

Thermal and RGB-D Sensor Fusion for Non-Intrusive Human Thermal Comfort Assessment

Da Li,

Ph.D. Candidate, Dept. of Civil and Environmental Engineering, University of Michigan, U.S.

(email: dliseren@umich.edu)

Carol C. Menassa,

Associate Professor, Dept. of Civil and Environmental Engineering, University of Michigan, U.S.

(email: menassa@umich.edu)

Vineet R. Kamat,

Professor, Dept. of Civil and Environmental Engineering, University of Michigan, U.S.

(email: vkamat@umich.edu)

Abstract

It is well established that thermal comfort is an influential factor in human health and wellbeing. Uncomfortable thermal environments can reduce occupants' comfort and productivity, and cause symptoms of sick building syndrome. To harness the built environment as a medium to support human health, wellbeing, and engagement, it is significantly important to understand occupants' thermal comfort in real time. To this end, this study proposes a non-intrusive method to collect occupants' facial skin temperature and interpret their thermal comfort conditions by fusing the thermal and RGB-D images collected from multiple low-cost thermographic and Kinect sensors. This study distinguishes from existing methods of thermal comfort assessment in three ways: 1) it is a truly non-intrusive data collection approach which has a minimal interruption or participation of building occupants; 2) the proposed approach can simultaneously identify and interpret multiple occupants' thermal comfort; 3) it uses low-cost thermographic and RGB-D cameras which can be rapidly deployed and reconfigured to adapt to various settings. This approach was experimentally evaluated in a transient heating environment (room temperature increased from 23 to 27 °C) to verify its applicability in real operational built environments. In total, all 6 subjects observed moderate to strong positive correlations between the ambient room temperature and subjects' facial skin temperature collected using the proposed approach. Additionally, all 6 subjects have voted different thermal sensations at the beginning (the first 5 minutes) and at the end (the last 5 minutes) of the heating experiment, which can be reflected by the significant differences in the mean skin temperature of these two periods ($p < .001$). Results of this pilot study demonstrate the feasibility of applying the proposed non-intrusive approach to real multi-occupancy environments to dynamically interpret occupants' thermal comfort and optimize the operation of building heating, ventilation and air conditioning (HVAC) systems.

Keywords: Infrared thermal imaging, RGB-D sensor, thermal comfort, skin temperature, HVAC control

1. Introduction

Heating, ventilation, and air conditioning (HVAC) systems represent the biggest energy end use in buildings, which account for approximately 48% of the total energy required to operate residential and commercial buildings (DOE 2017, Li et al. 2017a). Despite the significant energy footprint of building HVAC systems, the lack of thermal comfort is still a common problem where studies show that up to 43% of occupants are dissatisfied with the thermal environment in their workplace (Karmann et al. 2018).

The importance of thermal comfort cannot be overemphasized. Several studies have suggested that satisfying thermal environments can lead to a reduced number of complaints, absenteeism, and improved work productivity (Roulet et al. 2006). On the other hand, It is also not surprising that thermal comfort is an influential factor of occupants' health and well-being, especially given that people spend more than 90% of time indoors. For example, the reports of sick building syndrome symptoms, such as a headache, eye and throat irritation, have been found to be correlated with the high room temperature (Fang et al. 2004).

The crucial impacts of thermal comfort on human satisfaction, health, and productivity demonstrate the importance of understanding occupants' thermal comfort in real time to allow for autonomous control of indoor environments that enhance occupants' experience. Thermal comfort is defined as “*the condition of mind which expresses satisfaction with the thermal environment*” (ASHRAE 2010), which implies that it is one's subjective assessment of the environmental condition (e.g., air temperature, relative humidity). In addition, an individual's thermal comfort is also significantly affected by personal conditions including physiological (e.g., gender, age), psychological (e.g., expectation, stress), and behavioral factors (e.g., clothing and activity level) (Parsons 2014). As a result, both personal and temporal variations should be considered when assessing occupants' thermal comfort.

To date, researchers have proposed different methods for thermal comfort assessment. The most well-known approach is the Predicted Mean Vote (PMV) model, which was developed based on the heat transfer between the human body and environments (Fanger 1970). Brager and de Dear (2000) proposed the adaptive comfort models to account for the adaptive behaviors from occupants to maintain thermally comfortable states in naturally ventilated environments. Recently, researchers also investigated the personal comfort models in which the comfort condition of an occupant was exclusively predicted based on his/her previous feedback of thermal comfort under various environmental and human conditions (Kim et al. 2018, Li et al. 2017b). In general, personal comfort models demonstrated better predictive powers than the PMV and adaptive models as they account for the personal uniqueness and subjectivity in evaluating thermal comfort.

Human physiological data, such as skin temperature, heart rate, respiration rate, and activity level, are typically collected as the parameters of personal comfort models. These parameters can be measured using wearable bio-sensors, infrared thermometers, thermographic cameras, Doppler radars, and many other devices (e.g., Ghahramani 2016, Jung and Jazizadeh 2017, Li et al. 2017b and 2017c). However, one significant limitation of existing physiological data collection methods is that they can cause different levels of intrusiveness on the occupants. The “intrusiveness” mainly comes from 1) the continuous requirement of human participation for real-time thermal comfort evaluation, and 2) the dependence on wearable devices or personal equipment for physiological data collection. In our previous work (Li et al. 2017b), we observed that occupants' participation decreases with time as the novelty of the approach fades away, which emphasizes the needs to reduce human participation or interruption while collecting human physiological data. To this end, this study proposes a truly non-intrusive method to collect occupants' facial skin temperature and interpret their thermal comfort conditions by fusing the thermal and RGB-D images collected from multiple low-cost thermographic and Kinect sensors. Skin temperature is selected as the main physiological parameter as it has a strong correlation with the ambient room temperature and the

subjective thermal comfort (Li et al. 2018). The sensor fusion provides accurate detection of human occupants and allows robust extraction of facial skin temperature under flexible body postures and movements.

2. Background

The human body maintains its core internal temperature at around 37 °C. When thermoreceptors detect heat or cold stress, the hypothalamus will control body muscles, organs, and nervous system to adjust heat production and heat loss to maintain the homeostasis state, which can cause variations in skin temperature (Parsons 2014). Thus, skin temperature is commonly adopted as a proxy of the thermoregulation process. Moreover, the human face has a higher density of blood vessels than other body parts and it is not covered by clothing, which makes it an ideal location to measure the skin temperature variations.

Existing studies have explored various devices and approaches to collect skin temperature, which can be summarized into two main categories: the contact and non-contact approaches. The contact approaches typically use wearable bio-sensors or thermocouples that directly contact the skin surface. For example, Li et al. (2017b) adopted a wrist-worn fitness tracker to continuously measure occupants' wrist skin temperature under different thermal conditions and suggested the personal comfort model can achieve a higher prediction accuracy when incorporating skin temperature and other physiological parameters. On the other hand, the non-contact approaches typically use infrared thermometers or thermographic cameras to infer the skin temperature from infrared radiations, which reduce the intrusiveness caused by wearing body sensors. For example, Ghahramani (2016) developed an eyeglass which is equipped with infrared thermometers on its frame to collect the skin temperature of the front face, cheek, nose, and ear regions and observed significant differences in skin temperature under heat and cold stress conditions. Thermographic cameras gained attention in recent years as they can capture a full thermal image consisting of temperature values of each pixel and can also measure the temperature from a longer distance compared to infrared thermometers. However, commodity thermographic cameras are generally expensive (over \$ 5000) and cannot be directly incorporated in the building management system due to their large sizes and compatibility issues. To overcome these two limitations, Li et al. (2018) proposed a framework which adopts a low-cost thermographic camera (FLIR Lepton[®], cost: \$ 200, dimension: 8.5 x 11.7 x 5.6 mm) as an alternative to assess the thermal comfort of building occupants. The proposed framework can automatically and continuously detect human faces, measure the skin temperature of each facial region, clean and process raw skin temperature data, and interpret thermal comfort using personal comfort models. Results from this study suggested an 85% accuracy in predicting the three-point thermal preferences using the facial skin temperature.

Despite the contributions of exploring low-cost thermographic cameras in thermal comfort assessment, there are still two unaddressed issues which limit their applications in real operational environments. First, existing studies only focus on single occupant experiments rather than multi-occupancy scenarios. This is due to the fact that a single thermographic camera has a limited field of view which cannot cover a large indoor space. Second, thermographic cameras are generally placed in front of the occupants to measure the frontal face temperature. However, thermal images of frontal faces are not guaranteed in real operational environments especially in large open spaces where occupants do not remain static over time. To overcome these limitations, this study proposes a networked camera system to simultaneously interpret multiple occupants' thermal comfort with minimum intrusiveness of building occupants. The technical details and experimental validations are detailed in Sections 3 and 4.

3. Technical Approach

The proposed networked camera system consists of multiple camera nodes to collectively and simultaneously measure the facial skin temperature of each occupant in a multi-occupancy environment (see Figure 1). The camera nodes are placed at different locations to achieve comprehensive coverage of the environment such that each occupant can be seen (but not necessarily obtain a full frontal view of the face) by at least one camera node.

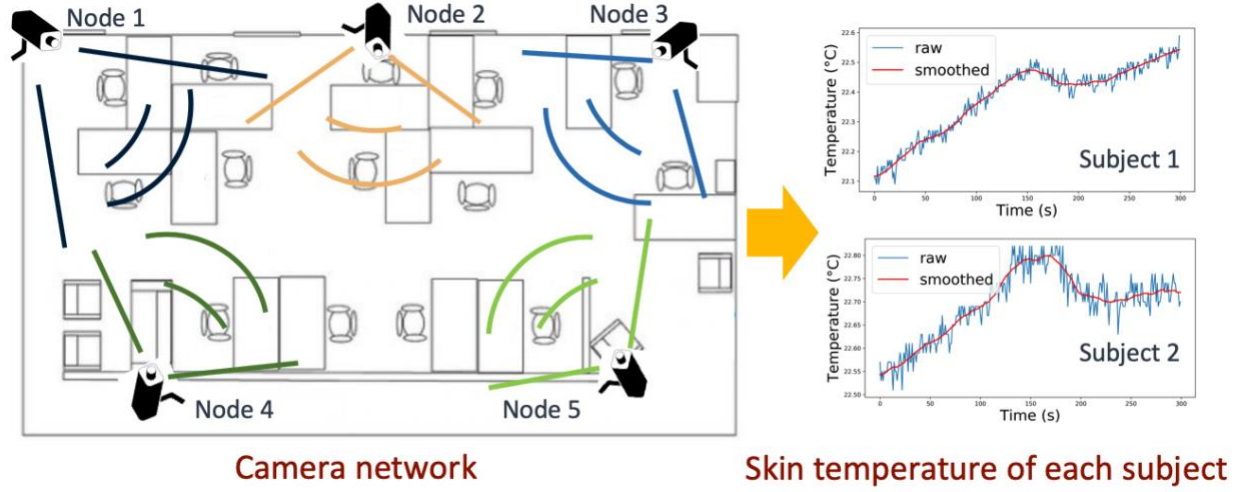


Figure 1: Overview of the Proposed Camera Network

3.1 Thermal and RGB-D dual camera system

Each camera node in the network is called a dual camera system. It comprises a FLIR Lepton 2.5 thermographic camera and a Microsoft Kinect (a type of RGB-D camera). As shown in Figure 2, the thermographic camera is rigidly mounted on top of the Kinect. In this dual camera system, the Kinect is responsible for detecting occupants in the RGB-D images and locating the coordinates that contain faces as our previous work showed that thermographic camera alone can only detect the frontal face at a short distance (1 meter) (Li et al. 2018), which is not ideal as occupants can have flexible distances and angles to the camera. The face coordinates are subsequently mapped to the thermal images to extract facial skin temperature measured by the thermographic camera. In this study, we adopted the deep neural network based face detectors in the OpenCV library. However, other algorithms such as the FastRCNN and DeepFace can also be applied to guide the thermographic camera in the face detection task.



Figure 2: The Dual Camera System

3.2 Kinect and thermographic camera registration

As the Kinect and the thermographic camera have different field of view and image resolutions, these two cameras need to be calibrated to find the point correspondences. Both the Kinect and the thermographic

camera can be modeled as a pinhole camera which projects the 3D world scene into a 2D image plane through the perspective transformation as shown in Eq. 1 (OpenCV 2018).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

Alternatively, in a more concise form,

$$s \, m = K[R|T]M$$

Where M is a 4×1 vector representing the homogeneous coordinate of a 3D point in the world coordinate space; m is a 3×1 vector representing the homogeneous coordinate of a 2D point in the image coordinate; K is the 3×3 intrinsic matrix of the camera consisting of the focal lengths (f_x, f_y) and principal points (C_x, C_y); $[R|T]$ is the 3×4 extrinsic matrix consisting of a rotation R and a translation T ; and s is a scaling factor.

In the dual camera system, the calibration process is to estimate the intrinsic matrix of the thermographic camera K_{IR} , the intrinsic matrix of the Kinect K_{RGB} , and the homogeneous transformation matrix $[R|T]$ between the two cameras. Once these three matrices are estimated, the point correspondences can be determined according to the pinhole camera model described in Eq. 1.

In practice, such a dual camera system can be calibrated in the stereo vision calibration process. The calibration requires both cameras to observe a planner and predefined pattern, such as a checkerboard or a square grid, from at least two different orientations to determine the unknowns using the maximum likelihood estimation (Zhang 1999). However, thermographic cameras typically cannot detect the black-and-white calibration patterns printed on the paper. Therefore, we made a special 6×7 checkerboard pattern from the aluminum foil and the vinyl polymer (see Figure 3). Each black or silver square in the checkerboard pattern has a dimension of 62.5 mm. Due to the color differences, the checkerboard pattern can be detected by the RGB camera to extract corner points (Figure 3a and 3c). On the other hand, as the aluminum foil has a higher emissivity, it emits more infrared energy and thus looks brighter in thermal images. As shown in Figure 3b and 3d, the checkerboard corner points can also be easily observed by a thermographic camera.

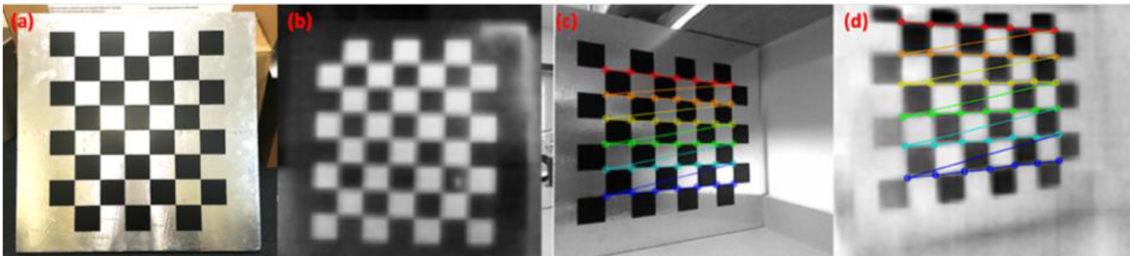


Figure 3: Special Checkerboard for the Kinect and Thermographic Camera Calibration. (a) Checkerboard Pattern in the Kinect; (b) Checkerboard Pattern in the Thermal Image; (c) Corner Detection using the Kinect; (d) Corner Detection using the Thermographic Camera

After the calibration process, face coordinates detected in the RGB image can be successfully mapped to the thermal image. Figure 4 shows the result of camera registration where faces in the RGB and thermal images are labeled in bounding boxes.



Figure 4: Dual Camera Face Detection. Left: RGB Images from the Kinect (for Face Detection); Right: Thermal Images from the Thermographic Camera (bounding boxes are mapped from the RGB image)

3.3 Skin temperature feature extraction

The facial skin temperature collected directly from each bounding box in thermal images are the raw data which contain several types of random noises such as the false detection of background as faces, inaccurate mapping of face coordinates due to occlusions, and interference of a high temperature object in the environment (e.g., hot water cup). These noises are typically shown as spikes (i.e., sudden changes) in the measurements. As a result, we applied the median filter to remove such noises before data analysis.

Moreover, as thermal images of frontal faces are not guaranteed in the camera network, unlike Li et al. (2018) which segmented the frontal face into the forehead, nose, cheeks, ears, and mouth regions and collected skin temperature from each local facial region, in this study we choose the global skin temperature features from the whole facial region. These global skin temperature features include the highest, lowest, first quartile, third quartile, and average temperature measurements of all pixels located in the detected facial region (including both frontal and profile faces), which can describe the overall distribution of skin temperature over a detected face.

4. Experimental Study

The proposed approach was experimentally evaluated in a transient heating environment to verify its applicability in real operational built environments. The experiment included a 20-minute preparation phase and a 50-minute data collection phase and was conducted in a research office at the University of Michigan (UM) during the heating season in 2018. The experiment office is equipped with a thermostat which can control the indoor temperature from a low level of 23 °C to a high level of 27 °C through two HVAC diffusers. As shown in Figure 5, two camera nodes were placed approximately 1.3 meters away from a table where subjects sat during the experiment to represent a non-intrusive distance. Two COZIR sensors (humidity accuracy: $\pm 5\%$, temperature accuracy: ± 1 °C) continuously monitored the ambient room conditions. The two sensors were placed at the waist level (0.65 meters above the floor) which was close to the specified height of 0.6 meters for seated occupants in ASHRAE standards 55. To represent a multi-occupancy scenario, 2 subjects were required to participate in the experiment at the time. In total, 6 subjects (3 male, 3 female) were recruited. All subjects were UM students and were healthy at the time of the experiment. This human subject experiment has been approved by the UM Institutional Review Board.

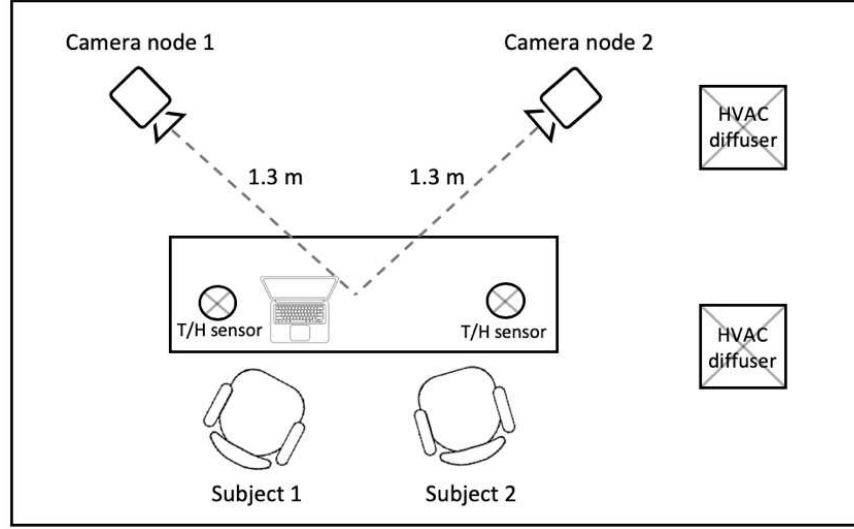


Figure 5: The Experimental Setup

The room temperature was initially set at a low level (approximately 23 °C) before the experiment started. During the 20-minute preparation phase, subjects remained seated to reach a steady-state skin temperature. Then, during the following 50-minute data collection phase, the thermostat was set at a high level (27 °C) to create a transient heating environment. During this period, subjects were asked to perform daily office activities such as reading, typing, browsing, or chatting with each other and report their five-point thermal sensation and three-point thermal preference through a phone application every five minutes. For more details about the phone application, please refer to Li et al. (2017b).

It is worth noting that unlike existing experimental studies which typically require subjects to remain in the same posture and refrain from movements, in this study the subjects were encouraged to move freely (e.g., stretching), change their posture and facing directions, or even move around in the room. The objective is to represent a real office environment and let subjects feel as comfortable as possible (to remove any intrusiveness caused by the system). We believe such an experimental protocol can verify the applicability of the proposed system to the greatest extent.

5. Results and Discussion

Figure 6 showed the room temperature and relative humidity during the heating experiment. For each subject, we calculated the mean facial skin temperature of the first five minutes (denoted as $\bar{T}_{start}$) and last five minutes (denoted as $\bar{T}_{end}$) in the transient heating experiment to represent one's starting and ending physiological states, respectively. As shown in Table 1, results of the two-tailed t-test showed that all 6 subjects experienced significantly different mean facial skin temperature ($p < .001$) in the heating experiment. Also, all subjects provided distinct thermal sensation votes at the beginning and the end of the experiment (denoted as TSV_{start} and TSV_{end}). This finding suggests that facial skin temperature captured by our proposed camera network can be used to interpret a subject's thermal sensation. This result is promising especially given that the room temperature only changes within a relatively small range.

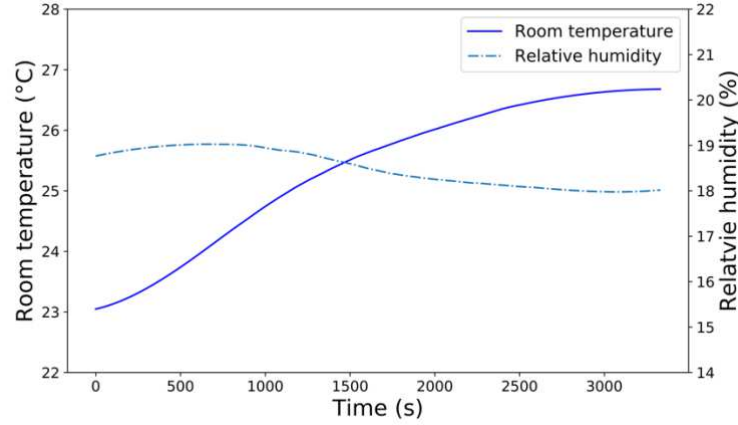


Figure 6: Room Temperature and Relative Humidity during the Heating Experiment

Table 1: T-tests between the Starting and Ending Mean Facial Skin Temperature in the Experiment

Subject		Camera Id	$\bar{T}_{start}$ (°C)	$\bar{T}_{end}$ (°C)	p -value	TSV_{start}	TSV_{end}
Experiment 1	Id1	1	31.6	33.4	< 0.001	Cold	Hot
		2	31.2	32.9	< 0.001		
	Id2	1	31.5	33.6	< 0.001	Cold	Warm
		2	31.4	32.8	< 0.001		
Experiment 2	Id3	1	31.8	33.6	< 0.001	Cold	Warm
		2	31.6	32.6	< 0.001		
	Id4	1	31.4	33.5	< 0.001	Cold	Neutral
		2	31.7	32.6	< 0.001		
Experiment 3	Id5	1	32.1	33.4	< 0.001	Cool	Warm
		2	31.5	33.5	< 0.001		
	Id6	1	32.1	33.2	< 0.001	Cool	Neutral
		2	31.3	32.9	< 0.001		

To visualize the variations of skin temperature in the heating experiment, the facial mean skin temperature was fitted using the polynomial regression. Specifically, the lower degree of polynomials was selected if the coefficient of determination R^2 did not increase significantly with a higher degree. As shown in Figure 7, all 6 subjects' facial mean skin temperature was fitted using polynomials of degree ranged from 1 to 3. All fitted curves (denoted in thin yellow lines) in Figure 7 demonstrated an increasing trend, which implies that facial skin temperature is increasing over time in the heat experiment.

Pearson correlation between the skin temperature of each subject and the corresponding room temperature in the experiment was also reported in Figure 7. In total, all 6 subjects observed moderate to strong positive correlations (R^2 ranged from 47.0 % to 82.2 %, $p < .001$) between the facial mean skin temperature (shown in blue lines) and the ambient room temperature (shown in thick orange curves), which suggests the proposed approach is capable of non-intrusively measuring multiple occupants' skin temperature variations in dynamic environments without any constraints on occupants' activities or engagements.

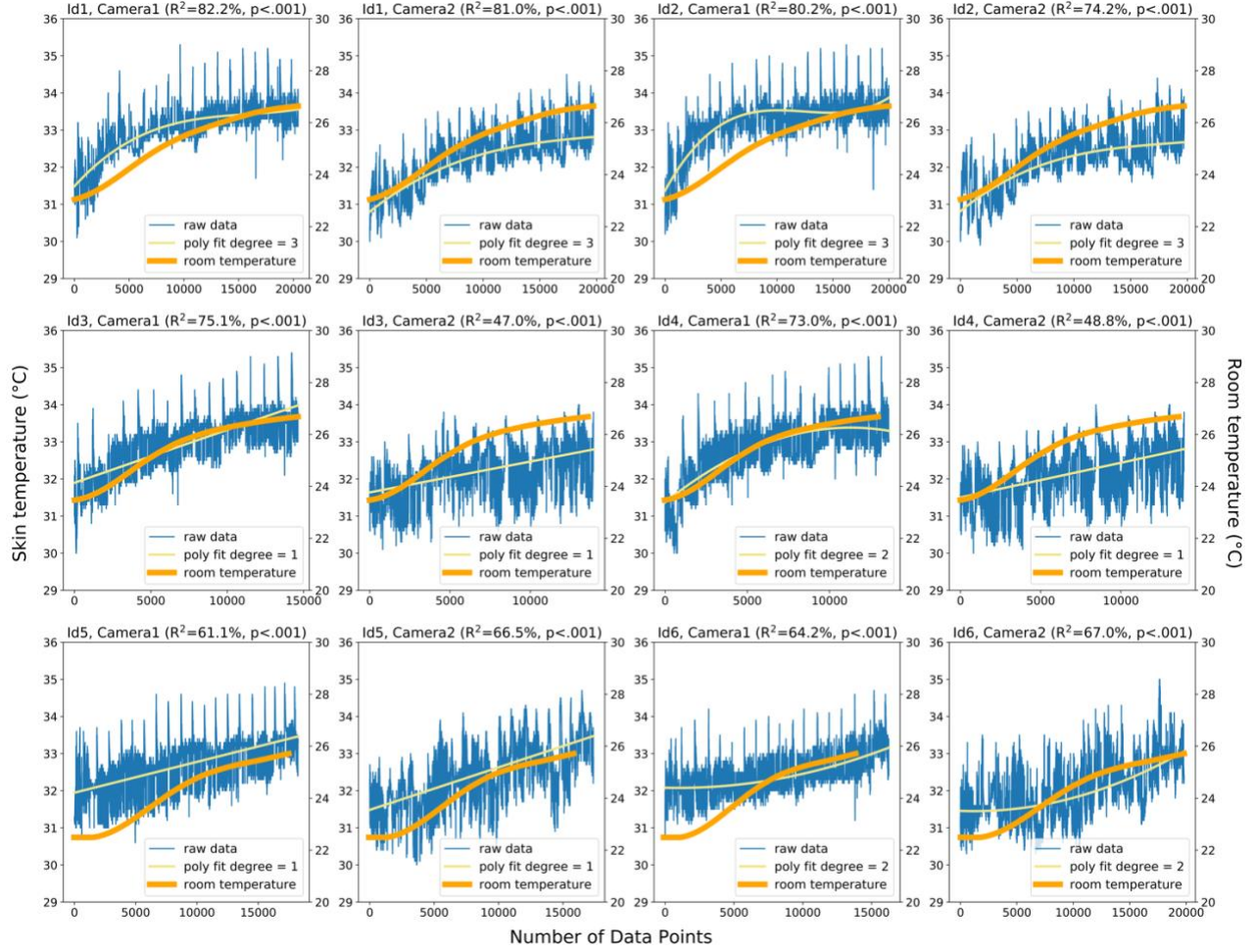


Figure 7: The Polynomial Fit of the Mean Facial Skin Temperature of Each Subject

6. Conclusions

In this study, we proposed a networked camera system to non-intrusively measure occupants' facial skin temperature in a multi-occupancy environment. Each camera node in the network fuses the RGB-D and thermal images collected from the Kinect and a low-cost thermographic camera. The experimental results from 6 subjects showed that the facial skin temperature collected under flexible body postures and movements has moderate to strong positive correlations with the ambient room temperature (ranged from 23 to 27 °C), which verified the capability of our proposed method to collect skin temperature for real-time thermal comfort interpretation. This pilot study has the potential to transition the human physiological data collection from an intrusive and wearable device driven approach to a truly non-intrusive and scalable approach, and also demonstrates its applicability in the operation of building HVAC systems to improve the indoor thermal environment.

Acknowledgments

The authors would like to acknowledge the financial support for this research received from the U.S. National Science Foundation (NSF) CBET 1349921 and CBET 1804321. Any opinions and findings in this paper are those of the authors and do not necessarily represent those of the NSF.

References

- ASHRAE, 2010. ANSI/ASHRAE Standard 55-2010. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.
- Department of Energy (DOE), 2017. Heating and Cooling. <https://energy.gov/public-services/homes/heating-cooling>.
- Fang, L., Wyon, D.P., Clausen, G. and Fanger, P.O., 2004. Impact of indoor air temperature and humidity in an office on perceived air quality, SBS symptoms and performance. *Indoor air*, 14, pp.74-81.
- Fanger, P.O., 1970. Thermal comfort. Analysis and applications in environmental engineering. *Thermal comfort. Analysis and applications in environmental engineering*.
- Ghahramani, A., Castro, G., Becerik-Gerber, B. and Yu, X., 2016. Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort. *Building and Environment*, 109, pp.1-11.
- Jung, W. and Jazizadeh, F., 2017, November. Towards integration of doppler radar sensors into personalized thermoregulation-based control of HVAC. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments* (p. 21). ACM.
- Karmann, C., Schiavon, S. and Arens, E., 2018. Percentage of commercial buildings showing at least 80% occupant satisfied with their thermal comfort.
- Kim, J., Zhou, Y., Schiavon, S., Raftery, P. and Brager, G., 2018. Personal comfort models: predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning. *Building and Environment*, 129, pp.96-106.
- Li, D., Menassa, C.C. and Karatas, A., 2017a. Energy use behaviors in buildings: Towards an integrated conceptual framework. *Energy Research & Social Science*, 23, pp.97-112.
- Li, D., Menassa, C.C. and Kamat, V.R., 2017b. Personalized human comfort in indoor building environments under diverse conditioning modes. *Building and Environment*, 126, pp.304-317.
- Li, D., Menassa, C.C. and Kamat, V.R., 2017c. A personalized HVAC control smartphone application framework for improved human health and well-being. In *Computing in Civil Engineering 2017* (pp. 82-90).
- Li, D., Menassa, C.C. and Kamat, V.R., 2018. Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography. *Energy and Buildings*, 176, pp.246-261.
- OpenCV, 2018. Camera Calibration and 3D Reconstruction. https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html.
- Parsons, K., 2014. *Human thermal environments: the effects of hot, moderate, and cold environments on human health, comfort, and performance*. CRC press.
- Roulet, C.A., Johnner, N., Foradini, F., Bluysen, P., Cox, C., de Oliveira Fernandes, E., Müller, B. and Aizlewood, C., 2006. Perceived health and comfort in relation to energy use and building characteristics. *Building research and information*, 34(5), pp.467-474.
- Zhang, Z., 1999. Flexible camera calibration by viewing a plane from unknown orientations. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on* (Vol. 1, pp. 666-673). Ieee.