Pixel-wise registration of close-range multispectral images acquired with multi-lens cameras

Antonio Maria Garcia Tommaselli [1], Beatriz Coêlho Silva [1], Wimerson Sanches Bazan[1,2], Andressa Yida Pinheiro Souza[1]. Letícia Rosim Porto[1]

[1] Department of Cartography, Faculty of Sciences and Technology, São Paulo State University (UNESP),
(a.tommaselli, beatriz.coelho-silva, wimerson.bazan, andressa.yida, leticia.porto)@unesp.br
[2] Espirito Santo Federal Institute

Keywords: Multispectral camera, Close-range photogrammetry, Calibration, Image registration, proximal sensing

Abstract

Lightweight multispectral cameras have been widely used for many applications, mainly agriculture. Suitable results in image analysis processes depend on the accurate registration of the image bands, collected by lenses that are mutually displaced. The problem of registering image bands of close-range objects collected by multispectral cameras is still lacking in effective solutions. Most of the existing approaches are not capable of accurately matching each pixel in the images since depth variations still cause displacements. In this paper, we proposed a solution by extracting a digital surface model (DSM) and generating orthoimages of each image band. We experimentally showed that image bands of the same frame shot, collected at the same station and time, can be used to generate a DSM and orthoimages, which will be mutually registered. To achieve this, a rigorous camera calibration must be previously performed to provide the orientation parameters for the DSM extraction and orthorectification. The experimental results showed that the standard deviation of the orthorectified pixels is approximately 1.5 pixels.

1. Introduction

The use of terrestrial mobile mapping systems for monitoring agricultural crops can generate accurate and detailed results at a low cost, helping in the detection of diseases and pests, yield prediction, and other applications (ElMasry et al., 2019). Lightweight multispectral cameras, such as Micasense (Micasense, 2024) and Agrowing (Agrowing, 2024), can collect significant data aiming at those tasks, but there are some drawbacks to be addressed.

Most of multispectral cameras in the market use multiple lenses and sensors to generate several spectral bands (Micasense, 2024; Parrot, 2022; DJI, 2022) as can be seen in Figure 1. Thus, the lenses’ perspective centres will not coincide in this type of camera, requiring more elaborate processes to generate registered image stacks.

(a) (b) (c)

Figure 1. Multiple lens multispectral cameras; (a) Micasense RedEdge-P; (b) Parrot-Sequoia; (d) DJI-P4.

Detailed and accurate multispectral analysis requires pixel-level band co-registration. Registration is straightforward for images acquired by aerial platforms since the flight height is much larger than the cameras’ offsets. For moderate steep terrain, ordinary 2D geometric transformations or even translations can be used with good results (Paul and Pati, 2021; Zitová and Flusser, 2003).

Close-range acquisition, however, poses further challenges due to the offsets of the cameras’ lenses, which will cause parallaxes between pixels in some spectral bands, depending on the object to the camera distances and the variations in depth (Jhan et al., 2017). This problem can happen also with low flight height acquisition, as reported by Hassanpour et al. (2019).

Several approaches were proposed to cope with this problem. A search in the Scopus database with keywords *multispectral AND image AND registration* returned 1088 documents. This reflects the huge amount of work on the general image registration problem in several areas. The search was then narrowed using keywords *multispectral AND image AND registration AND close-range* returning only 16 documents. However, only a few of these journal papers approached the problem of challenging scenes with huge depth variation, such as trees’ fields.

Jhan et al. (2017) were concerned with the parallaxes issues and proposed the RABBIT (Robust and Adaptive Band-to-Band Image Transform) technique which was divided into three steps: (1) system and camera calibration; (2) lens distortion correction and image transformation and (3) final optimization to improve registration accuracy. They focused on detailed and accurate system calibration and assessed the effects of parallax for close-range images, concluding that fine co-registration was achieved only for objects at the same distance.

Hassanpour et al. (2019) presented a registration technique that divides the image into patches, selecting the appropriate local window, aiming at reducing the effect of relief displacement on miss-registration error.

Fernández et al. (2021) assessed 2D geometric transformations for registration of multispectral images acquired at close range (1.5m) over greenhouse cucumber. They achieved an RMSE of less than 1 pixel with the similarity and affine transformations and of less than 2 pixels with the projective transformation. This acceptable result can be explained by the removal of outliers, probably caused by points farther or closer than the main group of points.

Dandrifosse et al. (2021) assessed eight registration methods for multimodal image fusion for wheat canopy. One of the evaluated methods was the DDTM (distance-dependent transformation

matrix), which considers a homography matrix for each distance of the camera to the object. Other techniques were also based on 2D geometric transformations. Elastix technique, originally proposed by Klein et al. (2010) was also assessed. The local image deformations were handled by Guo et al. (2022) with a piecewise local transformation based on triangular irregular networks.

From the literature review, it could be seen that the majority of approaches rely on full image geometric transformations or patch/local crops in which 2D transformations are applied. However, these approaches are not capable of accurately matching each pixel in the images since depth variations still cause displacements. The accurate registration is achieved only by properly correcting these displacements caused by the depth and viewpoint changes, as usually done with orthorectification in the aerial case.

In this paper, we propose an approach based on the extraction of a Digital Surface Model (DSM), followed by orthorectification of the spectral bands. The rationale under the proposed approach is to use images of the same frame shot, thus collected at the same station and time, to generate a DSM, considering the different viewpoints for each lens followed by orthorectification. The process relies on previous camera calibration to provide inner orientation and relative orientation with respect to a reference camera, in which reference system the images will be orthorectified.

2. Methodology

The registration of spectral bands acquired with multi-lenses multispectral cameras at close-range requires a strategy to compensate for the effect of differential parallax resulting from the proximity between the sensor and the depth variations.

This study used a multi-lens multispectral camera (Agrowing, 2024) and a commercial software for processing, Agisoft Metashape (Agisoft, 2023). The procedure encompasses four main steps: determining a suitable principal distance, calibrating the Interior Orientation Parameters (IOPs) and Exterior Orientation Parameters (EOPs), determining the Relative Orientation Parameters (ROPs) among the lenses, and correcting differential parallaxes based on the orthorectification of each image band.

2.1 The multispectral sensor

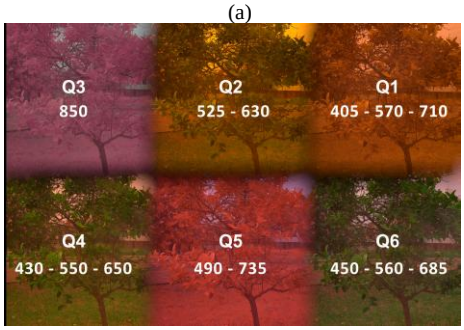
The sensor used in this work is an Agrowing Sony Alpha 7R2 Sextuple Multispectral camera (Figure 2.a) (Agrowing, 2024), which is a combination of a commercial digital camera with a special optical mounting with six lenses and filters that allow information to be collected in 14 spectral bands (Figure 2.b).

Characteristics	Specifications
Field of vision	45,90°
Lens distortion	< 1%
Effective focal length	21,6 mm
Diaphragm opening	f/5.6 (fixed)
Sensor weight (with battery)	845 g
Frane image size	9504x6336 pixels
Pixel size	0,0037 mm

Table 1. Camera specifications.

This camera optics projects six bundles over the sensor, and with further processing, 14 spectral image bands can be extracted, which are summarized in Figure 2.b.

(a)



(b)

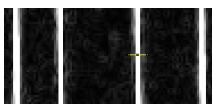
Figure 2. Multispectral sensor: (a) Front view of the Sony camera and attached Agrowing lenses; (b) arrangement of spectral bands over the full frame.

2.2 Setting the principal distance

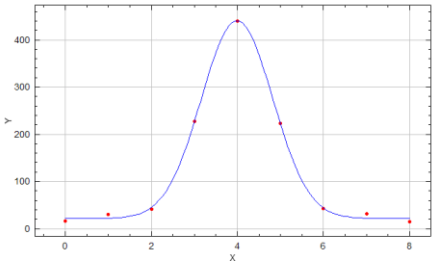
Because of the manufacturer modifications in the optical set, the most suitable focusing distance to the object depth cannot be automatically set in this camera, and, thus, the lenses must be manually adjusted. For close-range acquisition, visual-based focusing, or based on the existing distance scales markings over the lenses are not accurate enough to provide focused images. A millimetre scale was attached to the lens's focusing ring to serve as a more accurate reference (Figure 3.a). A binary bars target was placed vertically over a wall (Figure 3.b), and images were acquired at different object distances and with several focusing settings following the grades in the millimetric scale attached to the camera objective.

(a)

(b)



(c)



(d)

Figure 3. (a) Millimetric scale attached to the lens's focusing ring and (b) Binary bar target (d) A Gaussian fitted to the edge.

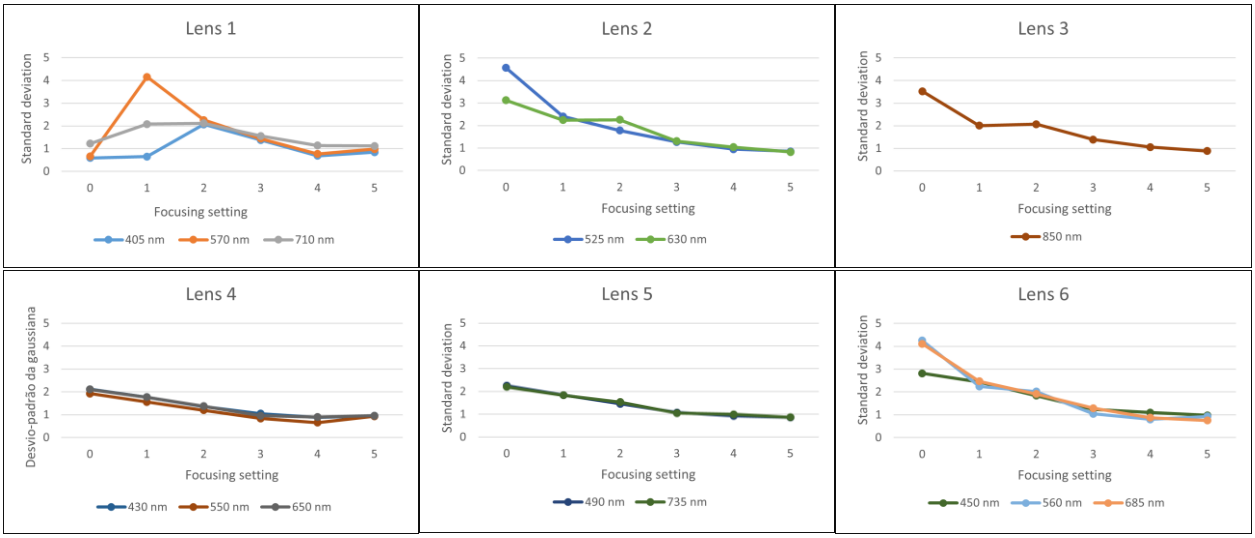
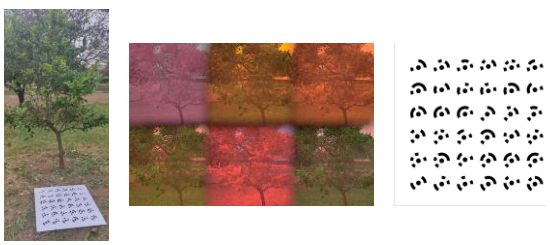


Figure 4. Standard deviations of the Gaussians adjusted over vertical edges for each group of image bands and for five focusing sets.

All image bands generated from each lens and filter set were extracted in order to analyse the focusing blur for each of the principal distance settings. Vertical and horizontal binary edges from the images of the target were extracted (Figure 3.c), and a Gaussian function was adjusted with ImageJ (Figure 3.d). Then, the standard deviations of this adjusted Gaussian function for each distance and for each image band were extracted and plotted (Figure 4). In Figure 4, it can be seen the standard deviation reduction when changing the focusing setting. The focus number 5 was then selected, with 0 corresponding to focusing on the infinite. The camera was then mechanically locked and used as the most suitable for the selected distance.

2.3 Data acquisition and camera calibration

The data can be acquired with the camera manually carried by the operator (Figure 5.a), or by a moving platform with the focusing setting locked to the principal distance selected and suitable to the average depth of the object.



(a) (b) (c)

Figure 5. Data acquisition: (a) details of the tree used in the experiments with the calibration targets; (b) an example of an original image from the camera;(c) the calibration plate.

The first step in the acquisition mission is to collect a set of images aiming at the camera calibration. This acquisition can be performed in the same environment as the field work to ensure consistent parameters or in previous days, provided that the stability of the camera is ensured. The camera calibration is done with self-calibrating bundle adjustment; a minimum of seven constraints are required to provide scale and a reference system, which is usually done by providing the coordinates of control points. An effective and practical solution is to use a calibration plate with coded targets, as presented in Figure 5. This calibration plate can be easily carried and inserted in the area to be assessed. The calibration panel shown in Figure 5 is composed of 36 circular coded targets with 10 cm of spacing over a metallic flat surface. The panel can be installed near the object and several oblique images with different locations and orientations are collected. To ensure accurate results the camera must be calibrated with the selected focusing distance locked and the calibration panel rigorously stable.

2.4 Calibration Processing and Determination of the Relative Orientation

The calibration can be performed with several techniques and software. Agisoft Metashape was used to automatically extract the targets’ centre image coordinates and further tie points over the image field. For the experiments presented in this paper, 24 targets were used as ground control points (GCP) and 12 targets were left as checkpoints (CkP). Additional tie points (approximately 30000 for each calibration set) were also extracted by Metashape and used in the camera calibration process. The interior and exterior orientation parameters for each lens were estimated with Metashape. Only one image band for each lens was selected for each image station collected, thus producing six calibration sets. The interior and exterior orientation parameters were estimated in the bundle adjustment using the reference system of the calibration plate. The average values of RMSEs in the GCPs and in the independent check points for the six sets are summarized in Table 2. It can be noted that the self-calibrating bundle adjustment produced coordinates with sub-millimetre errors.

Average RMSE				
	X error (mm)	Y error (mm)	Z error (mm)	Total (mm)
24 GCP	0.56	0.66	0.80	1.18
12 CkP	0.70	0.74	0.88	1.35

Table 2. Average values of the RMSEs for each component and the total, for the six calibrations sets, for control and check points.

After determining the IOPs and the EOPs for the calibration setup, using the calibration plate as a reference, it is required to estimate the ROPs of each lens with respect to the lens chosen as reference (Q6). This lens (Q6) was taken as a reference to compute the relative orientation parameters of each lens from the estimated exterior orientation parameters, resulting in the values shown in Table 3.

	Bx (mm)	By (mm)	Bz (mm)	ω	φ	κ
Q1	00.2	-11.7	0.2	0°0'59.38"	0°0'01.78"	0°0'16.31"
Q2	11.8	-12.4	7.0	0°5'27.60"	0°0'00.03"	0°0'31.33"
Q3	24.8	-11.6	5.4	0°6'24.08"	0°0'03.63"	0°0'25.94"
Q4	24.3	-0.2	-1.6	0°3'46.23"	0°0'04.66"	0°0'31.31"
Q5	12.7	-0.5	6.1	0°2'09.48"	0°0'13.58"	0°0'54.00"

Table 3. Relative orientation parameters, taking lens 6 as reference.

Analysing Table 3 it can be noted that the offsets among the lenses are larger in Bx (approximately 25 mm or 12 mm) and are compatible with the values directly measured with a calliper, which means that the indirect determination of the ROPs were successful. Although slight variations in the values of the principal distances are expected due to the different focus, depending on the wavelengths, mainly for the infrared band, the differences in Bz are more difficult to explain as physical values. Thus, it is likely that these variations are caused by errors in the estimation of the camera perspective centre coordinates in the bundle adjustment. As for the alignment of the optical axis, the larger values were achieved for X-axis rotation (ω). It is not clear whether these values are true misalignments or can be a result of the kind of calibration and setup. The discussion and mitigation of these remaining problems is left for future work.

The calibration step requires many images collected from different stations, to provide suitable geometry. The calibration must be performed only at the beginning of the acquisition session, or even, a previous calibration can be used, provided that the camera is stable. However, for the remaining steps for pixel-wise registration each frame must be treated independently, but using the previously determined IOPs and ROPs. For a typical acquisition session, when inspecting agricultural fields, for instance, two thousand frames are acquired, and thus, it would be unfeasible to generate a full photogrammetric image block and a 3D point cloud.

2.5 Estimating a Digital Surface Model and correcting differential parallaxes

The proposal of this work is the determination of a digital surface model from the image bands of a single frame shot, instead of using several frames from different positions and orientations. The generation of a 3D model considered the existing base distance of approximately 2.5 cm (see Table 3) between some lenses. For a distance of 2 m, this will result in a small base-to-depth ratio (0.0125), but it is still suitable to generate a depth model, although with larger errors when compared to those generated with a suitable base/depth ratio. For a single stereopair, error prediction (Kraus, 2011) indicates an error in the depth of 13.7 mm, considering only the parallax measurement error of 0.5 pixels, which can be considered acceptable for the existing conditions.

The first step for generating a DSM is to determine the relative orientation for each stereopair. This can be performed with many algorithms, such as the conventional relative orientation either with collinearity or coplanarity equations (Kraus, 2011). However, because of the short baseline and points at infinity in the object space, these algorithms can diverge in many situations or converge to incorrect values. The best alternative is the use of the previously calibrated IOPs and ROPs either as absolute values or as weighted values. The first option corresponds to the forward intersection problem, while the second one corresponds to a bundle adjustment with relative constraints. We selected the second option by considering a block with 6 images and performing a bundle adjustment with the previously determined IOPs as fixed parameters and the ROPs as weighted constraints.

The last step in this procedure is the generation of orthoimages using the generated digital surface model and the existing ROPs and IOPs.

2.6 Data acquisition for experimental assessment

For the validation experiments, we collected images of an isolated lima tree (Fig. 5.a), at approximately 2 m from the camera to the object, with the calibration plate installed horizontally in such a way that it can appear in some images. We selected a group of images for calibration and one frame for the validation of the pixel-wise DSM generation and registration.

As previously stated, even considering the short base length, we hypothesized that acceptable results for differential parallax correction could be achieved when using the images of the 6 lenses simultaneously.

The digital surface model was generated with the Agisoft Metashape software (Agisoft, 2023) setting the calibrated interior orientation parameters and constraining the exterior orientation with the known values of the relative orientation. The 6 image bands of a single shot were inserted in the same project, and the key points were extracted. Figure 6 shows the images of 6 bands, acquired from the same station and from the same frame. It can be seen some displacements among the bands due to the lens offsets and depth variations. In Figure 9.a, a colour composition using bands 550 nm and 560 nm can viewed with anaglyph glasses, showing that enough parallax exists to generate a 3D model. Then, the images were aligned, generating a sparse depth model and updating the relative orientation values within the inserted standard deviations. A dense depth model is then generated, and the images are orthorectified using this dense model.

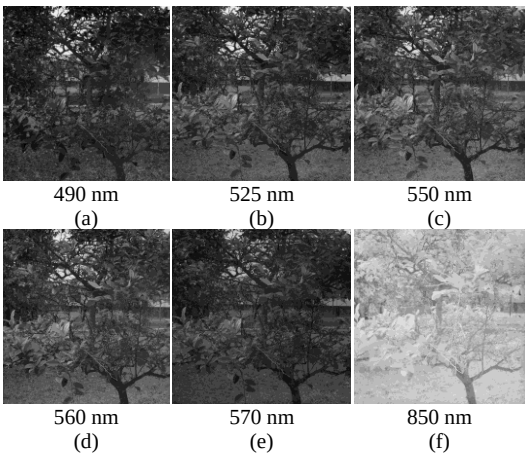
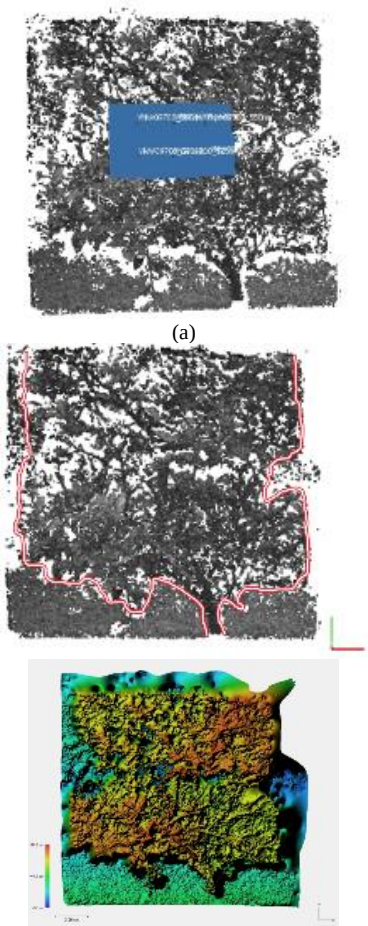


Figure 6. Examples of image bands used for the generation of the digital surface model.

3. Results and analysis

The images were imported and aligned in the Metashape following the procedures previously presented. A dense surface model was generated (Figure 7.a) indicating that the proposed process with the images from a single shot from the 6 lenses was satisfactory since the objects (leaves and branches) are clearly defined, as shown in Figure 7.b, although background object points were also modelled.



(a)

Figure 7. Generated digital surface model: (a) frontal view; (b) boundary of the object of interest; (c) colourized depth model.

The orthorectified images were then generated (Figure 8), which showed many empty areas due to occlusions. Nevertheless, with this result, it is possible to assess the result of the differential registration between the images of the 6 lenses.

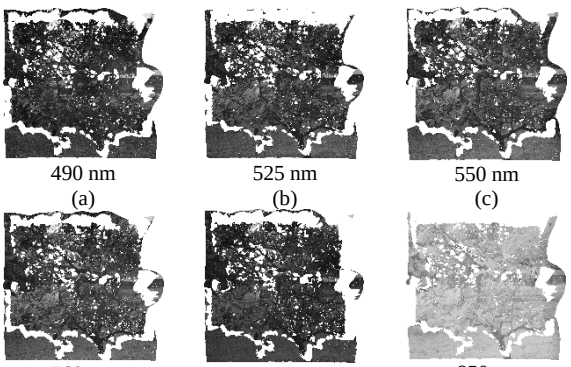
The result of the registration between the main pair of bands (550 nm - Q4 and 560 nm - Q6) can be analysed by comparing the overlap of two original images (Figure 9.a) and after orthorectified (Figure 9.c). These images are a colour composition of band 550 nm (in red) and band 560 nm (in green). Figures 9.b and 9.d show details of both orthorectified images to emphasize the differences for a bunch of leaves before and after rectification.

In addition to the visual comparison depicted in Figure 9, a quantitative analysis was performed by extracting keypoints in the five image pairs in relation to band 6, before and after orthorectification. The standard deviations of the discrepancies for these five image pairs were computed and are presented in Table 4.

	Lens pairs				
	1 - 6	2 - 6	3 - 6	4 - 6	5 -6
Original	2.0	9.9	7.3	18.9	3.7
Orthorectified	0.9	0.9	1.6	1.0	1.5

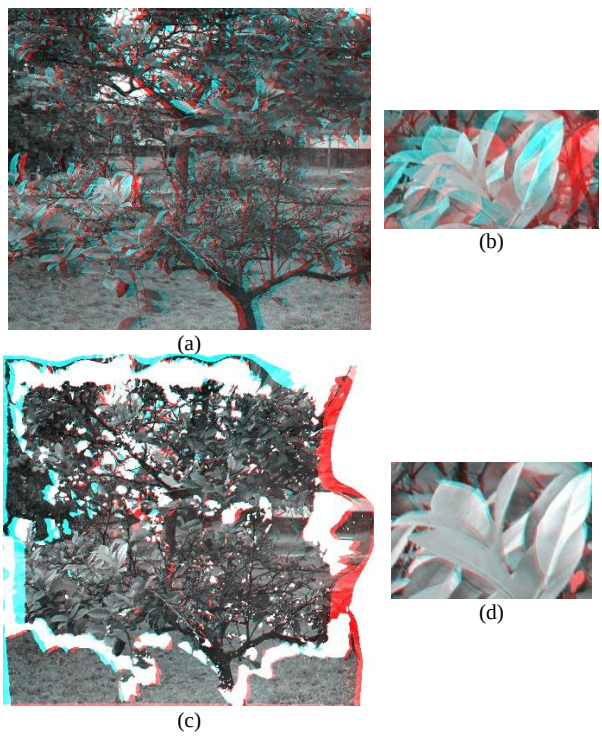
Table 4. Standard deviation (in pixels) of horizontal parallaxes between corresponding points, before and after orthorectification.

The results presented in Table 4 show that the parallaxes were significantly reduced after orthorectification, and thus, the pixels can be considered registered. The standard deviations of the residual discrepancies are approximately 1 pixel for pairs 1-6, 2-6 and 4-6 and 1.5 pixels for pairs 3-6 and 5-6. These differences can be caused by the previous processes, such as the determination of the parameters of Relative Orientation, digital surface model generated with weak geometry, keypoints determination, and orthoimages generation.



490 nm (a) 525 nm (b) 550 nm (c)
560 nm (d) 570 nm (e) 850 nm (f)

Figure 8. Example of orthorectified images, by spectral band.



(a) (b) (c) (d)

Figure 9. Registration between quadrants 4 and 6 and details for a bunch of leaves: (a) and (b) original; (c) and (d) after orthorectification.

4. Conclusions

Lightweight multi lens multispectral cameras have gained importance due to their widespread use with drones and with mobile terrestrial platforms for many applications, mainly agriculture. Suitable results in image analysis processes depend on the accurate registration of the image bands collected by lenses that are displaced from each other. The problem of registering image bands of close-range objects collected by multispectral cameras is still lacking in effective solutions.

In this paper, we proposed a solution by extracting a digital surface model and generating orthoimages. We experimentally showed that image bands of the same frame shot, collected at the same station and time, can be used to generate a DSM and orthoimages, which will be mutually registered. To achieve this, a rigorous camera calibration has to be previously performed. Having the IOPs and ROPs, the remaining processes (generating a DSM and orthoimages) can be run for each frame, with high efficiency. This process has to be executed automatically to enable the processing of a huge number of images.

This work was a feasibility study performed with commercial software and in particular conditions. For future work, it is expected to develop fully automatic processing and to test the approach for different conditions. Some parts of the process, such as the camera calibration, can be further developed to consider some constraints coming from direct measurements of the lens offsets.

Acknowledgements

This study was financed in part by the São Paulo Research Foundation, FAPESP (Grants 2021/06029-7 and 2022/12681-1) and by the National Council for Scientific and Technological Development, CNPq (grant n. 303670/2018-5).

References

Agisoft, 2023. MetaShape. URL https://www.metashape-la.com/en/home/ (accessed 2.11.23).

Agrowing, 2024. agrowing [WWW Document]. AGROWING. URL https://agrowing.com/ (accessed 11.30.24).

Dandrifosse, S., Carlier, A., Dumont, B., Mercatoris, B., 2021. Registration and fusion of close-range multimodal wheat images in field conditions. *Remote Sens.* https://doi.org/10.3390/rs13071380

DJI, 2022. DJI P4 Multispectral. URL https://aetha.global/product/dji-p4-multispectral/ (accessed 3.24.22).

ElMasry, G., Mandour, N., Al-Rejaie, S., Belin, E., Rousseau, D., 2019. Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. *Sensors*, 19, 1090.

Fernández, C.I., Haddadi, A., Leblon, B., Wang, J., Wang, K., 2021. Comparison between three registration methods in the case of non-georeferenced close-range multispectral images. *Int. Geosci. Remote Sens. Symp. IGARSS*. https://doi.org/10.1109/IGARSS47720.2021.9554517

Guo, H., Xu, H., Wei, Y., Shen, Y., Rui, X., 2022. Outlier removal and feature point pairs optimization for piecewise linear transformation in the co-registration of very high-resolution optical remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* 193, 299–313. https://doi.org/10.1016/j.isprsjprs.2022.09.008

Hassanpour, M., Dadras Javan, F., Azizi, A., 2019. Band to band registration of multi-spectral aerial imagery – relief displacement and miss-registration error. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII-4/W18, 467–474. https://doi.org/10.5194/isprs-archives-XLII-4-W18-467-2019

Jhan, J.P., Rau, J.Y., Haala, N., Cramer, M., 2017. Investigation of parallax issues for multi-lens multispectral camera band co-registration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch*. https://doi.org/10.5194/isprs-archives-XLII-2-W6-157-2017

Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J., 2010. elastix: A Toolbox for Intensity-Based Medical Image Registration. I*EEE Trans. Med. Imaging*, 29, 196–205. https://doi.org/10.1109/TMI.2009.2035616

Kraus, K., 2011. *Photogrammetry*. De Gruyter.

Micasense, 2024. MicaSense [WWW Document]. URL https://micasense.com/

Parrot, 2022. Parrot Sequoia [WWW Document]. Parrot. URL https://www.parrot.com/uk/shop/accessories-spare-parts/other-drones/sequoia (accessed 3.24.22).

Paul, S., Pati, U.C., 2021. A comprehensive review on remote sensing image registration. *Int. J. Remote Sens*. 42, 5396–5432. https://doi.org/10.1080/01431161.2021.1906985

Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S., 2016. 3-D Imaging Systems for Agricultural Applications - A Review. *Sensors*, 16, 618. https://doi.org/10.3390/s16050618

Zitová, B., Flusser, J., 2003. Image registration methods: a survey. *Image Vis. Comput.* 21, 977–1000. https://doi.org/10.1016/S0262-8856(03)00137-9