AUTOMATED CO-REGISTRATION OF INTRA-EPOCH AND INTER-EPOCH SERIES OF MULTISPECTRAL UAV IMAGES FOR CROP MONITORING

P. O. Mc'Okeyo[*], F. Nex , C. Persello, A. Vrieling

Dept. of Earth Observation, Faculty of Geoinformation Science and Earth Observation - ITC, University of Twente, The Netherlands
- mcokeyo.ochieng@gmail.com, f.nex@utwente.nl, c.persello@utwente.nl, a.vrieling@utwente.nl

KEY WORDS: Unmanned Aerial Vehicles, Multispectral, Co-registration, Intra-epoch, Inter-epoch, Image Matching, Orthophoto

ABSTRACT:

The application of UAV-based aerial imagery has advanced exponentially in the past two decades. This can be attributed to UAV operational flexibility, ultra-high spatial resolution, inexpensiveness, and UAV-based sensors enhancement. Nonetheless, the application of multitemporal series of multispectral UAV imagery still suffers significant misregistration errors, and therefore becoming a concern for applications such as precision agriculture. Direct image georeferencing and co-registration is commonly done using ground control points; this is usually costly and time consuming. This research proposes a novel approach for automatic co-registration of multitemporal UAV imagery using intensity-based keypoints. The Speeded Up Robust Features (SURF), Binary Robust Invariant Scalable Keypoints (BRISK), Maximally Stable Extremal Regions (MSER) and KAZE algorithms, were tested and parameters optimized. Image matching performance of these algorithms informed the decision to pursue further experiments with only SURF and KAZE. Optimally parametrized SURF and KAZE algorithms obtained co-registration accuracies of 0.1 and 0.3 pixels for intra-epoch and inter-epoch images respectively. To obtain better intra-epoch co-registration accuracy, collective band processing is advised whereas one-to-one matching strategy is recommended for inter-epoch co-registration. The results were tested using a maize crop monitoring case and the; comparison of spectral response of vegetation between the UAV sensors, Parrot Sequoia and Micro MCA was performed. Due to the missing incidence sensor, spectral and radiometric calibration of Micro MCA imagery is observed to be key in achieving optimal response. Also, the cameras have different specifications and thus differ in the quality of their respective photogrammetric outputs.

1. INTRODUCTION

Recently, the application of drone technology in crop monitoring has become rife. Nex & Remondino (2014) review the use of unmanned aerial vehicles for 3D mapping applications, and highlights agriculture as a domain that consumes digital surface models (DSM) and orthoimages to extract useful information on crop status. In addition, the ultra-high multispectral and multitemporal resolution of UAV imagery is undoubtedly an enabler of Precision Agriculture to obtain actionable crop properties (Elarab et al. 2015).

UAVs are embraced across domains because they are flexible low-altitude Remote Sensing (RS) platforms. Thus, they are not affected by cloud occlusion, and can achieve ground sampling distances (GSD) of up to 3cm or less depending on the flight parameters and the aim of the acquisition (Nex & Remondino, 2014). This is still ten times higher the spatial resolution of the best VHR satellite imagery. In addition, UAVs provide an inexpensive alternative to satellites and other platforms for aerial image acquisition; they increasingly offer tools and inspire innovations that seal the gap between terrestrial and aerial (high-altitude) platforms (Nex et al. 2015).

Conversely, UAVs face some drawbacks: regulatory constraints on the application of drones and licensing of drone pilots vary from country to country; limited areal coverage due to the battery endurance per flight; the instability of lightweight platforms; atmospheric elements such as strong winds and rain affect drone operations; the payload limit; image co-registration complexities, and difficulties in radiometric and geometric corrections (Freeman et al. 2015;Yang et al. 2017).

Accurate image co-registration is vital for reliable change detection assessment and accurate comparative analysis of crop phenology (Fytsilis et al., 2016; Tilly et al., 2014). Several models and algorithms that automate the co-registration process have been proposed. However, multispectral cameras with several lenses still suffer misregistration setbacks as demonstrated in related works of Jhan et al. (2017) and Rey et al. (2013). This is partly due to the fact that the technology space is dynamic and new camera sensors with different specifications and more abilities are continuously being engineered.

On the other hand, co-registration of multitemporal series is vital for reliable spatiotemporal analysis of crop's spectral properties. Misclassification of crop growth per pixel, vegetation index extraction errors, interpolation errors in values between available observations, and harvest index variation prediction are some of the inherent errors due to misregistration (Lobell, 2013).

The aim of this study was to provide a novel approach for accurate UAV-based multispectral and multitemporal monitoring of crops without repeated establishment of Ground Control Points (GCPs) which is laborious during photogrammetric processing. Usually, GCPs are meant to minimize systematic errors and deformations in images, stabilize bundle solutions, and determine correct 3D reconstruction (Nex & Remondino, 2014). However, the lack of GCPs did not hamper this study since the acquisition of the first epoch was assumed to be the reference epoch, and registration assessments of subsequent acquisitions were based on the first epoch.

2. RELATED WORKS

Misregistration of UAV multispectral and multitemporal imagery has attracted the need of researchers to propose and develop novel methodologies to resolve this problem. Kelcey & Lucieer (2012) suggest procedures to calibrate the six bands Mini MCA camera including radiometric correction, noise reduction and affine transformation for simultaneous image registration and correction of lens distortion. Turner et al. (2014) develop a semi-automated workflow for accurate spatial co-registration of a visible camera, six – band Micro MCA multispectral sensor, and a thermal infrared camera. Using the Scale Invariant Feature Transform (SIFT), a mean accuracy of 1.78pixels is achieved. This was deemed sufficient for monitoring of Antarctic moss beds.

Jhan et al. (2016) present a modified projective transformation model based on the principles of plane-to-plane projection to undertake accurate band-to-band registration (BBR) of RGB and Mini MCA 12 multispectral imagery. It is noted that feature matching for narrow band multispectral and hyperspectral sensor with no overlapping spectral range is difficult. An accuracy of 0.33 pixels is achieved for the proposed BBR method. However, co-registration errors of less 0.6 pixels was obtained between Mini MCA reference band and the RGB ortho-images.

A novel approach to automate the co-registration of UAV-based multi-temporal RGB image blocks without the use of GCPs is presented by Aicardi et al., (2016). The first acquisition is chosen as the reference dataset. The orientation parameters of the anchor images are fixed; this constrains the bundle block adjustment of the slave images to be aligned with the reference image. An array of tests to assess both manual and automatic registration approaches for the selected anchor images provides reliable results, which are quite comparable to a GCP-based strategy.

Onyango et. al. (2017) use keypoint descriptors to accurately estimate orientation parameters of UAV images through co-registration of oblique imagery. Using AKAZE, brute force is implemented to find putative correspondences and Lowe's ratio test used to discard wrong matches. Multiple homographies are computed using the putative correspondences to filter out remaining mismatches.

Recently, Banerjee, Raval, & Cullen (2018) optimize feature descriptors techniques to align UAV-hyperspectral images in a spectrally complex environment. It is observed that for band-to-band alignment, keypoint descriptors are inclined to spectral order vis a vis temporal order. In addition to spatial invariance, spectrally invariant descriptors will go a long way in improving the efficacy of the band-to-band alignment process.

Albeit novel and significant application-wise, the related works fail to offer solutions for band-to-band co-registration of multispectral UAV-imagery. This research therefore proposes an accurate sub-pixel co-registration approach for both intra-epoch and inter-epoch acquisitions using intensity-features-based algorithms invariant to scale, rotation, illumination and viewpoints such as SURF (Bay, Tuytelaars, & Van Gool, 2006), KAZE (Alcantarilla, Bartoli, & Davison, 2012), BRISK (Leutenegger, Chli, & Siegwart, 2011), and MSER (Matas, Chum, Urban, & Pajdla, 2004).

3. EQUIPMENT AND DATA

This research seeks to accurately co-register and assess the data quality of the images acquired by the Micro-MCA Tetracam camera mounted on the Matrice 600 UAV, and the Parrot Sequoia camera mounted on the Phantom 4 UAV.

3.1 The Parrot Sequoia and Micro MCA 6 Tetracam

The Parrot Sequoia (PS) multispectral sensor captures the electromagnetic spectrum in four separate parts: green, red, red-edge and Near Infrared (NIR). It incorporates the Global Positioning System (GPS), Inertial Measurement Unit (IMU) and magnetometer thus increased accuracy of data capture. The Parrot Sequoia also integrates an irradiance sensor to continuously record light conditions. Figure 1 shows the Parrot Sequoia camera system.

Fig 1. Parrot Sequoia

The micro MCA multispectral camera has six separate cameras. Each camera is synchronized with the other cameras so that each can capture the same scene at the exact same time of exposure. During each exposure instant, six separate channels of visible or NIR radiation move through each lens and filter to form separate monochromatic images on each sensor.

Figure 2 shows the Micro MCA Tetracam and an illustration of the architecture of each spectral sensor.
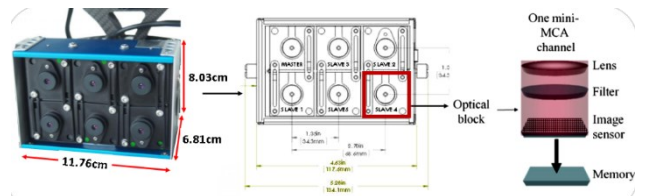

Fig 2. Micro MCA Tetracam

The spectral wavelength range of each UAV is shown in table 1.

Band	Parrot Sequoia	Micro MCA
Blue	-	410 - 490nm
Green	530 - 570nm	510 - 590nm
Red	640 - 680nm	630 - 710nm
Red	-	660 - 740nm
Red Edge	730 - 740nm	730 - 740nm
Near Infrared	770 - 810nm	760 - 800nm

Table 1. Wavelength specifications of the UAV sensors

Further comparison of Parrot Sequoia and Micro MCA is presented in table 2.

Specifications	Parrot Sequoia		Micro MCA
	RGB	Multispectral	Multispectral
Lenses	1	4	6
Focal Length	4.88mm	3.98mm	9.6mm
Spectral Range	400 – 700nm	530 – 810nm	450 – 800nm
Pixel Size	1.34μm	3.75μm	5.2 μm
Resolution (Pixels)	4608 x 3456	1280 x 960	1280 x 1024
FOV (H° ×V°) (m)	64.6 × 50.8	62.2 × 48.7	84 x 67
Camera Weight	135g		530g
Shutter type	Global		Rolling
Camera Size (cm)	6 × 4 × 3		12 x 8 x 7

Table 2. Specifications of the Parrot Sequoia and Micro MCA

The Parrot Sequoia was mounted on DJI Phantom 3 Pro, and the Micro MCA on the DJI Matrice 600 as shown in Figure 3 below. The two sensors were mounted on different UAV platforms due to their physical properties of size and weight, and the complexity of sensor and UAV integration in the case of the Micro MCA.

Figure 3. Showing sensors mounted on respective UAVs

3.2 Image Acquisition and Data properties

The maize field (approximately 25 acres) is located in the periphery of Gronau city, Germany (52° 10'N, 6° 55'E). Image acquisition was done in three time-steps for the Parrot Sequoia and only one acquisition for Micro MCA as shown in table 3.

Date	Camera	Flying Height	Forward Overlap	Side Overlap	GSD
08-08-2017	Sequoia	50m	80%	40%	5.01cm
		70m	80%	40%	6.84cm
11-08-2017	Sequoia	70m	80%	40%	6.69cm
19-09-2017	Sequoia	50m	80%	40%	5.02cm
	Micro MCA	100m	80%	40%	4.64cm

Table 3. Image acquisition details and image properties

4. METHODOLOGY

This section gives a detailed description of the approaches taken and methods used to realize the main aim of co-registration of multitemporal series of multispectral UAV imagery. The general overview of methods, processes, decisions, intermediate and final outputs are captured in the flowchart in Figure 4.

4.1 Photogrammetric Workflow

Pix4D software was used for most of the photogrammetric workflow involving the Parrot Sequoia images. Aligning the Micro MCA images was done using Tetracam's Pixel Wrench II (PW2), based on a calibration file that contains the relative orientation between the master and slave bands since the Micro MCA images are captured in a RAW file format. The RAW format images were converted to multipage TIFs using PW2 and subsequently processed in Agisoft Photoscan due to its aggressiveness to correct the rolling shutter effect.

Initial photogrammetric processing includes keypoints detection and matching of single images; estimation of interior and exterior orientation, aerial triangulation, bundle block adjustment, tie point generation, as well as georeferencing. Dense points sufficient enough to estimate planes and geometry of the image scene are then generated. The point clouds are used as an input to generate a Digital Surface Model (DSM), which is thereafter also used as an input to generate orthophoto bands for the whole scene. Each band therefore had its independent orthophoto.
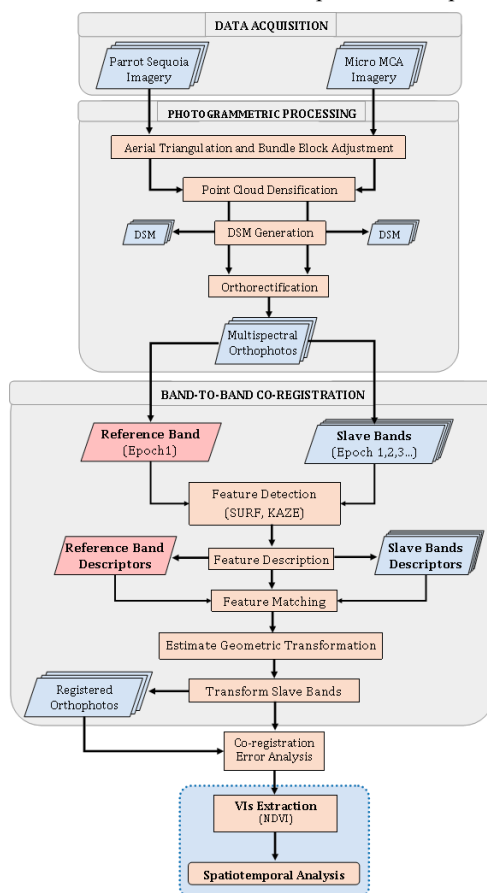

Figure 4. Flowchart showing an overview of the methodology

4.2 Orthophoto Image Co-registration

4.2.1 Feature Detection: The identification of corresponding keypoints between successive overlapping images is an integral part of image coregistration. Desirable keypoints must be devoid of noise, blurs, illumination variances and geometric differences. Experimentation revealed that for SURF, the higher the strongest feature threshold, *MetricThreshold*, the lesser the blobs, and the higher the octaves the larger the detected blobs.

For BRISK, the minimum contrast, *MinContrast'*, specifies the minimum intensity difference between a region and its immediate surrounding. It is a scalar in the range of zero (0) and one (1). An increase in this value would lead to a decrease in the number of blobs detected. Similarly, the minimum quality, *'MinQuality'*, ranges between zero (0) and one (1); it denotes the minimum accepted quality of detected regions. When the value tends towards one (1) erroneous blobs are removed.

In MSER, the size of the region is a two-element vector denoting the minimum and maximum areas of regions in pixels to be allowed in the detection process. At varying intensity thresholds, the maximum area variation between extremal regions is specified by a positive scalar between 0.1 to 1. An increase in this value results in detection of more external regions which may be less stable.

Finally, for KAZE, the local extrema is a function of the Hessian threshold, which is specified as a scalar greater than or equal to zero (0). An increase in this value excludes less significant local extrema. The multiscale detection factor and the scale levels are scalars in the range of 3 to 10. Larger features are detected by increasing the multiscale detection factor whereas smoother scale changes and additional intermediate scales between octaves are realized by increasing the scale levels.

4.2.2 Feature Description: Feature description is a function of the neighbouring pixels; it is done by extracting the intensity gradients of the neighbouring pixels and stored as a vector of numbers describing the center pixel. The vector size of the neighbourhood can vary between 64 and 128 pixels, and the descriptor can be considered rotation invariant if orientation of the feature vector is computed.

4.2.3 Feature Matching: To select strong matches, a matching threshold is specified according to different metrics (e.g. L1 or L2 norm); it represents a percent of the distance from a perfect match. Two feature vectors are a match when the distance between them is less than the set threshold. The higher the matching threshold, the more matches obtained (not necessarily 'good' matches).

4.2.4 Outlier Removal and Transformation matrix
In this study, the outliers in matched points were excluded using the M-estimator Sample Consensus (MSAC), which is a variant of the RANSAC (Torr & Zisserman, 2000). RANSAC suffers a setback; it is sensitive to the threshold that defines inliers and outlier. A very large threshold tends to rank all the hypotheses equally and qualify them as good for the fitted model. Conversely, a very small threshold tends to be unstable in estimating parameters. The MSAC (see equation 1) partially compensates for this undesirable effect. It penalizes the outliers equally but scores the inliers on how well they fit the data.

$$\rho_2(e^2) = \begin{cases} e^2 & e^2 < T^2 \\ T^2 & e^2 \geq T^2 \end{cases} \qquad (1)$$

Where ρ_2 is the robust error term, and T is the threshold for considering inliers.

A set of putative matches are taken in, and a random selection picks the best set of matches to fit the model, and computes the transformation matrix between the inlying points.
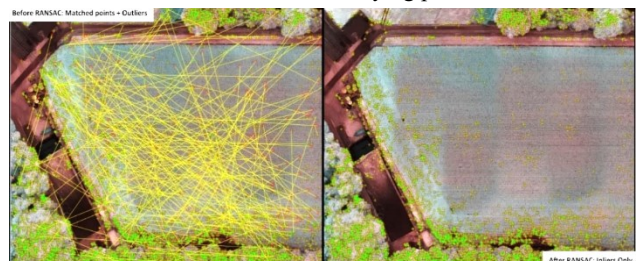

Figure 5. Showing putative matched plus outliers (left), and only the correct matches and inliers (right)

The outlier matches are defined by a distance threshold between features in band 'A' and band 'B' upon inverting the geometric transformation. Only points that meet this threshold were used to compute the transformation matrix. Estimation of the transformation matrix was done at the image and orthophoto level. At both levels, the transformation matrix was compared element by element, and quantified as a RMSE for comparison between different camera positions. The 2D similarity transformation method was used in this study because it retains angles and length ratios, and because the orthophotos are planimetric and geometrically similar. The Transformation matrix was further decomposed to fetch out the band-to-band rotation and translation (i.e. relative orientation).

4.3 Band-to-band Co-registration

4.3.1 Intra-Epoch registration: To examine misregistration within a single acquisition, it was necessary to establish the best band to use as the reference. The reference band should have sufficient keypoints to be matched with features extracted from other bands.

4.3.2 Inter-Epoch registration: Two approaches were evaluated as shown in Figure 6. The 'many-to-one' registration involved estimating the transformation matrix between all the subsequent bands of each acquisition with the red edge band of the first acquisition; geometrically transforming them; assessing their pairwise registration accuracy; and stacking them together per epoch. On the other hand, the 'one-to-one' approach involved using all bands in the first acquisition as the reference. Bands from subsequent acquisitions were considered slaves, and were thus aligned to spectrally corresponding bands of the first acquisition.
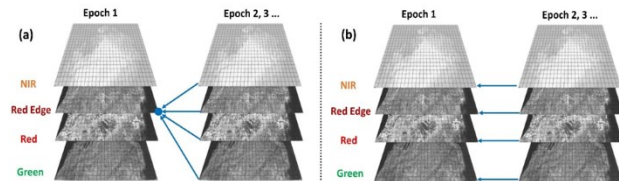

Figure 6 (a) Many-to-one registration (b) One-to-one registration

4.4 Co-registration Accuracy Assessment

The misregistration error between the spectral bands was measured by computing the projection distance between inlying point pairs. The horizontal positional positions of inliers before and after registration were used to assess the co-registration accuracy as illustrated in Figure 7. The positional accuracy and distance between detected features between bands is expected to be less than half a pixel after co-registration. For a perfect registration, the differential distance between conjugate pairs should be zero. Thus, values tending towards zero are desirable.
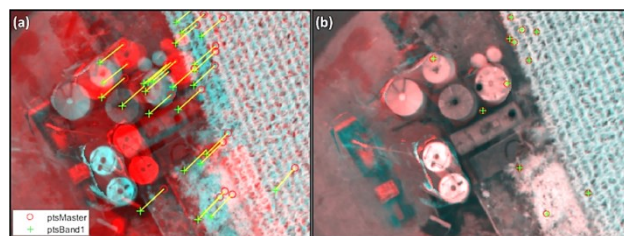

Figure 7 Showing distance between matched features (a) before and (b) after co-registration

The root mean square error (RMSE) of the horizontal displacement of the conjugate points was used to evaluate the registration accuracy. The RMSE was computed by finding root of the average of the set of squared differences between coordinate values of inlying slave and master keypoints, where the master refers to the first epoch. The positional RMSE was computed as shown in equations 2 – 4.

$$RMSE_x = \sqrt{\sum_i \frac{(Xslave - Xmaster)^2}{n}} \qquad (2)$$

$$RMSE_y = \sqrt{\sum_i \frac{(Yslave - Ymaster)^2}{n}} \qquad (3)$$

$$RMSE_r = \sqrt{RMSE_x^2 + RMSE_y^2} \qquad (4)$$

The $RMSE_x$ and $RMSE_y$ were used to evaluate systematic displacements in either direction. The combined $RMSE_r$ was used for overall registration accuracy assessment. The closer the value is to zero, the more accurate it is. The registration threshold was 0.5 of a pixel, therefore RMSEs less than 0.5 were considered 'good' and acceptable.

4.5 Vegetation Index and Spectral Analysis

Spectral indices are designed to give an approximate measure of vegetation status. The Normalized Difference Vegetation Index (NDVI) was used to characterize crop health in this research. NDVI is computed as shown below:-

$$NDVI = \frac{(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red})} \qquad (5)$$

To statistically assess and compare spectral variability between the two UAV cameras, intra-farm zonation was done. Spectral signatures of two classes of crops (photosynthetically active and less active) were extracted from corresponding composite images of Parrot Sequoia and Micro MCA Tetracam. This was aimed at explaining the spectral variability between the sensors.

5. EXPERIMENTATION, RESULTS AND ANALYSIS

In this section, a series of tests are run to inform coregistration decisions including selection of the master band, and the optimal parameters to use. All the algorithms are tested using default parameters, and further optimized for this particular vegetation scene and dataset combination. In addition, the experiments aim to compare the performance of the algorithms since they are architecturally different; SURF and KAZE use float descriptors, while MSER and BRISK use binary descriptors.

5.1 Master band Selection

5.1.1 Feature Detection using default parameters
The results indicated that KAZE outperformed SURF, BRISK and MSER by detecting three times the number of points detected. Also, the red edge was selected as the master band because it had the highest number of detected keypoints thus offering a higher chance of correct matches with other bands. The performance of the algorithms is shown in Figures 8.
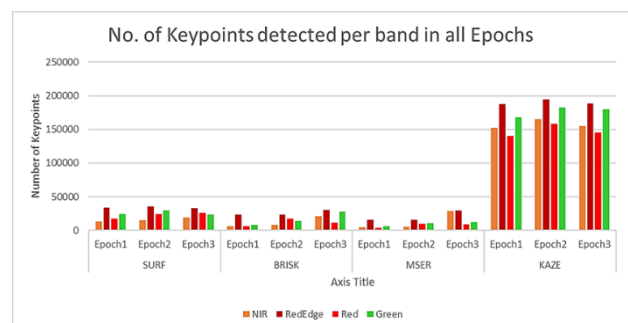

Figure 8 Showing feature detection per band for Parrot Sequoia

5.1.2 Feature Matching using default parameters
KAZE and SURF succeeded in finding correct matches within 1000 iterations for all the band combinations. Figure 9 shows algorithm performance for inter-epoch matching.
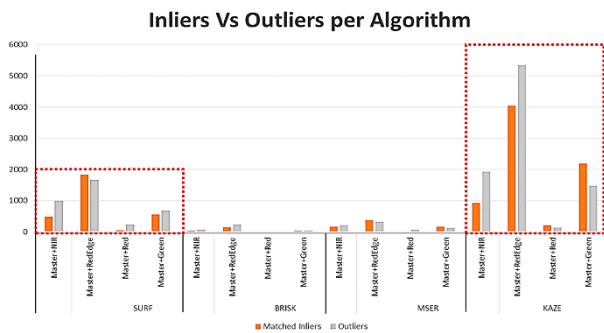
Figure 9. Illustrating inliers and outliers of matched Keypoints.

Despite being seen to have more outliers than inliers in the first and second band combinations, the overall performance of KAZE depicts more inliers than the other algorithms. On the other hand, BRISK and MSER failed to converge to find sufficient points after 1000 iterations for the master and red band combination. Thus, KAZE and SURF were selected for subsequent tests in this study.

5.2 Parameterization for feature detection

The results show that the higher the octave the more features are detected cumulatively. KAZE presents a sharp cumulative increase in points detected from the first to the second octave, but somewhat plateaus by the third octave.
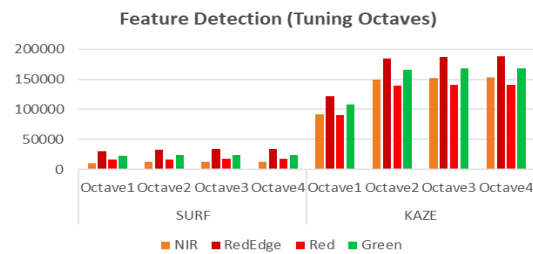

Figure 10. Showing the impact of tuning number of octaves

SURF is no match to KAZE in feature detection but is overwhelmingly fast as shown in table 4. SURF recorded a time difference per octave of less than one second. Conversely, KAZE doubles the time between the first to the second octave.

Time (Seconds)			Time (Seconds)		
Octaves	SURF	KAZE	Levels	SURF	KAZE
1	2.15	70.86	3	2.34	82.23
2	2.60	147.91	4	2.94	145.15
3	2.80	164.09	5	3.41	175.44
4	3.02	177.73	6	3.61	210.62

Table 4. Time taken per octave and per level for point detection

It was observed that a stable condition of feature detection is reached in scale level five and second octave because the number of features detected per band beyond these points are less than five percent of the total number of features detected. In addition, an increase in scale levels increases the computational time.

5.3 Parameterization for feature description and matching

A feature size of 128 provides a higher description accuracy but consequently decreased the number of matched and inlying features, which were insufficient for MSAC to fit the best model to estimate the transformation matrix. Therefore, misregistration errors were still evident after registration.

	Band Pairs	SURF		KAZE	
		Inliers	Outliers	Inliers	Outliers
64'	Master+NIR	190	1135	431	2420
	Master+RedEdge	661	2594	1822	7568
	Master+Red	25	250	122	234
	Master+Green	187	913	943	2735
128'	Master+NIR	22	80	107	401
	Master+RedEdge	75	264	946	2462
	Master+Red	7	5	13	10
	Master+Green	15	51	264	649

Table 5. Showing number of matches per descriptor size

On the other hand, with a descriptor size of 64, the number of matched and inlying features increased. The descriptor size of 64 was therefore selected since the number of inliers were enough for accurate estimation of the transformation matrix.

5.4 Intra-epoch band-to-band registration (Parrot Sequoia)

Using the Red edge as the master, image level analysis revealed a systematic displacement attributed to the basis distance of the cameras as presented graphically in Figure 11.
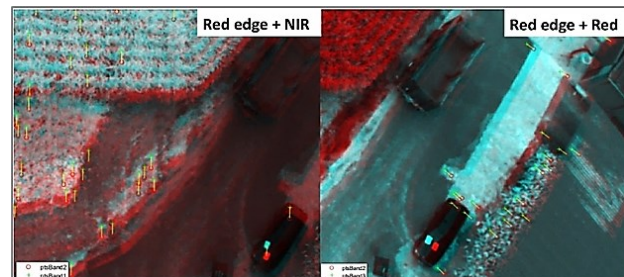

Figure 11. Showing image level systematic displacement

The red edge and NIR combination exhibited a displacement of about 11 pixels in the "Y" direction. The red edge and red showed a shift of about 6 pixels in both directions. The last combination, red edge and green, unveiled a uniform displacement of about 6 pixels in the "X" direction.

At the image level, misregistration is reduced from an average of 10 pixels to 0.28 pixel. The co-registration results at the image level within same epoch are presented in Table 6.

Stat.	RedEdge + NIR		RedEdge + Red		RedEdge + Green	
	Before	After	Before	After	Before	After
Max	12.37	0.7	14.06	0.57	8.54	0.89
Min	10.89	0.01	8.28	0.05	6.95	0.03
Mean	11.78	0.28	10.49	0.27	7.48	0.28
Std	0.28	0.15	1.64	0.14	0.33	0.16

Table 6. Image point pair distances before and after registration

On the other hand, misregistration of compositely processed orthophotos is at a minimal average of 0.3 pixels. This can be attributed to corrections during image level co-registration, triangulation, georeferencing, and orthorectification.

5.4.1 Intra-epoch accuracy assessment

At the orthophoto level, the results show that the bands are aligned to subpixel accuracy. The positional RMSE of the inliers is equal for epoch one and three, and decimal differences in epoch two, before and after registration. Since the registration procedure was intensity-based, the slight differences in epoch 2 could be attributed to randomness during outlier removal.

	Red edge + NIR		Red edge + Red		Red edge + Green	
	Before	After	Before	After	Before	After
Epoch1	0.17	0.17	0.22	0.22	0.18	0.18
Epoch2	0.18	0.18	0.20	0.16	0.21	0.18
Epoch3	0.18	0.18	0.21	0.19	0.20	0.20

Table 7. PS horizontal positional RMSE (pixels) using SURF

5.5 Inter-epoch band-to-band registration (Parrot Sequoia)

5.5.1 Many-to-One Registration

Subpixel accuracies were obtained using projection distance thresholds of 0.7 and 0.5 for SURF and KAZE respectively. SURF could not find sufficient point pairs to estimate geometric transformation between orthophotos of epoch 1 and 2 at a projection distance threshold of 0.5. However, the point pair distances obtained with SURF and KAZE after registration is presented in Figure 13. Misregistration errors in the range of 0.02 – 1.16 pixels for SURF, against KAZE's 0.02 – 1.37 pixels for all band combinations were evident.
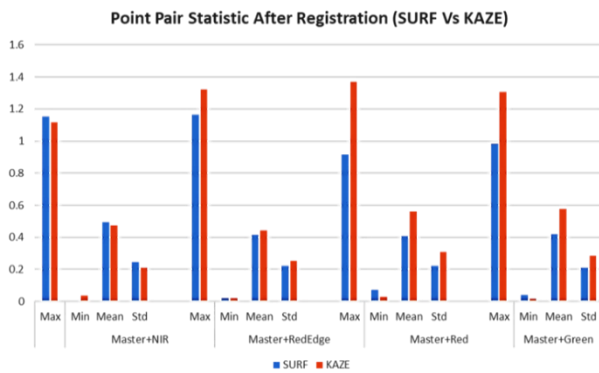
Figure 12. Point pair statistic of many-to-one band registration

5.5.2 One-to-One Registration

The displacement between corresponding bands was seen to be systematic across all the bands. It was observed that the red band combination recorded the lowest number of inliers whereas green had the highest. Nonetheless, the inliers allowed for accurate estimation of the transformation matrix, and thus subpixel accuracies were obtained. From the statistics presented in boxplots presented in Figure 13 (b), it is observed that the mean registration error is in the range of 0.26 – 0.38.
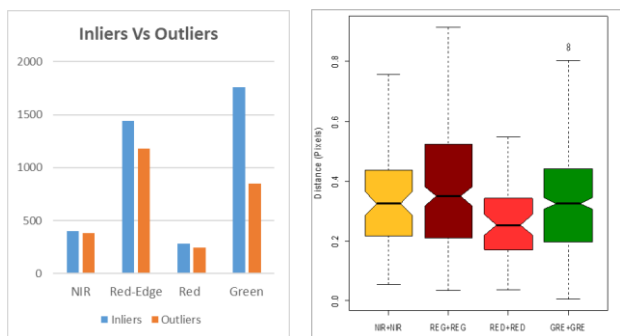

Figure 13 (a) Inliers vs outliers (b) Boxplots of paired distances

5.5.3 Inter-epoch accuracy assessment

The results presented in this section are those of aligning epochs two and three to epoch one using both matching strategies. As demonstrated in tables 8 and 9, both SURF and KAZE obtained subpixel registration accuracies. SURF however recorded lower RMSE values than KAZE.

Band Combination	Epoch 1 and 2		Epoch 1 and 3	
	SURF	KAZE	SURF	KAZE
Master + NIR	0.57	0.52	0.54	0.50
Master + Red edge	0.49	0.51	0.46	0.40
Master + Red	0.48	0.64	0.47	0.66
Master + Green	0.47	0.64	0.44	0.64

Table 8. Points pair RMSE (pixels) of many-to-one registration

Band Combination	Epoch 1 and 2		Epoch 1 and 3	
	SURF	KAZE	SURF	KAZE
NIR1 + NIR	0.36	0.34	0.39	0.37
Red edge1+Red edge	0.39	0.36	0.33	0.31
Red1 + Red	0.28	0.30	0.31	0.31
Green1 + Green	0.32	0.32	0.34	0.33

Table 9. Points pair RMSE (pixels) of one-to-one registration

The one-to-one registration approach yields better results than many-to-one approach. The many-to-one approach recorded an average RMSE of 0.5 against the one-to-one approach of 0.36 across all band combinations. The similarity in spectral properties per band combination in one-to-one approach is one of the possible reasons as to why band pairs are better aligned.

6. CASE STUDY: MONITORING OF MAIZE CROP

6.1 Parrot Sequoia Versus Micro MCA NDVI Analysis

Comparative NDVI zonal statistics were computed for both UAV images. Zones A, C, D, E, G, H and I are within the maize field but with different crop densities as shown in Figure 14.
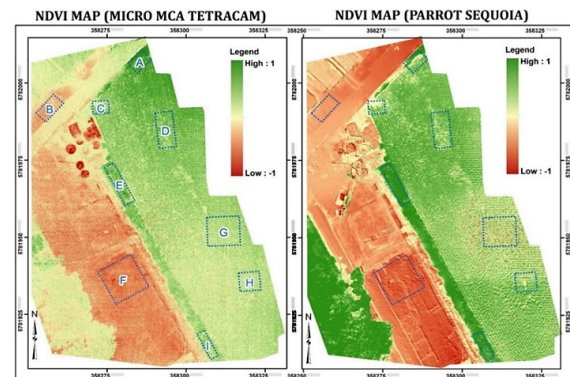

Figure 14. NDVI maps of Micro MCA and Parrot Sequoia

The Vegetation Index (VI) response as extracted with PS and Micro MCA are highly correlated; the spatial averages of each zone recorded a positive correlation of 0.86. Despite the differences in GSD and in image quality between the two cameras, the spatial variability of NDVI is highly comparable. The average deviation of the mean values between the two datasets is 0.16, with the PS registering higher scores than Micro MCA. As shown in Figure 15, the highest average deviations are seen in zones E and I; Micro MCA depicts comparatively low NDVI values than PS.
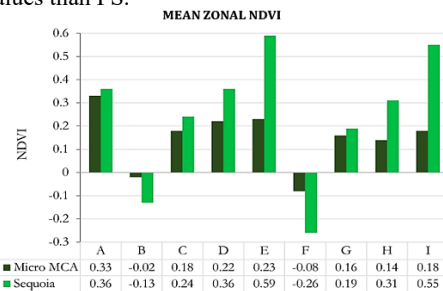

	A	B	C	D	E	F	G	H	I
Micro MCA	0.33	-0.02	0.18	0.22	0.23	-0.08	0.16	0.14	0.18
Sequoia	0.36	-0.13	0.24	0.36	0.59	-0.26	0.19	0.31	0.55

Figure 15. Zonal NDVI average for Micro MCA and Sequoia

6.2 Inter-epoch Analysis of NDVI

Having established the variations in NDVI from images acquired in the same day but with different spatial resolutions, an inter-epoch comparative analysis of vegetation performance was done for all the three epochs. Being a qualitative assessment, the zonal statistics presented in Figure 16 illustrate the observed spatiotemporal dynamics of NDVI values.

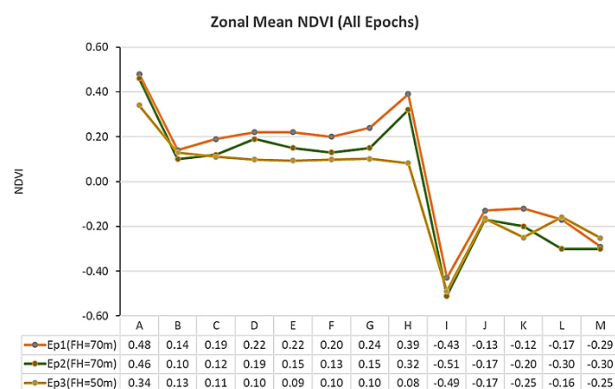|	A	B	C	D	E	F	G	H	I	J	K	L	M
Ep1(FH=70m)	0.48	0.14	0.19	0.22	0.22	0.20	0.24	0.39	-0.43	-0.13	-0.12	-0.17	-0.29
Ep2(FH=70m)	0.46	0.10	0.12	0.19	0.15	0.13	0.15	0.32	-0.51	-0.17	-0.20	-0.30	-0.30
Ep3(FH=50m)	0.34	0.13	0.11	0.10	0.09	0.10	0.08	0.08	-0.49	-0.17	-0.25	-0.16	-0.25

Figure 16. Statistical comparison of NDVI values (Epoch 1 to 3)

Spatiotemporal analysis of NDVI depicted a declining trend over time; this is possibly due reduced photosynthetic activity between fruit development (i.e. Silking) and ripening stages of maize. In addition, the heavy storm event experienced between

epoch two and three could be a possible reason for lower NDVI values in epoch three due to damaged crops (See Figure 17 for diminishing greenness of the crop).

Figure 17. Showing crop greenness between epoch 1 and 3

6.3 Spectral Variability Analysis

Both cameras can spectrally distinguish photosynthetically active and less active vegetation; the green band reflects more than the red which absorbs most reflectance, and a sharp transition is seen in the red edge. The average spectral deviation between the cameras is by a factor of 1.6 and 1.3 for photosynthetically active and less active crops respectively.
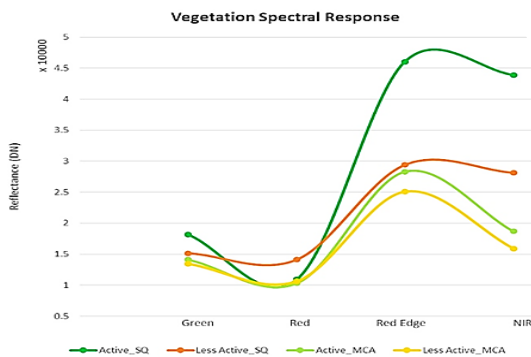

Figure 18. Spectral analysis of Sequoia and Micro MCA

The observed deviations could be attributed to image quality; difference in camera calibration, and spectral band width. More significantly, it can be attributed to the missing incidence sensor on the Micro MCA thus changes in irradiance during capture are not accounted for. Vegetation surfaces do not reflect light evenly in all directions i.e. Lambertian properties. The non-uniform reflectance hampers pixel-wise comparison and thus zonal comparison in this study.

7. DISCUSSION

The complex nature of vegetation surfaces calls for more stringent parameters as far as UAV-based multispectral sensing is concerned. In this study, a side overlap of 40% was used for image acquisition. This was observed to be insufficient to generate a stable photogrammetric block; artificial zonal variability corresponding to flight lines was evident on the orthophoto as shown on Figure 19 below.
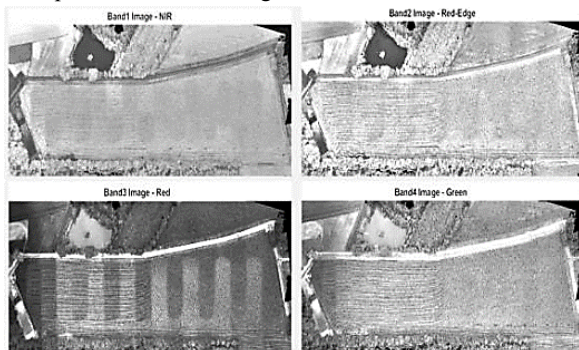

Figure 19. Per band artificial strips corresponding to flight lines

Larger overlaps of at least 60% are recommended. However, for UAV-based multispectral image acquisition, Assmann et al., (2019), recommend a minimum of 75% for both front and side overlap.

The results show that region detectors that use float descriptors are more robust than the ones using binary detectors. SURF and

KAZE detected and indexed correct matches in all the spectral channels while the binary descriptors, MSER and BRISK, failed to find enough keypoints to qualify as inliers for all the band combinations; binary descriptors compare pixel differences whereas float descriptors compute intensity gradients.

In this study, horizontal positional RMSE of the inlying conjugate points was used for accuracy assessment; another possible way to assess the co-registration accuracy could be to compute the epipolar geometry between band-pairs and compute the residual error of the distances of matched points from their corresponding epipolar lines (Onyango et al., 2017). It is however important to note that accuracy varies depending on the method used to estimate geometric transformation. Jhan et al. (2017), argue that image planes are not exactly parallel thus the use of similarity and affine transformation for band-to-band co-registration is unsuitable. On the flip side, image blocks in this study were orthorectified, thus assumptions of planarity, parallelism and similarity were made. As such, the use of similarity transformation method to estimate the geometric transformation.

SURF and KAZE obtained subpixel registration accuracies but SURF was observed to be faster. KAZE employs the additive operator splitting, which has been reported to be quite inefficient (Gerke, Nex, & Jende, 2016). On the other hand, KAZE is more rigorous in feature detection than SURF. The results of this study demonstrated that both KAZE and SURF are effective algorithms for co-registration of multispectral images.

Intra-epoch and inter-epoch coregistration achieved sub-pixel accuracy in the range of 0.16 – 0.22 and 0.28 – 0.39 pixels respectively, which is adequate for accurate crop monitoring. The results of Townshend et al. (1992), demonstrate that to obtain NDVI values with a 90% confidence level, intra-epoch co-registration accuracy of 0.2 pixels or less should be obtained.

8. CONCLUSION

The main objective of this research was to investigate intra-epoch and inter-epoch misregistration of multispectral UAV imagery, and to explore the potentials of unmanned aerial systems for crop monitoring. This study proposes an intensity-based feature detection and description method to automatically co-register both intra-epoch and inter-epoch multispectral imagery. SURF and KAZE were tested and both detectors demonstrated the ability to co-register multispectral imagery to subpixel accuracy.

In light of the results obtained in this study, SURF is equally robust, more efficient, and computationally inexpensive than KAZE. On the other hand, KAZE is more vigorous than SURF and always detects more keypoints hence increasing the chances to get more correct matches per band combination. Both algorithms can be used satisfactorily for intra-epoch and inter-epoch co-registration, although their performance will vary based on parameterization and image quality. The presented co-registration approaches of many-to-one and one-to-one can therefore be used but the one-to-one is preferred.

The analysis of NDVI results between Sequoia and Micro MCA demonstrated that due to differences in spectral regimes of multispectral UAV imagery, the use of one system throughout the monitoring period is prudent. In summation, both Parrot Sequoia and Micro MCA are applicable for crop monitoring; they both have the spectral bands vital for monitoring photosynthetic activities of crops. Although Micro MCA has two more bands and can therefore sense more spectral properties, these additional spectral features offer a basis for future research. Also, in this study, GCPs were not used to scale, georeferenced and estimate distortions within the photos. Thus, greater inter-epoch displacements. To investigate the misregistration error between orthophotos that have been processed using GCPs in all the epochs would go a long way in contributing to inter-epoch co-registration methods.

ACKNOWLEDGEMENTS

REFERENCES

Aicardi, I., Nex, F., Gerke, M., & Lingua, A. (2016). An Image-Based Approach for the Co-Registration of Multi-Temporal UAV Image Datasets. *Remote Sensing*, *8*(9), 779. https://doi.org/10.3390/rs8090779

Alcantarilla, P. F., Bartoli, A., & Davison, A. J. (2012). KAZE features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7577 LNCS, pp. 214–227). https://doi.org/10.1007/978-3-642-33783-3_16

Assmann, J. J., Kerby, J. T., Cunliffe, A. M., & Myers-Smith, I. H. (2019). Vegetation monitoring using multispectral sensors — best practices and lessons learned from high latitudes. *Journal of Unmanned Vehicle Systems*, *7*(1), 54–75. https://doi.org/10.1139/juvs-2018-0018

Banerjee, B. P., Raval, S. A., & Cullen, P. J. (2018). Alignment of UAV-hyperspectral bands using keypoint descriptors in a spectrally complex environment. *Remote Sensing Letters*, *9*(6), 524–533. https://doi.org/10.1080/2150704X.2018.1446564

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 3951 LNCS, pp. 404–417). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_32

Elarab, M., Ticlavilca, A. M., Torres-Rua, A. F., Maslova, I., & McKee, M. (2015). Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *International Journal of Applied Earth Observation and Geoinformation*, *43*, 32–42. https://doi.org/10.1016/j.jag.2015.03.017

Freeman, P. K., & Freeland, R. S. (2015). Agricultural UAVs in the U.S.: Potential, policy, and hype. *Remote Sensing Applications: Society and Environment*, *2*, 35–43. https://doi.org/10.1016/j.rsase.2015.10.002

Fytsilis, A. L., Prokos, A., Koutroumbas, K. D., Michail, D., & Kontoes, C. C. (2016). A methodology for near real-time change detection between Unmanned Aerial Vehicle and wide area satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 165–186. https://doi.org/10.1016/j.isprsjprs.2016.06.001

Gerke, M., Nex, F., & Jende, P. (2016). Co-registration of terrestrial and UAV-based images - Experimental results. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 40, pp. 11–18). https://doi.org/10.5194/isprsarchives-XL-3-W4-11-2016

Jhan, J. P., Rau, J. Y., Haala, N., & Cramer, M. (2017). Investigation of parallax issues for multi-lens multispectral camera band co-registration. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 42, pp. 157–163). https://doi.org/10.5194/isprs-archives-XLII-2-W6-157-2017

Jhan, J. P., Rau, J. Y., & Huang, C. Y. (2016). Band-to-band registration and ortho-rectification of multilens/multispectral imagery: A case study of MiniMCA-12 acquired by a fixed-wing UAS. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 66–77. https://doi.org/10.1016/j.isprsjprs.2016.01.008

Kelcey, J., & Lucieer, A. (2012). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, *4*(5), 1462–1493. https://doi.org/10.3390/rs4051462

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011). BRISK: Binary Robust invariant scalable keypoints. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2548–2555). https://doi.org/10.1109/ICCV.2011.6126542

Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, *143*, 56–64. https://doi.org/10.1016/j.fcr.2012.08.008

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. In *Image and Vision Computing* (Vol. 22, pp. 761–767). Elsevier Ltd. https://doi.org/10.1016/j.imavis.2004.02.006

Nex, F., Gerke, M., Remondino, F., Przybilla, H.-J., Bäumker, M., & Zurhorst, A. (2015). Isprs Benchmark for Multi-Platform Photogrammetry. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-3/W4*(March), 135–142. https://doi.org/10.5194/isprsannals-II-3-W4-135-2015

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: A review. *Applied Geomatics*. https://doi.org/10.1007/s12518-013-0120-x

Onyango, F. A., Nex, F., Peter, M. S., & Jende, P. (2017). Accurate estimation of orientation parameters of UAV images through image registration with aerial oblique imagery. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 42, pp. 599–605). https://doi.org/10.5194/isprs-archives-XLII-1-W1-599-2017

Rey, C., Martin, M., Lobo, A., Luna, I., Diago, M., Millan, B., & Tardaguila, J. (2013). Multispectral imagery acquired from a UAV to assess the spatial variability of a Tempranillo vineyard. *Precision Agriculture 2013 - Papers Presented at the 9th European Conference on Precision Agriculture, ECPA 2013*, *55*(June 2015), 617–624. https://doi.org/10.3920/978-90-8686-778-3

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., & Bareth, G. (2014). Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *Journal of Applied Remote Sensing*, *8*(1), 083671. https://doi.org/10.1117/1.JRS.8.083671

Torr, P. H. S., & Zisserman, A. (2000). MLESAC: A New Robust Estimator with Application to Estimating Image Geometry. *Computer Vision and Image Understanding*, *78*(1), 138–156. https://doi.org/10.1006/cviu.1999.0832

Townshend, J. R. G., Gurney, C., McManus, J., & Justice, C. O. (1992). The Impact of Misregistration on Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(5), 1054–1060. https://doi.org/10.1109/36.175340

Turner, D., Lucieer, A., Malenovský, Z., King, D. H., & Robinson, S. A. (2014). Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over antarctic moss beds. *Remote Sensing*, *6*(5), 4003–4024. https://doi.org/10.3390/rs6054003

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., … Yang, H. (2017, June 30). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science*. Frontiers Media S.A. https://doi.org/10.3389/fpls.2017.01111