Human Tracking Using a Far-Infrared Sensor Array and a Thermo-Spatial Sensitive Histogram

Takashi Hosono[1(✉)], Tomokazu Takahashi[2] , Daisuke Deguchi[3], Ichiro Ide[1], Hiroshi Murase[1], Tomoyoshi Aizawa[4], and Masato Kawade[4]

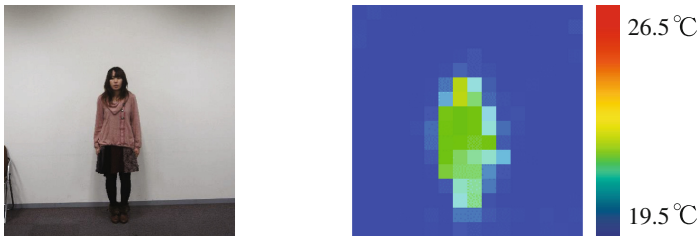
[1] Graduate School of Information Science, Nagoya University, Nagoya, Japan
hosonot@murase.m.is.nagoya-u.ac.jp,
{ide,murase}@is.nagoya-u.ac.jp
[2] Faculty of Economics and Information,
Gifu Shotoku Gakuen University, Gifu, Japan
ttakahashi@gifu.shotoku.ac.jp
[3] Information and Communications Headquarters, Nagoya University, Nagoya, Japan
ddeguchi@nagoya-u.jp
[4] Corporate R&D, OMRON Corporation, Kyoto, Japan
{aizawa,kawade}@ari.ncl.omron.co.jp

Abstract. We propose a human body tracking method using a far-infrared sensor array. A far-infrared sensor array captures the spatial distribution of temperature as a low-resolution image. Since it is difficult to identify a person from the low-resolution thermal image, we can avoid privacy issues. Therefore, it is expected to be applied for the analysis of human behaviors in various places. However, it is difficult to accurately track humans because of the lack of information sufficient to describe the feature of the target human body in the low-resolution thermal image. In order to solve this problem, we propose a thermo-spatial sensitive histogram suitable to represent the target in the low-resolution thermal image. Unlike the conventional histograms, in case of the thermo-spatial sensitive histogram, a voting value is weighted depending on the distance to the target's position and the difference from the target's temperature. This histogram allows the accurate tracking by representing the target with multiple histograms and reducing the influence of the background pixels. Based on this histogram, the proposed method tracks humans robustly to occlusions, pose variations, and background clutters. We demonstrate the effectiveness of the method through an experiment using various image sequences.

1 Introduction

In order to analyze human behavior, visible-light cameras are widely adopted for vision systems. However, they may cause uncomfortableness to the users because it could be easy to identify a person from the captured images. No matter how useful vision systems are, it is often difficult to install them in locations where users refuse them, e.g. toilets, bedrooms, and offices. Considering this problem, using systems based on sensitive floor [1,2] and far-infrared sensors network [3,4] could allow human tracking avoiding privacy issues. The former uses pressure

C.V. Jawahar and S. Shan (Eds.): ACCV 2014 Workshops, Part II, LNCS 9009, pp. 262–274, 2015.
DOI: 10.1007/978-3-319-16631-5_20



(a) Visible-light image. (b) Low-resolution thermal image.

Fig. 1. Example of an output of a 16 × 16 far-infrared sensor array. It is very difficult to identify the person from (b). Note that, this far-infrared sensor array outputs temperature lower than the actual body temperature. (a) is captured by a visible light camera put on the same position as the far-infrared sensor array.

sensitive floors that can detect human footsteps. The latter uses far-infrared sensors that are distributed and connected to each other by network. However, these solutions are not suitable for interactive systems such as gesture recognition because they cannot capture human body shapes.

A far-infrared sensor array is a device composed of a small number of far-infrared sensors, that is expected to be used for interactive systems [5]. Figure 1 shows an example of an output obtained from a 16 × 16 far-infrared sensor array. As we can see from this figure, it represents the spatial distribution of temperature as a low-resolution image. Since the image only shows the rough shape of an object, unlike visible-light cameras, far-infrared sensor arrays would cause less uncomfortableness to users because they are not easily identified. Accordingly, they can be installed in places where it is not preferable to install visible-light cameras due to privacy issues. In addition, far-infrared sensor arrays can capture the rough shape of a human body. Thus, they can also be used in an interactive system such as that for gesture recognition [6]. For the above reasons, there is a demand for a human behavior analysis system that uses a far-infrared sensor array. Among important technologies required to realize the human behavior analysis, in this paper, we propose a human body tracking method using a far-infrared sensor array.

There are some researches related to human body tracking using a far-infrared sensor array. Takahata *et al.* developed a method which localizes heat sources using a far-infrared sensor array [7]. This method simply selects a sensor which outputs the highest temperature to localize a heat source. Thus, it does not consider occlusions nor the effect of other heat sources. Meanwhile, various tracking algorithms which use visible-light cameras are developed until today. In general, tracking algorithms aim to extract and describe the feature of the target object efficiently. The feature points extraction approach is widely used in many tracking algorithms [8–10]. This approach extracts feature points from the target objects. This allows fast tracking with high accuracy by using only points effective for tracking the target object. However, in case of far-infrared sensor arrays, feature points extraction is difficult because the output thermal image is

in very low-resolution. The intensity histograms approach is also widely used to represent a target object. This feature is simple and robust to pose variation. In this approach, multiple local histograms [11] that are made from multiple regions in the target object is used in the state-of-the-art methods [12–14] because a single histogram could not preserve spatial information. These methods are robust to occlusion and can also be computed fast by using the integral histogram [15]. However, if the method using multiple local histograms is applied to the far-infrared sensor array output, each region becomes very small, e.g. 2×2 pixels. For the above reason, a feature that could represent the target object in low-resolution thermal images is needed.

The proposed method introduces a thermo-spatial sensitive histogram algorithm as a feature that represents the target object in a low-resolution thermal image. Unlike the conventional histograms, in case of the thermo-spatial sensitive histogram, a voting value is weighted depending on the distance to the target's position and the difference from the target's temperature. The proposed method constructs the histograms for multiple positions in the target human body region without reducing the number of pixels to be considered as in multiple local histograms [11]. In addition, this histogram reduces the influence of the background pixels by focusing to a particular temperature. Thus, this histogram can effectively represent the target human body in a low-resolution thermal image. Using this histogram, the proposed method can track humans robustly to occlusions, pose variations, and background clutters.

The remainder of this paper is organized as follows: Sect. 2 describes the thermo-spatial sensitive histogram. Next, Sect. 3 describes the human tracking algorithm using the thermo-spatial sensitive histogram. Then, results of an experiment to confirm the effectiveness of the proposed method are reported in Sect. 4. Finally, we conclude this paper in Sect. 5.

2 Thermo-Spatial Sensitive Histogram

Conventionally, when a histogram $\mathbf{H}$ is made from the whole image, the value of bin b ($b = 1, \ldots, B$) is defined as:

$$\mathbf{H}(b) = \sum_{i=1}^{N} Q(T_{\boldsymbol{x}_i}, b), \tag{1}$$

where N is the number of pixels, $\boldsymbol{x}_i$ is the location of the i-th pixel, B is the total number of bins, and $Q(T_{\boldsymbol{x}_i}, b)$ returns 1 if the intensity value $T_{\boldsymbol{x}_i}$ belongs to bin b, otherwise returns 0. In general, this histogram is created from the region of the target object. In case of tracking, histogram largely changes if occlusions occur because it describes the object with only one histogram.

To resolve this issue, He *et al.* developed the locality sensitive histogram [16] that preserves spatial information. In this paper, we call this histogram "spatial sensitive histogram" to distinguish it from other histograms introduced later.

Let $\mathbf{H}_{\boldsymbol{x}_p}$ denote the spatial sensitive histogram computed at a target pixel $\boldsymbol{x}_p$. It can be written as:

$$\mathbf{H}_{\boldsymbol{x}_p}(b) = \sum_{i=1}^{N} \alpha_1^{||\boldsymbol{x}_p - \boldsymbol{x}_i||} \cdot Q(T_{\boldsymbol{x}_i}, b), \tag{2}$$

where $0 < \alpha_1 < 1$ is a parameter that controls the influence of the distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_p$. In this way, this histogram can preserve the spatial information. By using multiple histograms constructed while moving $\boldsymbol{x}_p$ in the target region, tracking robust to occlusion can be achieved.

Spatial sensitive histogram can partially reduce the influence of the background pixels. In tracking, reducing the influence of the background pixels is very important because those around the target object change frame by frame, and degrades the tracking accuracy. In order to reduce the influence of the background further, we propose the "thermal sensitive histogram". In case of visible-light cameras, human bodies are observed as a set of pixels with various intensity values due to the variety of their appearance. Meanwhile, a far-infrared sensor array captures human bodies as similar temperature because it captures the object's temperature. Therefore, this histogram is constructed by voting values weighted by the distance to the target object's temperature. Let $\mathbf{H}^{T_m}$ denote the thermal sensitive histogram computed when the target object's temperature is T_m. It can be written as:

$$\mathbf{H}^{T_m}(b) = \sum_{i=1}^{N} \alpha_2^{|T_m - T_{\boldsymbol{x}_i}|} \cdot Q(T_{\boldsymbol{x}_i}, b), \tag{3}$$

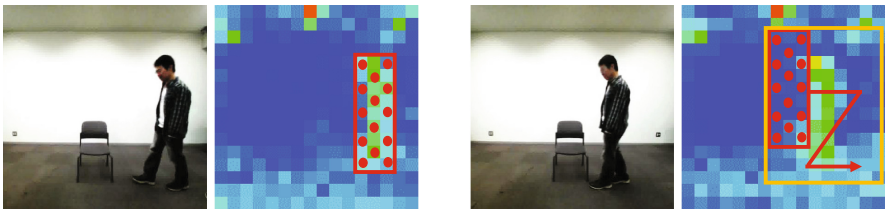
where T_m is the target object's temperature calculated by selecting the median of pixel values in the bounding box surrounding the human body, and $0 < \alpha_2 < 1$ is a parameter that controls the influence of the distance to the target object's temperature. In general, the temperature of the region that is occluded by an object is observed lower than that of the human body, and that of humans who are in front of the target human are observed higher due to the characteristic of the sensor. Thus, this histogram can reduce the influence of background pixels. However, it could not preserve the spatial information like in the conventional histogram. Therefore, we propose a "thermo-spatial sensitive histogram" that combines the above two histograms. The value of bin b of a thermo-spatial sensitive histogram $\mathbf{H}^{T_m}_{\boldsymbol{x}_p}$ made at pixel $\boldsymbol{x}_p$ is defined as:

$$\mathbf{H}^{T_m}_{\boldsymbol{x}_p}(b) = \sum_{i=1}^{N} \alpha_1^{||\boldsymbol{x}_p - \boldsymbol{x}_i||} \cdot \alpha_2^{|T_m - T_{\boldsymbol{x}_i}|} \cdot Q(T_{\boldsymbol{x}_i}, b), \tag{4}$$

The proposed method constructs multiple histograms while moving $\boldsymbol{x}_p$ in the target region like the method that uses the spatial sensitive histogram [16].

3 Human Tracking by a Far-Infrared Sensor Array

We assume that a target human body region in the initial frame is given manually as a rectangular region. First, multiple thermo-spatial sensitive histograms



(a) Template frame. (b) Current frame.

Fig. 2. Tracking using thermo-spatial sensitive histograms. (a) shows the template thermal image and a visible-light image taken at the same time. (b) is an example of the current frame of the thermal image and a visible-light image taken at the same time. The red rectangles represent the template and candidate regions. The red points indicate the sampling points to make the thermo-spatial sensitive histograms. The orange rectangle represents the search region (Color figure online).

are made at a number of sampling points in the target region in the initial frame as template histograms. Next, in the succeeding frame, the region that maximizes the similarity to the template is searched by a sliding window approach. Then, the template histograms are updated to adopt the change of the target's appearance. Details of each step are described below.

3.1 Tracking Using Thermo-Spatial Sensitive Histograms

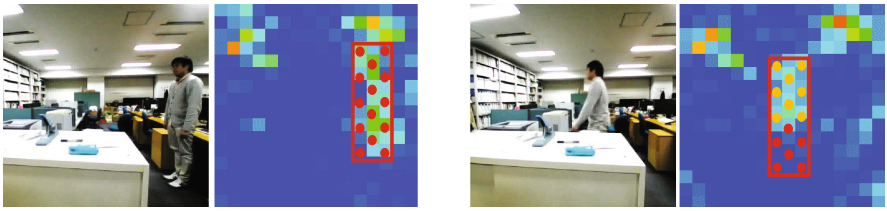
The proposed method makes the thermo-spatial sensitive histograms for each sampling point in the target region. This means the proposed method represents the target object with multiple histograms. The size of the target region and the relative position of the sampling points to the region are fixed, to compare each histogram to the template histogram. These histograms represent the target object region with multiple overlapping regions.

First, the temperature of the target human body represented as T_m in Eq. (4) is calculated as the median of pixel values in the region. In this way, it is possible to select the temperature of the human body without being influenced by background pixels in the region. To find a target human body region in the current frame, we define the similarity between the template and candidate regions as:

$$S(R_t, R_c) = \sum_{i=1}^{N} I(\mathbf{H}_{\boldsymbol{x}_i}^{T_m}, \mathbf{H}_{\boldsymbol{x}'_i}^{T_m}), \tag{5}$$

where R_t and R_c represent the template and candidate regions, respectively, N is the number of sampling points in the region, $\mathbf{H}_{\boldsymbol{x}_i}^{T_m}$ is the i-th histogram in the template and $\mathbf{H}_{\boldsymbol{x}'_i}^{T_m}$ is the i-th histogram in the candidate. To measure the similarity between the histograms, we use the histogram intersection function [17] defined as:

$$I(\mathbf{H}_{\boldsymbol{x}_i}^{T_m}, \mathbf{H}_{\boldsymbol{x}'_i}^{T_m}) = \sum_{b=1}^{B} \min(\mathbf{H}_{\boldsymbol{x}_i}^{T_m}(b), \mathbf{H}_{\boldsymbol{x}'_i}^{T_m}(b)). \tag{6}$$



(a) Template frame. (b) Current frame.

Fig. 3. Online template update. The red points in (a) indicate sampling points to make thermo-spatial sensitive histograms, and the orange points in (b) indicate the sampling points of the updated histograms that had stable similarities to the template. In this case, template histograms are updated only on histograms made from the orange points (Color figure online).

Thus, a candidate region R_c that maximizes the similarity $S(R_t, R_c)$ is determined as the human body region in the current frame. Note that candidate regions are restricted to the area around the tracking result of the previous frame to reduce computational cost. Figure 2 shows an example of the tracking procedure, where the red rectangles in (a) and (b) represent the template and candidate regions, respectively, the red points are the sampling points where the thermo-spatial sensitive histograms are made and the orange rectangle in (b) is the area restricted to find the target region in the current frame.

3.2 Online Template Update

In general, the target human's appearance continues to change. Therefore, the template histograms are required to be updated online to achieve stable tracking. The proposed method updates each histogram individually to decrease the effect of posture changes and occlusions. This online template update is performed by:

$$\mathbf{H}^{T_m}_{\boldsymbol{x}_i} = \mathbf{H}^{T_m}_{\boldsymbol{x}'_i} \quad \text{if } F_1 \cdot M < I(\mathbf{H}^{T_m}_{\boldsymbol{x}_i}, \mathbf{H}^{T_m}_{\boldsymbol{x}'_i}) < F_2 \cdot M, \tag{7}$$

where $\mathbf{H}^{T_m}_{\boldsymbol{x}'_i}$ is the histogram after the tracking, M is the median of all histogram similarities between the template and tracking results, and F_1, F_2 are parameters that control the ease of updating. A template histogram is updated only when the histogram similarity is stable by using this algorithm. Thus, this algorithm can avoid being affected by posture changes and occlusions. Figure 3 illustrates the online template update algorithm, where the red rectangles in (a) and (b) represent the template region and the tracking result, respectively, the red points represent the sampling points for the template, and the orange points in (b) are the sampling points whose histograms were updated because the similarity was stable in the current frame. In Fig. 3(b), since the lower body of the person is hidden, the similarity decreases. Therefore, the method updates the template only for the sampling points in the upper body that has stable similarities.

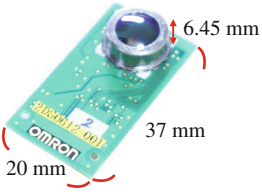


Fig. 4. Far-infrared sensor array used in the experiment. 16 × 16 far infrared sensors are integrated in one chip.

4 Experiment

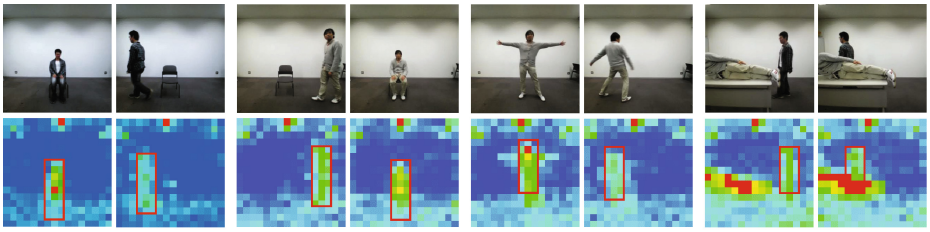
To confirm the effectiveness of the proposed method, we conducted an experiment by using actual data. Improvement in accuracy of the proposed method was evaluated by comparing it with two comparative methods. To conduct the experiment, we created a dataset using a 16 × 16 far-infrared sensor array. Figure 4 shows the far-infrared sensor array used in the experiment. We describe below the dataset and the experimental conditions, then discuss the results from the experiment.

4.1 Creation of a Dataset

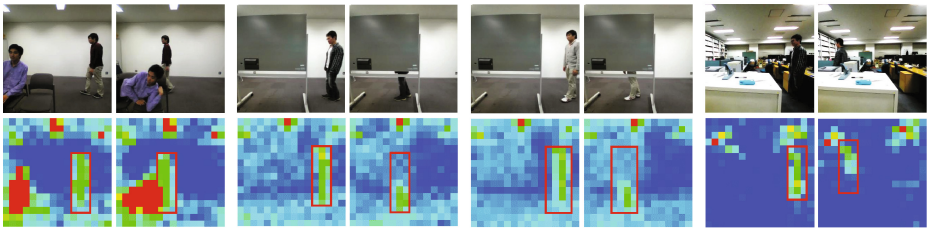
We captured various scenes using the far-infrared sensor array whose frame rate was 10 fps and temperature resolution was 0.1 degrees. The dataset consisted of 13 sequences composed of 1,729 frames in total. This dataset contains sequences that have many posture changes, occlusions, and background clutters. In case of thermal images, the background clutters were caused by various heat sources including humans, PC displays, and high room temperature. In addition, we prepared not only sequences capturing a human as a target, but also complex sequences captured in a room with other heat sources including other humans, in order to confirm that the proposed method can be used in actual situations. Figure 5 shows samples of the 13 sequences in the dataset, where the red rectangles indicate the ground-truth. For each sequence, we captured the whole body of the target human in the first frame and determined the target region manually to construct the initial template histograms. Visible-light image sequences were also captured at the same time in order to provide ground-truth regions frame-by-frame manually. Note that only the thermal image sequences were input to the program. The size of ground-truth regions were fixed in each sequence.

4.2 Experimental Conditions

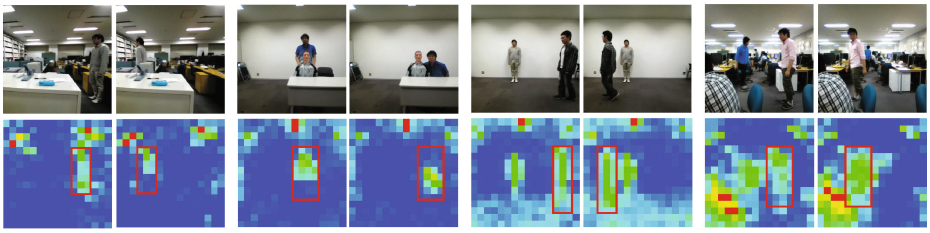
To confirm the effectiveness of using the proposed thermo-spatial sensitive histogram, we compared it with two methods using different types of histograms. One was the spatial sensitive histogram [16] that was used in the state-of-the-art



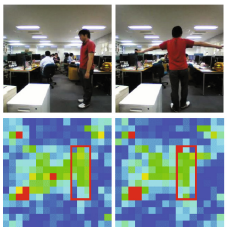
(a) Sitting on a chair 1. (b) Sitting on a chair 2. (c) Extending arms. (d) Occlusion by human 1.



(e) Occlusion by human 2. (f) Occlusion by whiteboard 1. (g) Occlusion by whiteboard 2. (h) Occlusion by locker 1.



(i) Occlusion by locker 2. (j) Occlusion by doll and table. (k) Passing in front of human. (l) Walking in a room 1.



(m) Walking in a room 2.

Fig. 5. Examples of sequences from the dataset used in the experiment. Visible-light images were taken at the same time to provide ground-truth regions. Red rectangles in the thermal images represent the ground-truth regions (Color figure online).

algorithm of tracking with a visible-light camera. Another was the thermal sensitive histogram. All three methods used the same tracking algorithm described in Sect. 3. However, the comparative method that used the thermal sensitive histogram tracked with only a single histogram in its template because the same histogram was constructed for all the sampling points. In our implementation, both parameters α_1 and α_2 in Eqs. (2), (3), and (4) were set to 0.5, F_1, F_2 in Eq. 7 were set to 0.98, 1.02, respectively. In addition, when making multiple histograms from a region, sample points were set at an interval of 1 pixel, and each histogram bin was set at an interval of 1 degree. Furthermore, the candidate regions were restricted within ± 2 pixels in horizontal and vertical directions from the previous tracking result when the human body region was searched in the current frame.

We used two types of criteria to evaluate each method. One was the success rate that is the percentage of the number of frames where the target regions were successfully tracked. We determined that tracking was successful if the tracking result satisfied the PASCAL criterion [18] defined as:

$$\frac{|B^i \cap G^i|}{|B^i \cup G^i|} \geq 0.5, \tag{8}$$

where B^i denotes the region of the tracking result in frame i, and G^i denotes the ground-truth region in frame i. Another evaluation criterion was the center location error that is the L_2 distance of center locations between the tracking result and the ground-truth. By these evaluation criteria, we evaluated the tracking accuracy both discretely and continuously.

4.3 Results and Discussion

Tables 1 and 2 show the tracking performance for the 13 sequences. As shown in this table, the proposed method using the thermo-spatial sensitive histogram provided the best performance for almost all sequences. In the success rate criteria, the average score improved approximately 20 % compared to the comparative methods. Furthermore, in the center location error criteria, the average score improved approximately 0.5 pixels. This may seem a small improvement, but the shift of one pixel in the 16×16 pixels template image is large. Actually, in this experimental setting, the error of 0.5 pixels was equivalent to approximately 0.1 m. From the above, the validity of the proposed method was confirmed quantitatively.

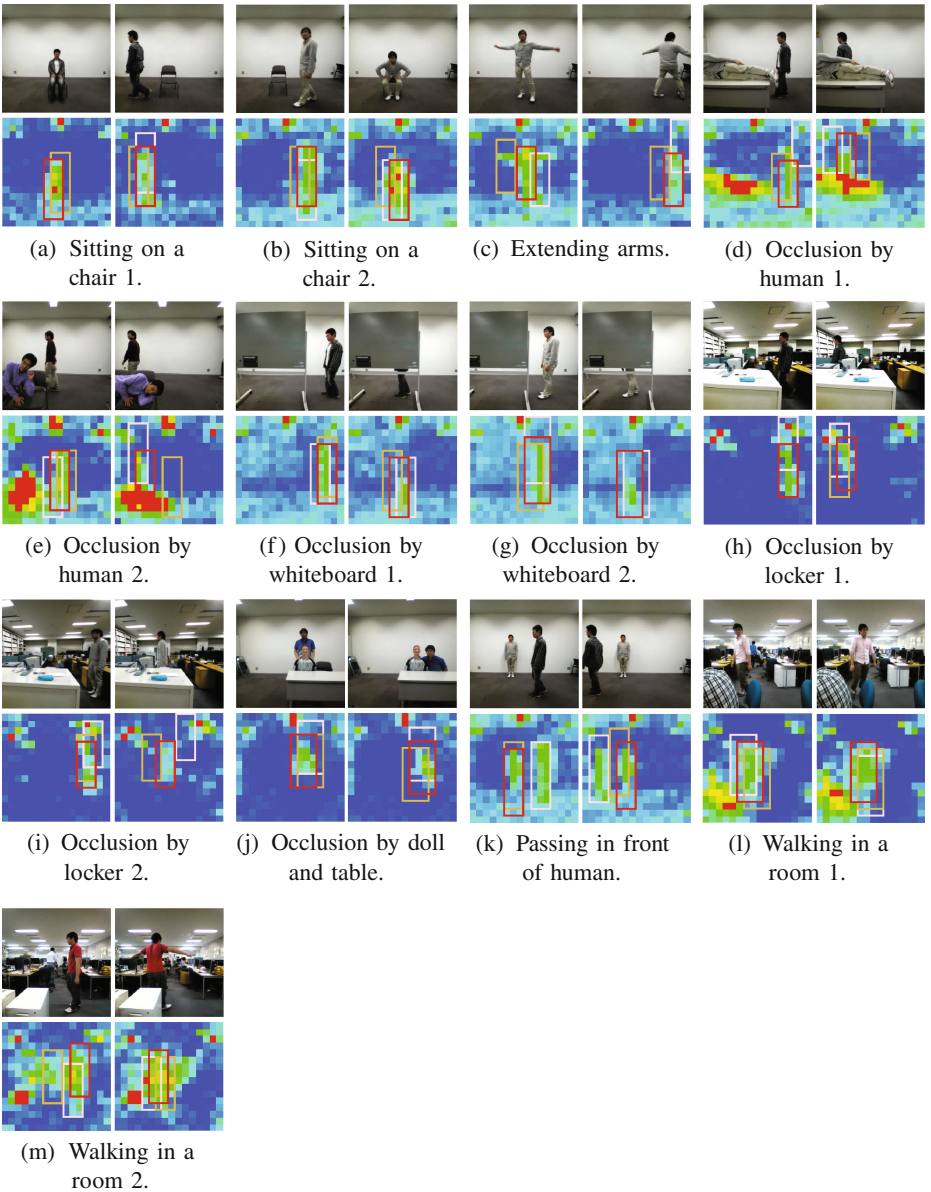
Figure 6 shows the screenshots of the tracking results, where the red, brown and pink rectangles represent the tracking result by the thermo-spatial sensitive histogram, the spatial sensitive histogram, and the thermal sensitive histogram, respectively. Figure 6(a), (b), and (c) show the results of sequences including posture changes. As shown in these figures, the proposed method using the thermo-spatial sensitive histogram can capture the human body stably because putting weight on pixels that had values near human temperature decreased the influence of background pixels which had lower values than the human temperature. Figure 6(d) to (j) show the results of sequences including occlusions.

Table 1. The success rates [%] of the 13 sequences. The highest value among the methods for each sequence is shown in bold. The total number of frames is 1,729.

Sequence	Thermo-Spatial (proposed)	Spatial	Thermal
Sitting on a chair 1	**85**	53	66
Sitting on a chair 2	**82**	12	81
Extending arms	**57**	10	46
Occlusion by human 1	34	**35**	9
Occlusion by human 2	**59**	45	41
Occlusion by whiteboard 1	72	**81**	73
Occlusion by whiteboard 2	**100**	99	81
Occlusion by locker 1	**89**	81	43
Occlusion by locker 2	**88**	48	12
Occlusion by doll and table	**75**	61	26
Passing in front of human	10	19	**35**
Walking in a room 1	**93**	86	76
Walking in a room 2	**70**	51	47
Average	**70**	52	49

Table 2. The average center location errors [pixels]. The smallest error among the methods for each sequence is shown in bold. The total number of frames is 1,729.

Sequence	Thermo-Spatial (proposed)	Spatial	Thermal
Sitting on a chair 1	**0.7**	1.2	0.9
Sitting on a chair 2	**0.7**	1.9	0.9
Extending arms	**1.3**	3.0	1.9
Occlusion by human 1	**2.1**	2.2	3.7
Occlusion by human 2	**1.4**	2.1	2.5
Occlusion by whiteboard 1	1.1	**0.7**	1.0
Occlusion by whiteboard 2	1.1	**1.0**	1.6
Occlusion by locker 1	**0.6**	0.8	2.2
Occlusion by locker 2	**1.0**	1.7	3.5
Occlusion by doll and table	**1.1**	1.2	2.6
Passing in front of human	2.9	2.9	**1.8**
Walking in a room 1	**0.6**	0.7	1.2
Walking in a room 2	**1.3**	**1.3**	1.5
Average	**1.2**	1.6	1.9



(a) Sitting on a chair 1. (b) Sitting on a chair 2. (c) Extending arms. (d) Occlusion by human 1.

(e) Occlusion by human 2. (f) Occlusion by whiteboard 1. (g) Occlusion by whiteboard 2. (h) Occlusion by locker 1.

(i) Occlusion by locker 2. (j) Occlusion by doll and table. (k) Passing in front of human. (l) Walking in a room 1.

(m) Walking in a room 2.

Fig. 6. Example of tracking results. The red, brown, and pink rectangles represent the tracking results by using thermo-spatial, spatial, and thermal sensitive histograms, respectively.

As shown in these figures, the proposed method can capture the human body accurately despite the occlusions. Especially in Fig. 6(e) and (i), the effect that reduces the influences of other heat sources in constructing the thermo-spatial sensitive histograms was effective. Furthermore, it can be seen that the proposed method can capture the human body stably under situations with background clutters in Fig. 6(l) and (m). From the above, it was demonstrated that the proposed human tracking using thermo-spatial sensitive histograms was robust to posture changes, occlusions, and background clutters. However, as shown in Fig. 6(k), the proposed method failed to track in the sequence where the target human passed in front of another human. In this case, the proposed method tracked the other human mistakenly because there was only a small difference in temperature between the target human and the other human. This problem could be solved by improving the histogram feature to distinguish more detailed differences. In addition, improving the tracking algorithm will also be needed.

5 Conclusion

In this paper, we proposed a human tracking method using a far-infrared sensor array. The proposed method uses the thermo-spatial sensitive histogram that is suitable to represent the human body in low-resolution thermal images. This histogram can reduce the influence of background pixels. Since the histogram can preserve the spatial information, it is possible to describe the feature of the target human body in detail by making multiple histograms for each sampling point in the target region. The experimental results show that the proposed method performed human tracking more robust to posture changes, occlusions, and background clutters than the state-of-the-art tracking method for visual images.

As future work, we will develop a tracking method which can track a human under a situation that several humans cross. In addition, we will develop a method which can also deal with scale change due to various distances between the human and the far-infrared sensor array. Furthermore, we will investigate interactive applications such as gesture recognition that uses far-infrared sensor arrays.

Acknowledgment. Parts of this research were supported by MEXT, Grant-in-Aid for Scientific Research.

References

1. Sousa, M., Techmer, A., Steinhage, A., Lauterbach, C., Lukowicz, P.: Human tracking and identification using a sensitive floor and wearable accelerometers. In: Proceedings of the 11th IEEE International Conference on Pervasive Computing and Communications, pp. 166–171 (2013)
2. Steinhage, A., Lauterbach, C.: Monitoring movement behavior by means of a large area proximity sensor array in the floor. In: Proceedings of the 2nd Workshop on Behaviour Monitoring and Interpretation, pp. 15–27 (2008)

3. Hao, Q., Brady, D., Guenther, B.D., Burchett, J., Shankar, M., Feller, S.: Human tracking with wireless distributed pyroelectric sensors. IEEE Sens. J. **6**, 1683–1696 (2006)
4. Zappi, P., Farella, E., Benini, L.: Tracking motion direction and distance with pyroelectric IR sensors. IEEE Sens. J. **10**, 1486–1494 (2010)
5. Ohira, M., Koyama, Y., Aita, F., Sasaki, S., Oba, M., Takahata, T., Shimoyama, I., Kimata, M.: Micro mirror arrays for improved sensitivity of thermopile infrared sensors. In: Proceedings of the 24th IEEE International Conference on Micro Electro Mechanical Systems, pp. 708–711 (2011)
6. Wojtczuk, P., Armitage, A., Binnie, T., Chamberlain, T.: PIR sensor array for hand motion recognition. In: Proceedings of the 2nd International Conference on Sensor Device Technologies and Applications, pp. 99–102 (2011)
7. Takahata, A., Shimada, Y., Yoshioka, F., Yoshida, M., Kimata, M., Ota, T.: Infrared position sensitive detector (IRPSD). In: Infrared Technology and Applications XXXIV, Proceedings of the SPIE, vol. 6940, pp. 694031-1–694031-11 (2008)
8. Baker, S., Matthews, I.: Lucas-Kanade 20 years on: a unifying framework. Int. J. Comput. Vis. **56**, 221–255 (2004)
9. Song, D., Zhao, B., Tang, L.: A tracking algorithm based on SIFT and Kalman filter. In: Proceedings of the 2012 International Conference on Computer Application and System Modeling, pp. 1563–1566 (2012)
10. Yan, Y., Wang, J., Li, C., Wu, Z.: Object tracking using SIFT features in a particle filter. In: Proceedings of the 3rd IEEE International Conference on Communication Software and Networks, pp. 384–388 (2011)
11. Adam, A., Rivlin, E., Shimshoni, I.: Robust fragments-based tracking using the integral histogram. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 798–805 (2006)
12. Cehovin, L., Kristan, M., Leonardis, A.: An adaptive coupled-layer visual model for robust visual tracking. In: Proceedings of the 2011 IEEE International Conference on Computer Vision, pp. 1363–1370 (2011)
13. Kwon, J., Lee, K.M.: Tracking of a non-rigid object via patch-based dynamic appearance modeling and adaptive basin hopping monte carlo sampling. In: Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1208–1215 (2009)
14. Shahed Nejhum, S.M., Ho, J., Yang, M.H.: Visual tracking with histograms and articulating blocks. In: Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
15. Porikli, F.: Integral histogram: A fast way to extract histograms in Cartesian spaces. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 829–836 (2005)
16. He, S., Yang, Q., Lau, R.W.H., Wang, J., Yang, M.: Visual tracking via locality sensitive histograms. In: Proceedings of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2427–2434 (2013)
17. Barla, A., Odone, F., Verri, A.: Histogram intersection kernel for image classification. In: Proceedings of the 2003 IEEE International Conference on Image Processing, vol. 3, pp. 513–516 (2003)
18. Everingham, M., Van Gool, L., Williams, C., Winn, J., Zisserman, A.: The PASCAL visual object classes (VOC) challenge. Int. J. Comput. Vis. **88**, 303–338 (2010)