

# A NORMALIZED SURF FOR MULTISPECTRAL IMAGE MATCHING AND BAND CO-REGISTRATION

J.P. Jhan <sup>1,2</sup>, J.Y. Rau <sup>1</sup>

<sup>1</sup> Department of Geomatics, National Cheng-Kung University, 701, No. 1, University Road, Tainan City, Taiwan - (jyunpingjhan, jyrau)@geomatics.ncku.edu.tw

<sup>2</sup> ITC, University of Twente, Hengelosestraat 99, 7514 AE Enschede - J.Jhan@utwente.nl

Commission II, WG II/1

**KEY WORDS:** Image Matching, Band Registration, Multispectral Camera, SURF

## ABSTRACT:

Due to the raw images of multi-lens multispectral (MS) camera has significant misregistration errors, performing image registration for band co-registration is necessary. Image matching is an essential step for image registration, which obtains conjugate features on the overlapped areas, and use them to estimate the coefficients of a transformation model for correcting the geometrical errors. However, due to the none-linear intensity of spectral response, performing feature-based image matching (such as SURF) can only obtain only a few conjugate features on cross-band MS images. Different to SURF that extracts local extremum in a multi-scale space and utilizes a threshold to determine a feature, we proposed a normalized SURF (N-SURF) that extracts features on single scale, calculates the cumulative distribution function (CDF) of features, and obtains consistent features from the CDF. In this study, two datasets acquired from Tetracam MiniMCA-12 and Micasense RedEdge Altum are used for evaluating the matching performance of N-SURF. Results show that N-SURF can extract approximately 2-3 times number of features, match more points, and have more efficient than original SURF. On the other hand, with the successful of MS image matching, we can therefor use the conjugates to compute the coefficients of a geometric transformation model. In this study, three transformation models are used to compare the difference on MS band co-registration, i.e. affine, projective, and extended projective. Results show that extended projective model is better than the others as it can compensate the difference of lens distortion and viewpoint, and has co-registration accuracy of 0.3 - 0.6 pixels.

## 1. INTRODUCTION

Multi-lens multispectral (MS) imaging system adopts highly synchronized multiple lenses, which can use different sensor and spectral filter combinations to obtain different image resolution, different wavelength, and different number of MS images. Some state-of-the-art multi-lens MS cameras and their specifications can be found respectively in Figure 1 and Table 1, which contains one twelve lenses (Tetracam MiniMCA-12) and one five lenses (Micasense RedEdge Altum) MS camera. Both cameras can acquire blue (BLU), green (GRE), red (RED), rededge (REG), and near infrared (NIR) spectral response. The advantages of small size and light weight make it suitable for mounting on a UAV (Unmanned Aerial Vehicle) to collect high resolution MS imagery. However, due to the slightly difference of perspective centers and viewing angles, the raw images of multi-lens MS camera have significant misregistration errors. Since multi-lens MS cameras are getting attraction on the applications of precision agriculture and environmental monitoring (Mulla, 2013; Torres-Sanchez et al., 2013), performing image matching and image registration (Mulla, 2013; Torres-Sanchez et al., 2013) to achieve the goal of band co-registration is important.

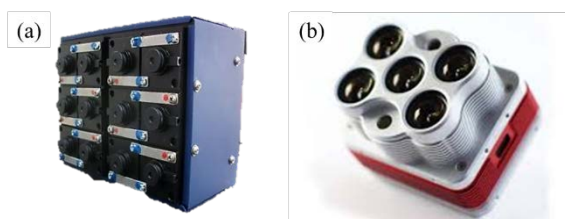


Figure 1. Multi-lens MS camera. (a) Tetracam MiniMCA-12, and (b) Micasense RedEdge Altum.

Manufacture	Tetracam	Micasense
Camera	MiniMCA-12	Rededge Altum
Number of Bands	12	5
Wavelength	450 – 950 nm	475 – 840 nm
Focal Length	9.8 mm	8 mm
Shutter Type	Rolling	Global
FOV (W° × H°)	38 × 30	48 × 37
Resolution	1280 × 1024	2064 × 1544
Dimensions (cm)	15.4 × 7.8 × 8.7	8.2 × 6.7 × 6.5
Weight	1.2 kg	0.4 kg

Table 1. Specifications of Multi-lens Camera in Figure 1.

Image matching is a crucial step to understand the space relationship of overlapped images, which matches conjugate features on the overlapped area, and uses them to reconstruct the scene orientation. The state-of-the-art feature-based matching method, such as scale and rotation invariant SIFT (Lowe, 2004) and SURF (Bay et al., 2008) features, have been widely applied to the applications of image registration. The goal of image registration is to using matched features for estimating the coefficients of a geometric transformation model. Therefore, the difference of scale, rotation, and translation between two images can be corrected and co-registered by image transformation.

For homogeneous image pairs, such as RGB vs. RGB, which have similar intensity distribution, it can have successful matching and obtain plenty number of conjugate points for sensor orientation and stitching (Rau et al., 2016). However, due to the none-linear intensity of spectral response, performing standard image registration on MS image is a difficult task, since it is hard to find enough conjugate features for coefficients calculation.

### 1.1 Motivation and Objective

In order to achieve MS band co-registration, researchers mainly adopt two steps approach to perform coarse-to-fine image registration (Jhan et al., 2018; Ye and Shan, 2014). Ye and Shan (2014) first used feature-based matching to minimize the misregistration errors by projective transformation, and later on another area-based matching is utilized to correct the remaining linear distortion effects. Jhan et al. (2018) directly obtain the coefficients of a modified projective transformation model by camera rig calibration, and another sequential image matching analysis is conducted in order to compensate the remaining systematic errors.

Since using image matching for misregistration errors correction and optimization is unavoidable, the objective of this study is to develop a robust MS image matching method in order to obtain sufficient features for directly calculate the coefficients of geometric transformation model.

### 1.2 Literature Review of MS Image Matching

Feature-based image matching generally has three steps that are feature extraction, feature description, and feature matching. Feature extraction is to extract local extremum in a multi-scale space and utilizes a threshold to determine if it is a robust interest point. In order to construct the descriptor for matching, each interest point is described by computing its intensity distribution of neighborhood pixels, and assigned a main direction by finding the maximum gradients in order to achieve rotation invariant. Therefore, a conjugate feature on overlap area can be matched by comparing if the Euclidian distance of descriptor in first and second interest points is small than a ratio.

Due to the low performance of MS image matching, Jhan et al. (2018) suggested using SURF to perform a sequential image matching on MS pairs that have closest spectral response. However, for low texture and low contrast images, sequential image matching remains fail as it is hard to find enough features for matching. On the other hand, Li et al. (2011), Saleem and Sablatnig (2014) and Ma et al. (2017) proposed different solution based on the modification of SIFT descriptor in order to increase the robust matching rate on MS images. However, we observe that the number of interest points and matched features varied with different image, without a point amount control will lead to an unstable matching result.

### 1.3 Research Methods and Experiments

Different to previous researches that are modified the descriptor to perform a robust matching, we notice that the failure of MS image matching is due to insufficient numbers of interest points, and the amounts are varied with image intensity and different spectral bands. In this study, based on SURF features, we proposed a normalized SURF (N-SURF) that can extract more interest points and has ability to control the number of points on each image. Therefore, we can use consistent number of features, like 5000 or 10000, to perform MS image matching, and acquire more number of successful matches. Meanwhile, by means of successful MS matching, the coefficients of a geometric transformation can be directly calculated and used to perform one-step image co-registration.

The following structure of this paper contains two part. First is to introduce the difference of N-SURF and original SURF, and utilize three evaluation indexes, i.e. Matching Rate (MS), Duplicate Rate (DR), and Correct Rate (CR), to evaluate their

performances on MS image matching. The used evaluation datasets are acquired from the multi-lens MS cameras in Figure 1. The second part is to perform MS image co-registration by using different transformation model, and analyze the differences by means of visual comparison and accuracy assessment.

## 2. RESEARCH MATERIAL

In order to perform MS image matching and image co-registration, two datasets as listed in Table 2, are respectively acquired from MiniMCA-12 and Rededge Altum. From the table, it shows that each dataset is collected from different UAV platforms, different flight heights, and different areas.



Camera	MiniMCA-12	RedEdge Altum
UAV	 Helicopter	 Multi-Rotor
Flight Height	100 m	50 m
Date (m, y)	9, 2018	3, 2019
Target Area	Mountain	Orchard
Group × Images	494 × 12	190 × 5

Table 2. Information of datasets.

One group of MS images from each dataset are demonstrated in Figure 2, while it also shows the color combinations of RED, GRE, and BLU band. From Figure 2 (c) and (d), we can observe the significant band misregistration errors of raw MS images, and there is no doubt that it requires band co-registration processing in order to achieve the spectral analysis.

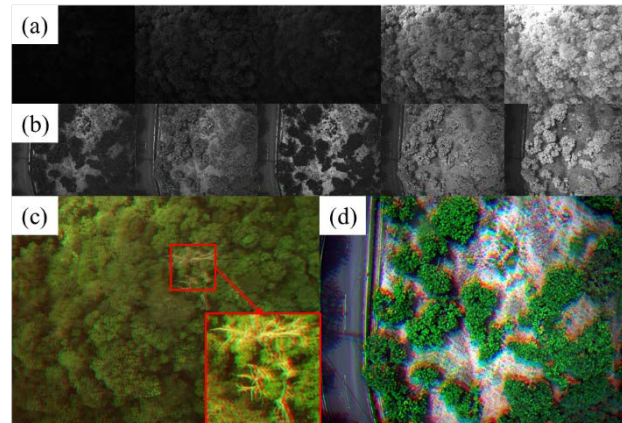


Figure 2. One group of MS images and its band misregistration effects. From left to right of the top two rows are respectively BLU, GRE, RED, REG, and NIR bands. (a) Only five images of MiniMCA-12 are demonstrated, (b) All five bands of RedEdge Altum, (c) and (d) are respectively the band misregistration errors of (a), and (b).

From Figure 2, it is obvious that for same ground object, the contrast and intensity of each image are varied with different spectral response. We also can observe that the intensity of BLU, GRE, and RED images of MiniMCA-12 is too low to see the details. On the other hand, the intensity histogram of each image from Figure 3 also shows the none-linear distribution among each image. Hence, we can expect the extracted number of features will vary with image intensity, and fail when the intensity is too low. To perform image matching for image registration is very tricky under this low brightness, contrast, and intensity environment.

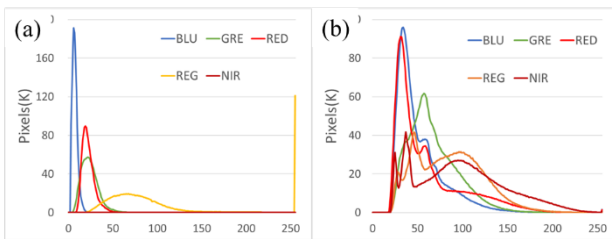


Figure 3. Intensity histogram of each image. (a) MiniMCA-12, and (b) RedEdge Altum.

### 3. METHODOLOGY

The workflow of this study is shown in Figure 4, which first to perform N-SURF matching on MS images and evaluate its performance. Then using the successful matched features for estimating the coefficients of three different geometric transformation models, and analysed the difference of MS band co-registration by accuracy assessment and visual comparison. Details are introduced in the following sections.

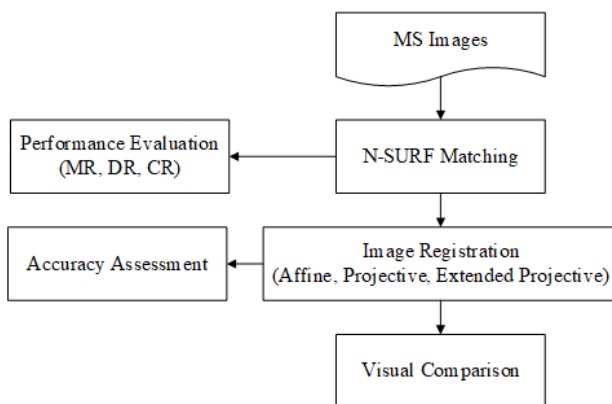


Figure 4. Workflow of N-SURF matching and MS band co-registration analysis.

#### 3.1 N-SURF

In order to increase the MS image matching rate, our concept is to extract as many features as possible and control its number for matching. Therefore, when image has different spectral, contrast, and brightness, we can obtain consistent number of features to perform image matching. Since the computation efficiency of SURF is 3-5 times than SIFT, we chose SURF and modify its feature extractor to implement our concept.

SURF utilizes integral image and Fast-Hessian (FH) box filter to increase the feature extraction efficiency. In order to achieve scale invariant, SURF utilizes different filter sizes of FH filter to construct a multi-octaves multi-scale space instead of reducing the resolution of image. For example, the first octave contains four different sizes of FH filter, i.e. FH(9), FH(15), FH(21), and FH(27). If the FH response of one point in a multi-scale space is a local extremum within  $3 \times 3 \times 3$  neighbouring points, it is treated as an interest point. By means of applying to a threshold, interest points who larger than the value are considered as robust features.

Although SURF can extract robust features via multi-scale approach, the potential maximum number of points is also reduced, and leads to there are not enough points when performing image matching. Meanwhile, the image contrast and brightness of MS image also influence the number of features. How to determine the threshold to extract the consistent number of points is a challenge.

**3.1.1 Single Layer Extraction:** In order to increase the number of features, N-SURF extracts local extremum in a single scale within  $3 \times 3$  neighbouring points. As shown in Figure 5, it shows the number of features in different scales and multi-scale SURF when the threshold value is 0, and two findings can also be observed. The different spectral of MS image leads to different number of features; Each single scale can extract more features than multi-scale SURF, while the higher scale has more features, likes FH(9) has approximate 3 times number of features than FH(27). Though the number of features are increased, we can also expect its robustness may decreased.

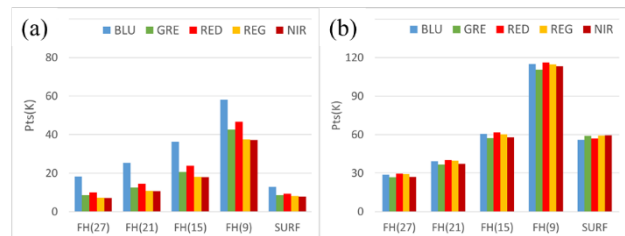


Figure 5. Number of extracted features in different scale. (a) and (b) are respectively the results of figure 2 (a) and (b).

**3.1.2 Cumulative Distribution Function (CDF):** Since a fixed threshold leads to different amounts of features, and it is also influenced by the image's intensity, we introduce the using of CDF to control the number of features. As depicted in Figure 6, which is the CDF of FH response of Figure 2 (a), it shows that the FH response among images are different to each other. Therefore, if we chose 1000 as a threshold, this will respectively acquire 0, 0, 1000, 9000, and 11000 number of features on BLU, GRE, RED, REG, and NIR image, which means performing image matching for BLU and GRE images is not possible. On contrary to using a fixed threshold, with the using of CDF, N-SURF can adaptively determine the threshold by the requirement of number of features. By querying the number of 10000 features, the threshold of BLU, GRE, RED, REG, and NIR is respectively determined at 10, 20, 100, 800, and 1000. The feature numbers will hence no longer be influenced by the intensity distribution of images.

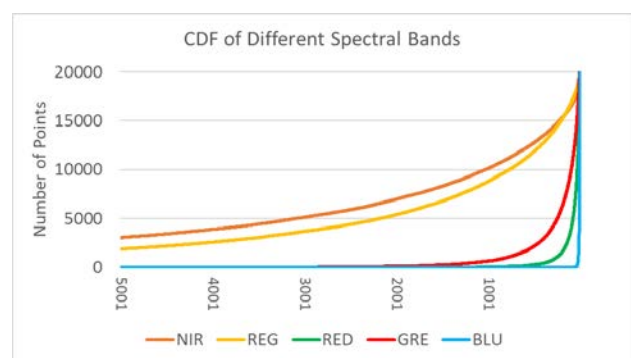


Figure 6. CDF of FH response of Figure 2 (a).

**3.1.3 Duplicates Removal:** Due to the robustness of N-SURF is decreased when applied on single scale feature extraction, single feature may have multiple correspondences. In order to avoid this situation, we searching all matches and remove all duplicate points, since there is no guarantee to which points is correct.

#### 3.2 Evaluation Indexes

In order to evaluate and compare the performance of N-SURF and original SURF, the formula of three evaluation indexes, i.e.

MR, DR, and CR, are shown in equations (1) - (3). MR represents how many points can be matched over total features on reference and target images, it is hence suitable for evaluate the matching efficiency. Since some features have multiple corresponds, DR calculating the ratio of duplicates in matches in order to evaluate the robustness. In order to evaluate the correctness, CR is to calculate the percentage of correct points in none duplicate matches after rigorous image registration is conducted

$$\text{MR (\%)} = 2 \times \text{Matches} / \text{Total Features} \quad (1)$$

$$\text{DR (\%)} = \text{Duplicates} / \text{Matches} \quad (2)$$

$$\text{CR (\%)} = \text{Correct points} / (\text{Matches} - \text{Duplicates}) \quad (3)$$

### 3.3 Image Registration

With the successful MS image matching, we can therefore estimate the coefficients of a geometric transformation model through least square adjustment, and directly using those coefficients to perform one-step image transform instead of coarse-to-fine image co-registration. In order to compare the difference of image transformation for multi-lens MS camera, three models are introduced here, i.e. affine, projective, and extended projective transformation.

**3.3.1 Affine Transform:** Affine transform describes the relationship of two parallel image that have two scales, two translations, one rotation, and one skew effects. As shown in Equation (4) and (5), this results in 6 unknown parameters ( $A - F$ ), in which  $(x, y)$  and  $(u, v)$  are respectively the coordinates of original image and transformed image.

$$u = A \times x + B \times y + C \quad (4)$$

$$v = D \times x + E \times y + F \quad (5)$$

**3.3.2 Projective Transform:** Unlike affine transform only describes the relationship of two parallel images, projective transform can describe the transformation among two unparallel image planes, which has three rotations, three translations, and two scales effects. Equation (6) and (7) shows that projective transform utilizes 8 unknown parameters ( $A_1 - A_3, B_1 - B_3,$  and  $C_1 - C_2$ ) to perform image transformation.

$$u = \frac{A_1 \times x + A_2 \times y + A_3}{C_1 \times x + C_2 \times y + 1} \quad (6)$$

$$v = \frac{B_1 \times x + B_2 \times y + B_3}{C_1 \times x + C_2 \times y + 1} \quad (7)$$

**3.3.3 Extended Projective Transform:** Except for the viewpoints difference of multi-lens camera, the difference of lens distortion should also be considered when performing image transformation. In order to correct the differences of lens distortion and viewpoints at the same time, we add additional lens distortion correction formulas in the projective transformation model. Therefore, an extended projective transformation model is formed by respectively adding equation (8) and (9) in equation (6) and (7). It utilizes additional five parameters to describe the radial distortion ( $K_1 - K_3$ ) and decentring distortion ( $P_1 - P_2$ ) effects, in which  $r$  represents the distance to the perspective center of image.

$$\Delta_x = x \times (K_1 \times r^2 + K_2 \times r^4 + K_3 \times r^6) + P_1 \times (r^2 + 2x^2) + 2P_2xy \quad (8)$$

$$\Delta_x = y \times (K_1 \times r^2 + K_2 \times r^4 + K_3 \times r^6) + P_2 \times (r^2 + 2y^2) + 2P_1xy \quad (9)$$

### 3.4 Accuracy Assessment and Visual Comparison

The image co-registration accuracy is represented by the RMSE value of least-square adjustment. During the computation, features have residuals larger than two times of RMSE are eliminated, and repeatedly perform least-square adjustment until there is no features can be removed. To compare the difference of each transformation model, we adopt visual comparison to exam the co-registration results.

## 4. RESULTS AND ANALYSIS

In this chapter, we first introduce the performance of N-SURF, and then analyze the image co-registration analysis through visual comparison and accuracy assessment.

### 4.1 Performance of N-SURF

As summarized in Table 3, we perform image matching on the five bands of MiniMCA-12 shown in Figure 2 (a), and compare the differences between original SURF and N-SURF. There are in total of ten different combinations of matching pair, and have four different matching methods, including two original SURF and two N-SURF matching. The thresholds of two different original SURF is 10 and 0, respectively represented as SURF(10) and SURF(0); Both N-SURF using FH(15) to extract features and fix the feature numbers to 10000 and maximum, which is respectively represented as N-SURF(10000) and N-SURF(MAX). The feature numbers of reference and target images, MR, DR, CR, correct points, and RMSE of co-registration accuracy can all be found in Table 3. Please notice that the transformation model used here is projective transform.

From the table, it shows that SURF(10) failed on most of the matching pairs, since it cannot obtain features on the BLU image, and only a few features can be extracted on GRE and RED images. Though the results of SURF(0) is better, it still failed on the matching pair of NIR vs. BLU. On the other hand, N-SURF has all successful matching on different matching pairs, and not only shows that the MR is higher than SURF(0), but have 2-3 times number of features more than SURF(0).

In order to better illustrate the content of table, we chose four matching pairs to demonstrate the trend of evaluation indexes, i.e. REG vs. NIR, REG vs. GRE, GRE vs. RED, and GRE vs. BLU. As shown in Figure 7, which depicts the trend of MR, DR, CR, and correct number of points. From the figure, we first notice the MR of N-SURF outperforms than SURF that is two times better, and with point numbers control, N-SURF(10000) can acquire better result than N-SURF(MAX). On the other hand, as expected, the robustness of N-SURF is decreased with single scale feature extraction, Figure 7 (b) and (c) both shows that DR and CR of N-SURF are worse than SURF(0). However, the loss of robustness is acceptable, since the most important thing is N-SURF have 3 times number of matches than SURF, which makes MS image registration possible.

In conclusion, N-SURF has better performance than SURF, can find more features and obtained more correct points. Except for the lowest image quality of BLU band, which has RMSE of image co-registration larger than one pixels, the other matching

pairs can achieve sub-pixels accuracy, meaning that N-SURF has ability to perform successful matching on images of non-linear intensity distribution.

Matching Method (Threshold)	Target Image (Feature Points)															
Target Image (Feature Points)	MR (%)				DR (%)				RR (%)				Correct Points RMSE(Pixels)			
SURF (10)	GRE (264)				RED (57)				REG (4544)				NIR (5830)			
BLU (0)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
GRE (264)	#				*	*	*	*	*	*	*	*	*	*	*	*
RED (57)	#				*	*	*	*	*	*	*	*	*	*	*	*
REG (4544)	#				*	*	*	*	*	*	*	*	32	3	86	$\frac{1378}{0.40}$
SURF (0)	GRE (8743)				RED (9427)				REG (8105)				NIR (7747)			
BLU (12868)	3	3	7	$\frac{239}{1.50}$	5	10	66	$\frac{319}{1.31}$	2	6	68	$\frac{131}{1.96}$	*	*	*	*
GRE (8743)	#				16	9	72	$\frac{977}{0.58}$	24	4	75	$\frac{1483}{0.41}$	11	4	76	$\frac{685}{0.67}$
RED (9427)	#				8	4	78	$\frac{532}{0.76}$	4	9	67	$\frac{211}{1.33}$				
REG (8105)	#								31	4	78	$\frac{1867}{0.39}$				
N-SURF (10000)	GRE (10022)				RED (10135)				REG (9989)				NIR (9994)			
BLU (10008)	14	8	76	$\frac{986}{1.03}$	20	11	72	$\frac{1291}{0.92}$	9	7	77	$\frac{619}{1.08}$	4	12	71	$\frac{271}{1.5}$
GRE (10022)	#				50	13	70	$\frac{3060}{0.53}$	56	5	71	$\frac{3770}{0.42}$	31	6	74	$\frac{2183}{0.56}$
RED (10135)	#								29	7	71	$\frac{1916}{0.56}$	13	8	74	$\frac{916}{0.82}$
REG (9989)	#								64	5	72	$\frac{4346}{0.40}$				
N-SURF (MAX)	GRE (20109)				RED (23811)				REG (18205)				NIR (17970)			
BLU (36276)	7	10	72	$\frac{1333}{1.04}$	10	16	66	$\frac{1639}{0.96}$	4	5	73	$\frac{702}{1.17}$	2	10	73	$\frac{348}{1.51}$
GRE (20109)	#				42	25	58	$\frac{4012}{0.56}$	48	10	61	$\frac{5031}{0.44}$	27	11	64	$\frac{2892}{0.58}$
RED (23811)	#								20	9	61	$\frac{2389}{0.59}$	9	8	69	$\frac{1203}{0.85}$
REG (18205)	#								59	11	61	$\frac{5747}{0.42}$				

Table 3. Statistics of original SURF and N-SURF.

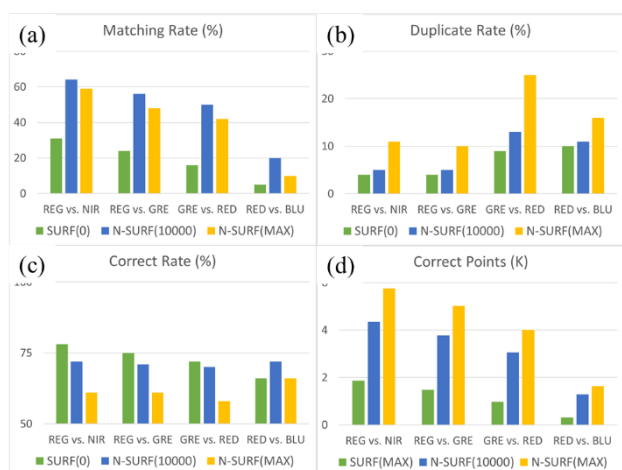


Figure 7. The trend of evaluation indexes of SURF and N-SURF. (a) MR, (b) DR, (c) CR, and (d) Number of correct points.

## 4.2 Visual Comparison of Image Co-registration

Before compare the results of image co-registration, the lens distortion curve of each band of multi-lens MS camera is demonstrated in Figure 8. Since each lens has same specification, the lens distortion curve is very similar to each other. But, we found that when moving far away from the center of image, the difference in lens distortion effect is also increased. Therefore, it will influence the band co-registration results of multi-lens MS camera, if the difference of lens distortion is not considered.

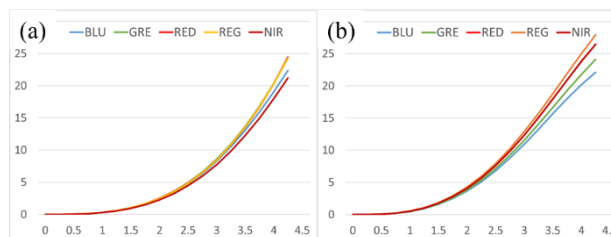


Figure 8. Lens distortion curve of multi-lens MS camera. (a) MiniMCA-12, (b) RedEdge Altum.

Since affine transform can only describe two parallel image plans, while it was not parallel for multi-lens camera, the misregistration errors after affine transformation in figure 9 (d) still exist and has poorest accuracy. Though projective transform is suitable for non-parallel images and has better result than affine transform, without considering the lens distortion differences of multi-lens, we still can observe the misregistration at the edge of figure 9 (e). On the other hand, since projective and lens distortion difference are considered, the extended projective transform shows the best result. There is no misregistration error in figure (f), and has image co-registration accuracy in 0.3-0.4 pixels.

Meanwhile, it also proves that there is no need to perform a camera pre-calibration for image co-registration, using extended projective transform can compensate the difference of viewpoints and lens distortion effect. However, please notice that the extended projective transform is to compensate the difference between two different lens distortion curves, the lens distortion effects still exists.

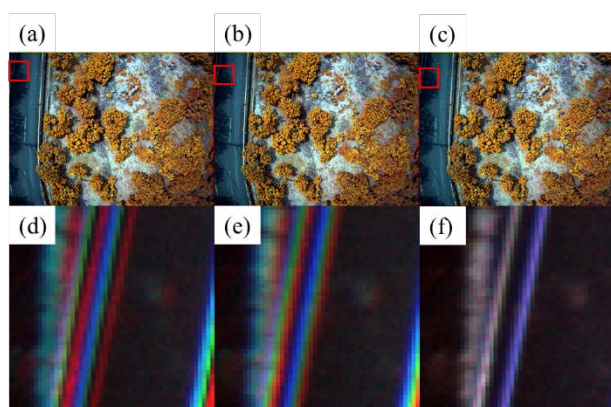


Figure 9. Image co-registration comparisons by using different transformation models. (a), (b), and (c) are respectively the results of affine, projective, and projective transform. In which, (d), (e), and (f) are the zoomed images of (a), (b), and (c).

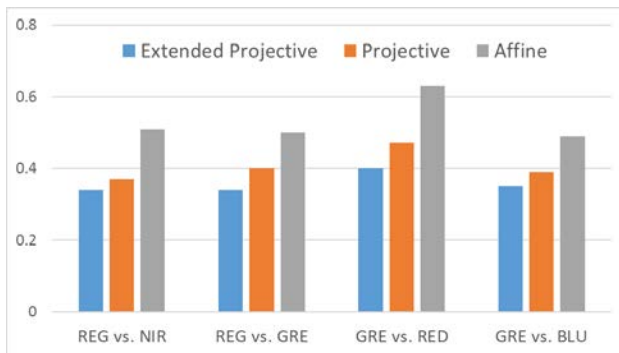


Figure 10. Image co-registration accuracy of different geometric models.

### 4.3 Accuracy Assessment of Image Co-registration

With the successful of MS image matching and the best co-registration results of using extended projective transformation, we can perform image co-registration on each group of images in the collected dataset. The co-registration accuracy of each group of images in table2 are respectively summarized in figure 11 and figure 12.

Except for the low contrast of BLU image in MiniMCA-12 leads to larger error, all results in MiniMCA-12 and RedEdge Altum show that the co-registration accuracy is in the range of 0.3-0.6 pixels, meaning that N-SURF matching and using extended projective transformation model is suitable for MS image matching and image co-registration.

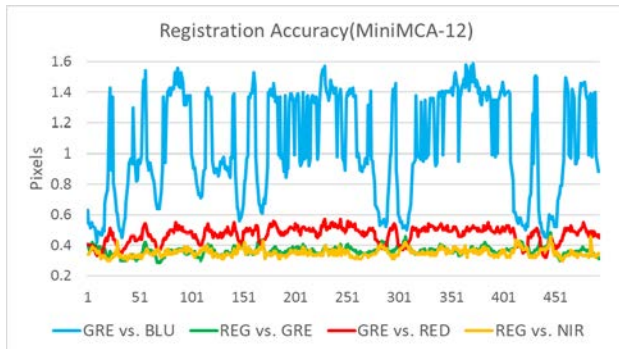


Figure 11. Image co-registration accuracy of all groups in MiniMCA-12.

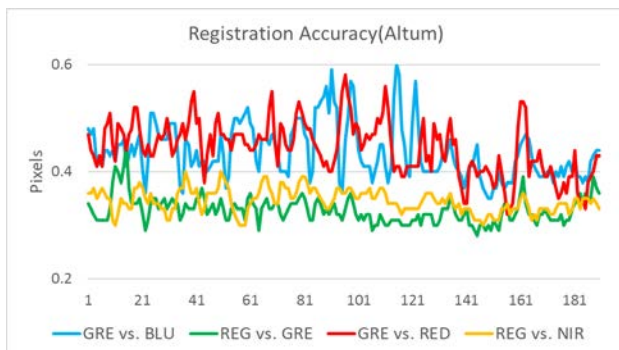


Figure 12. Image co-registration accuracy of all groups in RedEdge Altum.

## 5. CONCLUSIONS

In this study, we have developed a normalized SURF that has ability to perform a successful matching on MS images. It can

extract maximum number of features in a single scale space, and automatically determine the threshold via CDF in order to acquire a fixed number of features. By means of performance evaluation, we noticed that N-SURF can extract 2-3 times more features than original SURF, has more efficiency, and more ability to match correct point on different MS matching pairs. Though the robustness of N-SURF was decreased, the successful and sufficient matches makes image co-registration possible.

On the other hand, through the comparison of three different transformation models, it shows that the proposed extended projective transformation has the best results that can compensate the difference of viewpoints and lens distortion among two lenses at the same time. Meanwhile, we obtained a consistent co-registration accuracy in 0.3-0.6 pixels among all group of images. Therefore, the proposed N-SURF is robust and feasible to extract features and perform image matching on MS images.

## ACKNOWLEDGEMENTS

The authors are respectively grateful to Prof. Cho-Ying Huang of Department of Geography, National Taiwan University, and Rdata system CO., LTD., for providing a MiniMCA-12, and a RedEdge Altum. The authors are also grateful to Rdata system CO., LTD., for providing the UAV and assisting the sensor integration and data collection.

## REFERENCES

- Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110, 346-359.
- Jhan, J.-P., Rau, J.-Y., Haala, N., 2018. Robust and adaptive band-to-band image transform of UAS miniature multi-lens multispectral camera. *ISPRS Journal of Photogrammetry and Remote Sensing* 137, 47-60.
- Li, Q., Zhang, H., Wang, T., 2011. Multispectral Image Matching Using Rotation-Invariant Distance. *IEEE Geoscience and Remote Sensing Letters* 8, 406-410.
- Lowe, D., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60, 91-110.
- Ma, W., Wen, Z., Wu, Y., Jiao, L., Gong, M., Zheng, Y., Liu, L., 2017. Remote Sensing Image Registration With Modified SIFT and Enhanced Feature Matching. *IEEE Geoscience and Remote Sensing Letters* 14, 3-7.
- Mulla, D.J., 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* 114, 358-371.
- Rau, J.Y., Jhan, J.P., Li, Y.T., 2016. Development of a large-format UAS imaging system with the construction of a one sensor geometry from a multicamera array. *IEEE Transactions on Geoscience and Remote Sensing PP*, 1-10.
- Saleem, S., Sablatnig, R., 2014. A Robust SIFT Descriptor for Multispectral Images. *IEEE Signal Processing Letters* 21, 400-403.
- Torres-Sanchez, J., Lopez-Granados, F., De Castro, A.I., Pena-Barragan, J.M., 2013. Configuration and specifications of an

Unmanned Aerial Vehicle (UAV) for early site specific weed management. PloS one 8, e58210.

Ye, Y., Shan, J., 2014. A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. ISPRS Journal of Photogrammetry and Remote Sensing 90, 83-95.