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

Application to agronomic scenes for proxy-sensing.

Vayssade Jehan-Antoine
jehan-antoine.vayssade@inra.fr

February 19, 2020

AgroSup Dijon, Pole GestAd in Precision Farming group, France
Gawain Jones, Jean-Noël Paoli, Christelle Gee

Introduction

Working environment

New consumer and environmental requirements

- Reduction of plant protection products
- Government plans Ecophyto (I, II, II+)

How can these new requirements be met ?

- Localized control of weeds by imaging
- Agricultural Robotics (Pumagri / SITIA)

Working environment

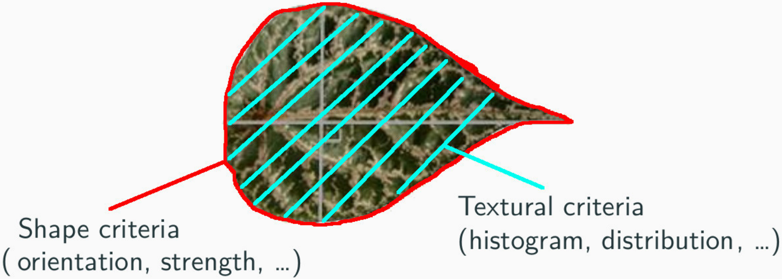


Figure 1: We search for criteria that can discriminate weeds and crops

Main Question

In precision farming, major mistakes are still done in image registration. There is no broad study of :

- the best spectral band to use as reference
- the most adapted key-points extractor algorithms

Focuses

This paper focuses on these two points and defines :

- a benchmark of different key-points detectors (time/number)
- a benchmark for each spectral reference.

Material

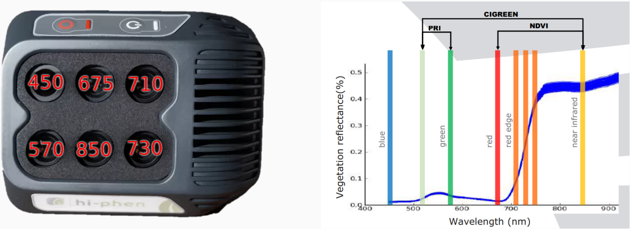


Figure 2: AIRPHEN camera

- interferential filters centered at 450/570/675/710/730/850 nm
- small width of each wavelengths (10 FWHM)
- focal lens is 8 mm for all wavelengths
- raw resolution 1280 $\times$ 960 px with 12 bit of precision.
- internal GPS antenna (3D position)

Data

Two datasets were taken, one for calibration, one for evaluation.

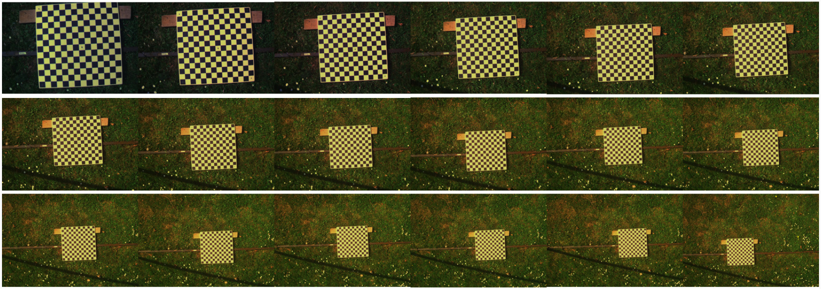


Figure 3: false color reconstruction of each acquisition height (18) for calibration dataset, from 1.2 to 5 meters.

Methods

Two steps : rough and full

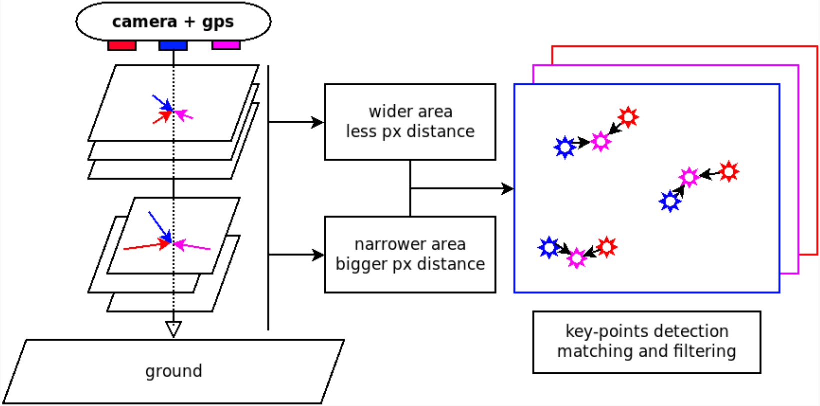


Figure 4: Rough registration through calibration and GPS (left) and refinement using key-point (right)

Affine Calibration, translation part

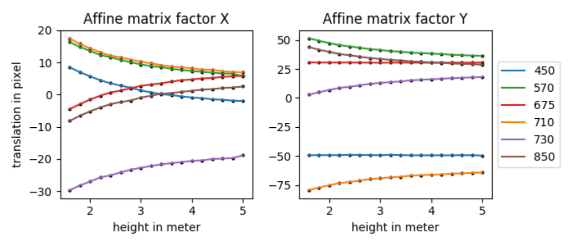


Figure 5: Translation factors from detected chessboard to "virtual" center chessboard at each acquisition height

Affine Calibration, rotation&scale part

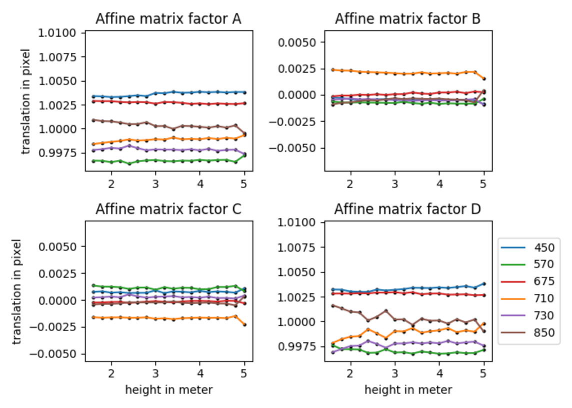


Figure 6: Rotation and scale factors from detected chessboard to "virtual" center chessboard at each acquisition height (precision depends on height but we can notice that these factors are likely invariant)

Affine Correction

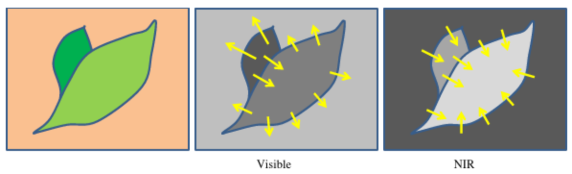
From that calibration an affine matrix model is built:

- For A, B, C, D the values at the most accurate height are used
- For X, Y factors an equation is fitted for each spectral band [1]
  - $t = \alpha h^3 + \beta h^2 + \theta h + \gamma$
- We use the height from the GPS to get the nearest correction
- We crop all images to remove the uncovered area

[1] Levenberg-Marquardt with linear least squiss regression

Methods : Perspective correction via key-points detector (refinement)

Gradient transform for key-points detection



To optimize the search of specific key-points such as gradient breaks, each spectral band is transformed :

- normalization using Gaussian blur $I/(G+1)*255$
- gradient is computed with the sum of absolute Sharr filter
- normalization using CLAHE to locally improve their intensity

Key-points detectors (9)

- (ORB) Oriented FAST and Rotated BRIEF
- (AKAZE) Fast explicit diffusion for accelerated features in nonlinear scale spaces
- (KAZE) A novel multi-scale 2D feature detection and description algorithm in nonlinear scale spaces
- (BRISK) Binary robust invariant scalable key-points
- (AGAST) Adaptive and generic corner detection based on the accelerated segment test
- (MSER) maximally stable extremal regions
- (SURF) Speed-Up Robust Features
- (FAST) FAST Algorithm for Corner Detection
- (GFTT) Good Features To Track

Perspective correction via Keypoint

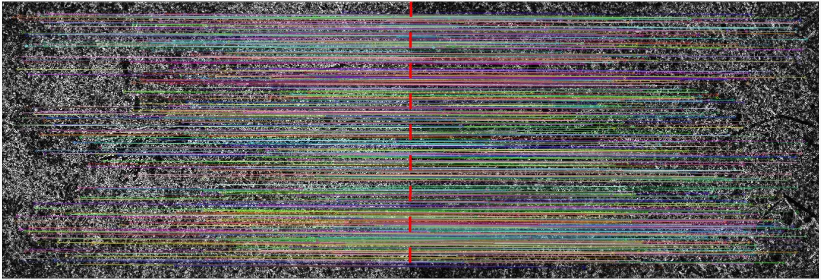


Figure 7: Bruteforce key-points matching in normalized gradient and filtering (reference 570nm left & 850nm right) using GFTT1

- detection inside the normalized gradient
- matching using texture properties of the gradient
- filtering all matches by bounding them (distance/angle)

Results

Results : benchmark of key-points extractors

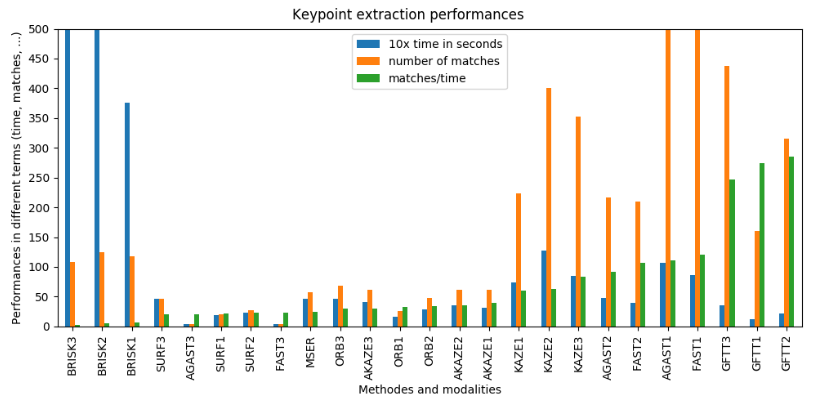


Figure 8: Performances of different key-points detectors

- number of matches → precision
- time → lower is better

Results : benchmark of spectral reference

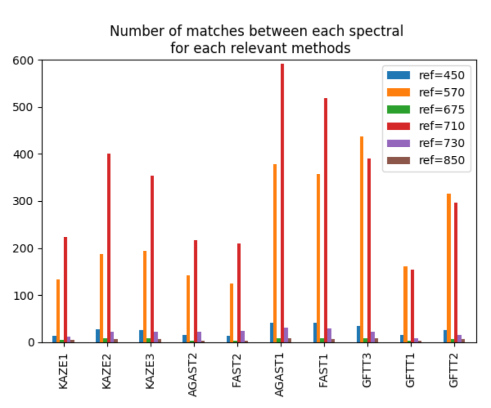


Figure 9: Minimum number of matches for each spectral band to the reference and using different algorithms

Results : precision

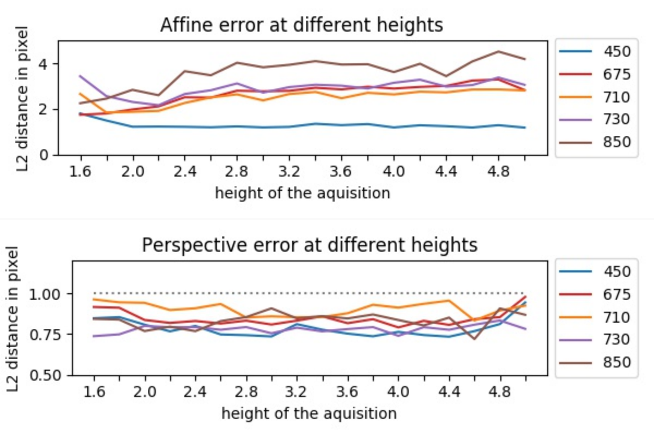


Figure 10: Performance evaluation with 570nm as reference using GFTT1

Results

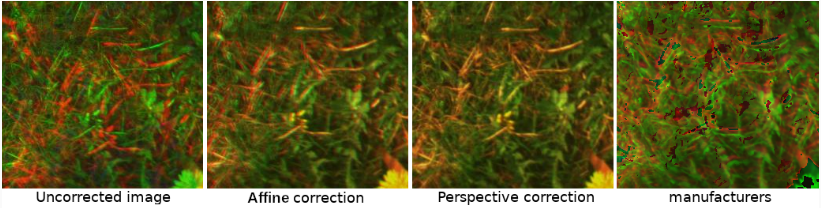


Figure 11: Result in pictures

- uncorrected $\approx 62px$
- manufacturer's $\approx 12px$
- affine 2 $\approx 3px$
- perspective $\approx 1px$

Conclusion

Conclusion

We have proposed a two-step method:

- A first rough correction, to get spatial properties
- A second full correction, using key-points

This methods is generalizable :

- Code open-source :

  `https://gitlab.com/phd-thesis-adventice/phd-airphen-alignment`
- Only need two datasets using your specific camera/scene
- Make the calibration and extract the benchmark
- Use the best combination of key-points detector and reference

Conclusion

For our specific use, such as our camera and scene, we have determined through a benchmark:

- the best spectral reference are 570 and 710nm (studies still define empirically 850nm).
- the best key-point detector is GFTT (time and number) (studies still use ORG or KAZE).

Better performances than the manufacturer's (12px to 1px)

Perspective

- Benchmark of key-points detectors using others cameras
- Benchmark of key-points detectors using others scenes
- Study about the spatial distribution of key-points
- Are the best spectral bands references almost the same ?
- Test of different deformations models
- More modalities about key-points detectors parameters
- Optimization with more efficient matching algorithms

Thank you for listening !

Questions ?