A Photogrammetry-Based Image Registration Method for Multi-camera Systems

H. Gan[1], W.S. Lee[1], V. Alchanatis[2]

[1] Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida

[2] Department of Sensing, Information and Mechanization Engineering, Institute of Agricultural Engineering, ARO - The Volcani Center, Bet Dagan, Israel

Abstract.

In precision agriculture, yield maps are important for farmers to make plans. Farmers will have a better management of the farm if early yield map can be created. In Florida, citrus is a very important agricultural product. To predict citrus production, fruit detection method has to be developed. Ideally, the earlier the prediction can be done the better management plan can be made. Thus, fruit detection before their mature stage is expected. This study aims to develop a thermal-visible camera system which will register thermal images with visible images, so that information fusion can be done later for detecting immature fruit. The registration method used in this study was based on photogrammetry that could be applied to register multiple cameras as well. The camera system used in the study consisted of two identical visible cameras and a thermal camera which were mounted on a single frame and their positions were fixed. Bundle adjustment was utilized to calibrate cameras' relative orientations with respect to each other and intrinsic parameters for each camera. Image registrations were conducted in real-time after each set of images were taken by the three cameras. Common points of interest in the two visible images were selected by running the random sample consensus (RANSAC). Coordinates of corresponding points in the thermal image were calculated by utilizing image intersection method. A transformation matrix was then solved based on selected corresponding points in the thermal image and one of the two visible images. Finally, image registration between the thermal image and the visible image was completed by applying the transformation to the thermal image. The stated method is expected to be fast and can be expanded for multiple camera systems.

Keywords. *Bundle adjustment, Photogrammetry, Registration, Thermal, Visible,*

Introduction

Immature citrus fruit detection has been a difficult task due to the color similarity between fruit and leaves. Recent years, researchers investigated multiple imaging techniques including multispectral imaging, hyperspectral imaging (Okamoto & Lee, 2009), (Annamalai & Lee, 2004) and thermal imaging (Bulanon, Burks, & Alchanatis, 2009), in addition to traditional imaging methods with visible cameras. Their work showed the effectiveness of these modern technologies, however, accuracies of those detections were mostly not satisfactory by using a single type of camera alone. Thermal cameras and visible cameras detect two well separated spectrum regions, so that features detected by those two types of cameras will be less correlated. Therefore, fusing images of thermal cameras and visible cameras together can provide richer information than that from either of them alone. One critical step for image fusion is to implement image registration.

Image registration can be classified into two types, software based automatic registration and hardware based image registration. Most works in literature focused on software based registration, especially in the area of medical imaging and facial recognition. To register computed tomography (CT) and magnetic resonance images (MRI) of brains, Maes, Vandermeulen and Suetens (1999) computed the mutual information of voxel intensities in both images and matched them by maximizing the value. This method required that both images had similar intensity features and the computation speed was usually a limitation. Instead of comparing features of images in multi-models, Chung, Wells and Norbash (2002) utilized a learning method by building the *a priori* knowledge of the expected joint intensity distribution estimated from aligned training medical images. The study aimed to find an optimal transformation that minimize the discrepancy between the observed and expected joint intensity distributions. The method was only tested on two clinical datasets, which had limited features. In addition to the applications in medical imaging, image registrations were often applied to facial recognition, more specifically, registering thermal images with visible images of human faces. (Kong et al., 2007) utilized the Gaussian Fields technique which measured both the spatial proximity and local feature similarity of two points in terms of a Gaussian function. This method was computationally expensive and was only tested on human facial images. All of those methods were shown to be effective in registering images in a certain field, while none of them has been applied to another field of application. One method was found in literature to register thermal and visible images of citrus trees. Bulanon et al. (2009) used a square frame, whose four corners were wrapped by aluminum foil, as a reference to register thermal and visible images. This method used affine transformation which was fast. However, the requirement of placing a reference for every image was not practical for most automation applications.

Hardware based image registration on the other hand had been rarely studied. This method could achieve high accuracy and fast speed. However, substantial preparations have to be done carefully before taking images. Chen and Warren (2013) did a project which registered two webcam with a thermal camera for fire-fighting robots. Camera calibrations were firstly done to calibrate cameras' intrinsic parameters and their relative orientations with respect to each other. World coordinates of each image points were then computed using the visible webcams, to generate a dense points cloud. At last, those points were back projected to the thermal image to find their coordinates.

This study utilized the hardware registration method. At the stage of camera calibration, it is similar to the work of Chen and Warren (2013). Improvements were expected when matching the corresponding points in visible and thermal images. Instead of generating a dense points cloud, world coordinates of only a few points were computed. A Random sample consensus (RANSAC) method will be applied for searching of an optimal affine transformation. The affine transformation will transform the thermal image onto the visible base image.

Materials and Methods

Hardware registration of thermal-visible cameras requires all cameras to be fixed on a single stable frame, because parameters calibrated from initial steps will be utilized for subsequent computations. Based on that requirement, a camera frame was designed and printed by a 3D printer for mounting two identical visible cameras and a thermal camera. The frame and the cameras are shown in Figure 1. All cameras were set to output image sizes of 480 × 640 pixels for the ease of computation.
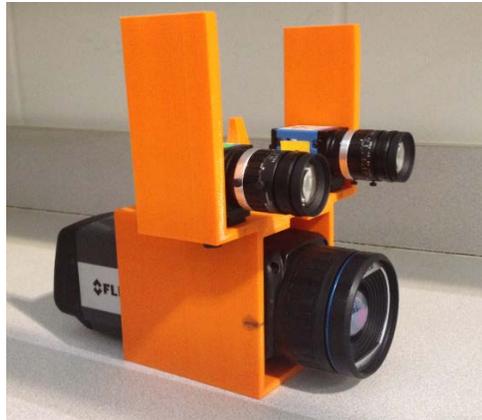

Figure 1 **Camera Frame for Visible and Thermal Cameras**

Camera Calibration

The first step of computation was to calibrate the intrinsic parameters of each camera and their relative orientations, also known as extrinsic parameters. The calibration was done by a bundle adjustment, which required matching common points in each image. In photography, it is usually a manual process. In order to automate this process, a custom made checkerboard was used. As shown in Figure 2, the black squares in the checkerboard were covered by electrical tapes. Two halogen lamps were turned on to heat up the checkerboard when taking images. Due to a higher heat absorption rate, electrical tapes presented clearly higher temperatures in thermal images. The entire setup is shown in Figure 3.

In preparation of bundle adjustment, 18 sets of images were taken. Each set contained three images of the checkerboard taken at the same time by these cameras. Corners of the checkerboard were automatically detected as pass points using MATLAB R2015a. MATLAB was also used to implement bundle adjustment, which calculated camera intrinsic and extrinsic parameters in one step under the collinearity condition. The output parameters were saved for further usage.
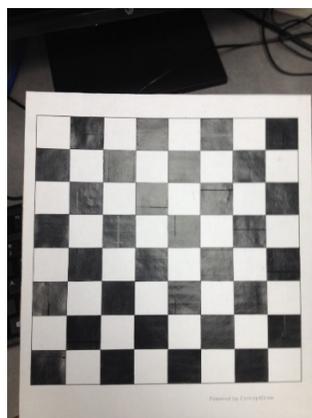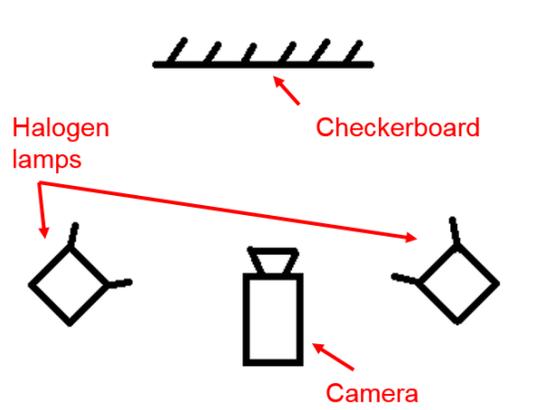

Figure 2 **Checkerboard**   **Figure** 3 **Imaging Setup**

World Coordinates Estimation

Calculated parameters were firstly utilized to compute objects' world coordinates using visible images in each set. Intrinsic parameters were used to rectify visible images, so that distortions were removed. The two rectified images were then combined by applying image intersection. For each pair of image points, the method of image intersection drew two lines connecting caremas' principle points and the image points, and calculated the world coordinates at the intersection of the two lines. In this step, multiple world coordinates were generated.

Registering Thermal Image with Visible Images

Parameters of thermal camera were input to a back-projection formula shown in Eq. (1) to calculate corresponding thermal image coordinates of the generated world coordinates. A random sample consensus (RANSAC) method was applied to iteratively select a best affine transformation between the thermal image and one of the visible images. The selected affine transformation was eventually used for registering the thermal image with the visible images.

$$Coordinates_{thermal} = Coordinates_{world} \times [R\ T]' \times K \qquad Eq.(1)$$

Where,

R = Rotational matrix of the thermal camera relative to the base camera;

T = Translational matrix of the thermal camera relative to the base camera;

K = Intrinsic matrix of the thermal camera.

Results and Discussion

Residuals of Camera Calibration

Camera calibration was done by a bundle adjustment, with 42 pass points in each of the 54 images. At convergence of the bundle adjustment, residuals for each set of images were presented. Image sets with large residuals were removed from the input and a second bundle adjustment were applied. Figure 4 shows the residuals for the first and second bundle adjustments. Parameters calculated from the second bundle adjustment include both intrinsic and extrinsic parameters. Figure 5 is the visualizations of their extrinsic parameters, as well as the positions of the checkerboard in each image.
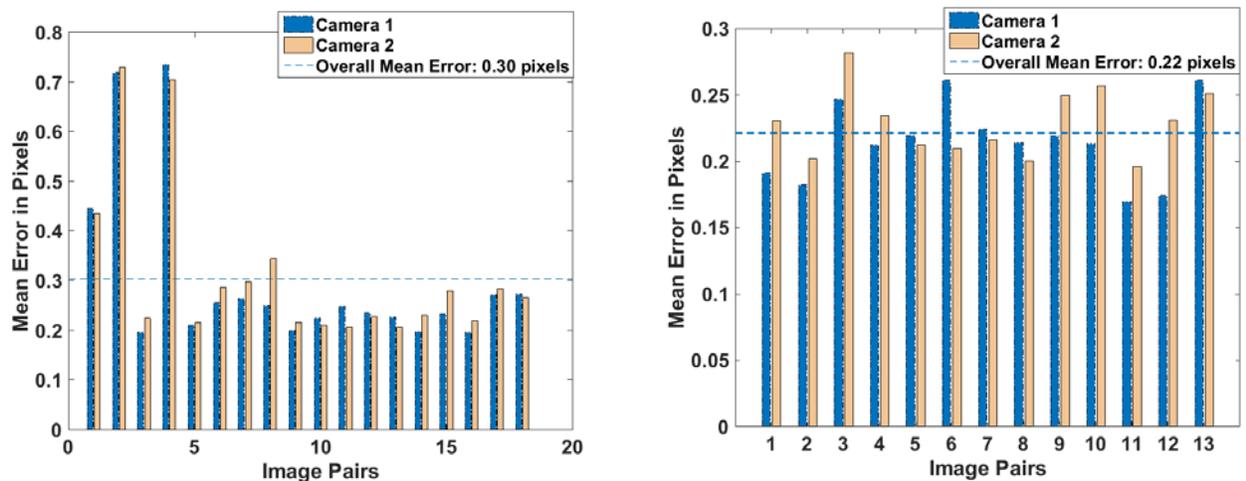

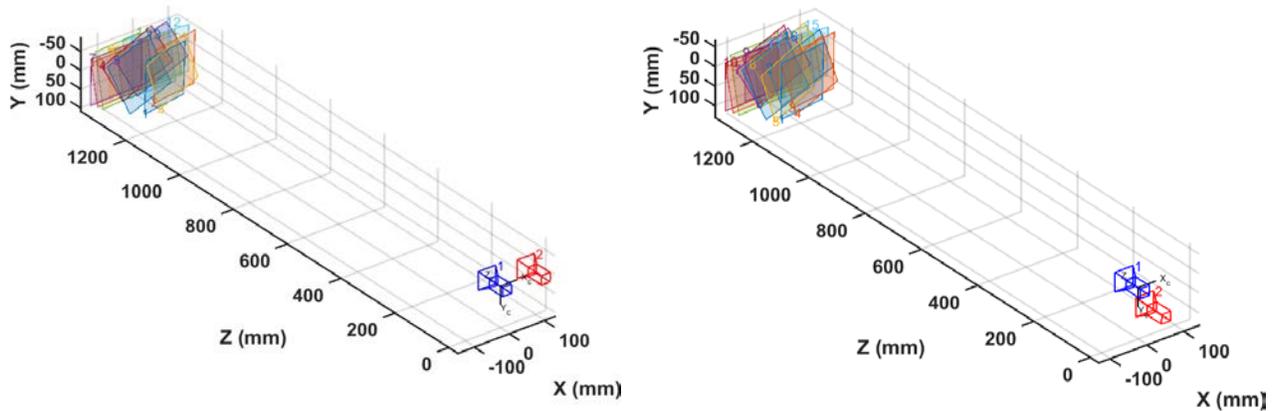**Figure 4 Mean reprojection error per image before (left) and after (right) removing big residuals**

Figure 5 Visualization of relative orientation of two visible cameras (left); visible and thermal cameras (right)

Results of the Registration

World coordinates for each pass point relative to the base camera were computed by image intersection of two visible images. Those points were then back projected to the thermal image as shown in Figure 6.

An evaluation of the back projection was performed by comparing the projected positions and the true positions. Figure 7 shows the histogram of the residuals of all 546 points. Residuals for all selected thermal images are listed below in Table 1. The average residual was 3.13 pixels in a 480 ✕ 640 pixels image.
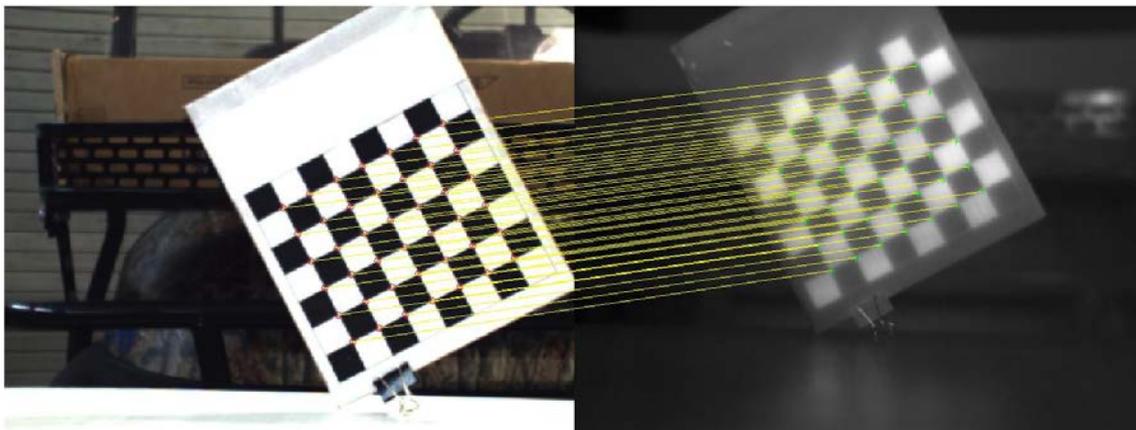

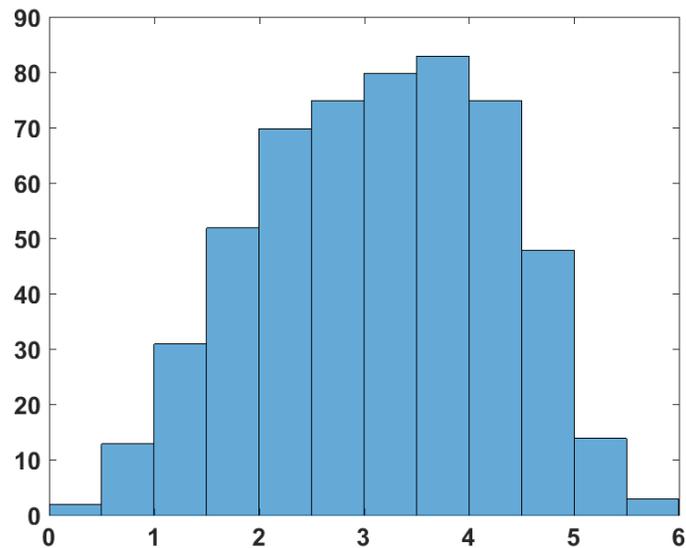**Figure 6 Projection of points from visible image to thermal image**

Figure 7 **Histogram of residuals of all 546 projected points**

Table 1 Residual in pixels for each image

Image #	Residual -Pixel	Image #	Residual -Pixel
1	3.09	8	2.82
2	2.89	9	2.59
3	3.00	10	3.51
4	3.47	11	3.29
5	3.15	12	2.27
6	3.07	13	3.73
7	3.85		

Conclusion

This study effectively registered visible images with thermal images utilizing a photogrammetry method. It is a fast and accurate approach that feature detection was not required for matching these two types of images, and that an accuracy of 3.13 pixels was achieved. This method can be expanded for registering images of different types and multiple imaging devices at the same time.

Acknowledgements

References

Annamalai, P., & Lee, W. S. (2004). Identification of green citrus fruits using spectral characteristics. *An ASAE Section Meeting Presentation*, Paper Number: FL04–1001. Retrieved from http://abe.ufl.edu/wlee/Publications/ASAE FL04-1001.pdf

Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2009). Image fusion of visible and thermal images for fruit detection.

Biosystems Engineering, *103*(1), 12–22. http://doi.org/10.1016/j.biosystemseng.2009.02.009

Chen, Y., & Warren, W. (2013). 3D Fusion of Infrared Images with Dense RGB Reconstruction from Multiple Views - with Application to Fire-fighting Robots, (Figure 1), 1–7.

Chung, a, Wells, W., & Norbash, A. (2002). Multi-modal image registration by minimising kullback-leibler distance. *Medical Image Computing and Computer-Assisted Intervention*, *2489*, 525–532. http://doi.org/10.1109/CVPR.2003.1211518

Kong, S. G., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B. R., Koschan, A., … Abidi, M. a. (2007). Multiscale fusion of visible and thermal IR images for illumination- invariant face recognition. *International Journal of Computer Vision*, *71*(2), 215–233. http://doi.org/10.1007/s11263-006-6655-0

Maes, F., Vandermeulen, D., & Suetens, P. (1999). Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, *3*(4), 373–386. http://doi.org/10.1016/S1361-8415(99)80030-9

Okamoto, H., & Lee, W. S. (2009). Green citrus detection using hyperspectral imaging. *Computers and Electronics in Agriculture*, *66*(2), 201–208. http://doi.org/10.1016/j.compag.2009.02.004