

Dynamic Offset Correction for Smartphone Thermal Cameras Using a Wristband Sensor

Hiroki Yoshikawa, Akira Uchiyama, Teruo Higashino
 Graduate School of Information Science and Technology, Osaka University
 {h-yoshikawa, uchiyama, higashino}@ist.osaka-u.ac.jp

Abstract—Thermal images are widely used to various health-care applications. However, thermal images captured by smartphone thermal cameras have insufficient accuracy to monitor human body temperature. In this paper, we propose an offset correction method for thermal images captured by smartphone thermal cameras. We fully utilize the characteristic which is specific to thermal cameras: the relative temperatures in a single thermal image are highly reliable although the absolute temperatures fluctuate frequently. Our method combines thermal images with a reliable absolute temperature obtained by a wristband sensor based on the above characteristic. The evaluation result shows that the mean absolute error and the standard deviation of face temperature decrease by 26.2% and 70.1%, respectively, highlighting the effectiveness of the proposed method.

Index Terms—thermal camera, wristband sensor, offset correction, skin temperature, smartphone

I. INTRODUCTION

Human skin temperature is one of the key vital signs to detect heat strokes and infections. In addition, psychological states, for example, cognitive load [1], thermal comfort [2], [3], stress [4] and emotion [5], are estimated by monitoring the skin temperature. Although some wearable devices with skin temperature sensors are available (e.g. E4 wristband [6] and Microsoft band 2 [7]), their applications are still limited due to the limitation of single point measurement. For this reason, thermal cameras are widely used to monitor the skin temperature because they can measure temperature distributions quickly without physical contact. Recently, thermal cameras are easy to use because low-cost and minimized models are available on the market. For example, FLIR ONE [8] is a low-cost thermal camera which is attached to a smartphone. Thanks to such smartphone thermal cameras, we can measure our skin temperature anytime anywhere. However, the accuracy of the low-cost thermal camera is insufficient to monitor the skin temperature compared with the high-end models.

Thermal cameras are categorized into two types: those with cooled infrared detectors and those with uncooled infrared detectors [9]. The performance of the cooled detectors is much higher than the other although they are bulky and expensive due to the cooling apparatus. Therefore, the smartphone thermal cameras are with uncooled infrared detectors. The uncooled infrared detector element converts its temperature rise to electric signals. The temperature of the object principally calculated from the signals and its emissivity. The measured temperature greatly fluctuates in accordance with the parameters configured by the user, the rising temperature

of the camera's body, the efficiency of the element, and the packaging method [9]. By these effects, the measurement error of the smartphone thermal camera is larger than the high end one. For example, the error range is $\pm 3^{\circ}\text{C}$ or $\pm 5\%$, which is larger than the range of human skin temperature changes in daily life. This is clearly not enough for various healthcare applications.

To overcome this problem, we propose an offset correction method for thermal images captured by smartphone thermal cameras. We fully utilize the key feature of the thermal cameras: the measurement fluctuation is mainly caused by the offset which is common in all the pixels in a single thermal image. In other words, we can measure the difference of temperature correctly between any pair of pixels in the same thermal image even using the smartphone thermal camera. Our method combines thermal images with a reliable absolute temperature obtained by a wristband sensor based on the above feature. First, we obtain a thermal image including a reference point (a wrist or a palm) of which absolute temperature is measured by a wearable device (a wristband sensor). Second, we estimate the offset at the reference point by comparing the temperature measured by the thermal camera and the wristband sensor. Finally, our method corrects temperature in the thermal image by adding the offset to all the pixels.

Since the wristband sensor covers the measurement point, the thermal camera cannot directly capture the temperature of the same point. Therefore, we define the reference point as the point which has a high correlation with the point measured by the wristband sensor. In this paper, we use a palm or a wrist for the reference point and compare the performance through the real experiment.

The main contributions of this paper are summarized as below:

- We propose an offset correction method for thermal images measured by a smartphone thermal camera by using a reference point combined with the specific feature of thermal cameras.
- We investigate the design of the reference point correlated with the wrist temperature obtained by the wristband sensor.
- We compare the two reference points and discuss their features through the real data with 871 samples collected from 9 subjects.

II. RELATED WORKS

A. Applications Using Skin Temperature

There are some research works to estimate human mental state using skin temperature. Chernyshov et al. [10] present the system for tracking positive cognitive and emotional states by using temperature sensors on eyeglasses. Genno et al. [4] use facial skin temperature to evaluate stress and fatigue. They revealed the fatigue is the load of the stress by assuming the estimation formula. Understanding the stress of human is a crucial issue in our society. Recently, Japanese companies are obligated to conduct the stress check test for the employees by law. Smartphone thermal cameras are one of the key technologies to record the stress levels ubiquitously, noninvasively and automatically.

B. Applications Using Thermal Camera

Thermal cameras are recently used for the estimation of human thermal comfort. Burzo et al. [3] divide the thermal comfort into three levels: "hot discomfort", "comfort" and "cold discomfort" and combine other biosensors with a thermal camera to estimate the thermal comfort. Ranjan et al. [2] estimate the thermal sense using thermal images and propose a method to reduce energy consumption in buildings. Also, Abdelrahman et al. [1] present an unobtrusive indicator of users cognitive load based on thermal images by monitoring forehead and nose temperature. Pavlidis et al. [11] detect lies based on bloodstream increase estimated by thermal images. López et al. [12] and Basu et al. [5] propose methods to estimate human emotion by facial temperature distribution. Thermal cameras are also useful to measure such psychological states since they do not disturb user behavior.

However, most of the existing works use high-end thermal cameras. In reality, we can not always use such high-end thermal cameras, causing the problem of frequent fluctuation in measurement by low cost (i.e. smartphone) thermal cameras.

C. Calibration of Thermal Cameras

There are several factors causing errors on the thermal camera. The emissivity of a target object is one of the factors, which is the efficiency of the surface in thermal energy emission. The emissivity is a specific parameter dependent on a component of substance. Mitchell [13] et al. reported the difference of the emissivity of human skin is 0.95 – 0.99, caused by the difference of the blood flow and skin color. This 4% difference leads to $\pm 0.25^{\circ}\text{C}$ error in the measurement of skin temperature. In order to determine the emissivity, it is necessary to measure the temperature of an object whose emissivity is known, such as black body tape [14]. However, the correction of the other factors is still challenging. Electrical noise and noise due to fluctuation in the scaling of thermal energy are the remaining major factors. To mitigate these effects, we essentially need high-end thermal cameras if we do not rely on any additional devices. Our method tackles with this challenge by a combination of a wristband sensor and a smartphone thermal camera.

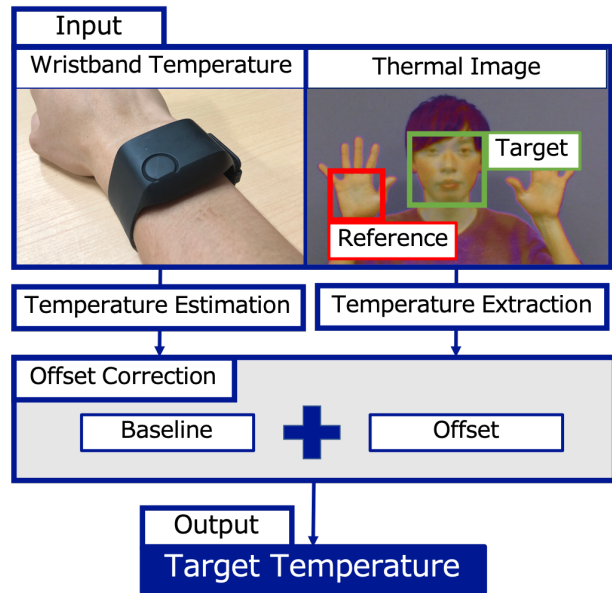


Fig. 1. Overview of the offset correction.

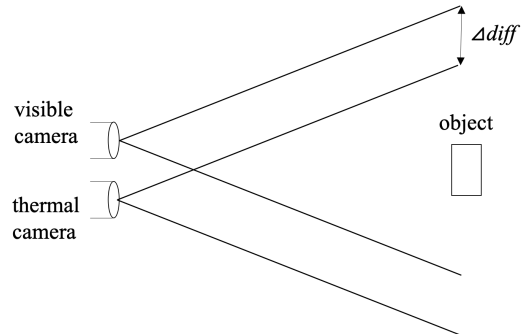


Fig. 2. Position difference of visible and thermal cameras.

III. CORRECTION METHOD

A. Overview

Fig. 1 shows an overview of our method. We assume a user wears a wristband sensor such as E4 sensor to measure her wrist skin temperature in real time. The input is wrist temperature measured by the wristband, a visible image, and a thermal image. We extract the skin temperature of body parts from the thermal image by using image processing techniques. To do this, we apply pre-processing to match each pixel in the visible image with the temperature in the thermal image. Based on linear regression, the reference point temperature is estimated from the temperature of the wrist. By comparing the reference point temperature measured by the wristband sensor and the thermal image, we estimate the offset. Finally, we obtain the target point temperature by adding the offset to the thermal image.

B. Image Pre-processing

As shown in Fig. 2 the visible image and the thermal image have slightly different views due to the difference between

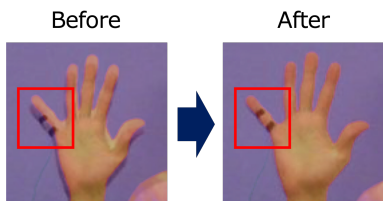


Fig. 3. Result of overlapping.

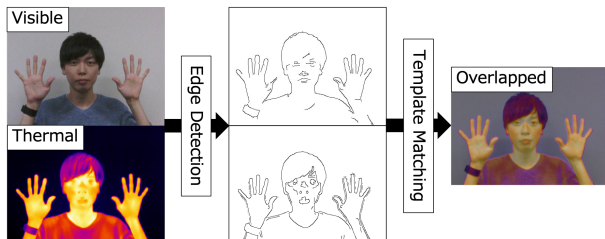


Fig. 4. Image pre-processing steps.

camera positions. This means, for example, the palm pixels in the visible image are not exactly equivalent to those in the thermal image. Therefore, if we recognize a target region (e.g. a palm) in the visible image and then obtain the temperature distribution of the area from the thermal image, the distribution wrongly contains temperature of different parts. Fig. 3 shows that this problem greatly affects temperature extraction especially for a small part such as a fingertip and a nose. Also, the position difference has a large impact on feature extraction from the distribution such as min and max. Therefore, we apply image pre-processing to match each pixel in the visible image with its correct temperature in the thermal image.

In the pre-processing, the key idea is that both images share almost the same edges. Fig. 4 shows the steps in the pre-processing. First, edge detection using the Canny algorithm [15] is performed on both images. Smoothing based on a Gaussian filter is applied to both images after edge detection. Finally, Normalized Cross-Correlation [16] is used for template matching.

C. Reference Point Temperature Extraction

The temperature of any pixel in the visible image can be extracted from the thermal image after the pre-processing. In this paper, we propose two reference points close to the wristband. One is a palm and the other is a wrist around the wristband. In the following sections, we describe the temperature extraction for each reference point.

1) *Palm Temperature Extraction*: We extract a palm from a visible image based on skin color as shown in Fig. 5 because it is difficult to detect hands from thermal images. This is because the skin is easily cooled by the atmosphere in the cold environment and it is assimilated into the background in the thermal image.

We use an object detector using OpenCV based on Haar-like feature [17] to detect a palm from a visible image. The position of the palm can be obtained as a rectangle. Since we

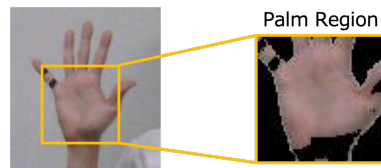


Fig. 5. Palm region extraction.

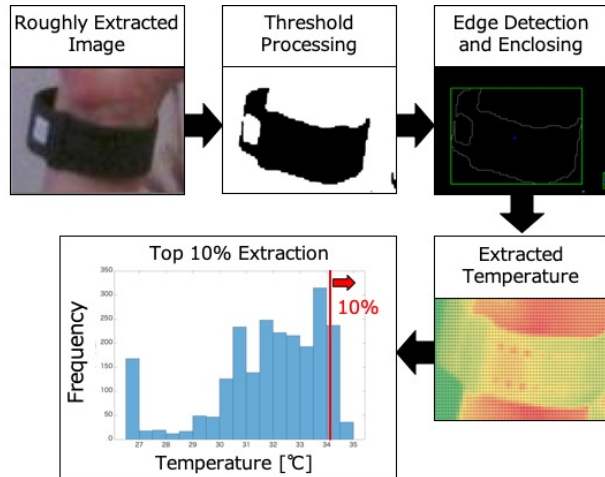


Fig. 6. Flow of wrist temperature extraction.

can obtain the area including other parts such as a background and clothing, we need to further extract the palm region. For this purpose, we use an approach for human skin detection proposed by Tan et al. [18]. We use the green value of RGB color space and the saturation value of HSV color space [19] to obtain a squared histogram of color values in the face. We regard a pixel whose distance from the center of the smoothed histogram [20] is within 3 times of the standard deviation as a skin pixel. We extract the temperature of the skin pixel and calculate the average of the temperatures as a palm temperature $T_{\text{palm_meas}}$. The definition is given as

$$T_{\text{palm_meas}} = \frac{\sum_{(x,y) \in P} T(x,y)}{|P|}, \quad (1)$$

where (x,y) is a coordinate in a visible image and P is a set of the extracted skin pixels. Also, $T(x,y)$ is the temperature of (x,y) and $|X|$ is the number of elements in the set X . P is defined as

$$P = \{(x,y) | (x,y) \in R_{\text{hand}} \cap C_{\text{hand}}\}, \quad (2)$$

where R_{hand} is a set of coordinates in the detected palm rectangle and C_{hand} is a set of coordinates whose colors are regarded as skin.

2) *Wrist Temperature Extraction*: We extract a wrist from a visible image based on the wristband color. Different from the palm detector implemented in OpenCV, we need to implement the wristband detector. In this paper, we manually extracted rough positions of the wristband. However, we note that this may be easily achieved by attaching a special marker or

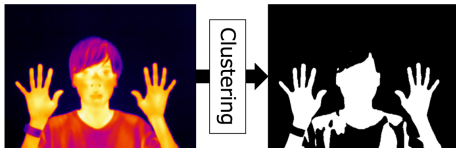


Fig. 7. Cluster thermal image pixels.

collecting training data. The flow of the wrist temperature extraction is shown in Fig. 6. First, to highlight the wristband, we conduct threshold processing based on the hue value from HSV color space. Next, we conduct edge detection, enclose the detected edges in a rectangle, and extract temperatures in the enclosed box from corresponding pixels of thermal images. Finally, we extract wrist temperature $T_{\text{wrist_meas}}$ by calculating the average of the top 10% in the temperature distribution. $T_{\text{wrist_meas}}$ is defined as

$$T_{\text{wrist_meas}} = \frac{\sum_{(x,y) \in W} T(x,y)}{|W|}, \quad (3)$$

where W is a set of the wrist coordinates which is defined as

$$W = \{(x,y) | (x,y) \in R_{\text{wrist}} \cap C_{\text{wrist}}\}. \quad (4)$$

In the above equation, R_{wrist} is a set of coordinates in the enclosed box of the wristband and C_{wrist} is a set of coordinates whose temperature is in the top 10%.

D. Reference Point Temperature Estimation

As we mentioned earlier, thermal cameras cannot capture the temperature of the point measured by the wristband sensor since it is covered by the wristband sensor itself. Therefore, we need to correlate the wristband sensor temperature with the temperature of the reference points. For this purpose, we construct regression models given as

$$y = a + bx, \quad (5)$$

where x is an explanatory variable which is a temperature reported by the wristband sensor, y is a target variable as a temperature of the reference point (T_{est}) which is a palm ($T_{\text{palm_est}}$) or a wrist ($T_{\text{wrist_est}}$). a and b are parameters determined by training data.

E. Target Part Extraction

In this paper, we select faces as the target part because the face temperature is used in many applications [1], [2], [4], [10]. We note that our method can be also applied to other target parts. We use the face detector of OpenCV based on Haar-like feature [17]. Since we can obtain a face position as a rectangle including a background, we need to extract the temperatures of facial skins from the thermal image. For this purpose, k-means++ clustering [21] is applied to the thermal image. We set the cluster size k to 2 (the face region and the other region) and used the cluster with the higher temperature as a face. Fig. 7 shows an example of the clustering. We

TABLE I
SPECIFICATION OF FLIR T540 AND FLIR ONE.

	FLIR T540	FLIR ONE 2
IR Sensor Resolution	464 × 348	160 × 120
Accuracy	±1°C or ±1% (10–35°C)	±3°C or ±5% (0–35°C)
Thermal Sensitivity	0.04°C	0.15°C

extract face temperature T_{face} by calculating the average of face pixels. The definition is given by

$$T_{\text{face}} = \frac{\sum_{(x,y) \in F} T(x,y)}{|F|}, \quad (6)$$

where R_{face} is a set of coordinates in the detected face rectangle and C_{face} is a set of coordinates in the face cluster.

F. Offset Correction

We obtain the corrected temperature T_{corr} of the target by adding the offset C to the temperature T_{target} of the target in the thermal image as below.

$$\begin{aligned} T_{\text{corr}} &= T_{\text{target}} + C \\ C &= T_{\text{est}} - T_{\text{meas}}. \end{aligned} \quad (7)$$

The offset C is calculated for the reference point (either a palm or a wrist) by subtracting the temperature T_{meas} measured by the thermal camera from the temperature T_{est} estimated by the regression. Namely, T_{est} is either $T_{\text{palm_est}}$ or $T_{\text{wrist_est}}$ and T_{meas} is either $T_{\text{palm_meas}}$ or $T_{\text{wrist_meas}}$.

IV. EVALUATION

A. Evaluation Settings

For evaluation, we used FLIR ONE 2 and E4 wristband as a smartphone thermal camera and a wristband sensor, respectively. We also used a high-end thermal camera, FLIR T540, for the ground truth. The specifications of FLIR T540 and FLIR ONE 2 are shown in the TABLE I. E4 wristband measures the temperature at the sampling rate of 4 Hz. We use the mean temperature per minute for the evaluation. The accuracy of the temperature sensor is 0.2°C within 36–39°C and its thermal sensitivity is 0.02°C.

We collected the real data from nine male subjects aged the twenties for 6 hours. In the experiment, we captured both visible and thermal images 7 times with a 10-second interval every 30 minutes. During the experiment, the subjects were asked to wear E4 wristbands and worked as usual in the laboratory. They stood in front of the thermal cameras with the palms of both hands facing toward the cameras without any overlap as shown in Fig. 7. After removing some images which are incorrectly captured, 871 samples (i.e. pairs of the visible and thermal images) were collected in total. The maximum and the minimum numbers of samples per subject are 105 and 77, respectively. The highest and the lowest air temperatures in the room were 30.6°C and 21.9°C, respectively.

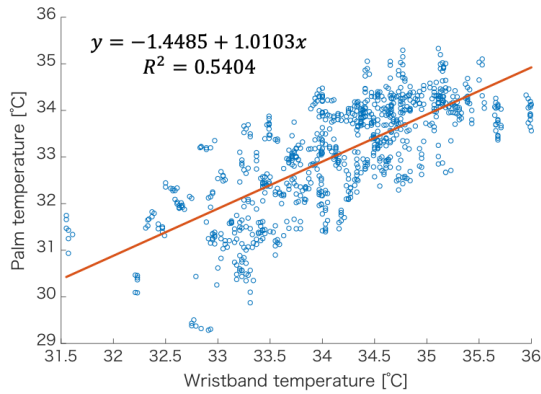


Fig. 8. Relation between wristband temperature and palm temperature.

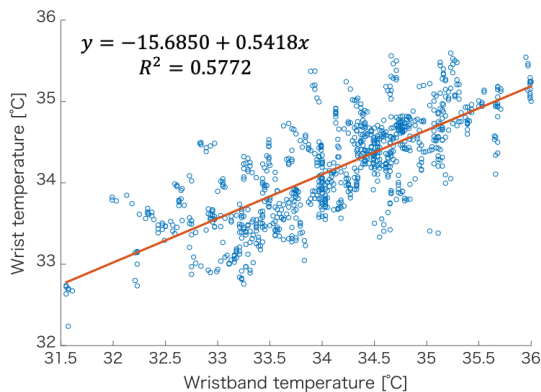


Fig. 9. Relation between wristband temperature and wrist temperature.

B. Result

1) *Reference Point Temperature Estimation*: We evaluated the performance of the temperature estimation of the reference points by linear regression. The purpose of the experiment is to build models to estimate the reference point temperature from the wristband temperature. Therefore, we used the thermal images captured by the high-end thermal camera to collect the training data.

Fig. 8 shows the relation between the temperature measured by the wristband and the temperature of the palm. The regression function is given as

$$T_{\text{palm_est}} = -1.4485 + 1.0103 T_{\text{wristband}}. \quad (8)$$

The mean absolute error was 0.6460°C .

On the other hand, Fig. 9 shows the relation between the wrist temperature and the temperature around the wristband. The regression function is given as

$$T_{\text{wrist_est}} = -15.6850 + 0.5418 T_{\text{wristband}}. \quad (9)$$

The mean absolute error was 0.3143°C .

From the above results, we see that the temperature around the wristband can be estimated more accurately than the palm. This is natural because of the closeness to the point

TABLE II
RESULT OF DYNAMIC OFFSET CORRECTION.

Method	MAE	SD
Baseline	0.8839	1.6157
Palm referenced	0.6527	0.7071
Wrist referenced	0.7232	0.4835

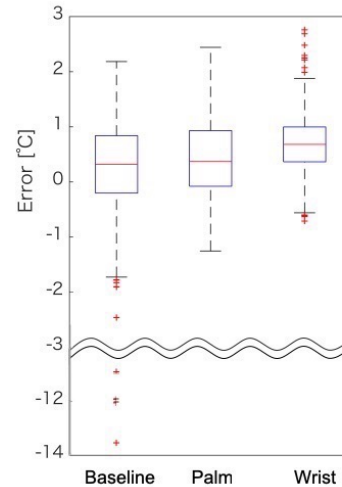


Fig. 10. Error distributions of facial temperature.

measured by the wristband sensor. We further investigate the performance of our method based on the above regression functions in the next section.

2) *Dynamic Offset Correction*: Fig. 10 shows the error distributions of the baseline (without correction), our palm referenced method, and our wrist referenced method. We see that the distribution approaches 0 in the palm referenced method while the error dispersion becomes smaller in the wrist referenced method.

Also, TABLE II shows the mean absolute error (MAE) and the standard deviation (SD) of the baseline and our methods. It is obvious that both of our methods reduce MAE and SD compared to the baseline. The palm referenced method achieves 0.2312°C smaller MAE than the baseline, which is also 0.0705°C smaller than the wrist referenced method. This is because the resolution of the smartphone thermal camera is much smaller than the high-end one. We found that the average temperature in a small region (e.g. a wrist) tends to become higher when the resolution is low. The low resolution leads to the difficulty in capturing the correct temperature distribution in a small region. On the other hand, the wrist referenced method achieves smaller SD than the other. This is because the accuracy of the reference point temperature estimation for the wrist is higher than the palm.

We note that the outliers of the baseline are difficult to remove by filtering such as smoothing over time. Fig. 11 shows the change of temperature measurement over time. We continuously recorded the images of a subject for 10

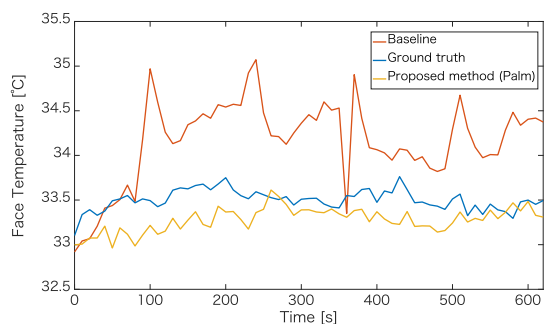


Fig. 11. Temperature change over time.

minutes. Even if we remove some large fluctuations of the baseline as outliers, there still remains larger error than our method. This is because the error characteristic of smartphone thermal cameras is unpredictable and sometimes biased due to environmental factors such as the camera body temperature itself.

From the above results, the palm referenced method is suitable for accuracy while the wrist referenced method is suitable for high precision. In practice, the accuracy may be more important for many applications. For example, the applications for cognitive load estimation [1] and thermal comfort estimation [2] need to capture temporal changes and differences among different persons. However, if we can mitigate the low-resolution effect of the smartphone thermal camera, the wrist referenced method may achieve better in terms of both accuracy and precision.

V. CONCLUSION

In this paper, we presented dynamic offset correction for a smartphone thermal camera using a wristband sensor for low cost and accurate temperature monitoring. The design of our method is based on the key feature that the measurement fluctuation of the thermal cameras is due to the offset which is common in all the pixels in a single thermal image. Our method estimates the temperature of the reference point by regression from the wristband temperature measurement. We selected a palm and a wrist for the reference points for comparison. Through the real experiment with 871 samples from 9 subjects, we confirmed that the palm referenced method is better than the wrist referenced method, showing 0.65°C MAE.

Our future work includes the improvement of the wrist referenced method by investigating the difference of thermal cameras especially in terms of resolution. We are also planning to apply our method for some applications to show its effectiveness.

REFERENCES

- [1] Y. Abdelrahman, E. Velloso, T. Dingler, A. Schmidt, and F. Vetere, "Cognitive heat: Exploring the usage of thermal imaging to unobtrusively estimate cognitive load," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 33:1–33:20, Sep. 2017. [Online]. Available: <http://doi.acm.org/10.1145/3130898>
- [2] J. Ranjan and J. Scott, "Thermalsense: Determining dynamic thermal comfort preferences using thermographic imaging," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, September 2016.
- [3] M. Burzo, M. Abouelenien, Verónica Pérez-Rosas, Cakra Wicaksono, Yong Tao, and Rada Mihalcea, "Using infrared thermography and biosensors to detect thermal discomfort in a building's inhabitants," in *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition*. ASME, November 2014.
- [4] H. Genno, K. Ishikawa, O. Kanbara, M. Kikumoto, Y. Fujiwara, R. Suzuki, and M. Osumi, "Using facial skin temperature to objectively evaluate sensations," *International Journal of Industrial Ergonomics*, vol. 19, no. 2, pp. 161 – 171, 1997, kansei Engineering and Comfort. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016981419600011X>
- [5] A. Basu, A. Routray, S. Shit, and A. K. Deb, "Human emotion recognition from facial thermal image based on fused statistical feature and multi-class svm," *2015 Annual IEEE India Conference (INDICON)*, pp. 1–5, 2015.
- [6] Empatica, "Real-time physiological signals — E4 EDA/GSR sensor," 2015. [Online]. Available: <https://www.empatica.com/research/e4/>
- [7] Microsoft, "Microsoft Band — Official Site," 2014. [Online]. Available: <https://www.microsoft.com/microsoft-band/en-us>
- [8] FLIR, "FLIR ONE Pro — FLIR Systems," 2017. [Online]. Available: <http://www.flir.jp/flirone/>
- [9] M. Kimata, *Infrared Sensor Principles and Technologies*. Kagakujyoho shuppan Co., Ltd., 2018.
- [10] B. Tag, G. Chernyshov, and K. Kunze, "Facial temperature sensing on smart eyewear for affective computing," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ser. UbiComp '17. New York, NY, USA: ACM, 2017, pp. 209–212. [Online]. Available: <http://doi.acm.org/10.1145/3123024.3123084>
- [11] I. Pavlidis and J. Levine, "Thermal image analysis for polygraph testing," *IEEE Pulse*, vol. 21, no. 6, pp. 56–64, 11 2002.
- [12] E. Salazar-López, E. Domínguez, V. J. Ramos, J. de la Fuente, A. Meins, O. Iborra, G. Glvez, M. Rodríguez-Artacho, and E. Gmez-Miln, "The mental and subjective skin: Emotion, empathy, feelings and thermography," *Consciousness and Cognition*, vol. 34, pp. 149 – 162, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S105381001500080X>
- [13] D. Mitchell, C. H. Wyndham, and T. Hodgson, "Emissivity and transmittance of excised human skin in its thermal emission wave band." *Journal of Applied Physiology*, vol. 23, no. 3, pp. 390–394, 1967. [Online]. Available: <https://doi.org/10.1152/jappl.1967.23.3.390>
- [14] FLIR Systems Japan K.K., "Infrared thermography guidebook for researchers," https://www.flirmedia.com/MMC/THG/Brochures/T559243/T559243_JP.pdf.
- [15] J. F. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679–698, 1986.
- [16] OpenCV, "Template matching opencv 2.4.13.7 documentation," https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html.
- [17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 2001, pp. I–511–I–518.
- [18] W. R. Tan, C. S. Chan, P. Yogarajah, and J. Condell, "A fusion approach for efficient human skin detection," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 138–147, Feb 2012.
- [19] A. R. Smith, "Color gamut transform pairs," *SIGGRAPH Comput. Graph.*, vol. 12, no. 3, pp. 12–19, Aug. 1978. [Online]. Available: <http://doi.acm.org/10.1145/965139.807361>
- [20] P. H. C. Eilers and J. J. Goeman, "Enhancing scatterplots with smoothed densities," *Bioinformatics*, vol. 20, no. 5, pp. 623–628, 2004. [Online]. Available: <http://dx.doi.org/10.1093/bioinformatics/btg454>
- [21] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1283383.1283494>