Detection of Control Points for UAV-Multispectral Sensed Data Registration through the Combining of Feature Descriptors

Jocival Dantas Dias Junior, André Ricardo Backes and Maurício Cunha Escarpinati
Faculty of Computing, Federal University of Uberlândia, Uberlândia/MG, Brazil

Keywords: Image Registration, Unmanned Aerial Vehicle, Multispectral Image, Feature Descriptors.

Abstract: The popularization of the Unmanned Aerial Vehicle (UAV) and the development of new sensors has enabled the acquisition and use of multispectral and hyperspectral images in precision agriculture. However, performing the image registration process is a complex task due to the lack of image characteristics among the various spectra and the distortions created by the use of the UAV during the acquisition process. Therefore, the objective of this work is to evaluate different techniques for obtaining control points in multispectral images of soybean plantations obtained by UAVs and to investigate if combining features obtained by different techniques generates better results than when used individually. In this work Were evaluated 3 different feature detection algorithms (KAZE, MEF and BRISK) and their combinations. Results shown that the KAZE technique, achieve better results.

1 INTRODUCTION

It is becoming increasingly more common to see imaging technologies used to aid agriculture in terms of providing precision based tasks, either to estimating crop growth or when identifying characteristics of agronomic interest (Sankaran et al., 2015). In this scenario, the use of unmanned aerial vehicles (UAVs) has gained more and more space due to the reduction of operational costs regarding the use of such technology (Zecha et al., 2013). According to the latest economic report by the Association of Unmanned Aerial Vehicles International, precision agriculture occupies the largest portion of the potential worldwide market for UAVs. (AUVSI, 2013).

Sensors represent a fundamental part of the imaging process, a variety of sensors are being used to scan plants for health problems, record growth rates and hydration, and locate disease outbreaks. The first UAVs used regular comercial cameras that operated in the red, green and blue bands (RGB) and / or in regions near the infrared (Hunt et al., 2010). The newly developed sensors offered the UAVs the possibility of obtaining multispectral and hyperspectral images (Berni et al., 2009).

However, despite the growing use of UAVs to obtain low and medium altitude images (100 to 400 m), the techniques of image processing used requires specialized software. The reason for this is that conventional methods applied to remote sensing image processing are not applied to images obtained by UAV as these methods have been developed to perform data processing on more stable images with a much larger spatial extent than the images obtained by UAVs (Soares et al., 2018).

Using UAVs for image acquisition in precision agriculture requires hundreds and in some cases thousands of overlapping images to cover an area. After acquiring aerial images, it is necessary to perform the registration process of the acquired images, in order to extract agronomic characteristics. In RGB images this process presents some difficulties that are easily identified and resolved, such as changes in lighting, rotations and changes in scale from unforeseen events along the UAV path. However, in addition to the previously mentioned problems for RGB images, we have that most multispectral cameras use different physical sensors to obtain different spectra, which causes a spatial misalignment between the spectra due to their physical displacement. The variation of the analyzed spectrum also leads to a loss of characteristics between the bands which hinders the process of detection of common characteristics between bands.

According to (Banerjee et al., 2018) the registration of two channels is achieved by inferring the necessary transformations from a set of control-point correspondences that pair identical points in the scene on each of the two images. In general, the greater the

Dias Junior, J., Backes, A. and Escarpinati, M.
Detection of Control Points for UAV-Multispectral Sensed Data Registration through the Combining of Feature Descriptors.
DOI: 10.5220/0007580204440451
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 444-451
ISBN: 978-989-758-354-4

number of identified points between the images, the better the alignment of the channels. The traditional approach for multispectral image registration is to designate one channel as the target channel and register all other image channels to the target. There is currently no comparative assessment of the best possible way to perform such a registering procedure.

A framework for the registration of multispectral images in spectral complex environments within the temporal and spectral order is proposed in (Banerjee et al., 2018). The descriptors Harris-Stephens Features (HSF), Min Eigen Features (MEF), Scale Invariant Feature Transformation (SIFT), Speeded-Up Robust Features (SURF), Binary Robust Invariant Scalable Keypoints (BRISK) and Features from Accelerated Segment Test (FAST) were evaluated for this problem. The registration of these images in the spectral order obtained a superior result to the temporal registration, where the best result was obtained by the SURF method. However, the authors state that the use of other descriptors can significantly improve results.

In (Yasir, 2018), an automatic framework was proposed for the registration of multispectral images that define the target channel based on the assumptions that a minimum number of control-points correspondences between two channels is needed to ensure low-error registration, and a greater number of such correspondences generally results in higher registration performance. Basically, this work consists of analyzing all spectra in pairs and identifying the best way to perform the registration so that the steps for the registration of all bands have on average the largest set of control points possible.

In (Junior et al., 2018) the authors performed a comparative analysis between the main descriptors of local characteristics in the context of multispectral registration of images obtained by UAVs. In this work, Harris-Stephens Features (HSF), Min Eigen Features (MEF), KAZE Features (KAZE), Speeded-Up Robust Features (SURF), Binary Robust Invariant Scalable Keypoints (BRISK) and Features from Accelerated Segment Test (FAST) were analyzed. The authors concluded that algorithms that use corner features provide better alignment of multispectral images of crops, and MEF was considered the best algorithm for this process.

In (Faria, 2018) a combination approach of different descriptors is proposed for improving the classification of the interesting cells. Results showed that the union of different features descriptors generates a better classification of cells than the individual application of the same ones.

In this paper we propose the aplication of the approach proposed by (Faria, 2018) for the registration process of aerial images. Our objective is to investigate if the union of descriptors can be used with the framework proposed by (Yasir, 2018) in order to improve multispectral image registration. To accomplish this task we used the following set of descriptors: Binary Robust Invariant Scalable Keypoints (BRISK), Min Eigen Features (MEF), Kaze Features (KAZE) and a combination between MEF and BRISK, MEF and KAZE, BRISK and KAZE, because, according to (Junior et al., 2018), these algorithms obtained on average a superior result in the registration of multispectral crop images.

The authors of the present paper conducted experiments on aerial images of soybean plantations. These images were chosen due to their peculiar characteristics that hinder the registration process. Soybean images have a very similar texture and do not usually contain much information (e.g. roads, lines, trees) that can be used as control points for later alignment.

The remainder of this paper is organized as follows. In section 2, the authors describe the dataset and the main concepts used for the development of this work. Section 3 presents the experiments and results obtained. Section 4 presents conclusions, limitations and future work.

2 METHODS

In this section, the authors of the present paper discuss the datasets used in this work and the characteristics observed during the acquisition of these images. A description of the techniques of extraction and detection of features is carried out. Finally, a brief explanation is presented of the framework used to find the best way to register each dataset.

2.1 Dataset

Three datasets were used, in all cases the datasets are from soybean plantations with a size of 1280×960, a resolution of 96 dpi and average 75% overlap between the images. The channels present in the datasets are, respectively, blue, green, red, near-IR (NIR) and red-edge (REDEG). The datasets were obtained in different soybean plantations located at the following decimal coordinates (-20.379918, -46.242159), (-20.448603, -46.308684) and (-18.730114, -48.772294) respectively. Images from each dataset were obtained on a single flight without any kind of pre-processing. The images were obtained by a MicaSense Red-Edge (see Figure 1) (MicaSense Inc. Seattle, WA, USA) camera coupled in a Micro UAV SX2 (see Figure 2) (Senxis Innovations in Drone Ltda, Uberlândia, MG, Brazil) at an average height of 100 meters. The datasets contains respectively 565 (113 scenes and 5 channels), 670 (134 scenes and 5 channels) and 200 (40 scenes and 5 channels) images. Figure 3 shows an example image scene containing all channels obtained by the Red-Edge MicaSense camera coupled to the Micro UAV SX2.

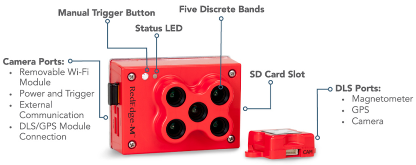

Figure 1: MicaSense Red-Edge camera by MicaSense.

Figure 2: Micro UAV SX2 by Sensix.

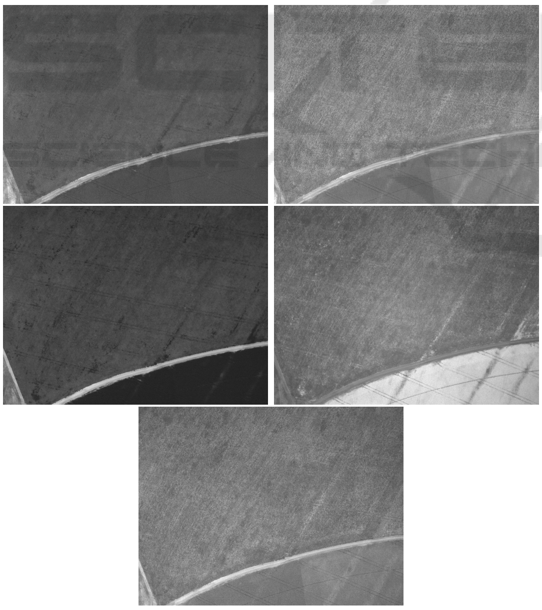

Figure 3: Example image scene containing all channels (Blue, Green, Red, near-IR, red-edge respectively) obtained by the Red-Edge MicaSense camera.

2.2 Image Registration

According to (Zitov and Flusser, 2003) image registration can be defined as a process that overlaps two or more images from various imaging devices or sensors taken at different times and angles, to geometrically align the images for analysis.

In this section, the authors present a review of the main concepts related to the work developed in this paper.

2.2.1 Feature Descriptors

Features, in the context of image registration, can be defined as a pattern that occurs in one location of the image and differs from its closest neighbors. Usually this pattern is associated with a sudden change in one or more properties of an image (e.g., texture, color or intensity). These features may or may not be located at the same location of the change and are usually small areas of the image, corners or points. Descriptors are obtained by performing some type of processing on the region where a feature is present (Kumar, 2014) (Tuytelaars and Mikolajczyk, 2008).

Several techniques for obtaining features descriptors were proposed. In this work, as previously described, the following techniques were analyzed: Min Eigen Features (MEF) (Shi and Tomasi, 1994), Kaze Features (Alcantarilla et al., 2012), Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger et al., 2011), along with the combination of the BRISK and MEF, KAZE and MEF, BRISK and KAZE techniques. In the following, a brief description of these techniques is provided.

Min Eigen Features are obtained by using the Shi-Tomasi Corner Detector algorithm. This algorithm was proposed by (Shi and Tomasi, 1994) and is based on the Harris Corner Detector (Harris and Stephens, 1988) algorithm with a small change in the selection criterion. The change in the selection criterion occurred on the fact that while in the Harris Corner Detector, the eigenvalues are passed to a function that returns the score for the determination as to whether the analyzed pixel is or not a corner, in the algorithm of Shi-Tomasi that function has been removed and only the eigenvalues are considered. As discussed in (Shi and Tomasi, 1994) this change proved to be experimentally superior to the selection criterion proposed in the Harris Corner Detector algorithm. In addition to the best selection method, this algorithm is also invariant to illumination, scale and rotation changes.

The Kaze algorithm was proposed by (Alcantarilla et al., 2012) for the purpose of detecting 2D features in a nonlinear scale space to obtain greater accuracy of localization and distinctiveness. The Gaussian blurring method used to generate the space scale in other algorithms does not maintain the natural edges of the analyzed image and also the noise is smoothed at all scaling levels. In order to solve this problem, the KAZE algorithm uses non-linear diffusion filtering in conjunction with the Additive Operator Splitting (AOS) (Andersson and Marquez, 2016) method. This algorithm is also invariant to illumination, scale and rotation changes.

The Binary Robust Invariant Scalable Keypoints (BRISK) was proposed by (Leutenegger et al., 2011) in order to be an algorithm with a high performance, however, with a drastic reduction of computational cost when compared to algorithms like SIFT or SURF. To obtain the localization of characteristics, the BRISK algorithm uses the AGAST Corner Detector (Mair et al., 2010) technique, which holds a performance improvement over the FAST algorithm. To deal with scale changes, the BRISK algorithm finds the points of interest within a space of scales, applying the technique of non-maximum suppression (NMS) and by performing an interpolation between all the scales. The BRISK algorithm is invariant to scale and rotation changes.

The combination between techniques is accomplished by extracting all the features of both images with the two techniques to be combined. Subsequently, the feature vectors are vertically concatenated forming only one vector. Following this, the removal of duplicate features is performed.

2.3 Registration Framework

The authors here used the data-driven framework for multispectral image registration proposed by (Yasir, 2018), and summarized in Figure 4. Two changes were made in step 2 of this framework. The first change is due to the fact that this work aims at the analysis of the image registry using only one technique (be it single or a combination of techniques). Therefore, after the extraction of the control points, a graph will be constructed for each technique described above. The second is the removal of equal control points and the removal of outlier points via random sample consensus (RANSAC) algorithm (Fischler and Bolles, 1981).

This framework consists of the construction of a complete graph where the nodes of the graph are the channels to be registered and the weights the quantity control points obtained by the algorithms between those channels (see Figure 5). Then, using the Kruskal (Kruskal, 1956) algorithm, a Maximum Spanning Tree is constructed. To find the channel to be used as the target for the other alignments, the weights between the nodes are replaced by 1 and the Floyd-Warshall all-pairs-shortest-path (Floyd, 1962) algorithm is used. The node with the smallest sum of distances from itself to all the other nodes is selected as the target channel for the registration scheme.

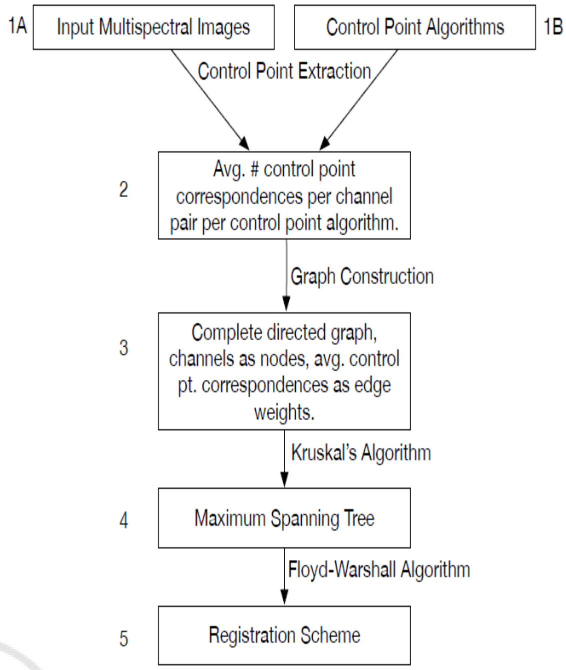

Figure 4: Registration Scheme proposed by (Yasir, 2018).

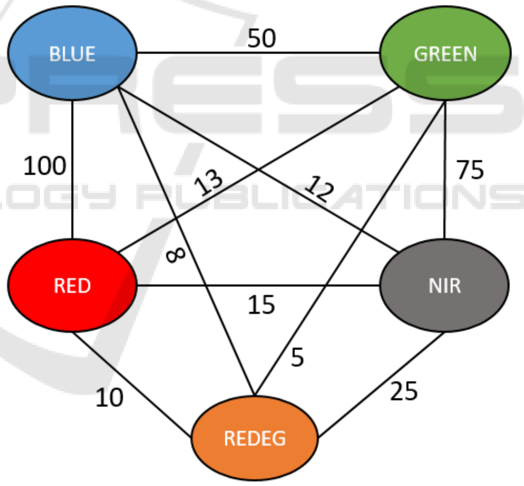

Figure 5: Example of a complete graph generated by the framework.

With the aim to perform the comparison between techniques, the authors use the number of distinct control points found after the removal of the outlier. The number of control points to be detected by each technique are not limited.

An example of a result of the framework described above is demonstrated in figure 6.

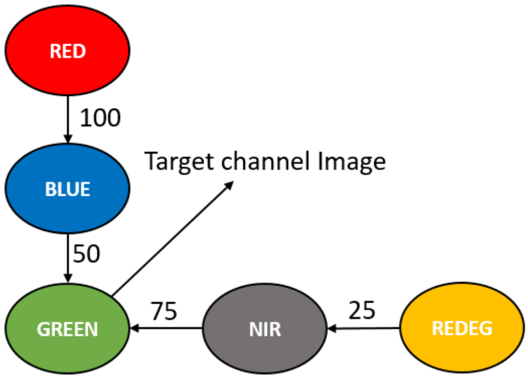

Figure 6: Example of output from Registration Scheme.

Figure 7: Example image of Soybean Dataset 01.

3 EXPERIMENTS

Each technique, or combining, (MEF, BRISK, KAZE, MEF and BRISK, MEF and KAZE, BRISK and KAZE) was applied individually over the datasets and the results obtained are presented in this section. As previously described, in this work the techniques are being evaluated by the number of control points found.

3.1 Soybean Dataset 01

Soybean Dataset 01 consists of 113 scenes in five distinct channels (blue, green, red, near-IR and red-edge) resulting in a total of 565 images. Some images of this dataset have few elements that favor the process of alignment between the bands (e.g. plantations lines and roads), it can be seen in the figure 7. For this dataset the technique that generated the highest average number of control points was Kaze Features (see figure 8). This technique recognized on average 459 points on each image. The schema for the registry after all steps of the framework is shown in Figure 9. As previously described, this framework generates a scheme for registration between the channels of an image, in order that the greatest number of control points between the channels are obtained. This scheme also demonstrates that the red and near-IR channels have to be first registered in the blue and red-edge channels respectively, and then later recorded in the green channel.

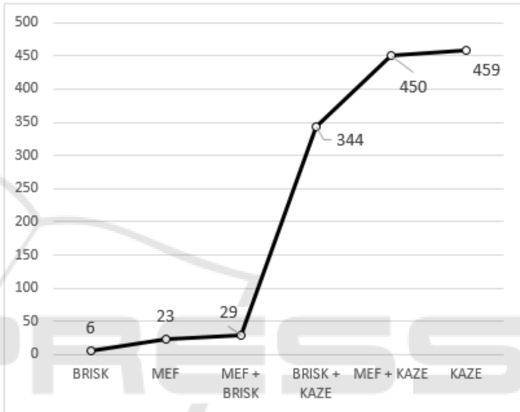

Figure 8: Average points per image in dataset 01.

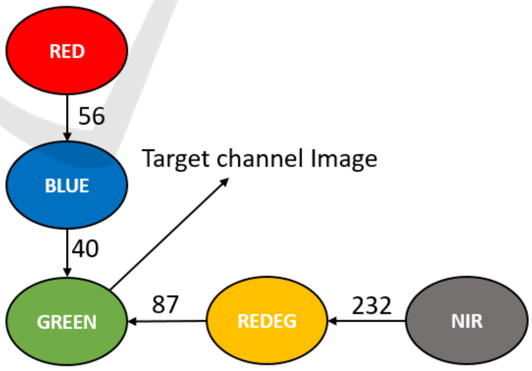

Figure 9: Dataset 01 registration scheme using the KAZE technique.

3.2 Soybean Dataset 02

Soybean Dataset 02 consists of 134 scenes in five distinct channels (blue, green, red, near-IR, red-edge) resulting in a total of 670 images. Differently from dataset 01, the images present in this dataset present several elements that favor the alignment process between the bands (eg planting lines, trees, roads, planting failures), an image exemplifying these elements can be seen in the figure 10. Due to these characteristics, a higher number of control points was obtained when compared to the other datasets. For this dataset the best from among the techniques, when evaluating the average number of control points, was obtained throught the combining of those features obtained by the KAZE technique with those from the MEF technique (see figure 11). A total of 2406 control points were recognized for each image. The large number of control points available in the dataset images allowed algorithms such as MEF to identify features not detected by the KAZE technique. For this reason, the combination of features obtained by MEF and KAZE achieved a superior result. The schema for the registration after all framework steps is shown in Figure 12. This scheme demonstrates, as in Dataset 01, that red and near-IR channels have to be first registered in the blue and red-edge channels respectively and subsequently registered in the green channel.

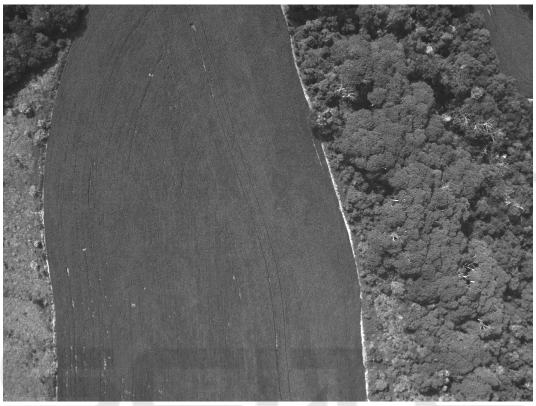

Figure 10: Example image of Soybean Dataset 02.

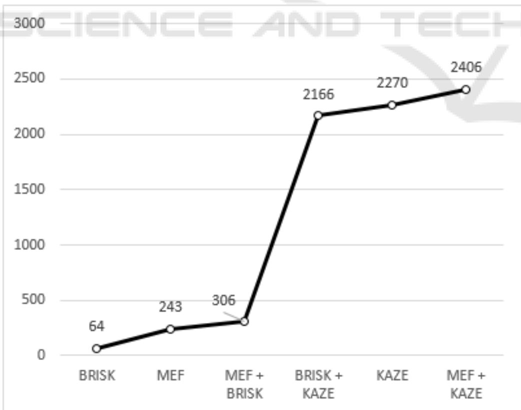

Figure 11: Average points per image in dataset 02.

3.3 Soybean Dataset 03

Soybean Dataset 03 consists of 40 scenes in five distinct channels (blue, green, red, near-IR, red-edge) resulting in a total of 200 images. Dataset 03 presents only a few planting lines to be used by the algorithms as control points. An image exemplifying the dataset 03 is shown in the figure 13. For this dataset we have the technique that generated the highest average

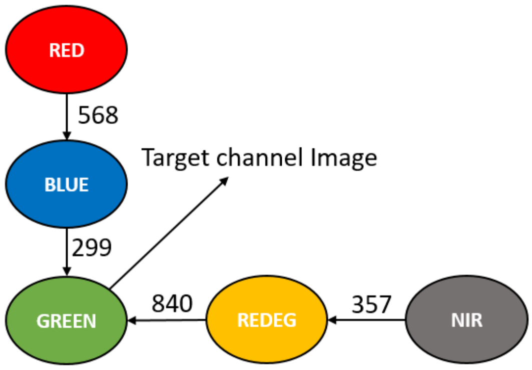

Figure 12: Dataset 02 registration scheme using the MEF+KAZE technique.

Figure 13: Example image of Soybean Dataset 03.

number of control points as being Kaze Features (see figure 14). This technique recognized on average 400 points on each image. The schema for the registry after the conclusion of all steps in the framework is shown in Figure 15. This structure presents a change in the resulting scheme for datasets 01 and 02, unlike the previous schemes, the near-IR channel is no longer be registered in the red-edge channel. All channels, with the exception of red, which continues to be registered in blue, is recorded directly on the target channel. For this dataset, just as with the prior sets, the target channel is green.

4 CONCLUSION

In this work was explored the application of different techniques for the registration of multispectral images obtained by UAVs in soybean plantations. The authors also evaluated how the combination of features obtained by different descriptors impacts on the number of control points obtained.

As seen in the experiments section, on average, the KAZE technique provided more control points for

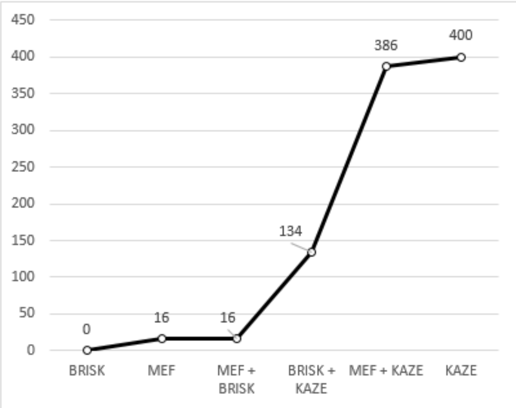

Figure 14: Average points per image in dataset 03.

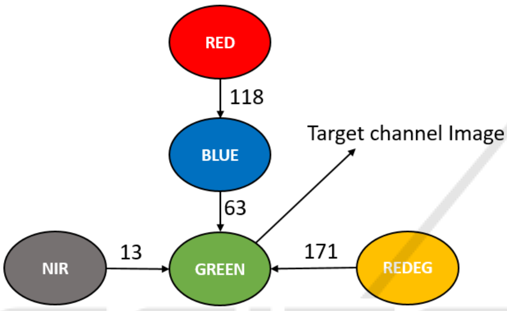

Figure 15: Dataset 03 registration scheme using the KAZE technique.

the registration process than other techniques. In Dataset 02, the combination of the features obtained by the MEF and KAZE techniques achieved the greatest number of control points among the others techniques. However, throught an analysis of the difference obtained between the results, when compared to other factors (e.g. complexity to obtain features of two different techniques), one concludes that the KAZE technique was superior to other techniques evaluated in this paper.

One of the factors that led to the KAZE technique to obtain a large number of control points than MEF and BRISK techniques is mainly due to how the KAZE technique performs the scale invariance process. The KAZE algorithm detects features in a nonlinear scale space by means of nonlinear diffusion filtering. Thus, the KAZE algorithm makes blurring locally adaptive to the image data, reducing noise but retaining object boundaries, while obtaining superior localization accuracy and distinctiveness (Alcantarilla et al., 2012). This is quite different to BRISK that uses Gaussian blurring, which, according to (Alcantarilla et al., 2012), does not respect the natural boundaries of objects and smooths to the same degree both the details and noise, reducing localization accuracy and distinctiveness. The MEF algorithm, although invariant to scale, presents a considerable degradation in the repeatability of the features when the scale changes, as demonstrated by (Schmid et al., 2000). The scale invariance is extremely important when we are evaluating images obtained by UAVs since, as previously demonstrated, there are several factors (e.g. wind speed, wind direction) that influence the path of the UAV leading to distortions of scale and rotation.

Noteworthy is that althought dataset 2 presents a small modification, the scheme for channel registration is very similar among all datasets. The green channel is always used as the final target for the other channels, where red and near-IR are first recorded in the blue and red-edge channels respectively. Through a comparison of the registration scheme proposed by (Banerjee et al., 2018), the spectral order (blue, green, red, red-edge, NIR) was verified as not being the best in any of the datasets evaluated.

The main limitation of this work is to have evaluated only the number of control points obtained between channels. The quality of the alignment was not evaluated in this work due to the lack of previously aligned datasets containing multispectral images obtained by UAVs at low/medium altitude (100 to 400m). Satellite images cannot be used for this purpose due to particular characteristics, such as higher spatial resolution and greater stability. The datasets presented in this work are manually aligned for analysis purposes.

Future work proposals are based on the evaluation of bioinspired algorithms (e.g. Genetic Algorithms, Swarm Intelligence, etc.) to perform the alignment of multispectral crop images. Also, as previously described, another study proposal that is under development is the creation of an aligned dataset containing multispectral images obtained by UAVs to evaluate the algorithms.

ACKNOWLEDGEMENTS

The authors gratefully acknowledges CAPES (Coordination for the Improvement of Higher Education Personnel), CNPq (National Council for Scientific and Technological Development, Brazil) (Grant #302416/2015-3), FAPEMIG (Foundation to the Support of Research in Minas Gerais (Grant #APQ-03437-15) for the financial support and the company Sensix (http://sensix.com.br) for providing the images used in the tests.

REFERENCES

Alcantarilla, P. F., Bartoli, A., and Davison, A. J. (2012). Kaze features. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., and Schmid, C., editors, *Computer Vision – ECCV 2012*, pages 214–227, Berlin, Heidelberg. Springer Berlin Heidelberg.

Andersson, O. and Marquez, S. R. (2016). A comparison of object detection algorithms using unmanipulated testing images: Comparing sift, kaze, akaze and orb. Master's thesis.

AUVSI (2013). The economic impact of unmanned aircraft systems integration in the united states march. Technical report, The Association for Unmanned Vehicle Systems International.

Banerjee, B. P., Raval, S. A., and Cullen, P. J. (2018). Alignment of uav-hyperspectral bands using keypoint descriptors in a spectrally complex environment. *Remote Sensing Letters*, 9(6):524–533.

Berni, J. A. J., Zarco-Tejada, P. J., Suarez, L., and Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3):722–738.

Faria, L. C., R. L. F. M. J. F. (2018). Cell classification using handcrafted features and bag of visual words. In *Proceedings of XIV Workshop de Visão Computacional*.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395.

Floyd, R. W. (1962). Algorithm 97: Shortest path. *Commun. ACM*, 5(6):345–.

Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151.

Hunt, E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S. T., and McCarty, G. W. (2010). Acquisition of nir-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sensing*, 2:290–305.

Junior, J. D. D., Backes, A. R., Abdala, D. D., and Escarpinati, M. C. (2018). Assessing the adequability of image feature descriptors on registration of uav-multispectral sensed data. In *Proceedings of XIV Workshop de Visão Computacional*.

Kruskal, J. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. In *Proceedings of the American Mathematical Society*, pages 48–50.

Kumar, R. S. M. (2014). A survey on image feature descriptors.

Leutenegger, S., Chli, M., and Siegwart, R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. In *2011 International Conference on Computer Vision*, pages 2548–2555.

Mair, E., Hager, G. D., Burschka, D., Suppa, M., and Hirzinger, G. (2010). Adaptive and generic corner detection based on the accelerated segment test. In Daniilidis, K., Maragos, P., and Paragios, N., editors, *Computer Vision – ECCV 2010*, pages 183–196, Berlin, Heidelberg. Springer Berlin Heidelberg.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., Miklas, P. N., Carter, A. H., Pumphrey, M. O., Knowles, N. R., and Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. 70:112–123. Exported from https://app.dimensions.ai on 2018/11/20.

Schmid, C., Mohr, R., and Bauckhage, C. (2000). Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172.

Shi, J. and Tomasi (1994). Good features to track. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600.

Soares, G., Abdala, D., and Escarpinati, M. (2018). Plantation rows identification by means of image tiling and hough transform. pages 453–459.

Tuytelaars, T. and Mikolajczyk, K. (2008). Local invariant feature detectors: A survey. *Found. Trends. Comput. Graph. Vis.*, 3(3):177–280.

Yasir, R. (2018). Data driven multispectral image registration framework. Master's thesis, University of Saskatchewan.

Zecha, C. W., Link, J., and Claupein, W. (2013). Mobile sensor platforms: categorisation and research applications in precision farming. *Journal of Sensors and Sensor Systems*, 2(1):51–72.

Zitov, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11):977 – 1000.