An Accurate Calibration Method of a Multi Camera System

Song Han, Xiaojing Gu, and Xingsheng Gu[(✉)]

Key Laboratory of Advanced Control and Optimization for Chemical Process,
Ministry of Education, East China University of Science and Technology, Shanghai, China
{xjing.gu,xsgu}@ecust.edu.cn

Abstract. In this paper, we proposed a novel method of geometric calibration and synchronization for a multi camera system. Traditional calibration methods of visible cameras can't be applied to thermal cameras. According to the imaging characteristics of thermal cameras, we designed a new calibration board using materials with different emissivity for calibration. Our calibration board can accurately calibrate RGBD cameras and thermal cameras. In general, thermal cameras have regular non-uniformity corrections, which will result in camera interruption about 1.5 to 2 s and can impact synchronization. In this respect, we adopted the method named nearest adjacent time using timestamp to solve the problem of non-uniformity corrections and synchronization. We evaluated our methods and the experiments showed that our methods had an ideal result for camera calibration and synchronization.

Keywords: Multi camera system · Thermal cameras · RGBD cameras · Calibration · Synchronization

1 Introduction

Consumer level RGBD cameras, such as Kinect [1], are popular in the field of computer vision in recent years. A large number of Kinect performance tests have been completed by [2–4]. Kinect is applicable to many fields, such as robotic field, 3D reconstruction field and so on [5]. Recently, with the decrease of thermal camera price, infrared imaging has become a helpful and useful tool in the field of computer vision. Thus, the idea of combining a RGBD camera with a thermal camera has started to become more popular. More and more people design multi camera systems including Kinect and a thermal camera to perform a 3D reconstruction of a thermal scene [6]. The 3D reconstruction of thermal scene can be applied to building inspection [7] and fire rescue [8].

Geometric calibration and temporal calibration are the precondition of a multi camera system. Without above calibrations, the multiple image information from a multi camera system can't be integrated together. The infrared image represents the temperature information of the object in the scene. Therefore, traditional calibration boards can't calibrate thermal cameras. In this paper, we designed a novel calibration board to calibrate RGBD and thermal cameras. The temporal calibration means synchronizing scene images from different cameras. The scene information will be confusing without

temporal calibration. We adopted the method of nearest adjacent time to synchronize RGBD and thermal cameras. The contribution of this paper is as follows:

1) We designed a novel calibration board which can calibrate both RGBD and thermal cameras. The advantages and disadvantages between the circular calibration plate and the checkerboard calibration plate are compared.
2) We considered the effect of the interruption from non-uniformity corrections of the thermal camera on synchronization. And we adopted the method of nearest adjacent time by using timestamps to solve camera synchronization.

The rest of the paper is organized as follows: In Sect. 2 we provide an overview of prior works in camera calibration and synchronization of multi camera systems. And then, we introduce our methods of camera calibration and synchronization in detail in Sect. 3 and 4. Next, we carry out some experiments and evaluation in Sect. 5 and then provide conclusion and future work in Sect. 6.

2   Related Work

Camera calibration is a core problem in the field of 3D reconstruction and robot navigation [9]. The most widely accepted calibration method is the strategy from Zhang [10]. The traditional calibration boards usually use printed chessboards and the calibration points are easily located in a visible image, but they cannot be accurately located in an infrared image. This is because the thermal camera mainly acquires the temperature information of objects and the printed chessboard generally maintains near-uniform temperature with low contrast. In recent years, many researches have designed different calibration boards in order to calibrate thermal cameras. In [11–13], authors cut rectangular holes through a board and put it in front of different temperature objects. The method can calibrate both RGBD and thermal cameras. However, the thickness of holes will result in inaccurate calibration. A thermal calibration rig with 42 small LED lights located on the intersections of the conventional checkerboard was designed. Which can calibrate thermal cameras when the LED lights were on [14]. However the light will affect the precise positioning of feature points in infrared images. Kim [15] presented a line-based grid pattern board, which included a line based grid of regularly sized squares pattern. Michael [16] adopted a black and white checkerboard with one resistor mounted in the center of each checkerboard square. In [17], the authors used rubber heater to warm up a plastic mask with a grid of holes for thermal camera calibration. It worked well but the price of the calibration board is a little high. We proposed a novel calibration board made of aluminum foil and cardboards, which is easy to make and low cost. Besides, the calibration board only need to be heat one minute by a hair drier before calibrating cameras. We designed square and circle pattern calibration boards and carried out experiments to study which one is more accurate.

The process that all images from all sensors of a multi camera system have corresponding timestamps is called camera temporal calibration or synchronization. Many previous studies adopted hardware to solve synchronization. Soonmin [18] used the master-slave synchronization technique to synchronize visible and thermal cameras. The

visible and thermal cameras are synchronized by sending trigger signals from master to slaves. In [19], the author adopted beam splitter to synchronize cameras. The synchronization result is ideal with the method of hardware, but most cameras don't have trigger generators or beam splitters. Therefore, the cost of synchronization will increase. Some other works such as [20] dealt with extrinsic calibration and synchronization jointly.

The previous works have taken little account of the impact of non-uniformity corrections from thermal cameras on synchronization. Most thermal cameras have regular non-uniformity corrections, which will lead to camera interruption about 1.5 to 2 s. To overcome this problem, we adopted the method of nearest adjacent time, which can quickly synchronize cameras and can solve the impact of non-uniformity corrections.

3   Calibration

3.1   Intrinsic Calibration

Pinhole model is one of the simplest types of camera models, which can mathematically describe the projection of points in 3D space onto an image plane. P is the point in 3D space, p is the perspective projection point from P to a 2D image plane. In homogeneous coordinates, the 3D world point P and perspective projection point p can be respected as:

$$P = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix} \tag{1}$$

$$p = \begin{bmatrix} u & v & 1 \end{bmatrix} \tag{2}$$

The geometric relation between P and p can be written as:

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

Where z is the distance from the point P to the camera, R and T are the spatial relation matrixes between the world coordinate system and the camera coordinate system. The intrinsic matrix consists of focal length f_x, f_y and the principal point (u_0, v_0).

The traditional calibration boards can't calibrate the thermal camera. The thermal camera acquires the information of the scene through the reflected and radiate infrared rays of the objects. Materials with high emissivity reflect most of the infrared rays back because they don't absorb much infrared energy. Therefore we can calibrate the thermal camera with the calibration board made of different materials. We use the second generation Kinect as the RGBD camera. The second generation Kinect contains both a color camera and a thermal camera. And the spatial depth information is detected by the thermal camera. This means that the depth camera and the thermal camera are the same camera. Thus the calibration of the depth camera can be converted to calibrated the thermal camera.

In this paper, we proposed two different calibration pattern boards which can cali-brate both thermal cameras and RGBD cameras. One is the square planar pattern cali-bration board. The other is the circle pattern calibration board. The square planar pattern calibration board was consisted of white cardboard squares and aluminum foil squares. The aluminum squares were made by sticking aluminum foil on the white cardboard squares. The circle pattern calibration board was made of circle aluminum foil and a thick white cardboard. Both square planar pattern and circle pattern calibration boards are suitable, convenient and low cost. In order to know which kind of calibration boards has higher calibration accuracy, some tests was carried out by using OpenCV. Figure 1 shows the two types of calibration boards and the results of calibration. The accuracy of calibration can be analyzed by calculating the mean reprojection error (MRE), which indicates the distance between a detected point and a reprojection point in a photo. The unit of MRE is pixel.
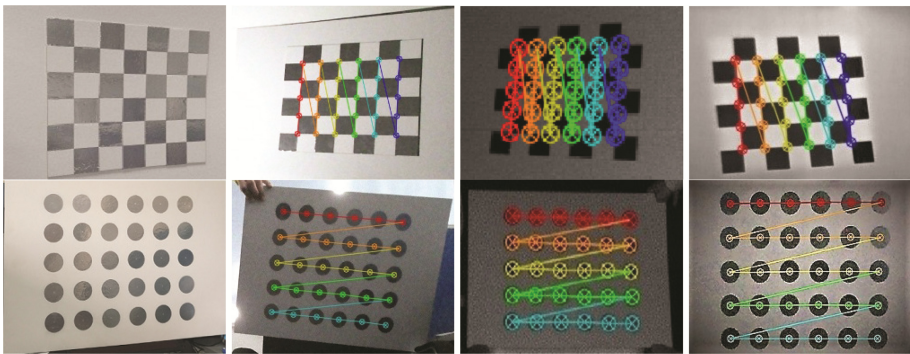
Fig. 1. Calibration boards and the results of calibration. First column: square and circle planar pattern calibration boards. Second column: RGB camera. Third column: near-infrared (depth) camera. Last column: thermal camera.

Table 1 shows a comparison of the proposed calibration patterns and the pattern presented by Skala et al. [13]. From Table 1 we can see that our method is more accurate than the method from Skala et al. Besides, the MRE of circle pattern is less than square pattern. It is suggested that the circle pattern is more accurate than the square pattern.

Table 1. The MRE of different calibration patterns.

Cameras Types	MRE of Circle	MRE of Square	MRE of Square with holes [13]
Thermal Camera	0.0183	0.0230	0.1312
Depth Camera	0.0272	0.0281	0.1861
Color Camera	0.0294	0.0334	0.1464

The main reason of affecting calibration precision is the precise localization of feature points. The localization of feature points is very difficult when the calibration board is non-front parallel to the camera imaging plane [21]. The calibration board is usually not parallel to the camera imaging plane. As a result, the square in the photo becomes a diamond and the circle becomes an ellipse. The algorithm detecting the center of a circle is divided into the three steps. Firstly, an input image is converted into a set of binary images by a series of continuous threshold values. Secondly, the connected region of each binary image is extracted by detecting their boundaries. The connected region is the blob of corresponding a binary image. Finally, the blob is fitted to a circle and the center of the circle can be calculated. And it can also detect the center of an ellipse by fitting the ellipse edge. The method is very simple and efficient in detecting the center of a circle or an ellipse. For square pattern, the algorithm is detecting the corners of a chessboard. When the square in the photo becomes a diamond, the localization error of feature points becomes larger. According to Table 1, the calibration accuracy of circle pattern is higher than square's. So we adopted the circle pattern calibration board to calibrate cameras.

3.2 Stereo Calibration of Thermal and RGBD Cameras

In order to obtain fusion information from Kinect and the thermal camera, it is require to make them in a common geometric coordinate system. The process of acquiring the rotation and translation relationships between different cameras is known as stereo calibration. The new calibration board that is presented in this paper can calibrate thermal camera, RGB camera and depth camera. Thus we are able to use calibration-based to carry out stereo calibration.

According to pinhole projection model, a spatial point in scene can be project onto the camera image plane. And there is an infinite number of spatial points can be project onto the same point on the camera image plane. All these spatial points lie on a straight line that goes through the camera center. The points on the camera image plane can't be reprojected into spatial scene because they lost the distance information. Nevertheless, the pixels of a depth image represent the distance between the spatial object and the depth camera. So the depth information is able to be used to reproject points from the camera image plane to the spatial scene.

$$z_d p_{depth} = K_{depth} \left[R_{depth} \ t_{depth} \right] P \tag{4}$$

$$z_t p_{thermal} = K_{thermal} \left[R_{thermal} \ t_{thermal} \right] P \tag{5}$$

Where P is the point in a 3D space scene, p_{depth} and $p_{thermal}$ are the projection points from P to the image plane of the depth camera and the thermal camera. K_{depth} and $K_{thermal}$ are the intrinsic matrixes of the depth camera and the thermal camera. $\left[R_{depth} \ t_{depth} \right]$ and $\left[R_{thermal} \ t_{thermal} \right]$ are the extrinsic matrixes of the depth camera and the thermal camera, including the rotation matrix and the translation vector. z represents the distance from P

to camera in the camera coordinate system. The projection point P_{depth} from P in the camera coordinate system can be written as:

$$P_{depth} = z_d K_{depth}^{-1} p_{depth} = R_{depth} P + t_{depth} \tag{6}$$

The spatial position relationship between the thermal camera and the depth camera can be written as:

$$P_{thermal} = R P_{depth} + T \tag{7}$$

Where R is the rotation relationship between two cameras, T is the translation relationship between two cameras. The following equation can be acquired by combining Eqs. (4), (5) and (6).

$$P_{thermal} = R_{thermal} R_{depth}^{-1} P_{depth} + t_{thermal} - R_{thermal} R_{depth}^{-1} t_{depth} \tag{8}$$

Comparing Eqs. (7) and (8), R and T are able to be calculated.

$$R = R_{thermal} R_{depth}^{-1} \tag{9}$$

$$T = t_{thermal} - R t_{depth} \tag{10}$$

From Eqs. (9) and (10), we can come to a conclusion that we only need to obtain the extrinsic matrixes of the thermal camera and the depth camera in the same scene, and then the transformation matrixes of two cameras can be calculated. Similarly, the relationship between visible and depth cameras can obtain using the same method.

4   Synchronization

It is very important for synchronization of a multi camera system to integrate various information. The authors in [12] adopted motion statistic alignment to synchronize cameras. They used the TV-L1 optical flow to calculate the average flow vector magnitude of RGBD and thermal images streams. However, it cannot be used in low-frame cameras and the thermal camera with non-uniformity corrections (NUCs). In general, most thermal cameras need regular intensity correction in order to increase the accuracy of receiving the infrared radiation signal, which is known as NUCs. When the thermal camera performs NUCs, a baffle of uniform temperature will temporarily block the imaging sensor. Therefore, this leads to a data stream interruption about 1.5 to 2 s. NUCs are significant for thermal images. If a thermal camera doesn't have NUCs, the pixel values of thermal images can't represent an accurate estimate of infrared radiance in the scene. On the other hand, the sampling frequency of Kinect and thermal camera is different, which maybe cause an infrared photo can't match a corresponding depth photo.

In order to solve above problems, we improved the method from [22]. We used the nearest adjacent time between an infrared image and a depth image to quickly find the corresponding frames between the Kinect and the thermal camera. The RGB and depth

image frames from the Kinect are already synchronized. It only need to synchronize the depth and thermal cameras. We first obtain the timestamps of depth and thermal frame streams $\{t_n^d, n = 1, 2, 3 \cdots\}$ and $\{t_m^{th}, m = 1, 2, 3 \cdots\}$. After that, we calculate the time difference between depth frames and corresponding thermal frames:

$$\Delta t = \min\left\{ \left| t_{n-1}^d - t_m^{th} \right|, \left| t_n^d - t_m^{th} \right|, \left| t_{n+1}^d - t_m^{th} \right| \right\} \tag{11}$$

The frame rate of the Kinect is almost 30 frames per second and frame rate of the thermal camera is almost 20 frames per second. So every infrared frame maybe correspond more than one depth frames. Considering that, we select the three depth frames adjacent to the infrared frame to calculate and keep the minimum time difference as Δt. The threshold σ represents half of the inter-frame interval of the camera with the slowest frame rate. So the value of σ is set to 25 ms. If Δt is less than the threshold σ, the depth frame and corresponding thermal frame are the synchronized frames and we will keep two fames. Otherwise, we will abandon them temporarily. Considering NUCs of thermal camera, many depth frames don't have corresponding thermal frames during NUCs. When Δt exceeds 1.5 s, we can consider that the thermal camera is carrying out NUCs. Thus, the depth frame should be abandoned and the corresponding infrared frame is kept to search next depth frames. If Δt is greater than 25 ms but no more than 1.5 s, relative infrared and depth frames are both abandoned.

5   Experiment and Evaluation

In our experiment, we adopted the second generation Kinect as RGBD camera and FLIR T420 as thermal camera. We first carried out the experiment of camera calibration including single camera calibration and stereo camera calibration. For the visible camera, the calibration needs to be carried out in a weak light environment. And a light source needs to be arranged on one side of the calibration board. After we obtained the intrinsic parameters and extrinsic parameters of cameras, we registered depth images with infrared images and RGB images. Figure 2 shows the results of registration. We can find that results are particularly desirable and indicate that we can obtain an accurate camera calibration using our calibration board.

Fig. 2. The results of registration. First row: left column is original infrared images, middle column is original depth images, right column is registered images. Second row: left column is original color images, middle column is original depth images, right column is registered images.

And then we evaluated the efficacy of synchronization using infrared and depth image frames. Figure 3 shows the offset between thermal camera and the depth camera. Figure 4 shows the results of synchronization. We selected a key frame as a start frame. From Fig. 3, as time goes by, the offset is increasing because of different frame rates between two cameras. We can clearly see two curve pulses, which are NUCs operation of the thermal camera. In Fig. 4, the depth frames those in NUCs time are abandoned. Comparing with pictures from Figs. 3 and 4, the nearest adjacent time has an ideal effect on synchronization.
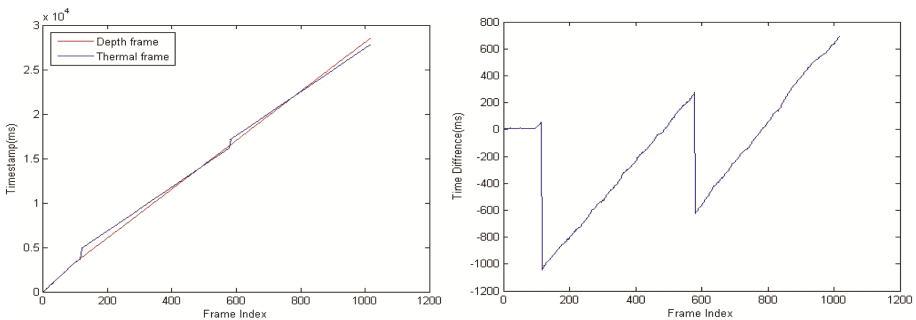
Fig. 3. The offset between the thermal camera and the depth camera. Left image shows the timestamps of two cameras. Right image shows the time difference of thermal frames and depth frames.
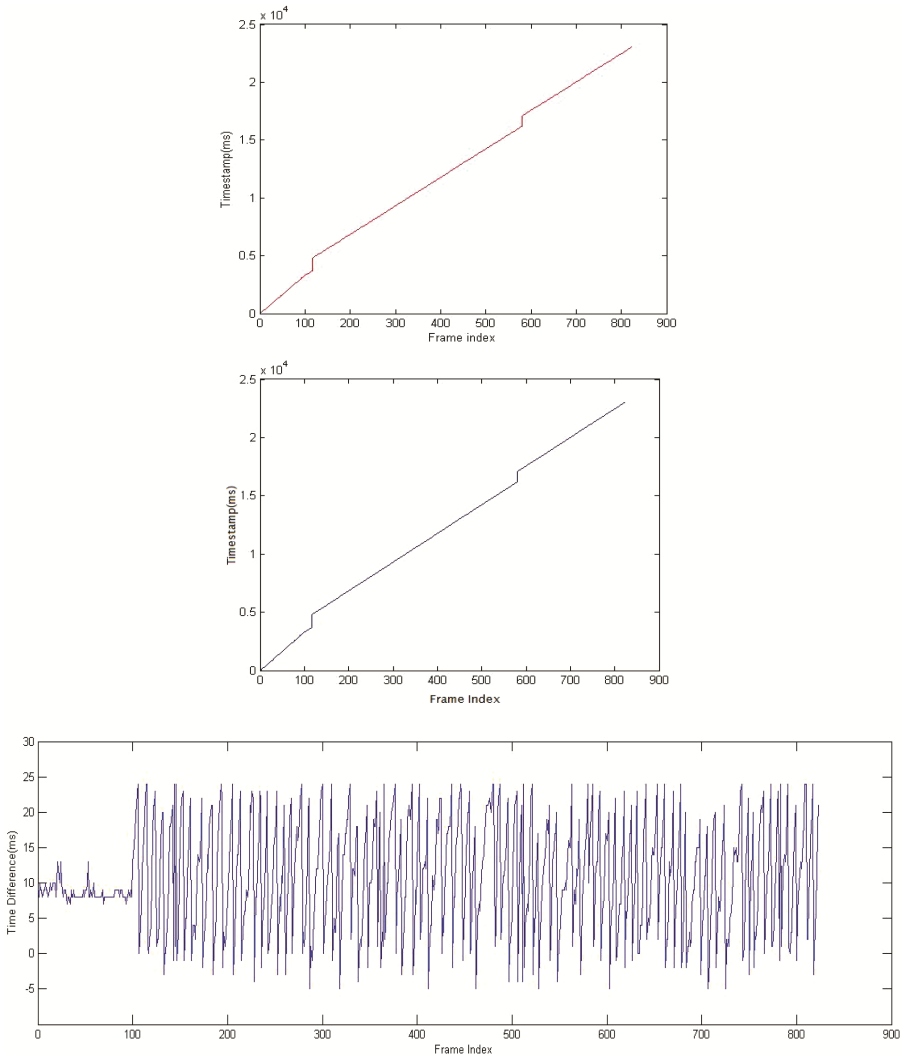
Fig. 4. The results of synchronization using our method. First row: left image shows the result of depth frames after synchronization; right image shows the result of thermal frames after synchronization. Second row shows the time difference between thermal frames and depth frames after synchronization.

6   Conclusion

In this paper we presented novel methods for geometric calibration and temporal calibration of the Kinect and the thermal camera. We designed a new calibration board for camera calibration, which can be used to calibrate thermal camera, depth camera and RGB camera at the same time. Comparing with existing methods that require high cost

and complex operation, our calibration board is easy to product and use. The calibration board only consists of aluminum foil and a cardboard, so the manufacturing cost is lower than other kinds of calibration boards. In addition, the method of the nearest adjacent time we adopted can synchronize the image streams from a multi cameras system. By calculating the time difference and eliminating invalid frames, the method of the nearest adjacent time can solve the problem of NUCs quickly.

In the future, we will focus on the fusion of depth datas, thermal datas and RGB datas from the Kinect and the thermal camera. Besides, we will study indoor three dimensional temperature field reconstruction. In addition, we will explore a real time reconstruction system using the Kinect and the thermal camera.

References

1. Zennaro, S., Munaro, M., Milani, S., Zanuttigh, P., Bernardi, A., Ghidoni, S., et al.: Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications. In: International Conference on Multimedia and Expo, pp. 1–6. IEEE, Turin (2015)
2. Butkiewicz, T.: Low-cost coastal mapping using Kinect v2 time-of-flight cameras. In: Oceans, pp. 1–9. IEEE, St. John's (2015)
3. Gonzalez-Jorge, H., Rodríguez-Gonzálvez, P., Martínez-Sánchez, J., González-Aguilera, D., Arias, P., Gesto, M., et al.: Metrological comparison between kinect I and kinect II sensors. Measurement **70**, 21–26 (2015)
4. Corti, A., Giancola, S., Mainetti, G., Sala, R.: A metrological characterization of the Kinect V2 time-of-flight camera. Robot. Auton. Syst. **75**(8), 584–594 (2015)
5. Lachat, E., Macher, H., Landes, T., Grussenmeyer, P.: Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. Remote Sens. **7**(10), 13070–13097 (2015)
6. Moghadam, P., Vidas, S., Lam, O.: Spectra: 3D multispectral fusion and visualization toolkit. In: Australasian Conference on Robotics and Automation. ACRA, Melbourne (2014)
7. Borrmann, D., Nüchter, A., Djakulovi'C, M., Maurovi'C, I., Petrovi'C, I., Osmankovi'C, D., et al.: The Project thermalmapper-thermal 3D mapping of indoor environments for saving energy. Ifac Symp. Robot. Control **45**(22), 31–38 (2012)
8. Nauer, C., Vonach, E., Gerstweiler, G., Kaufmann, H.: 3D building reconstruction and thermal mapping in fire brigade operations. In: Augmented Human International Conference, pp. 1–2. IEEE, Lake Buena Vista (2013)
9. Borrmanna, D., Nüchtera, A., Đakulovic, M., Maurovic, I., Petrovic, I., Osmankovic, O., et al.: A mobile robot based system for fully automated thermal 3D mapping. Adv. Eng. Inform. **28**(4), 425–440 (2014)
10. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)
11. Vidas, S., Moghadam, P., and Bosse, M. (2013).: 3D thermal mapping of building interiors using an RGB-D and thermal camera. In: International Conference on Robotics and Automation, pp. 2311–2318. IEEE, Karlsruhe (2013)

12. Lussier, J. T., Thrun, S.: Automatic calibration of RGBD and thermal cameras. In: International Conference on Intelligent Robots and Systems, pp. 451–458. IEEE, Chicago (2014)
13. Skala, K., Lipić, T., Sović, I., Gjenero, L., Grubišić, I.: 4D thermal imaging system for medical applications. Periodicum Biologorum **113**(4), 407–416 (2011)
14. Ham, Y., Golparvar-Fard, M.: An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery. Adv. Eng. Inform. **27**(3), 395–409 (2013)
15. Kim, N., Choi, Y., Hwang, S., Park, K.: Geometrical calibration of multispectral calibration. In: International Conference on Ubiquitous Robots and Ambient Intelligence, pp. 384–385. IEEE, Goyang (2015)
16. Gschwandtner, M., Kwitt, R., Uhl, A., Pree, W.: Infrared camera calibration for dense depth map construction. In: Intelligent Vehicles Symposium, pp. 857–862. IEEE, Baden-Baden (2011)
17. St-Laurent, L., Prévost, D., Maldague, X.: Fast and accurate calibration-based thermal/color sensors registration. In: 10th International Conference on Quantitative Infrared Thermography, QIRT, Québec, Canada, pp. 27–30 (2010)
18. Hwang, S., Choi, Y., Kim, N., Park, K.: Low-cost synchronization for multispectral cameras. In: International Conference on Ubiquitous Robots and Ambient Intelligence, pp. 435–436. IEEE, Goyang (2015)
19. Baar, J.V., Beardsley, P., Pollefeys, M., Gross, M.: Sensor fusion for depth estimation, including TOF and thermal sensors. In: 2nd International Conference on 3d Imaging, Modeling, Processing, Visualization and Transmission, pp. 472–478. IEEE, Zurich (2012)
20. Miller, S., Teichman, A., Thrun, S.: Unsupervised extrinsic calibration of depth sensors in dynamic scenes. In: Intelligent Robots and Systems, pp. 2695–2702. IEEE, Tokyo (2013)
21. Datta, A., Kim, J.S., Kanade, T.: Accurate camera calibration using iterative refinement of control points. In: 12th Computer Vision Workshops, pp. 1201–1208. IEEE, Kyoto (2009)
22. Bahnsen, C.: Thermal-visible-depth image registration (2017). http://projekter.aau.dk/projekter/en/studentthesis/termiskvisueldybde-billedregistering(8564c183-bdc1-4a9c-8bda-c71878cb0814).html