A strategy for multimodal canopy images registration

Clément Douarre, Carlos Crispim-Junior, Anthony Gelibert, David Rousseau, Laure Tougne

A strategy for multimodal canopy images registration

Clément Douarre[1,2,3], Carlos F. Crispim-Junior[1], Anthony Gelibert[3], Laure Tougne[1], and David Rousseau[2]

[1]Univ Lyon, Université Lyon 2, LIRIS, UMR CNRS 5205, F-69676, Lyon, France
[2]Laris, UMR INRA IRHS, Université d'Angers, 62 Avenue Notre Dame du Lac, 49000 Angers, France
[3]Carbon Bee, Rue du Commerce, ZA Les Plaines, 26320 St Marcel-Lès-Valence, France

Abstract

Registration of complex and self-similar images such as plant canopy images is a challenge in plant sciences. Yet, this is often a required step for multimodal imaging, where unaligned sensors yield unregistered image pairs. We propose a pipeline adapted to such constraints, applied to apple tree canopies. Specifically, we apply an intensity-based registration on downscaled and/or Gaussian blurred versions of the targeted images. This helps to eliminate spurious details, which smooths the optimization landscape and also helps to reduce differences between the modalities. Results show better registration than with standard feature-based or intensity-based methods.

Keywords: Registration, Multimodal Imaging, Apple scab, Infrared Imagery.

1 Case study

Apple scab is one of the most serious fungal infections of the apple tree. RGB imaging has shown success for automatic scab detection [2], but to improve its treatment, early detection, *i.e.* before visible symptoms, would be valuable. Infrared (IR) imagery has been shown to be suitable for this task [1].

Our goal is to perform automatic detection of scab on apple plants images acquired both in RGB and in IR. We acquired such multimodal images of apple plants inoculated with scab, from a canopy point of view. Acquisition was done in the LARIS laboratory (Angers) greenhouses in the 2018-19 period. The sensor was a multimodal camera developed by the company Carbon Bee. However, like in many multimodal acquisitions cases, RGB and IR sensors were not exactly aligned and thus, the two modalities of a given acquisition were shifted (Fig. 1).

Figure 1: An example of a RGB (left) and a IR (right) image pair. A yellow rectangle is drawn on both images at the same position. This shows the offset, which seems a feasible objective for a registration algorithm.

This offset was an important problem for us as we needed the images to be aligned for the rest of our analysis to work. Therefore, we needed to perform registration within image pairs.

2 State of the art

Image registration techniques may be grouped in two families [6]: feature-based and intensity-based.

Feature-based methods use matching features in image pairs to find the transformation. The most well-known features are SIFT [4]. However, even with careful tuning of the SIFT algorithm, we find that our images are not suitable for this kind of registration. We can see in Fig. 2 that keypoints in one modality are numerous and incorrectly matched with keypoints in the other, and that there are almost no correct matches. The facts that (i) there are highly complex and self-similar structures in our images, and (ii) leaves have a different aspect depending on the modality, make these images a bad fit for such a registration.

Figure 2: Keypoints found by SIFT and the 20 strongest matches on an image pair, drawn as black lines.

Intensity-based methods consist in warping one image while keeping the other fixed, guided by the optimization of some similarity metric between the two images. This kind of registration applied to multimodal grapevine canopies has been done by *e.g.* [5]. Our method is an adaptation of such a registration, adapted to the specificities of our images.

3 Method

To perform an intensity-based registration, one must choose the warp type and the similarity metric. Since the images represent the same scene acquired through two cameras, the transformation between the images resembles a homography (it is not exactly so, as the scene is non-planar: leaves are at different heights in the canopy). Accordingly, the warping was set to a homography rigid transform. Concerning the similarity metric, we chose the Enchanced Correlation Coefficient (ECC) [3]. It is a measure of similarity between normalized images ($ECC \in [0, 1]$). Hence, it can interestingly, like Mutual Information, work with multimodal images.

Having checked visually that a high ECC between two images correlates with an adequate registration, we used this metric to assess registration quality. Our first registration attempts using ECC were sometimes quite poor: We hypothesized that this was because of the numerous details in the scene, yielding a highly non convex similarity metric optimization landscape.

Consequently, the contributions of this paper are the following modifications on the intensity-based registration: perform the registration on images whose resolution are lowered, through downscaling and/or Gaussian blur.

4 Results

We applied our method to a dataset composed of 50 image pairs of 2592 × 1944 pixels. Image pairs were acquired at different times and orientations, leading to a relatively diverse dataset (Fig. 3). In particular, the offset between images varied from pair to pair. For each image pair, we applied the same downscaling/blur transformations to the two images, registered those images, and adapted the resulting registration homography matrix back to the original images.

Figure 3: Examples of RGB images from the dataset.

When using SIFT, registrations did not improve ECC for any of the image pairs in the dataset. With the proposed strategy, a combination of strong downscale and a strong blur yielded the best results (Table 1, Fig. 4). Those were the cases where some details were "blurred out" and where the only remaining high frequencies were ones from leaf borders, which are robust features for such a multimodal registration. The details blurred out could even be misleading: inner structures in IR images could strongly differ from the ones in RGB (*e.g.* scab vs veins) while leaf contours stayed similar. In other words, our strategy enabled us to control the scales our registration worked at.

	$\sigma = 0$	$\sigma = 2$	$\sigma = 4$	SIFT
$d = 1$	8.4	9.4	10.0	0.0
$d = 0.5$	9.4	10.5	10.8	0.0
$d = 0.1$	10.9	11.5	12.0	0.0

Table 1: Results of our registration method on our database : difference of ECC (in %) between the registered pair of images and the original pair, averaged over all pairs in the dataset. d refers to the scale applied to each dimension of the images. σ refers to the Gaussian blur s.d. The last column refers to registration done by SIFT with different d values.

Figure 4: An example of registration from a pair of our dataset, with $d = 0.1$ and Gaussian blur, $\sigma = 4$.

5 Perspectives

Further work will focus on multiple resolution registration: once images have been registered at a lower resolution, perform another registration at a higher level of detail, starting from the first registration. Once the registration will satisfy us, we will pursue our analysis of automatic scab detection, taking full advantage of both modalities.

References

[1] É. Belin et al. Thermography versus chlorophyll fluorescence imaging for detection and quantification of apple scab. *AGR*, 90:159–163, 2013.

[2] S. Delalieux et al. Hyperspectral indices to diagnose leaf biotic stress of apple plants, considering leaf phenology. *IJRS*, 30(8):1887–1912, 2009.

[3] G. Evangelidis et al. Parametric image alignment using enhanced correlation coefficient maximization. *TPAMI*, 30(10):1858–1865, 2008.

[4] D. Lowe. Object recognition from local scale-invariant features. In *ICCV*, pages 1150–1157. Ieee, 1999.

[5] X. Wang et al. Efficient registration of optical and ir images for automatic plant water stress assessment. *AGR*, 74(2):230–237, 2010.

[6] B. Zitova et al. Image registration methods: a survey. *IVC*, 21(11):977–1000, 2003.