Lens Distortion Correction for Thermal Cameras to Improve Aerial Imaging with Small-Scale UAVs

Saeed Yahyanejad Jakub Misiorny Bernhard Rinner
Institute of Networked and Embedded Systems
Alpen-Adria Universität Klagenfurt, AUSTRIA
<firstname.lastname>@aau.at

***Abstract*—Lens distortion as a result of the shape and construction of a photographic lens is a common problem in image acquisition. Thermal cameras are no exception to this artifact. So far many methods have been developed to formulate the distortion model and almost all of them exploit the patterns in visible range to calibrate the lenses in RGB cameras. A checkerboard is among the most common and well-defined patterns for RGB camera calibration. Unfortunately, most of those patterns will not be easily visible in images taken by a thermal camera. Furthermore, since the thermal cameras measure the infrared radiation (heat), the conductivity of the heat to the bordering objects in the pattern might mitigate sharp edges, which will make detection of relevant features within the pattern harder and less precise.**

In this paper we propose an algorithm to construct a calibration pattern visible for the thermal infrared cameras. We show how to extract robust features out of the sensed checkerboard pattern which is used afterward for lens distortion correction. Further, we evaluate our method and compare it to results obtained from well established algorithms for visible-light lens calibration. We also demonstrate how distortion correction improves the image registration between thermal and RGB aerial images taken by small-scale unmanned aerial vehicles (UAVs).

I. Introduction

Over the years, the use of thermal cameras has been expanded. Initially mainly used by military and government agencies for surveillance and security, the thermal camera technology has now migrated to many other exciting areas. A rapidly developing area for thermal cameras is multi-spectral aerial imaging based on small unmanned aerial vehicles (UAVs). The UAVs serve as convenient robotic platform to position the cameras up in the air. Figure 1 shows two examples of such UAVs. Potential applications of these aerial robots include surveillance, disaster management or environmental monitoring (e.g., [5], [4]).

A few researches have focused on jointly registering thermal and visible-light images for the purpose of disaster site reconnaissance [7], [6]. Registration of visible-light (also referred to RGB images) and thermal (specifically, radiation from mid-/far- infrared part of the electromagnetic spectrum) images is a challenging task. It is further complicated due to the distortions generated by the different camera lenses. Thus, compensating the lens distortion is a very important prerequisite for successful multi-spectral image processing.

While the lens distortion correction is a well established field in RGB images (e.g., [2], [3]), it is mostly uncharted territory for thermal imaging. The only thermal lens calibration method we could find was proposed by Rudol and Doherty [6]. Their approach is hard to re-implement since they make use of specially built and not-readily available components. In contrast, the approach proposed by us require only a printed calibration pattern and an infrared radiation (most conveniently generated by a heat lamp).

(a) AscTec Pelican drone with an FLIR Photon 640 thermal camera. (b) MD4-200 drone with an RGB camera.

Figure 1. Two drones used for acquiring thermal and visible-light images.

Using methods already developed for the lens calibration of RGB cameras might not be useful for thermal cameras because:

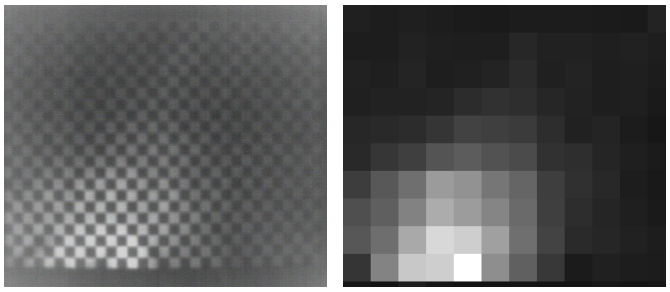
- patterns used for RGB cameras are not usually visible in the infrared part of the spectrum,
- lenses used for thermal cameras are different than ones used for RGB imaging (considering the shape, material, etc.).

In order to take advantage of the readily available calibration tools we adopt the work-flow of the RGB lens calibration method which consists of the following steps:

1) acquire a picture (or pictures) of a calibration pattern.
2) detect relevant features of the pattern.
3) calculate intrinsic and extrinsic parameters of the lens based on features extracted in previous point, mathematical models of the lens and known properties of the used pattern (such as straight or parallel lines, right angles, etc.).

The calibration method presented this paper replaces the steps 1 and 2 that are tailored to thermal imaging. For the remaining steps we apply the standard RGB calibration methods which have been proven over the years.

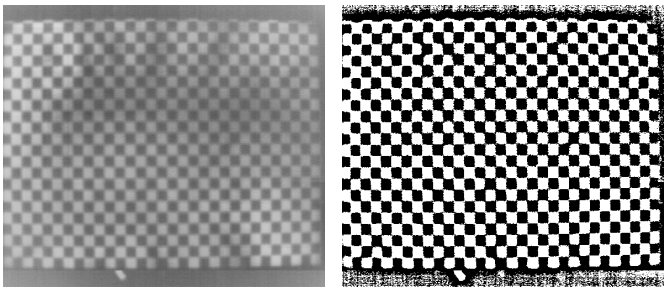
The rest of the paper is organized as follows. In Section II we give a brief description of the proposed method and



(a) A sample frame out of all set of frames.

(b) The weight function over the corresponding frame.

Figure 2. Extracting the target area.



(a) Sample integration of 92 individual frames.

(b) The result after adaptive thresholding.

Figure 3. Constructed integration of frames.

used rig. Sections III and IV give detailed description of the algorithm for constructing pattern visible to thermal camera and extracting relevant features from it. In Section V we describe test rig used in out experiments. Sections VI and VII give details about evaluation method used and experimental results, respectively. Finally Section VIII concludes this paper.

II. Proposed method setup

In our method we aim to make the conventional checkerboard pattern visible for the thermal cameras by radiating the infrared light over the pattern. Radiation makes the black squares heat up more than the white squares, hence we are able to sense the reflected IR radiation by thermal sensors. So first step is to heat up a normal, checkerboard pattern printed on a $100\,g/m^2$ with IR radiation. Our first test were made while the sun was shining on the pattern with a great success. However, to be independent from the weather condition and to achieve a higher precision and a more stable rig, we switched to IR radiation lamps which works as good as sun and can be precisely controlled. The main problem with the IR lamps is that their field of operation is not wide enough to heat up the whole $841 \times 1189\,mm$ pattern uniformly. The reason that we have chosen such a big pattern is that the minimum focus distance of our thermal lens is $2\,m$ (For detailed description of the test rig see Section V).

The following idea has been implemented to resolve that issue. A completely fixed thermal camera takes series of images while the IR emitting lamp is moved across the pattern. Then images are analyzed and the final image is assembled from pieces of those input images with highest contrast. This procedure will be explained in detail in the Section III.

III. Maximum contrast image assembly

The input of the assembly is series of images $\{\mathbf{I}_i | i = 1 \dots n\}$ taken exactly from the same area as explained in the Section II. Figure 2(a) shows a sample frame of thermal image while the pattern is partially heated up by IR radiation.

In order to extract the position of the squares in checkerboard we perform the following steps over the set of frames:

- First we need to crop a region of each frame which has the required information. In other words, since the infrared light heats up only a part of the pattern we need to extract that specific area. As you can see in Figure 2(a) inside this target area the checkerboard pattern has more contrast and therefore it produces a more clear edge at each square. Since the bimodality of the checkerboard is a good characteristic of this area we use the metric used by Yahyanejad and Strom [10] which favors an image intensity histogram with two peaks where intra-peak variance is small and inter-peak distance is large:

$$\lambda(C) = \frac{\sigma_{C'} + \sigma_{C''}}{(\mu_{C'} - \mu_{C''})^2}$$
$$\text{where} \quad \begin{cases} C' = \{C(x) \mid C(x) < \mu_C\} \\ C'' = \{C(x) \mid C(x) > \mu_C\} \end{cases}, \tag{1}$$

  $C \subset I$ is the subregion of I which we are measuring the bimodality over. $\mu_{()}$ and $\sigma_{()}$ imply the mean and the variance respectively. We partition each image into a grid. The size of the grid depends on the size of the black and white squares. To avoid aliasing we obey the Nyquist sampling rate so that the minimum grid size is twice the size of each black or white square. Afterwards we compute the metric over each subregion C_{ij} of the grid. Figure 2(b) shows the corresponding weight function which is constructed as follows:

$$\lambda(I) = \begin{bmatrix} \lambda(C_{11}) \cdots \lambda(C_{1n}) \\ \vdots \quad \ddots \quad \vdots \\ \lambda(C_{m1}) \cdots \lambda(C_{mn}) \end{bmatrix} \tag{2}$$

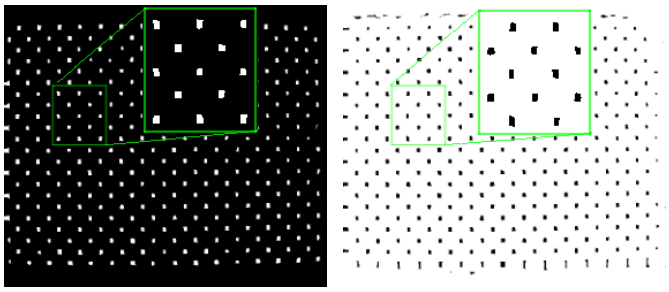
- In second step we integrate all frames together and construct a single frame with maximum contrast O:

$$O = \frac{\sum_i I_i \cdot \lambda(I_i)}{\sum_i \lambda(I_i)} \tag{3}$$

  A sample integration of 92 individual frames is shown in Figure 3(a).

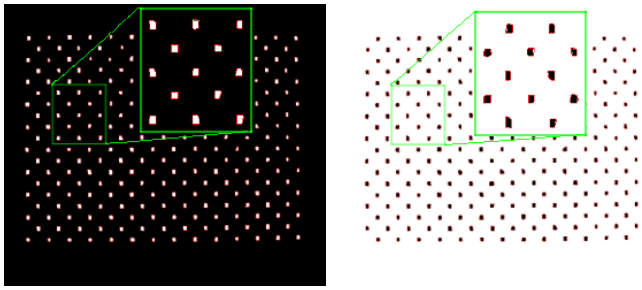
IV. Relevant feature detection

Since the images obtained from the assembly described in previous section are not characterized by very sharp/precise edges, usual RGB-camera lens calibration methods cannot be yet used. In this stage we will describe process of extracting



(a) The binary image after erosion. (b) The binary image after dilation.

Figure 4. Constructed integration of frames.



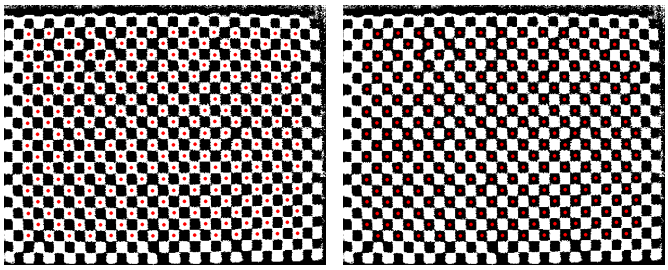
(a) The convex hull around the eroded blobs. (b) The convex hull around the dilated blobs.

Figure 5. Bounding the blobs by convex hulls.

precise and robust features which can be fed to standard checkerboard calibration algorithms.

Using the integrated image from the previous step we continue with the following operations:

- Since we know an ideal checkerboard pattern should be bimodal, in this step we convert our obtained result to a binary image. Due to the intense variation of the gain level in thermal cameras and also because of the non-uniform infrared heating ratio, the constructed integration also has a non-uniform contrast in its different parts. In this case it would be impossible to separate the black and white squares from each other by using a global thresholding. Instead we utilize a quick local adaptive thresholding algorithm proposed by Wellner [9]. The result of thresholding is shown in Figure 3(b).
- In this step we define the robust features out of the obtained checkerboard pattern. In most of the calibration methods features of choice are corners of the checkerboard. As mentioned before the thermal camera may sense the conducted heat from black square to its adjacent pixels and therefore the size of the black squares might become bigger than its original size. To avoid this error we mark the center of squares as our features. The center extraction is done by performing first a dilation and erosion over the binary image in a way that we obtain a explicitly distinguished blobs instead of black and white squares as you can see in Figure 4.
- To approximate the center of the squares we need to find the center of the mass of blobs obtained in the previous step. Initially we need to distinguish between individual blobs. In order to do that, we perform a simple contour detection based on method proposed by Suzuki [8]. Desired output of this step is a list of individual contours corresponding to all blobs which is fed to the second step.
- Each contour, which is a list of points defining the outer border of the blob, is processed separately. A convex hull is computed as having smallest possible dimensions as to include all points of each contour. Example of convex hulls obtained in this manner can be seen in Figure 5. As you can see, we also have removed some rows/columns of squares on each margin to avoid the marginal errors.
- Finally we calculate the centroid of each blob within its convex hull. Since all pixels have the equal mass, the center of mass is calculated simply by averaging over all pixel locations of each blob: $(\mu(X), \mu(Y))$. The projection of these centers over the checkerboard pattern is shown in Figure 6.



(a) Centers of eroded squares. (b) Centers of dilated squares.

Figure 6. Square centers visualized as circles.

V. Test Environment

A. Hardware

Our test were carried out with FLIR Systems[1] Photon 640[2] with a $25\,mm$, $f\backslash 1.4$ lens which has a Field of View (FOV) of $36°$ x $29°$.

The resolution of the camera in analog PAL mode is 640×512.

It is a Mid-Wave Infrared (MWIR) camera operating in a $7.5 - 13.5\,\mu m$ wavelength range, with a noise equivalent differential temperature (NEdT) of $< 50\,mK$. The images have been acquired through analog interface of the camera, through a consumer video grabber card.

The lens has a minimum focus distance of $2\,m$.

B. Detailed test rig setup

The final set-up for the image acquisition has been as follows:

Calibration image is a checkerboard pattern printed on an A0[3], $100\,g/m^2$, matte paper. It has been attached to a wall with special care taken for it to be as flat as possible. The

[1] http://www.flir.com/

[2] http://www.flir.com/cvs/cores/uncooled/products/photon640/

[3] dimensions: $841 \times 1189\,mm$ (ISO 216)

pattern itself is made out of squares with a side of $38\,mm$. It has been printed with a laser printing technology, which has been chosen along with matte paper, for their superior radiation absorbance/emittance properties.

Camera has been mounted on a tripod, $2.5\,m$ from a calibration pattern. It was positioned so the pattern is fitted within the FOV in a way that the optical axis is approximately along the center of the pattern.

VI. Evaluation

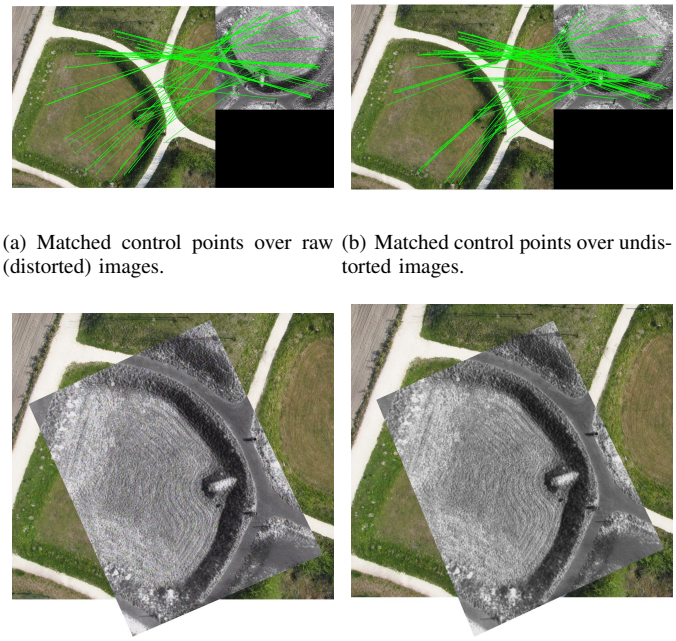
For evaluation we need to perform a calibration over the outcome of our algorithm. To do that we exploit the well-known MATLAB calibration toolbox developed by Bouguet [2]. Most of the existing calibration toolboxes including the mentioned toolbox is designed in a way that they calculate the corners of checkerboard pattern as their input. In order to be able to feed the centers of the square we have modified the code in a way that it accepts the center of the squares, which is the output of our algorithm. Hence we combined the centers of the squares obtained from erosion and dilation (cf. Figure 6). The combinational result is shown in Figure 7(b). In Figure 7(a) the automatic corner extraction result (from Bouguet toolbox) over Figure 3(a) is depicted.

The main goal of our research was to construct a visible pattern for thermal cameras which is possible to perform a calibration over it. This task is fulfilled since the existing calibration toolbox accepts the previously mentioned corner extraction result (cf. Figure 7(a)). The corresponding calibration result is shown in Figures 8(a), 9(a) and 10(a).

Brown distortion model [3] can tackle the radial and tangential distortion which is used in our calibration method. But since the error tolerance for $4th$ order of radial distortion and for all orders of tangential distortion were higher than their distortion coefficient so we ignored them. In result we only consider the second order of the radial distortion including the principal point estimation which is usually sufficient for narrow fields of view such as our thermal camera. The calibration result performed by our algorithm is depicted in Figures 8(b), 9(b) and 10(b). By feeding our extracted center of squares to the calibration toolbox with the same parameters we will obtain a more accurate result in a sense that the reprojection error depicted in Figures 10(a) has lower variance and mean comparing to Figure 10(b). The average pixel error, $(\mu(X-\hat{X}), \mu(Y-\hat{Y}))$, in our algorithm is $(0.36, 0.38)$ while that of the automatic corner detection method is $(0.49, 0.51)$.

VII. Practical Results

Now we will see how lens distortion correction will affect a real case scenario. In our practical scenario, we aim to register the thermal and RGB aerial images taken from the same scene by exploiting the UAVs shown in Figure 1. We made a test set of some pairs of thermal and RGB aerial images. This test set consists of both distorted and undistorted aerial images. Hence, over each pair of images we extract the matching control points by using SURF feature extraction method [1]. Afterwards registration procedure is performed once over the raw (distorted) images and once over the undistorted images (cf. Figure 11). Green lines in Figures 11(a) and 11(b) are implying the corresponding control points. In Figures 11(c) and 11(d) the registered result after homography is shown.



(a) Matched control points over raw (distorted) images.

(b) Matched control points over undistorted images.

(c) Registered after homography of raw (distorted) images.

(d) Registered after homography of undistorted images.

Figure 11. Registration of thermal and RGB aerial images.

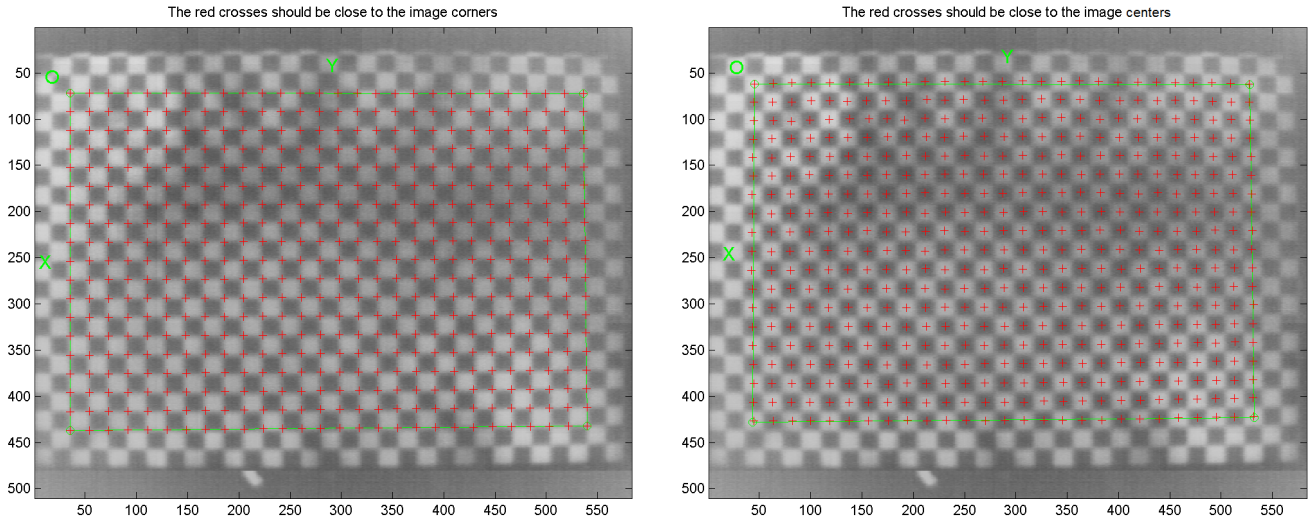
To be able to quantify the improvement of the results after distortion correction, we use the Mean Squared Error (MSE) as our error function:

$$MSE(\hat{P}) = \frac{\sum_{i=1}^{n} (\hat{p_i} - p_i)^2}{n} \tag{4}$$

where $P = \{p_i | i = 1 \dots n\}$ is the set of the control points in the base image, $\hat{P}$ is the estimation of P which is obtained by performing the homography over the corresponding control points of the transformed image and n is the number of control points. This error for raw images in Figure 11(a) with 38 control points is $MSE = 13.2$, and for undistorted images in Figure 11(b) with 54 control points is $MSE = 11.7$. We have calculated this error with 5 different sets of control points over 4 pairs of images and the statistics shows that the MSE error when registering the undistorted images is reduced in average by 17% comparing to raw image registration. Obviously using a wider lens will produce a higher barrel distortion and consequently in that case the mean squared error will show a higher reduction after distortion correction.

VIII. Conclusion

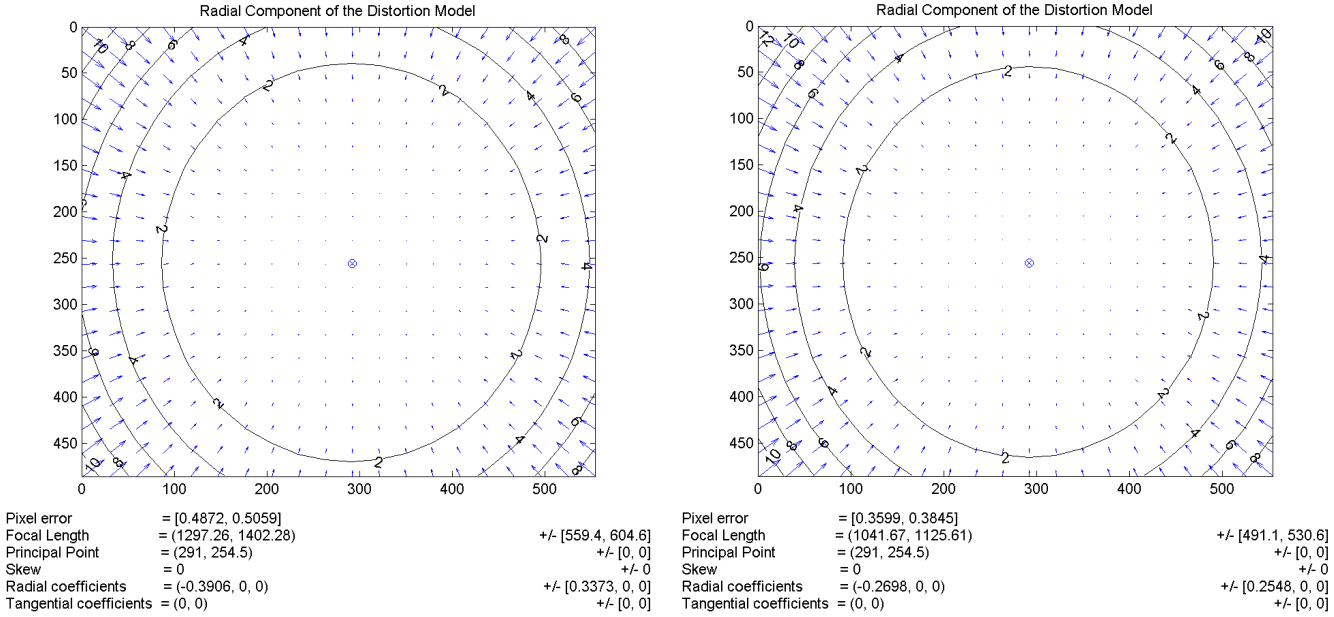
Use of thermal cameras has grown rapidly over the past years while a well-defined lens calibration method is missing. The calibration and lens distortion correction will help us in many fields of computer vision such as image registration,



(a) Set of corners extracted by MATLAB calibration toolbox [2].

(b) Set of all centers from our algorithm.

Figure 7. Results of extracted centers vs extracted corners.



(a) Calibration with automatic corner detection.

(b) Calibration with our center of squares input.

Figure 8. Radial component of calibration.

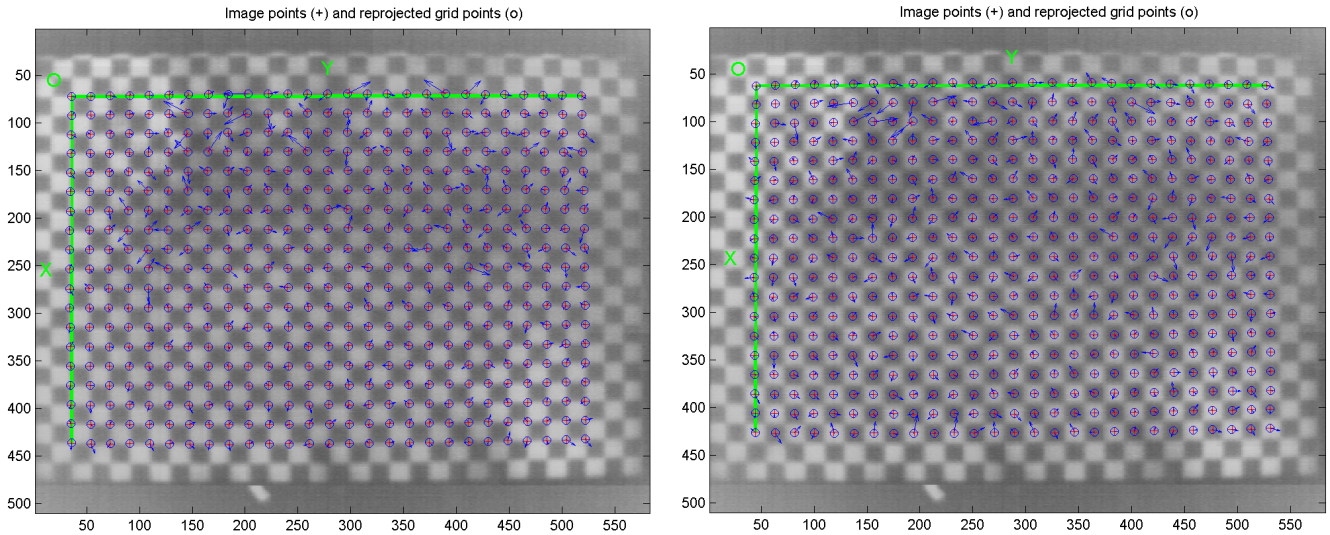
stitching, mosaicking and georeferencing. In this paper we proposed an approach to construct a visible pattern for thermal cameras which can be used to calibrate their lens distortion. We evaluated the calibration result and we showed later on how this result help us to improve the registration of thermal and RGB images taken by small scale UAVs.

ACKNOWLEDGMENT

This work was performed in the project *Collaborative Microdrones (cDrones)* of the research cluster Lakeside Labs and was partly funded by the European Regional Development Fund, the Carinthian Economic Promotion Fund (KWF), and the state of Austria under grant KWF-20214/17095/24772. Jakub Misiorny has been supported by the Austrian agency for international cooperation in education and research, financed by the scholarship foundation of the Republic of Austria.

REFERENCES

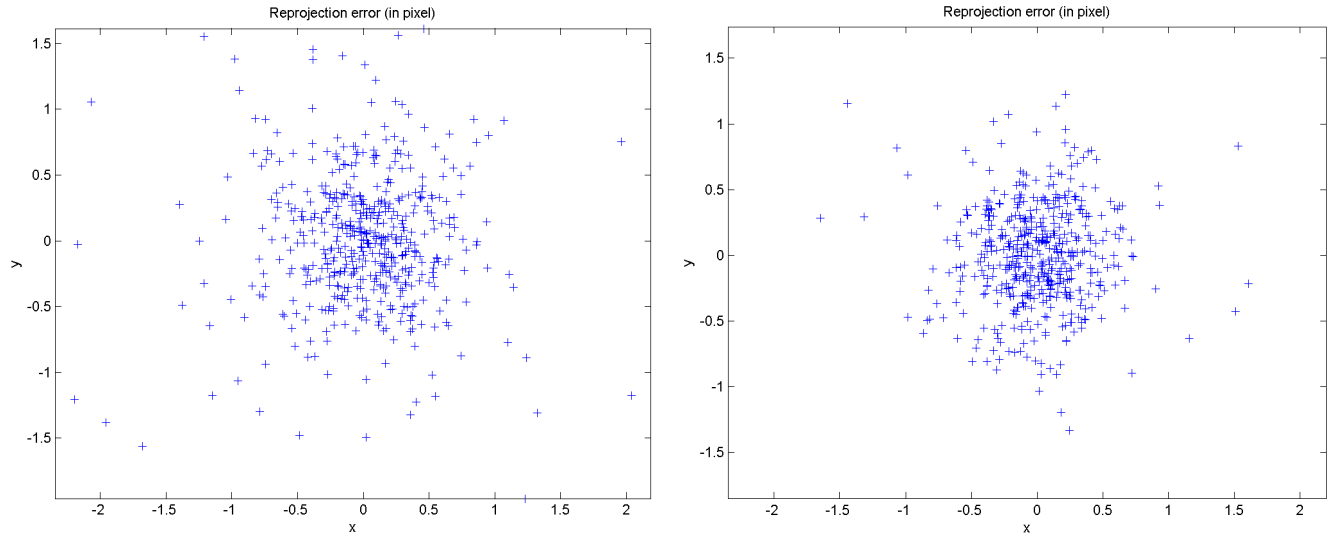
[1] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.*, 110(3):346–359, 2008.



(a) Output of calibration with automatic corner detection.

(b) Output of calibration with our center of squares input.

Figure 9. Reprojection error over the pattern.



(a) Output of calibration with automatic corner detection.

(b) Output of calibration with our center of squares input.

Figure 10. Chart of reprojection error.

[2] J.-Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/, June 2011.

[3] D. C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3):444–462, 1966.

[4] M. Quaritsch, R. Kuschnig, H. Hellwagner, and B. Rinner. Fast Aerial Image Acquisition and Mosaicking for Emergency Response Operations by Collaborative UAVs. In *8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011)*, Lisbon, Portugal, 2011.

[5] M. Quaritsch, E. Stojanovski, C. Bettstetter, G. Friedrich, H. Hellwagner, B. Rinner, M. Hofbauer, and M. Shah. Collaborative microdrones: Applications and research challenges. In *Proceedings of the Second International Conference on Autonomic Computing and Communication Systems*, 2008.

[6] P. Rudol and P. Doherty. Human body detection and geolocalization for uav search and rescue missions using color and thermal imagery. In *Aerospace Conference, 2008 IEEE*, pages 1–8, march 2008.

[7] G. Rufino and A. Moccia. Integrated vis-nir hyperspectral/thermal-ir electro-optical payload system for a mini-uav. *Infotech@ Aerospace*, pages 1–9, 2005.

[8] S. Suzuki and K. Abe. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46, 1985.

[9] P. D. Wellner. Adaptive thresholding for the digitaldesk. Technical report, 1993.

[10] S. Yahyanejad and J. Strom. Removing motion blur from barcode images. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 41–46, 2010.