much the distribution of matched points can affect the homogeneous distribution metric. As shown in Figure 9, 3<sup>th</sup> and 5<sup>th</sup> images have 45 and 46 inliers at the 1<sup>st</sup> iteration which is quite close. However homogeneous distribution metric for the 5<sup>th</sup> is twice of the 3<sup>rd</sup> image's metric. This situation is also observed when 1<sup>st</sup> and 2<sup>nd</sup> iteration results are compared. The homogeneous distribution metric is not directly proportional to number of matches in the same image.

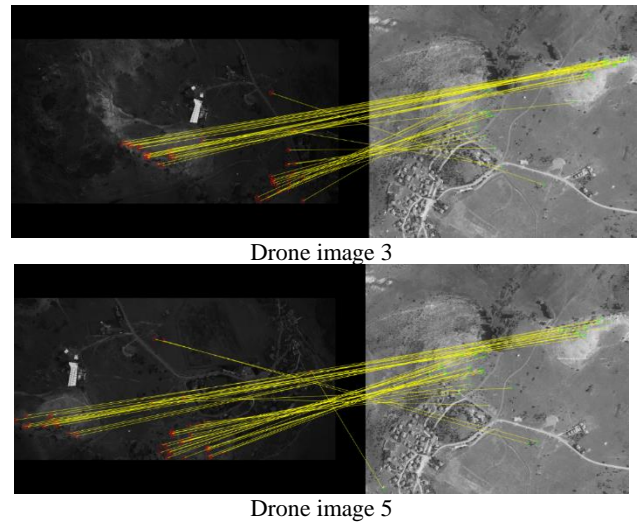


Figure 9 Inliers at the 1<sup>st</sup> iteration for the 3<sup>rd</sup> and 5<sup>th</sup> drone images

Image Index	Homogeneous Distr. Metric for the 1 <sup>st</sup> iteration	Homogeneous Distr. Metric for the 2 <sup>nd</sup> iteration
1	14.66	30.21
2	16.81	34.81
3	8.79	16.92
4	16.04	23.05
5	16.33	24.84
6	10.49	23.49
7	6.22	17.60
8	5.35	16.25
9	4.80	18.72
10	1.39	15.04
Avg.	10.09	22.09

Table 2. Homogeneous distribution metric, as percentage, for inliers after outlier rejection for the 1<sup>st</sup> and 2<sup>nd</sup> iterations

Image Index	1 <sup>st</sup> iteration		2 <sup>nd</sup> iteration	
	Matches	Inliers	Matches	Inliers
1	95	48	360	248
2	90	52	366	294
3	77	45	142	114
4	83	56	221	161
5	88	46	164	118
6	60	28	177	124
7	49	20	136	101
8	38	11	95	54
9	33	8	113	64
10	26	11	148	71
Avg.	63.9	32.5	192.2	134.9

Table 3. Number of matches and inliers for the 1<sup>st</sup> and 2<sup>nd</sup> iterations

## 5 CONCLUSION

A simple yet effective method to increase the number and homogeneous distribution of matched points is proposed in this paper. The proposed method can be used with any feature detection & description method. The contribution of the algorithm is preventing elimination of correct matches because of ratio test through iterative matching. This is achieved while keeping the ambiguous match preventing ability of ratio test. Consequently, number of matches are increased, and the homogeneous distribution of the feature points matched between the image pair is enhanced. The proposed method is shown to be effective with experiments conducted on aerial images.

## ACKNOWLEDGEMENTS

This study was partially supported in the framework of TUBITAK-TEYDEB 1511 program project no: 1170179.

## REFERENCES

Alcantarilla, P., Bartoli, A., & Davison, A. (2012). KAZE features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7577 LNCS, pp. 214-227.

Bansal, P., & Ardell, A. (1972). Average nearest-neighbor distances between uniformly distributed finite particles. *Metallography*, 5(2), 97-111.

Baumberg, A. (2000). Reliable feature matching across widely separated views. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*. 1, pp. 774-781. IEEE Comput. Soc.

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008, 6). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3), 346-359.

Brox, T., & Malik, J. (2011, 3). Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3), 500-513.

Canclini, A., Cesana, M., Redondi, A., Tagliasacchi, M., Ascenso, J., & Cilla, R. (2013). Evaluation of low-complexity visual feature detectors and descriptors. *2013 18th International Conference on Digital Signal Processing (DSP)* (pp. 1-7). IEEE.

Cheng, G., & Han, J. (2016, 7 1). A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117, 11-28.

Deriche, R., Zhang, Z., Luong, Q.-T., & Faugeras, O. (1994). Robust recovery of the epipolar geometry for an uncalibrated stereo rig. In R. Deriche, Z. Zhang, Q.-T. Luong, & O. Faugeras.

Doucette, P., Antonisse, J., Braun, A., Lenihan, M., & Brennan, M. (2013). Image georegistration methods: A framework for application guidelines. *Proceedings - Applied Imagery Pattern Recognition Workshop*. Institute of Electrical and Electronics Engineers Inc.

Ghosh, D., & Kaabouch, N. (2016, 1 1). A survey on image mosaicing techniques. *Journal of Visual Communication and Image Representation*, 34, 1-11.

Goshtasby, A. (1986, 1). Piecewise linear mapping functions for image registration. *Pattern Recognition*, 19(6), 459-466.

Goshtasby, A. (1988, 11). Image registration by local approximation methods. *Image and Vision Computing*, 6(4), 255-261.

Li, Y., Wang, S., Tian, Q., & Ding, X. (2015, 2 3). A survey of recent advances in visual feature detection. *Neurocomputing*, 149(PB), 736-751.

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*.

- Lu, Y., Xue, Z., Xia, G., & Zhang, L. (2018, 1 2). A survey on vision-based UAV navigation. *Geo-Spatial Information Science*, 21(1), 21-32.
- Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22, pp. 761-767. Elsevier Ltd.
- Mishkin, D., Matas, J., & Perdoch, M. (2015, 12 1). MODS: Fast and robust method for two-view matching. *Computer Vision and Image Understanding*, 141, 81-93.
- Morel, J.-M., & Yu, G. (2009, 1). ASIFT: A New Framework for Fully Affine Invariant Image Comparison. *SIAM Journal on Imaging Sciences*, 2(2), 438-469.
- Moreno-Noguer, F. (2011). Deformation and illumination invariant feature point descriptor. *CVPR 2011* (pp. 1593-1600). IEEE.
- Mur-Artal, R., & Tardos, J. (2017, 10). ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics*, 33(5), 1255-1262.
- Salahat, E., & Qasaimeh, M. (2017). Recent advances in features extraction and description algorithms: A comprehensive survey. *2017 IEEE International Conference on Industrial Technology (ICIT)* (pp. 1059-1063). IEEE.
- Sanfourche, M., Delaune, J., Besnerais, G., Plinval, H., Israel, J., Cornic, P., . . . Plyer, A. (2012, 5 1). Perception for UAV: Vision-Based Navigation and Environment Modeling. *AerospaceLab*(4), p. 1-19.
- Sedaghat, A., & Mohammadi, N. (2018, 1 1). Uniform competency-based local feature extraction for remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135, 142-157.
- Taketomi, T., Uchiyama, H., & Ikeda, S. (2017, 12 2). Visual SLAM algorithms: a survey from 2010 to 2016. *IPSJ Transactions on Computer Vision and Applications*, 9(1), 16.
- Yahyanejad, S., & Rinner, B. (2015, 6 1). A fast and mobile system for registration of low-altitude visual and thermal aerial images using multiple small-scale UAVs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104, 189-202.
- Yu, M., Yang, H., Deng, K., & Yuan, K. (2018, 5 19). Registrating oblique images by integrating affine and scale-invariant features. *International Journal of Remote Sensing*, 39(10), 3386-3405.
- Zhang, Q., Wang, Y., & Wang, L. (2015). Registration of images with affine geometric distortion based on Maximally Stable Extremal Regions and phase congruency. *Image and Vision Computing*, 36, 23-39.
- Zhang, Z., Deriche, R., Faugeras, O., & Luong, Q.-T. (1995, 10). A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2), 87-119.
- Zhu, Q., Wu, B., & Xu, Z.-X. (2006, 4). Seed Point Selection Method for Triangle Constrained Image Matching Propagation. *IEEE Geoscience and Remote Sensing Letters*, 3(2), 207-211.
- Zitová, B., & Flusser, J. (2003, 10 1). Image registration methods: a survey. *Image and Vision Computing*, 21(11), 977-1000.