Accuracy of Reconstruction with Short baseline from a Single-frame Multispectral Camera

Beatriz Coêlho Silva [1], Antonio Maria Garcia Tommaselli [1], Rahuan Miguel da Silva [1],
Leticia Ferrari Castanheiro [2], Wimerson Sanches Bazan [1,3]

[1] São Paulo State University (UNESP), Presidente Prudente, São Paulo, Brazil,
{beatriz.coelho-silva, a.tommaselli, rahuan.miguel, wimerson.bazan}@unesp.br
[2] Embrapa Digital Agriculture, Campinas, São Paulo, Brazil, leticia.castanheiro@colaborador.embrapa.com
[3] Espírito Santo Federal Institute, Vitória, Espírito Santo, Brazil

Keywords: Multispectral Camera, Close-range Photogrammetry, Image Registration, 3D Reconstruction.

Abstract

Close-range acquisition of multispectral images is becoming widespread in many applications, such as forestry and agriculture monitoring. However, some challenges still need to be addressed, such as the co-registration of the spectral bands generated with multi-lens cameras. Depth variations can introduce differential parallax effects depending on the lens position, which makes the coregistration more challenging. Therefore, a rigorous registration of this multispectral imagery requires pixel-wise parallax compensation, based on a depth model. This work presents a detailed evaluation of the accuracy of depth estimation with images acquired from a single-frame camera with a very short baseline. The methodology involves accurate camera calibration and three-dimensional reconstruction using Agisoft Metashape software. Field validation was conducted using a high-resolution laser scanner to produce reference data. Results show that even a single image frame, composed of six sub-images from different lenses, enables 3D reconstruction with reasonable (centimetre) accuracy, while a full image set achieves sub-centimetre accuracy. Comparison with a laser scanner reference point cloud showed the spatial consistency of the reconstructed surface. The results demonstrate the viability of reconstruction with short baseline images, aiming for pixel-wise registration.

1. Introduction

Multispectral images have been used to solve many problems in different applications. The availability of many lightweight and affordable cameras has made it possible to obtain high-resolution images, suitable for applications such as crop monitoring. However, close-range acquisition, poses different challenges for multispectral multi lens cameras; one of these issues is the registration of the different image bands, since the different viewpoints can introduce differential parallaxes, depending on the objects' depths, as it can be seen in Figure 1.a, which shows how the use of images taken from different viewpoints leads to the introduction of a differential parallax between the images of the same object.. The distance between the sensor and the object directly influences the observed parallax (Wolf and Dewitt, 2000), as shown in Figure 1.b, where point A has a larger parallax than point B because it is closer to the sensor.

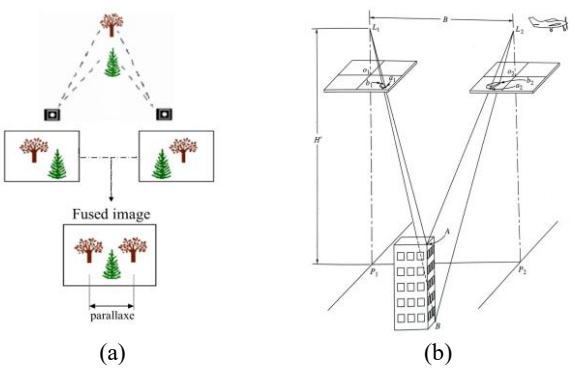


Figure 1. Parallax effect: (a) Observing objects from different viewpoints (Adapted from Raposo, 2004); (b) Photographs from two exposure stations with a building in a common overlap area (From Wolf, 2020).

Several solutions have been proposed to address the multispectral image registration problem, including patch-based registration (Hassanpour et al., 2019), distance-dependent transformation matrix (Dandrifosse et al., 2021), local image deformations (Guo et al., 2022), and differential registration (Tommaselli et al., 2025). However, these methods still face some limitations.

Despite considerable progress, obtaining robust pixel-level registration in multispectral near-field scenarios with significant depth variations remains an open challenge. Patch-based approaches (Hassanpour et al., 2019) were introduced to mitigate local relief displacement but are inherently limited by window size selection and the non-uniform distribution of corresponding points in areas with high depth variability. In response to the need for greater efficiency, distance-dependent transformations (Dandrifosse et al., 2021) seek to provide a near real-time solution but assume a simplified scene geometry and are insufficient to deal with the complex local deformations induced by parallax in objects distributed over different depth ranges. To overcome this rigidity, piecewise local deformation techniques (Guo et al., 2022) explicitly model non-linear distortions, offering greater flexibility.

This complexity increases the computational cost and imposes severe requirements on the accuracy of the feature points and the structure of the local triangular network. Finally, recognising that 2D transformations (even local ones) do not rigorously correct for depth-dependent displacements, the differential registration approach (using orthorectification) (Tommaselli et al., 2025) proposes a more rigorous solution for pixel-by-pixel alignment. However, its effectiveness depends critically on rigorous camera calibration and the quality of the Digital Surface Model (DSM) generated. Despite these advances, ensuring pixel-to-pixel alignment under the effect of significant near-field parallax remains a challenge, which motivates the present work.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-327-2026

This complexity increases the computational cost and imposes severe requirements on the accuracy of the feature points and the structure of the local triangular network. Since 2D transformations (even local ones) do not correct depth-dependent displacements, the differential registration approach using orthorectification proposed by Tommaselli et al. (2025) offers a more robust solution for pixel-by-pixel alignment. However, a short baseline results in a low base/range ratio, which can lead to high errors in the estimated depth due to this poor geometry, making the construction of a DSM complicated.

However, its effectiveness depends critically on rigorous camera calibration and the quality of the Digital Surface Model (DSM) generated. Despite these advances, ensuring pixel-to-pixel alignment under the effect of significant near-field parallax remains a challenge, which motivates the present work.

Aiming to cope with the limitations presented in Tommaselli et al. (2025), in this work, we propose the reconstruction of a single scene, composed of leaves and control targets, using sub-images from a single frame (a single station) of a multispectral camera with multiple lenses. A bundle adjustment of images from different lenses is used to generate a DSM as proposed by Tommaselli et al. (2025). Then, a point cloud using single-station and short-baseline multispectral images is generated. This generated multiespectral point cloud is compared with a point cloud obtained with a high-resolution FARO Focus terrestrial laser scanning (TLS) system.

2. Methodology

A multi-lens multispectral camera (Agrowing, 2024) was used to acquire the images, which were processed in Agisoft Metashape (Agisoft, 2023). For ground reference, point clouds were collected with the Faro Focus Premium laser scanner (FARO Technologies Inc., 2022). The point clouds generated in this work were visualised and compared using Cloud Compare software (Girardeau-Montaut, 2020).

The main procedure comprises four main steps: determination of a suitable principal distance, calibration of the Interior Orientation Parameters (IOPs) and Exterior Orientation Parameters (EOPs), determination of the Relative Orientation Parameters (ROPs) between the lenses, three-dimensional reconstruction using images from a single frame. The next sections present the details of the sensors and software used, as well as the proposed approach.

2.1 The multispectral sensor

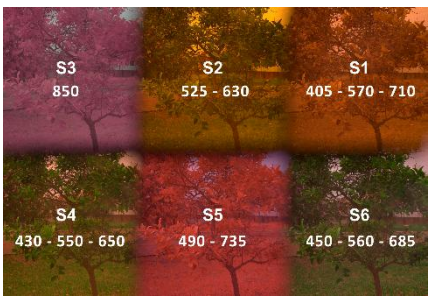
The optical sensor used in this work is an Agrowing Sony Alpha 7RIV Sextuple Multispectral camera (Figure 2.a) (Agrowing, 2024), which is a combination of a commercial digital camera with a special optical mounting with six lenses and filters that allow collecting image bands centred at the following wavelengths: close to Blue (405, 430, 450 and 490 nm), Green (525, 550, 560 and 570 nm), Red (630, 650 and 685 nm), *Red edge* (710 nm) and Near infrared (735 and 850 nm), each of which is assigned to one of the six lenses, as shown in Figure 2.b. The specifications of the camera are shown in Table 1.

The camera optics projects six bundles over the single CMOS sensor, and with further processing, 14 spectral image bands can be extracted, which are summarised in Figure 2.b. In this work, the processing chain is performed using a single frame from which six sub-images are extracted, corresponding to the centre bands of each sub-images.

Characteristics	Camera specifications
Field of view	45.90°
Lens distortion	< 1%
Focal length	21.8 mm
fstop	f/5.6 (fixed)
Sensor weight (with battery)	845 g
Full frame image size	9504x6336 pixels
Pixel size	0.0037 mm

Table 1. Specifications of the Agrowing Alpha 7RIV Sextuple Multispectral camera.

(a)



(b)

Figure 2. Multispectral sensor: (a) Front view of the Sony camera and attached Agrowing lenses (From AGROWING, 2024); (b) arrangement of spectral bands over the full frame.

2.2 Data acquisition with multispectral sensor

Due to changes made by the manufacturer to the optical assembly, the most suitable focal length for the depth of the object cannot be set automatically on this camera and must therefore be adjusted manually. For short-distance shooting, visual focusing or focusing based on the distance scale markings on the lenses is not accurate enough to provide focused images. A study was conducted to define the best focus setting by testing different ranges based on a millimetre scale fixed to the lens focus ring. The best focus setting identified as optimal for 2 m was used in all experiments conducted for this distance.

Data collection in close-range photogrammetry can be done on a mobile platform, a tripod, or manually by the operator. For this experiment, the images were captured by an operator holding the camera around a Myrtle shrub (Figure 4.a), over which coded targets were also placed, both on the Myrtle and on tripods, for better control of the analysis, as shown in Figure 4.b. In addition, a calibration plate (Figure 4.c) was placed next to the Myrtle tree to enable an accurate camera calibration.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-327-2026 | © Author(s) 2026. CC BY 4.0 License.

(a) (b) (c)

Figure 4. Data acquisition: (a) details of the field layout, with the bust and targets; (b) main scene detail; (c) calibration plate.

The first step in the acquisition mission was to collect a set of images aiming at camera calibration. This acquisition can be carried out in the same environment as the field work, to guarantee consistent parameters, or on previous days, if the stability of the camera is ensured. Camera calibration is performed with a self-calibrating bundle adjustment. At least seven constraints are required to provide scale and a reference system, which is usually done by fixing the coordinates of control points.

Using a calibration plate with coded targets can be an effective and practical solution, as shown in Figure 4.c. This calibration plate consists of 36 circular coded targets spaced 10 cm apart on a flat metal surface. The panel can be installed close to the object; then, several oblique images are collected at different locations and orientations. To ensure accurate results, the camera must be calibrated with the selected focusing distance locked and the calibration panel maintained in a rigorously stable position.

The second step in this acquisition stage was to collect images of the object of interest (in this case, the Myrtle tree), maintaining an average distance of 2 m from the object and using the same camera settings. At this stage, the operator can choose to take images of the entire object of interest. However, due to the multi-lens arrangement, it is also possible to reconstruct part of the object with just one frame.

2.3 Laser scanning system: Ground reference data

To generate accurate ground reference data, a third acquisition was conducted in this experiment, consisting of a laser scan of the scene with a Terrestrial Laser Scanner (TLS). In this work, we used the FARO Focus Premium 70 TLS (Figure 3, FARO Technologies Inc., 2022), which provides higher spatial resolution, thus being suitable to collect point clouds to be used as ground reference data.

Figure 3. Laser scanner: Faro Focus Premium 70. Adapted from FARO User.

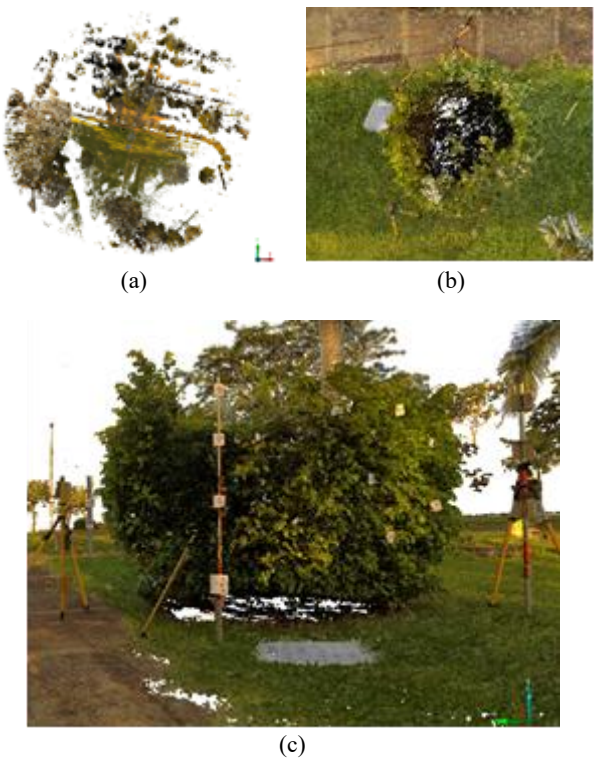
FARO Focus Premium 70 TLS has a field of view of 300° x 360° and operates at an infrared wavelength of 1553 nm. This TLS was designed to produce data with millimetre precision, and includes further features such as colour capture, automatic registration of scans and integration with processing software, allowing for an efficient and simplified workflow. Table 2 shows the technical information of FARO Focus Premium TLS.

Characteristics	Laser scanner specifications
Wavelength	1553.5 nm
Max. Range	350 m (White targets) 150 m (Black targets)
Pulse power	Approx. 800 mW
Beam divergence	0.3 mrad (0.024°) (1/e)
Beam exit diameter	2.12 mm (1/e)
Field of view	360° × 300°

Table 2. LiDAR specifications. Source: Adapted from FARO User Manual (2022).

For data collection, the TLS was configured to collect up to 1,000,000 points per second (1/2 of the maximum possible), with a 2x quality setting to store the point coordinates.

The laser scanner was placed 3 m apart from the Myrtle tree, at 4 well-distributed stations. Automatic co-registration of the point clouds was performed using the equipment's proprietary software (FARO Technologies Inc., 2022), generating a three-dimensional point cloud, as shown in Figure 5.a. In Figures 5.b and 5.c, sections of the point cloud generated by the laser scanner are shown for the area of interest. In Figure 5.a, it is possible to see the entire area in which the Myrtle tree is located; Figure 5.b shows a top view of the point cloud, and in Figure 5.c, a frontal view of the scanned Myrtle tree. This frontal view is the same as the one from which the images were collected using the Agrowing Sextuple Multispectral camera.



(a) (b)

(c)

Figure 5. 3D reconstruction by laser scanning: (a) complete point cloud in top view; point cloud of the area of interest (b) in top view and (c) in frontal view.

https://doi.org/10.5194/isprs-annals-X-3-W4-2025-327-2026

2.4 Camera Calibration and Determination of the Relative Orientation

Calibration can be carried out using various techniques and software. In this work, Agisoft Metashape software was initially used to automatically extract the image coordinates of the coded targets and other key points in all collected images. A total of 21 image frames were collected, and the corresponding 14 bands were approximately registered and extracted with the Agrowing software.

For the experiments presented in this article, of the 36 targets present on the calibration plate (Figure 4.c), 24 were used as ground control points (GCP) and 12 were left as check points (CP). Additional tie points (approximately 10,000 for each calibration set) were also extracted by Metashape and used in the camera calibration process.

The interior and exterior orientation parameters (IOPs and EOPs) for each of the six lenses were estimated using self-calibrating bundle adjustment, with the calibration plate serving as a reference system. Only one image band for each lens was selected from the images collected at each image station, resulting in six calibration sets.

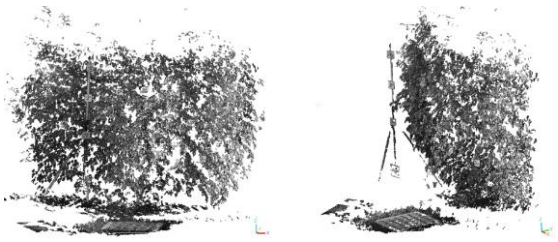
After determining the IOPs and EOPs with the calibration configuration, using the calibration plate as a reference, the ROPs of each lens were estimated in relation to the lens chosen as the reference (S6). The ROPs were estimated as the average values of the parameters estimated for each pair. For the Euler angles, however, the average does not make sense, since the angles are dependent; thus, the set of angles that are closest to the average values was chosen.

The calibration process, typically carried out at the beginning of the session, involves acquiring numerous images from various stations to obtain a suitable geometry. However, for the remaining stages of pixel-wise registration, each frame with the 6 displaced sub-images must be treated independently, but using the previously determined IOPs and ROPs. In a typical agricultural real field operation, the acquisition of a complete set of convergent images and the related processing for each tree would be unfeasible, since some farms have thousands of trees. Thus, in real-world scenarios, pixel-wise band co-registration should utilise a depth map generated from a single frame to produce registered multispectral images.

2.5 3D reconstruction of the object of interest with a complete set of images

To enhance the analysis of this work, a 3D point cloud was generated (using Agisoft Metashape software) using the entire set of images collected with the Agrowing Sextuple Multispectral camera, enabling a comparison of the results obtained by the three methods of reconstruction: 1. Laser scanning (Figure 5.c), 2. Reconstruction with a full set of convergent images and 3. Reconstruction from a single frame with six sub-images.

To generate this point cloud, 116 frames were used, including images of the calibration plate and images acquired around the Myrtle tree. This processing was also carried out in Metashape, and the resulting point cloud (BBA – point cloud from the Bundle Block Adjustment) for the 550 nm spectral band is shown in Figure 6.



(a) (b)

Figure 6. 3D reconstruction from optical images, complete point cloud: (a) in frontal view; (b) at side view.

2.6 3D reconstruction of the object of interest with a single frame

Reconstructing the object from a single frame with six sub-images is only possible in this case because this sensor has 6 lenses slightly displaced, as shown in the previous sections. These displacements will not provide a large base to distance ratio, but considering the simultaneous use of six images from 6 viewpoints, a point cloud can be generated from multiple correspondences of key points found in these sub-images.

In this work, the Bundle adjustment technique implemented in Metashape was used. For each station to be analysed individually, the sub-images of the image are specified as images from “different sensors” (Figure 7). However, in the data configuration part, the IOPs, previously calculated were inserted and considered as fixed constraints while the relative orientation parameters (ROPs) are inserted in the “slave offset” tab, setting the adjust location option and the adjust rotation option, so that in the three-dimensional model generation stage, the value would be taken as weighted constraints. After configuration, the “normal” photo alignment processing is applied, and the results obtained are a point cloud of the scene.

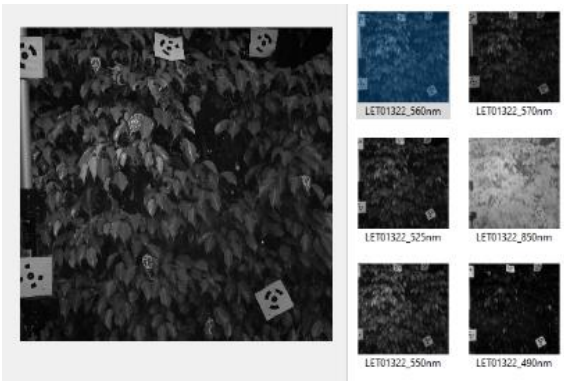


Figure 7. Image chosen for the experiment, the 6 sub-images corresponding to each lens.

3. Results and analysis

3.1 Relative orientation (ROPs) results

The process of estimating the ROPs between the perspective centres of the six sub-images depends on estimating the EOPs, which is why the first results presented in this section refer to calibration processing discussed in Section 2.4. After the processing performed for each of these calibrations, reports were generated, from which the RMSE values of the GCPs and CPs were extracted.

https://doi.org/10.5194/isprs-annals-X-3-W4-2025-327-2026

From these values, the mean for each component (X, Y, Z) and the resultant were calculated, as shown in Table 3. Analysis of the RMSE of the discrepancies in the Check points after the bundle adjustment with self-calibration, shows that the planimetric coordinates have millimetric accuracy. For the altimetry, the accuracy was slightly lower (approximately 3 mm), which can be attributed to a small deformation of the calibration plate.

	Average RMSE			
	RMSE_X	RMSE_Y	RMSE_Z	Total
24 GCPs	0.66 mm	0.78 mm	2.65 mm	2.84 mm
12 CkPs	0.82 mm	0.83 mm	2.77 mm	3.01 mm

Table 3. Average RMSEs values on control and check points in X, Y and Z coordinates, and in the resultant, using the six calibration datasets.

After determining the IOPs and EOPs with the calibration configuration, using the calibration plate as a reference, the ROPs of each lens were estimated in relation to the lens chosen as a reference (S6). The results of the calculated ROPs considering the entire set of images used for calibration are shown in Table 4. The ROPs were estimated as the average values of the parameters estimated for each pair, and for the Euler angles, the set of angles that are closest to the average values was chosen.

	Bx (mm)	By (mm)	Bz (mm)	Δω	Δφ	Δκ
S1	-0.04	-11.90	0.30	0°4'35.57"	0°0'02.01"	0°0'12.83"
	±0.47	±0.29	±0.13	±0°0'42.40"	±0°0'1.38"	±0°0'20.45"
S2	12.01	-12.35	-1.02	0°5'52.66"	0°0'01.38"	0°0'17.22"
	±0.36	±0.43	±0.17	±0° 1' 4.71"	±0°0'1.06"	±0°0'33.04"
S3	24.98	-11.55	-1.27	0°4'44.50"	0°0'01.29"	0°0'44.15"
	±0.69	±0.56	±0.23	±0°1'35.83"	±0°0'1.83"	±0°0'27.41"
S4	24.25	-00.41	-1.28	0°1'25.81"	0°0'01.41"	0°0'1.703"
	±0.49	±0.44	±0.26	±0° 1' 2.24"	±0° 0' 1.15"	±0°0'14.63"
S5	12.69	-00.68	-0.80	0°1'19.77"	0°0'00.90"	0°0'16.47"
	±0.50	±0.56	±0.19	±0°1'21.97"	±0°0'1.25"	±0°0'12.47"

Table 4. Average relative orientation parameters, with respective standard deviations, taking lens 6 as reference.

The analysis of Table 4 reveals that the displacements among the lenses were accurately determined, since the nominal displacements are approximately 0, 12 or 25 mm, depending on the lens set. Small variations in the estimated principal distances, which also reflected in the Bz component, were expected due to different lens' focus at various wavelengths and the errors in the EOPs determination. The X axis rotation (ω) showed the highest alignment values, posing uncertainties on whether these indicate misalignments or calibration issues. Resolving these discrepancies and improving alignment are suggested as future research areas for further exploration and enhancement.

3.2 Visual comparison of clouds from a frame with clouds from the complete set and from the laser

In addition to the full block reconstruction presented in subsection 2.5, we also tested the reconstruction using sub-images from individual stations. Figures 8.a and 8.c show the point clouds generated from single frames (F1 and F2, respectively), while Figures 8.b and 8.d present the corresponding results obtained with the Bundle Block Adjustment (BBA) with the full set of convergent images. Here, F1 and F2 refer to two independent acquisition stations of the Myrtle tree with a single frame, and BBA denotes the reconstruction using the complete image block. It can be observed that, although the single-frame approach is capable of reconstructing most of the objects (leaves, branches, and calibration targets), the point clouds are less dense and include more spurious points. In contrast, the BBA reconstruction with the full set of images yields denser and more homogeneous results, with finer structural details and reduced noise.

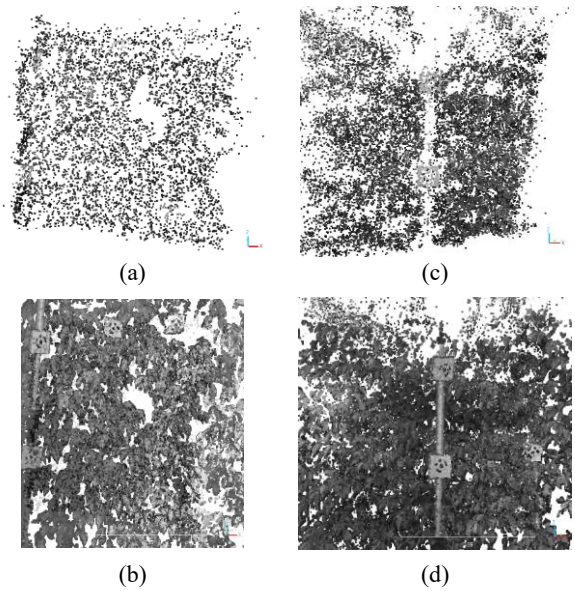


Figure 8. 3D reconstruction of a part of the Myrtle tree: station 1 generated by (a) a single frame (F1) and (b) full image dataset (BBA); station 2 generated by (c) a single frame (F2) and (d) full image dataset (BBA).

To better highlight the differences in the projection of the points belonging to the objects in the scene (leaves, branches, and targets), coloured visualisations of the four generated point clouds are presented in Figure 9. A front view of the Myrtle tree is shown in Figure 9.a, with the point cloud generated by the terrestrial laser scanner (TLS) (in light green), the point cloud from bundle block adjustment (BBA) (in dark green), the point cloud from frame 1 (F1) (in orange), and the point cloud from frame 2 (F2) (in pink). In addition, cross sections of parts of this scene were generated, with the complete profile shown in Figure 9.b, a profile focused only on the Myrtle shown in Figure 9.c, and a profile focused only on the targets shown in Figure 9.d.

As shown in Figure 9.b, the point cloud generated from a single image shows a greater dispersion, but even though the reconstruction generated a consistent point cloud, especially because, in this case, only six sub-images with very short baselines were used.

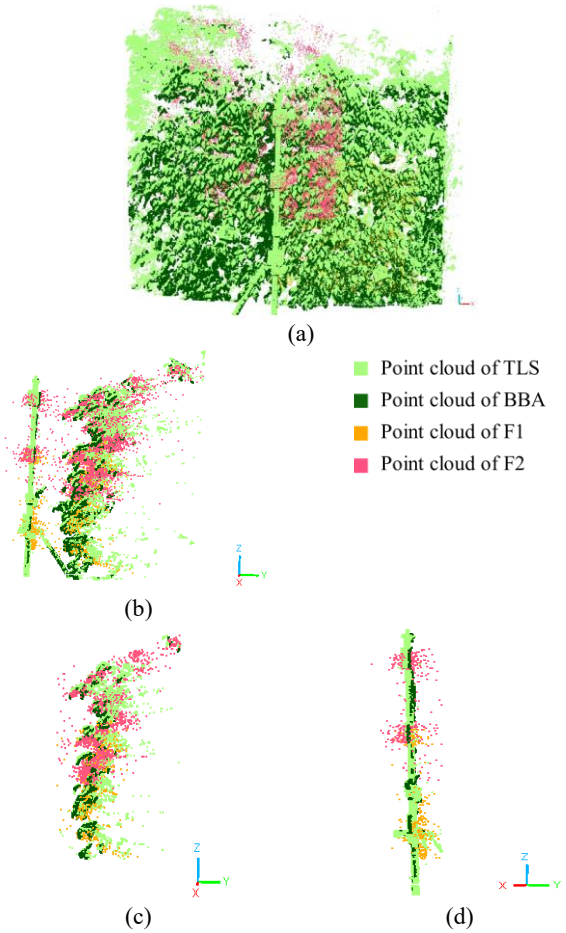


Figure 9. Comparing the 3D reconstructions for all clouds generated: (a) complete cloud, fusing the points from the three techniques; (b) general profile; (c) profile of the Myrtle tree; (d) profile of the pole with targets

3.3 Accuracy of point clouds from a single frame

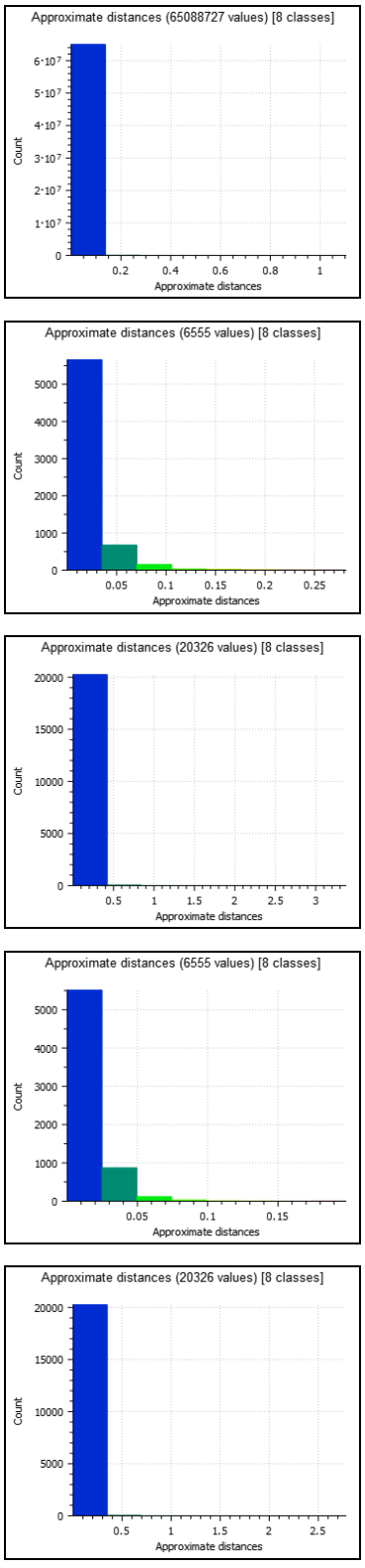
For quantitative analysis, some homologous points were manually measured in the four point clouds using the CloudCompare software, allowing for the estimation of errors in the depths of these specific points (Table 5), taking the point cloud generated with BBA of the full images dataset as a reference.

Errors in point depth		
Point	Point Cloud	Depth error
1	F1	12.01 mm
	F2	51.13 mm
2	F1	8,86 mm
	F2	101,42 mm

Table 5. Errors in two points in the point clouds generated by a single frame, using the point cloud from the complete set as a reference.

In addition, the Cloud Compare tool was used to calculate the distance between points in the point clouds, for each pair. In this tool, it is possible to select two point clouds to be compared and assign which one will be the reference. The algorithm calculated the distances between the point clouds, and the result was an average distance, the standard deviation, and a histogram.

Figure 10 shows the histograms, average distance, and standard deviation for each comparison made, taken as reference either LiDAR or the one generated with the complete set of optical images.



(a) TLS and BBA
mean_dist = 7 mm
std = 12 mm

(b) TLS and F1
mean_dist = 16 mm;
std = 14 mm

(c) TLS and F2
mean_dist = 36 mm;
std = 68 mm

(d) BBA and F1
mean_dist = 20 mm;
std = 22 mm

(e) BBA and F2
mean_dist = 50 mm;
std = 66 mm

Figure 10. Histograms and statistics of the distance between the clouds.

The accuracy of the point cloud generated with BBA (the complete image dataset), is approximately 1 cm, which is more accurate than the other two point clouds generated from individual frames. However, it should be noted that in the reconstruction with a single frame, the image geometry is much poorer, since the sub-images have a very short base.

The point cloud generated with F1 (Figures 10.b and 10.d) presented better results when compared with the point cloud generated with F2 (Figures 10.c and 10.e). However, these results were also dependent on the registration performed between the point clouds, to ensure they were in the same reference frame. For some points, this may have worsened the result.

4. Conclusions

The use of multispectral cameras with multiple lenses has become usual in ground-based applications, especially on mobile platforms and drones for agricultural monitoring. In these applications, the quality of spectral analysis is directly related to the accuracy of the alignment between the bands captured by different lenses, which are displaced from each other. Despite some advances, accurate registration of images acquired at short distances remains a challenge, especially in scenes with huge depth variations, where parallaxes between bands become more significant. In this work, it was emphasised the assessment of the point cloud generated with short baseline images from the same frame/station.

The registration of spectral bands acquired with multi-lens multispectral cameras at close-range requires a strategy to compensate for the effect of differential parallax. This effect results from the depth variations, and it is increased in close-range acquisition, making it essential to apply a technique to correct these effects in order to generate registered spectral bands for spectral analysis.

This paper presents a rigorous methodology for reconstructing 3D scenes using a single frame. By combining precise camera calibration, processing in Agisoft Metashape software and field validation with a high-resolution laser scanner, we demonstrate that even an image composed of six slightly displaced sub-images can produce a reasonably accurate point cloud.

The proposed methodology, based on rigorous camera calibration using a calibration plate (Figure 4.c) with coded targets, can be an effective and practical solution. This calibration plate can be easily transported and inserted into the area to be assessed. This calibration has proved effective in generating three-dimensional models with multispectral images, both with a complete set of frames and from single images.

Reconstruction with only one frame per station, although less accurate (with an average error of less than 5 cm), showed promising results, allowing the partial reconstruction of nearby objects. Reconstruction with the complete set of images resulted in a point cloud with sub-centimetre accuracy, demonstrating the technique's potential for applications that require high geometric and spectral accuracy, but at the cost of more complex data acquisition and processing.

Comparison with reference data obtained by laser scanning (LiDAR) reinforced the reliability of the proposed approach by demonstrating spatial consistency between the different methods and allowing objective quantification of errors in the reconstructions. In addition, the strategy of configuring the sub-images as distinct sensors in the processing proved to be a practical and efficient solution for exploiting the geometry of multi-lens cameras.

However, the workflow remains technically complex, requiring manual focus adjustments and rigorous calibration, which can make widespread adoption difficult. Single-frame reconstructions also produce lower point density and exhibit greater dispersion.

Future work should focus on improving the processing to make it simpler and more automatic, increasing the efficiency of the technique, as well as investigating ways to mitigate small inaccuracies in alignment. Further steps, such as smoothing and enhancement of the density of the point clouds, are also suggested.

5. Acknowledgements

This study was financed by the São Paulo Research Foundation (FAPESP) – Brazil (grants n. 2021/06029-7, 2023/15204-2, 2025/01739-7, 2025/05985-2), by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil, CAPES (grant n. 88887.147492/2025-0) and by the National Council for Scientific and Technological Development, CNPq (grant n. 303670/2018-5).

References

Agisoft (2020) *Metashape Agisoft*. Available at: https://www.agisoft.com/.

Agrowing (2024) *Agrowing Sensors*, *AGROWING*. Available at: https://agrowing.com/.

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. *et al.* (2019) 'Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview', *Sensors*, 19(5), p. 1090. Available at: https://doi.org/10.3390/s19051090.

FARO Technologies Inc. (2022) Focus Premium Laser Scanner User Manual. Lake Mary, FL: FARO Technologies, Inc.

Girardeau-Montaut, D. (2020) *CloudCompare*. Available at: https://www.cloudcompare.org/main.html.

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 multispectral 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. *et al.* (2017) Investigation of parallax issues for multi-lens multispectral camera band co-registration, *The*

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W6, pp. 157–163. Available at: https://doi.org/10.5194/isprs-archives-XLII-2-W6-157-2017.

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

Raposo, A.B., Szenberg, F., Gattass, M., Celes, W., 2004. Visão Estereoscópica, Realidade Virtual, Realidade Aumentada e Colaboração. XXIII JAI–Jornada de Automatização em Informática, Capítulo 8.

Tommaselli, A.M.G., Silva, B.C., Bazan, W.S., Souza, A.Y.P., Porto, L.R., 2025. Pixel-wise registration of close-range multispectral images acquired with multi-lens cameras. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-G-2025, 1441–1446. https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-1441-2025

Wolf, P.R., Dewitt, B.A., 2000. Elements of photogrammetry: with applications in GIS, 3rd. ed. ed. McGraw-Hill, New York.