

# Two-step multi-spectral registration via key-point detector and gradient similarity

Jehan-Antoine VAYSSADE<sup>1</sup>, Christelle Gee<sup>1</sup>

<sup>a</sup>*Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France*

---

## Abstract

The potential of multi-spectral images is growing rapidly in precision agriculture, and is currently based on the use of multi-sensor cameras. However, their development usually concerns aerial application and their parameters are optimized to high altitudes per drone ( $\approx 50$  meters) to ensure surface coverage and reduce technical problems. With the recent emergence of terrestrial robots (UGV), their use is diverted for nearby agronomic applications. It is possible to explore new application maximizing specific traits extraction (spectral index, shape, texture ...) which requires high spatial resolution

The problem with these cameras is that all sensors are not aligned and the manufacturers' methods are not adapted for close-field acquisition, resulting in offsets between images and degrading the quality of extractable information. We therefore need a solution to align the different images accurately in such condition. The objective of this article is to propose a simple method that allows to define the ideal conditions for matching multi-spectral images from the multi-sensor camera at low heights of use applied on agronomic scenes.

In this study we propose a two-step method applied to the six-bands Airphen multi-sensors camera with (i) affine correction using pre-calibrated matrix at different heights, the closest transformation can be selected via internal GPS and (ii) perspective correction to refine the previous one, using key-points matching between enhanced gradients of each spectral bands. Nine types of key-point detection algorithms (ORB, GFTT, AGAST, FAST, AKAZE, KAZE, BRISK, SURF, MSER) with three different modalities of parameters was evaluated for their benchmark and their corresponding best reference spectra.

*Keywords:* Registration, Multi-spectral imagery, Precision farming, Feature descriptor

---

---

*Email addresses:* [jehan-antoine.vayssade@inra.fr](mailto:jehan-antoine.vayssade@inra.fr) (Jehan-Antoine VAYSSADE), [christelle.gee@agrosupdijon.fr](mailto:christelle.gee@agrosupdijon.fr) (Christelle Gee)

## 1. Introduction

Modern agriculture is changing towards a system that is less dependent on pesticides [15] (herbicides remain the most difficult pesticides to reduce) and digital tools are of great help in this matter. The development of imaging and image processing have made it possible to characterize an agricultural plot [24] (crop health status or soil characteristics) using non-destructive agronomic indices [3, 11, 13] replacing traditional destructive and time-consuming methods. In recent years, the arrival of miniaturized multi-spectral and hyper-spectral cameras on Unmanned Aerial Vehicles (UAVs) has allowed the plot spatio-temporal monitoring of the field. These vision systems have been developed for precise working conditions (flight height 50m). Although, very practical to use, they are also used for proxy-sensing applications. However, the algorithms for correcting and matching multiple single-band images of different spectral range, offered by manufacturers, are not optimal for low heights. It thus requires a specific close-field image registration.

Image registration is the process of transforming different images of one scene into the same coordinate system. The spatial relationships between these images can be rigid (translations and rotations), affine (shears for example), homographic, or complex large deformation models (due to the difference of depth between ground and leaves) [14]. The main difficulty is that multispectral cameras have low spectral coverage between bands, resulting in a loss of characteristics between them. Which is caused by (i) plant leaves have a different aspect depending on the spectral bands (ii) there are highly complex and self-similar structures in our images [10]. It therefore affects the process of detecting common characteristics between bands for image registration. There are two types of registration, feature based and intensity based [32]. (i) Feature based methods by extract point of interest and using feature matching, in most cases a brute-force matching is used, making those techniques slow. Fortunately these features can be filtered to reduce the matching cost depending of the spatial properties we have, and a GPGPU implementation can reduce the comparisons cost. (ii) Intensity-based automatic image registration is an iterative process, and the metrics used are sensitive to determine the numbers of iteration, making such method computationally expensive for precise registration. Furthermore multi-spectral need different metrics for each registered bands which is hard to achieve.

Different studies of images alignment using multi-sensors camera can be found for acquisition using UAV at medium and high distance. Some show good performances (in term of number of key-points) for feature based [8, 29] with strong enhancement of feature descriptor for matching performances. Other do not and prefer to use intensity based [10] on better convergence metrics [7] (in term of correlation), which is slower and not necessarily robust against light variability [30]. And their optimization can also fall into a local minimum.

Traditional approach to multi-spectral image recording is to designate one channel as the target channel and record all the others on the selected one. Currently, only [8] show a methods for selecting the best reference, but there is no studies who as defined the best spectral reference in agronomic scene. In all cases NIR (850nm) or the middle range spectral reference are conventionally used without studying the others on precision agriculture. In addition those studies mainly propose features matching without large methods comparison [8](less than 4) of their performance (time/precision), without showing the importance of the spectral reference and the normalized gradients transformation like in Intensity-based methods.

However, despite the growing use of UGVs and multi-spectral imaging, the domain is not very well sourced, and no studies have been found under agricultural and external conditions in near field of view (less than 10 meter) for multi-spectral registration. Thus, this study propose a benchmark of popular feature extractor inside normalized gradients transformation and the best spectral reference was defined for each of them. Moreover a pre-affine registration is used to filter the feature matching, evaluated at different spatial resolution. So this study shows the importance of the selection of the reference and the features extractor on normalized gradients in such registration.

## 2. Material and Method

### 2.1. Material

The multi-spectral imagery is provided by the six-band multi-spectral camera Airphen<sup>1</sup>. AIRPHEN is a scientific multi-spectral camera developed by agronomists for agricultural applications. It can be embedded in different types of platforms such as UAV, phenotyping robots, etc.

AIRPHEN is highly configurable (bands, fields of view), lightweight and compact. It can be operated wirelessly and combined with complementary thermal infrared channel and high resolution RGB cameras. The camera was configured using 450/570/675/710/730/850 nm with FWHM<sup>2</sup> of 10nm, the position of each bands is referenced on table 2.1. The focal lens is 8mm. The raw resolution for each spectral band is 1280x960 px with 12 bit of precision. Finally the camera also provides an internal GPS antenna, that can be used to get the distance from the ground.

450	675	710
570	850	730

Table 1: Disposition of each bands on the Airphen multi-sensors camera

Two datasets were taken at different heights. We have used a metallic gantry for positioning the camera at different heights. The size of the gantry is  $3m * 5m * 4m$ . Due to the size of the chessboard, the limited focus of the camera and the gantry height, we have bounded the acquisition height from  $1.6m$  to  $5m$  with  $20cm$  steeps, which represent 18 acquisitions.

The first dataset is for the calibration. A chessboard is taken at different heights, the corresponding data can be found in *data/step-chess/*. The second one is for the alignment verification under real conditions. One shot of an agronomic scene is taken at different heights, the corresponding data can be found in *data/step/* with a bias of  $\pm 10cm$  to be in the worst case (most complex).

### 2.2. Methods

Alignment is refined in two stages, with (i) affine registration approximately estimated and (ii) perspective registration for the refinement and precision. As example the figure 1 show each correction step, where the first line is for the (i) affine correction (section 2.2.1), the second and third is for (ii) perspective correction. More precisely the second are per-channel pre-processing where feature detectors are used to detect key points (section 2.2.2). And the last line each channel key points are associated to compute the perspective correction through homography, to the selected spectral band (section 2.2.2). Those steps are explained on specific subsections.

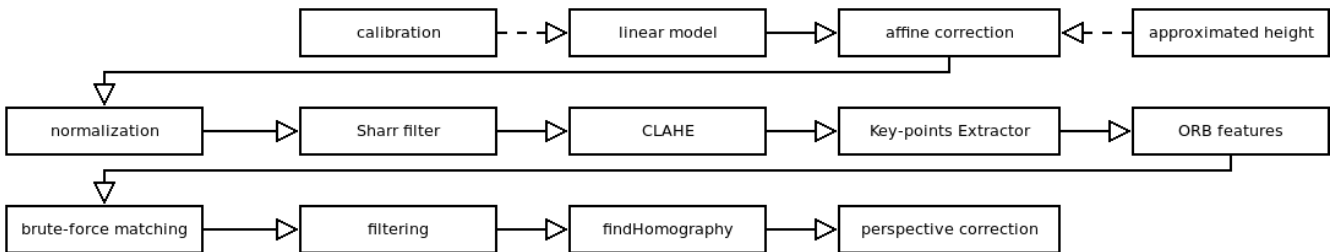


Figure 1: Each stages

#### 2.2.1. Affine Correction

It is important to notice the closer we take the snapshot, the bigger is the distance of the initial Affine Correction. On the other hand at a distance superior or equals to  $5m$ , the initial affine correction become stable (figure 4). Then a calibration is used to build a linear model, which makes the affine correction to work at any height. The main purpose of this step is to reduce the offset between each spectral band, which allows the similarity to be spatially delimited within a few pixels, making feature matching more effective.

<sup>1</sup><https://www.hiphen-plant.com/our-solutions/airphen/>

<sup>2</sup>Full Width at Half Maximum

*Calibration.* We use opencv 4 `findChessboardCorners` for each spectral image (normalized) at different heights (from 1.6m to 5m). The function attempts to determine whether the input image is a view of the chessboard pattern and locate the internal chessboard corners. The detected coordinates are roughly approximated, and to determine their positions accurately we use the function `cornerSubPix` as explained in the documentation <sup>3</sup>.

*Linear model.* Using all the points detected for each spectral band, we calculate the centroid grid (each point average). The affine transform from each spectral band to this centroid grid is estimated. It appear that the rotation and the scale ( $A, B, C, D$ ) do not depend on the distance to the ground, but only on the translation. On the other hand, the translation ( $X, Y$ ) depends on the height. Thus a Levenberg-Marquardt curve fitting algorithm with linear least squares regression [19] can be used to fit an equation for each spectral band again  $X$  and  $Y$  independently to the centroid grid. We have chosen to adjust the following curve  $t = \alpha h^3 + \beta h^2 + \theta h + \gamma$  where  $h$  is the height,  $t$  is the resulted translation and factors  $\alpha, \beta, \theta, \gamma$  are the model parameters.

*Correction.* To make the affine matrix correction, we used the rotation and scale factors at the most accurate height (1.6m where the spatial resolution of the chessboard is higher), because it does not depend on the height. For the translation part, the curve model is applied for each spectral bands at the given height, it can be provided by (i) roughly known by the user (ii) using the internal GPS sensor or (iii) estimating the height by detecting inter-row distance [6]. Each spectral band are warped using the corresponding affine transformation. Finally, all spectral bands are cropped to the minimal bounding box (minimal and maximal translation of each affine matrix).

### 2.2.2. Perspective correction

Each spectral band has different properties and values by nature (figure 2) but we can extract the corresponding similarity by transforming each spectral band into its absolute derivative, to find similarities in gradient break among them.

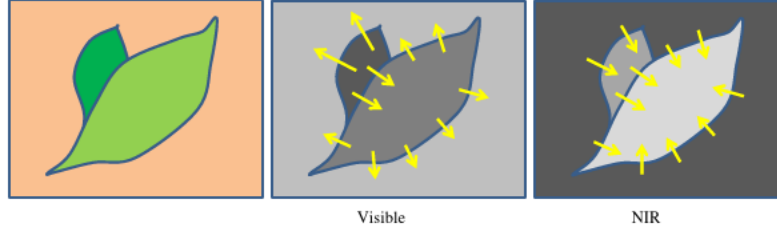


Figure 2: Gradient orientation in spectral band [21], we can see that the orientation of the gradient is not the same depending to the spectral band, making the absolute derivative an important step for matching between different spectral band.

The affine correction attempts to help the feature matching by adding properties of epipolar lines (close). Thus, the matching of extracted features can be spatially bounded, (i) we know that the maximum translation is limited to a distance of a few pixels (less than 10px), and (ii) the angle between the initial element and the matched one is limited to  $[-1, 1]$  degree.

*Computing the gradient.* To compute the gradient of the image with a minimal impact of the light distribution (shadow, reflectance, specular, ...), each spectral band is normalized using Gaussian blur [23], the kernel size is defined by  $\lfloor [image\_width^{0.4}]/2 \rfloor * 2 + 1$  (19 in our case) and the final normalized images are defined by  $I/(G+1)*255$  where  $I$  is the spectral band and  $G$  is the Gaussian blur of those spectral bands. This first step minimizes the impact of the noise on the gradient and smooths the signal in case of high reflectance. Using this normalized image, the gradient  $I_{grad}(x, y)$  is computed with the sum of absolute Sharr filter [25] for horizontal  $S_x$  and vertical  $S_y$  derivative, noted  $I_{grad}(x, y) = \frac{1}{2}|S_x| + \frac{1}{2}|S_y|$ . Finally, all gradients  $I_{grad}(x, y)$  are normalized using CLAHE [33] to locally improve their intensity and increase the number of key points detected.

*Key-points Extractor.* A key point is a point of interest. It defines what is important and distinctive in an image. Different types of key point extractors were available and the following more popular were tested : (ORB) Oriented FAST and Rotated BRIEF [22], (AKAZE) Fast explicit diffusion for accelerated features in nonlinear scale spaces [1], (KAZE) A novel multi-scale 2D feature detection and description algorithm in nonlinear scale spaces [20],

<sup>3</sup>[https://docs.opencv.org/2.4/modules/calib3d/doc/camera\\_calibration\\_and\\_3d\\_reconstruction.html](https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html)

(BRISK) Binary robust invariant scalable key-points [16], (AGAST) Adaptive and generic corner detection based on the accelerated segment test [18], (MSER) maximally stable extremal regions [9], (SURF) Speed-Up Robust Features [4], (FAST) FAST Algorithm for Corner Detection [28] and (GFTT) Good Features to Track [26]. These algorithms were largely summarized across multiple studies [8, 27, 31, 2], they are all available and easily usable in OpenCV, thus we have studied them by varying the most influential parameters for each of them with three modalities, the table 2 in appendix shows all modalities and methods. All the results can be found in “figures/\*”.

*Key-point detection.* We use one of the key-point extractors mentioned above between each spectral band gradients (all extractors are evaluated). For each detected key-point, we extract a descriptor using ORB features. We match all detected key-points to a reference spectral band (all bands are evaluated). All matches are filtered by distance, position and angle, to eliminate false positives along the epipolar line. Finally we use the function *findHomography* between the key points detected/filtered with RANSAC [12], to determine the best subset of matches to calculate the perspective correction.

*Correction.* The perspective correction between each spectral band to the reference is estimated and applied. Finally, all spectral bands are cropped to the minimum bounding box, the minimum and maximum points are obtained by applying a perspective transformation to each corner of the image.

### 3. Results and discussion

Firstly the results will focus on affine corrections and then on the effects of the perspective correction. Figure 3 shows a zoom inside unregistered image, registered image of each correction steps at 1.6m and the manufacturer results.

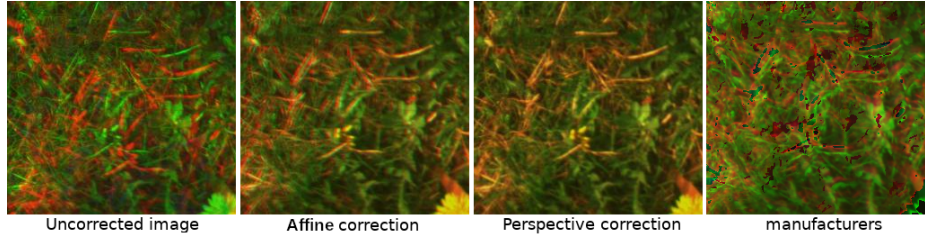


Figure 3: Example of each correction and the manufacturers results

#### 3.1. Affine correction

The 6 coefficients ( $A, B, C, D, X, Y$ ) of the affine matrix were studied according to the height of the camera in order to see their stability. It appears that the translation part ( $X, Y$ ), depends on the distance of the field (appendix figure 4). On this part the linear model is used to estimate the affine correction from an approximated height.

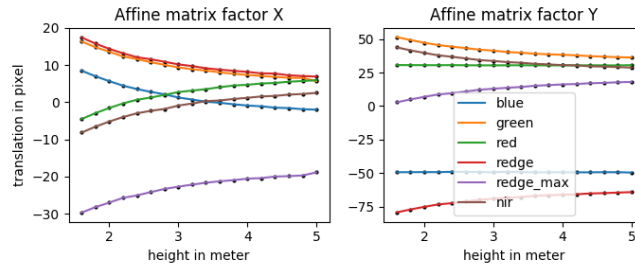


Figure 4: Affine matrix value by height

Rotation and scale do not depend on the ground distance (figure 5). These factors ( $A, B, C, D$ ) are quite stable and close to identity, as expected (accuracy depends on the spatial resolution of the board). As result, single calibration can be used for this part of the matrix, and the most accurate are used.

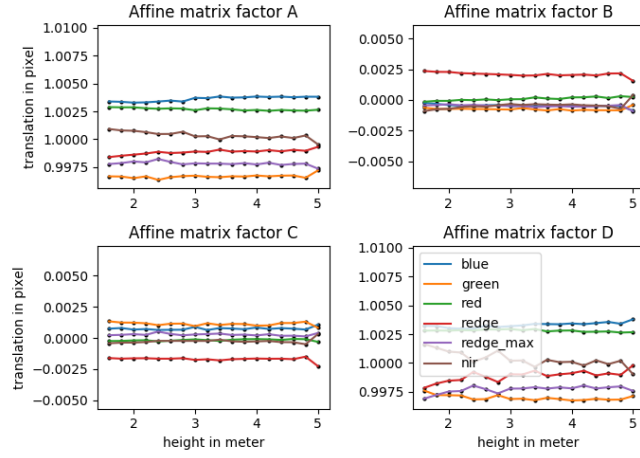


Figure 5: Affine matrix value by height

The remaining distance between each spectral band varies according to the distance between the real height and the nearest selected. These residual distances can be seen in the figure 6, remember that a bias of  $\pm 10cm$  was set, implying superior error.

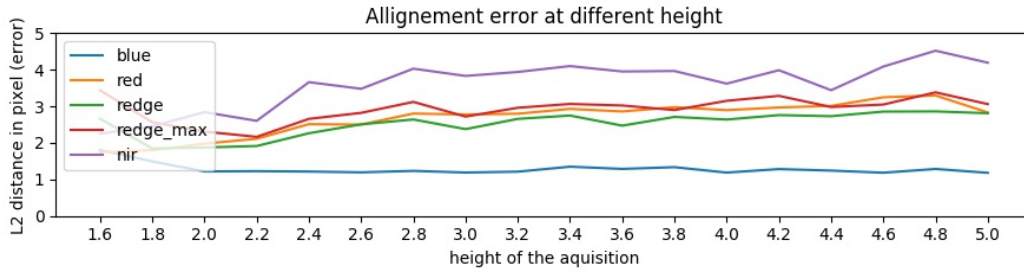


Figure 6: The mean distance of detected key-point before perspective correction with 570nm as spectral reference

### 3.2. Perspective correction

The figures 7 shows the numbers of key-points after filtering and homographic association (minimum of all matches) as well as the computation time and performance ratio (matches/time) for each method.

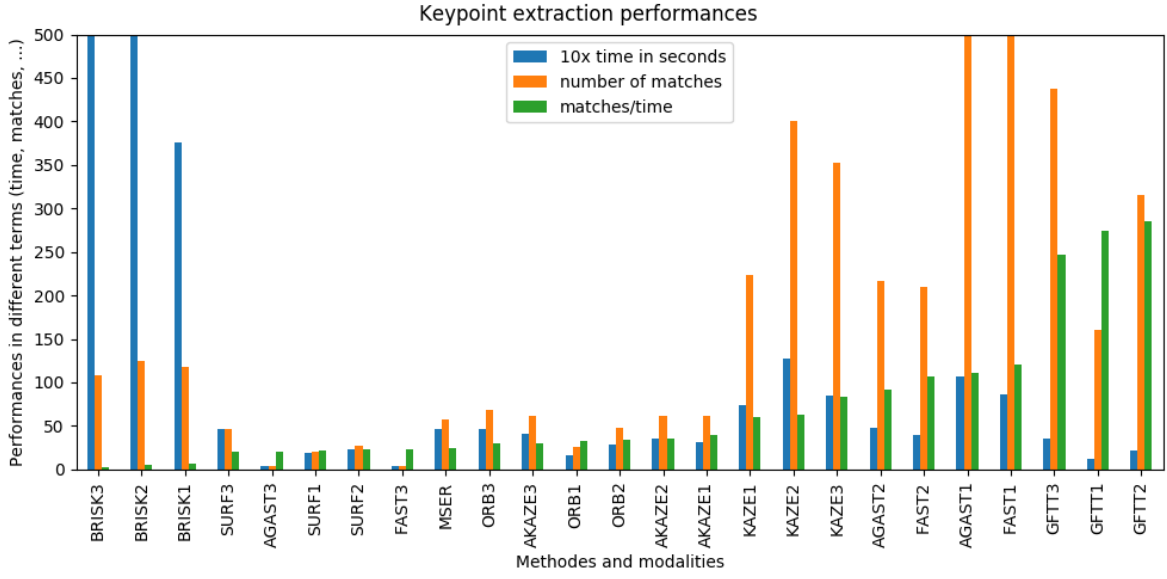


Figure 7: features extractor performances after filtering and homography association

All these methods offer interesting results, the choice of method depends to the wanted balance between computation time and accuracy:

- GFTT shows the best overall performance in both computation time and number of matches
- FAST and AGAST1 are quite suitable too, balanced in time and with greater matches performances.

The other ones did not show improvement in term of time or matches:

- KAZE show correct number of matches ( $>200$ ) but it's also 2.5 times slower than FAST or AGAST.
- AKAZE and MSER did not show benefits comparing to FAST.
- ORB and SURF could be excluded, the number of matches is too small to ensure the precision.
- BRISK shows good number of matches, at the cost of an important computation time (79 sec).

Increasing the number of key points matched allows a slightly higher accuracy [8? ]. For example, switching from SURF (30 results) to FAST (130 results) reduces the final residual distances from 1.2 to 0.9 pixel but increases the calculation time from  $\approx 5$  to  $\approx 8$  seconds.

All methods show that the best reference spectrum is  $710nm$ , with the exception of SURF and GFTT which is  $570nm$ . The figure 8 shows the minimum number of matches between each reference spectrum and all others using the GFTT algorithm. Choosing the right spectral reference is important, as we can see, no correspondence is found in some cases between  $650 - 850nm$ , but correspondences are found between  $675 - 710nm$  and  $710 - 850nm$ , making the  $710nm$  more appropriate, the same behavior can be observed for the other bands and  $570nm$ . This is visible on the "all\_min" cols,  $570nm$  and  $710nm$  have the best minimum number of matches. Other best spectral reference are available in supplementary material.



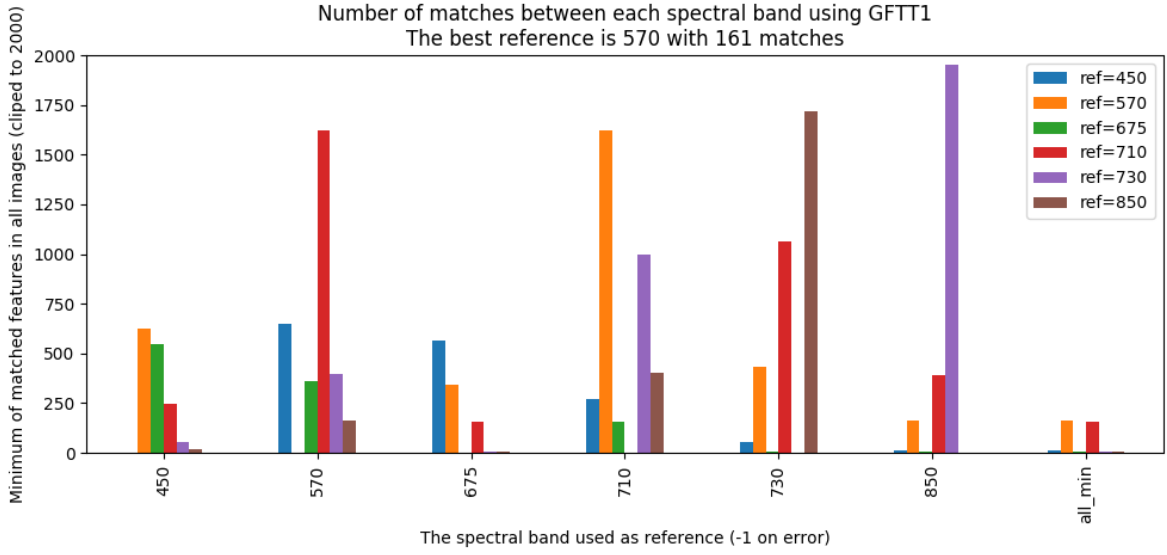


Figure 8: feature GFTT performances

Residues of the perspective correction show that we have correctly registered each spectral band with a residual error of less than 1 pixel, the figure 9 shows the residual distance at different ground distances.

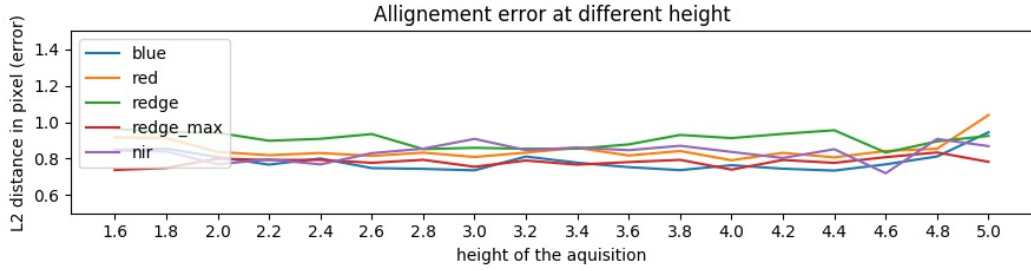


Figure 9: Perspective re-projection error with GFTT1 and 570nm as reference

The figure 10 shows the difference between detected points for two bands (red-green) before (left) and after (right) the perspective correction, and show that the residual errors are spatially uniform.

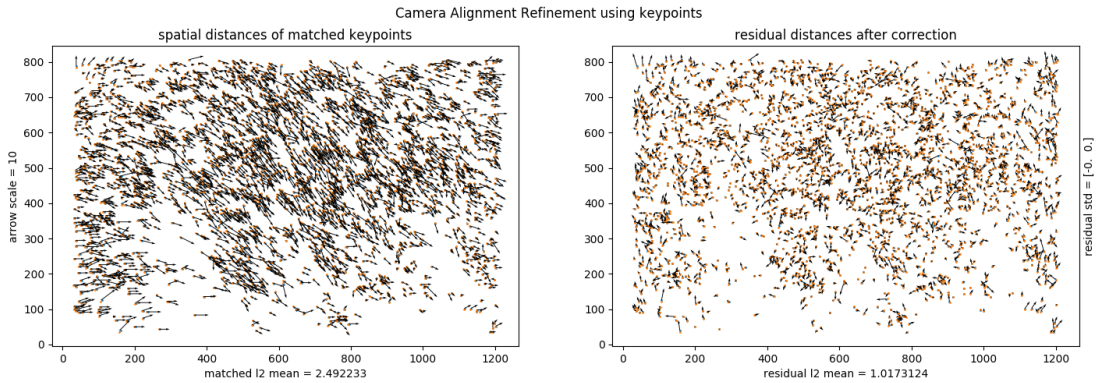


Figure 10: perspective-features-matching-scatter



The decomposition of the residual distances by angles, visible in the figure 11 is interesting.

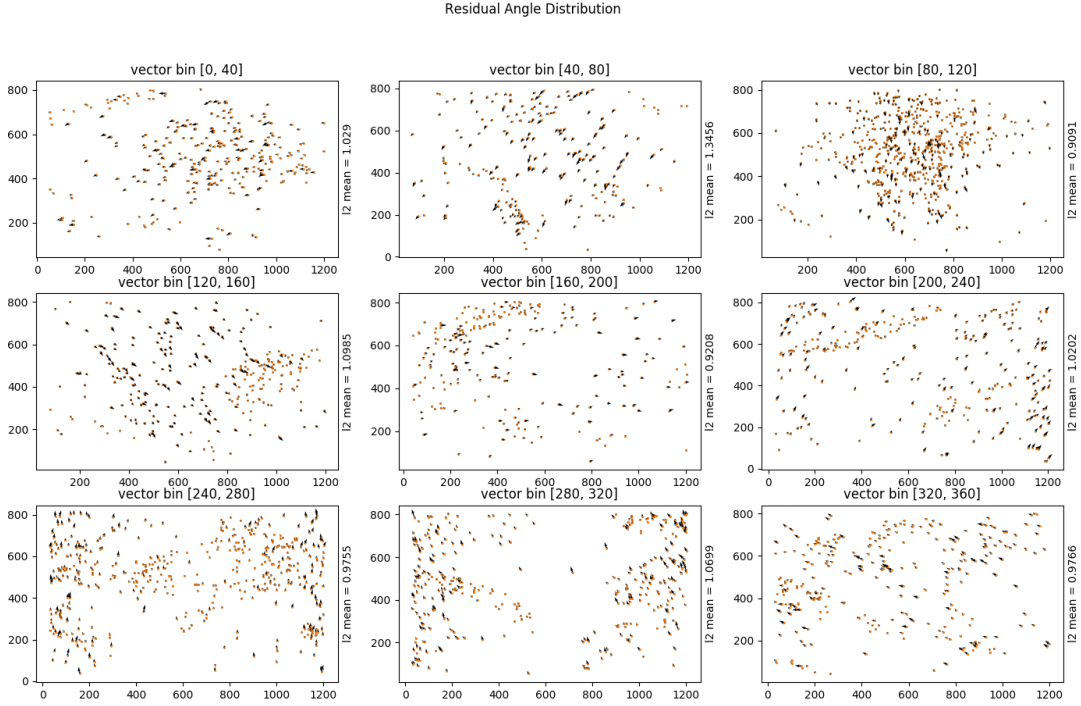


Figure 11: Residual Distribution Again Angle

We can notice that the spatial distribution of the residues, for each different angle, is equally distributed. Our hypothesis is that the nature of the base information (spectral band + different lens) makes a small difference to the gradient break, which is detected by the features detector and propagated until the final correction (observed residue). This is interesting because these residues uniformly distributed by angle in space tend to minimize the resulting correction of its center (gradient), thus the detected residual error are overrated and should be less than 0.4 pixel.

### 3.3. General discussion

Even if the relief of the scene is not taken into account due to the used deformation model, in our case, with flat ground, no difference arise. However, more complex deformation models [17, 5] can be used to improve the remaining error. But can also, in some cases, create large angular deformations caused by the proximity of key-points, of course, it's possible to filter these key-points, which also reduces the overall accuracy.

Further researches can be performed on each parameter of the feature extractors, for those who need specific performance (time/precision), We invite anyone to download the dataset and test various combinations. Otherwise feature matching can be optimized, at this stage, we use brute-force matching with post filtering, but a different implementation that fulfill our epipolar line properties should greatly enhance the number of matches by reducing false positives.

Finally, the method was tested on more than 8000 images in real conditions (not present in the study), randomly taken between 1.6 and 2.2 meters without registration error (always a minimum number of matches, without observed error, less than 0.9px). All were acquired manually using a wheelbarrow and the camera was mounted on a pole to scan the crops along the row in the field.

## 4. Conclusion

In this work, the application of different techniques for multi-spectral image registration was explored for Airphen camera. We have tested nine type of available OpenCV4 methods (ORB, GFTT, AGAST, FAST, AKAZE, KAZE, BRISK, SURF, MSER) for key-points extractors at different heights and the number of control points obtained. As seen in the method, the most suitable method is the GFTT (regardless of modalities 1, 2 or 3) with a significant number of matches and a reasonable calculation time.

Furthermore, the best spectral reference was defined for each method, for example 570 for GFTT. We have observed a residual error of less than 1 px, supposedly caused by the difference (spectral range, lens).

## 5. Acknowledgments

The authors acknowledge support from European Union through the project H2020 IWMPRAISE <sup>4</sup> (Integrated Weed Management: PRActical Implementation and Solutions for Europe) and from ANR Challenge ROSE through the project ROSEAU <sup>5</sup> (Robotics SEnsorimotor loops to weed AUtonomously).

We would like to thanks Jones Gawain, Combaluzier Quentin, Michon Nicolas, Savi Romain and Masson Jean-Benoit for the realization of the metallic gantry that help us positioning the camera at different heights.

## 6. Supplementary material

The additional data and source code associated with this article can be found in the online version at the following address [gitlab.com/phd-thesis-adventice/phd-airphen-alignment](https://gitlab.com/phd-thesis-adventice/phd-airphen-alignment) the access is limited, and we invite you to send an email to the author for full access.

## 7. Appendix

ABRV	1	2	3
ORB	nfeatures=5000	nfeatures=10000	nfeatures=15000
GFTT	maxCorners=5000	maxCorners=10000	maxCorners=15000
AGAST	threshold=71	threshold=92	threshold=163
FAST	threshold=71	threshold=92	threshold=163
AKAZE	nOctaves=1, nOctaveLayers=1	nOctaves=2, nOctaveLayers=1	nOctaves=2, nOctaveLayers=2
KAZE	nOctaves=4, nOctaveLayers=2	nOctaves=4, nOctaveLayers=4	nOctaves=2, nOctaveLayers=4
BRISK	nOctaves=0, patternScale=.1	nOctaves=1, patternScale=.1	nOctaves=2, patternScale=.1
SURF	nOctaves=1, nOctaveLayers=1	nOctaves=2, nOctaveLayers=1	nOctaves=2, nOctaveLayers=2
MSER	None	None	None

Table 2: list of algorithms with 3 variation of their parameters

## 8. References

- [1] Alcantarilla, P. F. and Solutions, T. (2011). Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell.*, 34(7):1281–1298.
- [2] Ali, F., Khan, S. U., Mahmudi, M. Z., and Ullah, R. (2016). A comparison of fast, surf, eigen, harris, and msr features. *International Journal of Computer Engineering and Information Technology*, 8(6):100.
- [3] Bannari, A., Morin, D., Bonn, F., and Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, 13(1-2):95–120.
- [4] Bay, H., Tuytelaars, T., and Van Gool, L. (2006). Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer.

---

<sup>4</sup><https://iwmpraise.eu/>

<sup>5</sup><http://challenge-rose.fr/en/projet/roseau-2/>

- [5] Bookstein, F. L. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585.
- [6] Bossu, J. (2007). Segmentation d’images pour la localisation d’adventices : application à la réalisation d’un système de vision pour une pulvérisation spécifique en temps réel.
- [7] Chen, S., Shen, H., Li, C., and Xin, J. H. (2018). Normalized total gradient: A new measure for multispectral image registration. *IEEE Transactions on Image Processing*, 27(3):1297–1310.
- [8] Dantas Dias Junior, J., Backes, A., and Escarpinati, M. (2019). Detection of control points for uav-multispectral sensed data registration through the combining of feature descriptors. pages 444–451.
- [9] Donoser, M. and Bischof, H. (2006). Efficient maximally stable extremal region (mscr) tracking. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, volume 1, pages 553–560. Ieee.
- [10] Douarre, C., Crispim-Junior, C. F., Gelibert, A., Tougne, L., and Rousseau, D. (2019). A strategy for multimodal canopy images registration. In *7th International Workshop on Image Analysis Methods in the Plant Sciences*, Lyon, France.
- [11] Filella, I., Serrano, L., Serra, J., and Penuelas, J. (1995). Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Science*, 35(5):1400–1405.
- [12] Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395.
- [13] Jin, X.-l., Diao, W.-y., Xiao, C.-h., Wang, F.-y., Chen, B., Wang, K.-r., and Li, S.-k. (2013). Estimation of wheat agronomic parameters using new spectral indices. *PLOS ONE*, 8(8):1–9.
- [14] Kamoun, E. (2019). Image registration: From sift to deep learning.
- [15] Lechenet, M., Bretagnolle, V., Bockstaller, C., Boissinot, F., Petit, M.-S., Petit, S., and Munier-Jolain, N. M. (2014). Reconciling pesticide reduction with economic and environmental sustainability in arable farming. *PLOS ONE*, 9(6):1–10.
- [16] Leutenegger, S., Chli, M., and Siegwart, R. (2011). Brisk: Binary robust invariant scalable keypoints. In *2011 IEEE international conference on computer vision (ICCV)*, pages 2548–2555. Ieee.
- [17] Lombaert, H., Grady, L., Pennec, X., Ayache, N., and Cheriet, F. (2012). Spectral demons – image registration via global spectral correspondence. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., and Schmid, C., editors, *Computer Vision – ECCV 2012*, pages 30–44, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [18] Mair, E., Hager, G. D., Burschka, D., Suppa, M., and Hirzinger, G. (2010). Adaptive and generic corner detection based on the accelerated segment test. In *European conference on Computer vision*, pages 183–196. Springer.
- [19] Moré, J. J. (1978). The levenberg-marquardt algorithm: Implementation and theory. In Watson, G., editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 105–116. Springer Berlin Heidelberg.
- [20] Ordonez, A., Arguello, F., and Heras, D. B. (2018). Alignment of hyperspectral images using kaze features. *Remote Sensing*, 10(5).
- [21] Rabatel, G. and Labbe, S. (2016). Registration of visible and near infrared unmanned aerial vehicle images based on Fourier-Mellin transform. *Precision Agriculture*, 17(5):564–587.
- [22] Rublee, E., Rabaud, V., Konolige, K., and Bradski, G. (2011). Orb: An efficient alternative to sift or surf. In *Proceedings of the 2011 International Conference on Computer Vision, ICCV ’11*, pages 2564–2571, Washington, DC, USA. IEEE Computer Society.
- [23] Sage, D. and Unser, M. (2003). Teaching image-processing programming in java. *IEEE Signal Processing Magazine*, 20(6):43–52. Using “Student-Friendly” ImageJ as a Pedagogical Tool.

- [24] Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., Miklas, P. N., Carter, A. H., Pumphrey, M. O., Knowles, N. R., and Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *European Journal of Agronomy*, 70:112 – 123.
- [25] Seitz, H. (2010). Contributions to the minimum linear arrangement problem.
- [26] Shi, J. et al. (1994). Good features to track. In *1994 Proceedings of IEEE conference on computer vision and pattern recognition*, pages 593–600. IEEE.
- [27] Tareen, S. A. K. and Saleem, Z. (2018). A comparative analysis of sift, surf, kaze, akaze, orb, and brisk. *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pages 1–10.
- [28] Trajković, M. and Hedley, M. (1998). Fast corner detection. *Image and vision computing*, 16(2):75–87.
- [29] Vakalopoulou, M. and Karantzalos, K. (2014). Automatic descriptor-based co-registration of frame hyperspectral data. *Remote Sensing*, 6.
- [30] Vioix, J.-B. (2004). Conception et réalisation d’un dispositif d’imagerie multispectrale embarqué : du capteur aux traitements pour la détection d’adventices.
- [31] Zhang, H., Wohlfeil, J., and Griebach, D. (2016). Extension and evaluation of the agast feature detector.
- [32] Zitová, B. and Flusser, J. (2003). Image registration methods: A survey. *Image and Vision Computing*, 21:977–1000.
- [33] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In *Graphics gems IV*, pages 474–485. Academic Press Professional, Inc.