Investigation of Material Properties for Thermal Imaging-Based Interaction

Yomna Abdelrahman, Alireza Sahami Shirazi, Niels Henze, Albrecht Schmidt
VIS, University of Stuttgart
Stuttgart, Germany
{firstname.lastname}@vis.uni-stuttgart.de

ABSTRACT

Recent work demonstrated the exciting opportunities that thermal imaging offers for the development of interactive systems. It was shown that a thermal camera can sense when a user touches a surface, performs gestures in the camera's direct field of view and, in addition, performs gestures outside the camera's direct field of view through thermal reflection. In this paper, we investigate the material properties that should be considered for detecting interaction using thermal imaging considering both in- and outdoor settings. We conducted a study to analyze the recognition performance for different gestures and different surfaces. Using the results, we derive guidelines on material properties of surfaces for detecting on-surface as well as mid-air interaction using a thermal camera. We discuss the constrains that should be taken into account using thermal imaging as the sensing technology. Finally, we present a material space based on our findings. The space depicts surfaces and the required properties that enable the different interaction techniques.

Author Keywords

thermal imaging; roughness; conductivity; interaction

ACM Classification Keywords

H.5.2 Information interfaces and presentation: User Interfaces

INTRODUCTION

There is an ongoing trend to extend the interaction with computers to a broad range of situations. Different situations require different interaction techniques. As a result new means to sense the user need to be developed. An example is the Kinect that was initially designed to increase the immersion when playing games by enabling interaction without a physical controller. The Kinect can sense a user and derive a 3D model of the scene by exploiting near-infrared light patterns projected on the scene.

One of the main reasons that makes the Kinect suitable for usage in daily life is the use of a non-visible light spectrum. The user cannot see the projected pattern and therefore it does not intervene with the experience. Recent work proposed to go even further in the light spectrum from near-infrared light to the far-infrared band [7, 12]. Thermal cameras are used to sense heat traces that are left behind when a user moves a finger across a surface [7] or to sense mid-air interaction besides and behind a thermal camera's field-of-view (FOV) through thermal reflection [12]. As a novel sensing technology for human-computer interaction, thermal imaging offers exciting opportunities for the development of interactive systems. Prior work only provided details about detection accuracy for interaction on surfaces but not for mid-air interaction. Further, prior work did not assess the properties of materials which can be used for creating interactive surface using thermal imaging. For a sensing system to be competitive compared to current RGB and near-infrared depth cameras the individual advantages and limitations have to be identified.

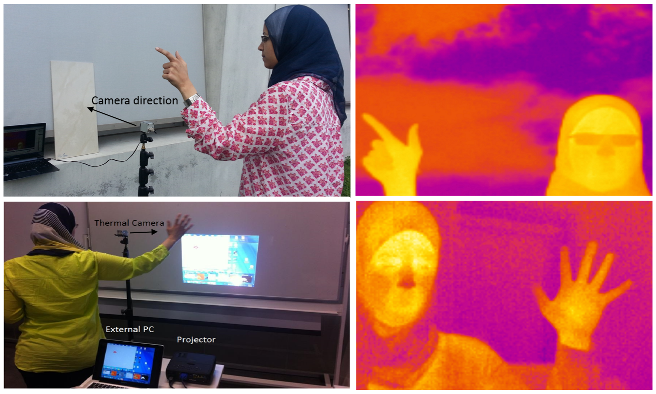

Figure 1. (top) the outdoor setup with the thermal view of tile as the reflective surface (bottom) the indoor setup with the thermal view of glass as the reflective surface.

In this paper, we present a study that assesses the recognition accuracy of mid-air gestures sensed through thermal reflection using different surfaces in indoor and outdoor setups to complement the work presented by Sahami Shirazi et al. [12]. Further, we provide holistic insights on the surfaces' properties that should be considered when it comes to sense interactions on the surface and/or mid-air using a thermal camera. Based on the identified properties, we provide a material space describing surfaces that enable interaction on the surface and/or mid-air interaction through thermal reflectivity. The guidelines enable to select materials for creating interactive surfaces using thermal imaging.

RELATED WORK

A large body of work combined projectors and sensing cameras to build interactive projected surfaces. Initially, RGB cameras were used to detect hands and fingers [6, 8]. Such systems typically use skin color detectors or template match-

CHI 2015, April 18–23, 2015, Seoul, Republic of Korea.

http://dx.doi.org/10.1145/2702123.2702290

ing to segment the hand and then calculate contour and convexity defects to identify fingers. A major challenge of such systems is the sensitivity for different light conditions.

Research prototypes have used infrared imaging [3] and depth cameras to enable multi-touch and mid-air gestures when interacting with projection screens and tabletop setups. In such systems, the space behind the screen is typically illuminated with an infrared source. Using an infrared-pass filter, all lights except the infrared light is blocked for the infrared camera. Using a depth camera, the depth information can be used to detect touch and hand gestures on projected screens [9, 14]. Such systems generally utilize either a 2D view above the surface [9] or a selective 2D projection of 3D sensed data [14] for processing users' input on or above the surface using common 2D computer vision techniques.

Using existing computer vision techniques, thermal imaging can be used to detecting interaction on surfaces as well as mid-air gestures. Two thermal properties have been leveraged for monitoring interactions. First, heat traces that are left on a surface due to the temperature difference and the heat transfer between hands and surfaces. Such traces have been used to detect interactions and pressures on surfaces [5, 4, 7]. The second property is thermal reflectivity which is the result of radiations' reflection when striking a surface. Sahami Shirazi et al. [12] propose to use specular reflectivity for extending the mid-air interaction space behind the camera's direct field-of-view and detecting mid-air gestures. However, they did not report the recognition accuracy of mid-air gestures in the extended space. In contrast, this paper investigates the recognition accuracy of mid-air interaction using thermal reflectivity. Further, we provide a holistic overview on surface properties which should be considered for creating an interactive setup using thermal imaging as the sensing technology.

RECOGNITION ACCURACY OF MID-AIR GESTURES

We conducted a study in both indoor and outdoor settings to assess the detection accuracy of mid-air gestures performed in the reflected space sensed using thermal reflectivity. We replicated the system reported in [12]. We recruited 30 participants (11 female, with an average age of 26 years, SD=3.8) using our university's mailing lists. All participants were students in different majors. Three participants were left handed. None had experience with thermal cameras. The participants were divided in two groups of 15, one group used the indoor setup, and the other used the outdoor setup.

Apparatus

The indoor and the outdoor setup were identical, including a projector and an Optris PI160 contactless thermal camera with 23° x 17° field of view(Figure 1). The projector was connected to a PC and displayed the tasks on a surface. The camera was mounted on a tripod and faced toward the surface from a distance of 50 cm to cover and capture the thermal reflection of the surfaces with a dimension of 30x60 cm. Since the setup was stationary no dynamic calibration was required. No special light source or illumination conditions were requisite as the thermal camera operates independent of illumination. We considered four common surfaces, also used in [12], for the experiment: glass, tile, MDF, and aluminum.

To detect mid-air gestures we implemented the algorithms described in [7, 12]. We used OpenCV for the hand and finger tips extraction. The frame extraction from the thermal imaging is done through the dynamic link library (DLL) the Optris camera provides. The image analysis covered the steps reported in [12] including pre-processing, noise and background removal, and thresholding. The feature extraction, i.e., the hand and finger information, was computed from the hand contour, convex hull, and convexity defects. The mid-air gesture detection was based on a view-based approach where the fingers positions and their relative distance were used to match the predefine gesture.

Tasks & Procedure

To assess the detection accuracy we considered three different gestures: (1) mid-air pointing interaction using one finger, (2) continuous interaction using one finger, and (3) mid-air hand postures using two fingers. We considered different tasks for each gesture. For the mid-air pointing 5 points were randomly projected on the surface. Participants were asked to point at the points in-air. We considered the dragging task for the continuous interaction. A pair of points were projected and users were asked to drag an object from the start point to the end point. Users repeated this task three times. For the mid-air posture, users were asked to perform a pinch and a pan gestures. For this task, users had to use two fingers.

After welcoming participants, we described the purpose of the study. We showed them where they had to stand and where to perform the gestures. Each participant performed the tasks for four different surfaces resulting in 12 tasks (3 tasks x 4 surfaces). The order of the tasks and surfaces were randomized per participant. The light condition of the indoor setup was constant during the whole study. The outdoor setup was in a shadow for the projection to be visible and the temperature was constant (24°C). The study took approximately 30 minutes per participant. We recorded the study on video and stored the thermal video for later analysis as ground truth. The time between the camcorder and the thermal camera was synchronized. To calculate the accuracy, an experimenter watched the videos and counted the number of times a gesture was correctly recognized by the software.

Results & Discussion

Table 1 includes the recognition accuracy of all tasks in indoor and outdoor setups. We found very similar results for both setups. Using the glass surface resulted in the highest accuracy followed by tile, MDF, and aluminum. Considering the roughness of the surfaces reported in [12] the Pearson coefficient revealed a strong inverse correlation between the accuracy and the roughness (r= -.98). Surfaces with lower roughness result in more mirror-like reflectivity, respectively, sharper images and higher recognition accuracy. Whereas, surfaces with higher roughness have hazy reflectivity, hence, lower accuracy.

In this study, we investigated the recognition performance of the gestures performed by single and multiple fingers. However, our system could support any arbitrary gestures by feeding the tracked fingers positions to either \$1 or \$N gesture recognizers which recognize arbitrary gestures formed of single or multiple strokes [1, 15].

Surface	Setup	R_a	Task						$b_{surface}$	$T_{surface}$	T_c
			Pointing		Continuous		Posture				
Glass	Indoor	.004	97%	SD=.17	95%	SD=.22	93%	SD=.26	1288	17	21
	Outdoor		96%	SD=.20	95%	SD=.21	93%	SD=.26		22.7	24.9
Tile	Indoor	.04	86%	SD=.35	88%	SD=.33	89%	SD=.31	3852	18	21
	Outdoor		87%	SD=.34	88%	SD=.33	89%	SD=.31		20.5	22.9
MDF	Indoor	.11	85%	SD=.35	86%	SD=.33	88%	SD=.36	365	19	22
	Outdoor		83%	SD=.37	86%	SD=.35	87%	SD=.36		27.05	27.9
Aluminum	Indoor	.33	56%	SD=.50	67%	SD=.48	46%	SD=.51	22265	14	14.07
	Outdor		58%	SD=.50	64%	SD=.48	47%	SD=.51		24	24.03

Table 1. It depicts two information: (1) the recognition accuracy of mid-air interactions for indoor and outdoor setups based on the experiment conducted (2) temperature at the contact point (T_c) calculated using Equation 1 with T_{skin} = 30$^\circ C$ and b$_{skin}$= 1000 $JS^{-1/2}m^{-2}K^{-1}$. The difference between T_c and $T_{surface}$ should be $>$ 0.08$^\circ C$ to be detected by the Optris PI160 thermal camera.

THERMAL IMAGING & INTERACTIVE SURFACES

Thermal imaging as sensing technology has shown promising potentials for interactive surfaces beyond traditional imaging systems. The robustness of different lighting conditions and the thermal properties allows further flexibility. The results of our study and a review of prior work reveal that surfaces with specific properties should be used to create interactive surfaces using thermal imaging. Based on both findings, we derive properties of surfaces that should be considered to support these interactions. We divide the interaction with a surface into two spaces: (1) interaction on the surface through touch, (2) mid-air gesture interaction in the extended FOV using thermal reflectivity as discussed in [12].

On surface interaction using heat traces

Tracking interaction on a surface using a thermal camera relies on heat traces left behind by the contact of the finger with the surface. The thermodynamic laws state that heat goes from a warm object to a cold object. Hence, the heat transfer between the finger and surface occurs as far as their temperature differs. The amount and direction of heat transferred principally relies on the surfaces material property knowing as the thermal contact conductance [2]. The thermal contact conductance refers to the conductivity of heat between two objects in contact. The amount of heat transferred (conducted) between the hand and the surface in contact could be either reflected or absorbed by the surface.

To determine on-surface interaction, we are interested in the heat trace and the temperature change at the point of contact. Ray suggests a simple model that calculates the temperature at the contact point [11]. Hence, T_c ($^\circ C$) the temperature at the contact point between human's skin and the surface is as follows:

$$T_c = \frac{b_{skin}T_{skin} + b_{surface}T_{surface}}{b_{skin} + b_{surface}} \qquad (1)$$

$$b = \sqrt{K.P.C} \qquad (2)$$

The T_c depends on the temperature of the two contact points (T_{skin} and $T_{surface}$) as well as their thermal penetration coefficient (b). The b depicts the amount of heat penetrated and absorbed by a surface. It is expressed in terms of thermal conductivity (K), thermal density (P), and specific heat capacity (C) [10] (Equation 2). The b of human skin for short contact is 1000 $JS^{-1/2}m^{-2}K^{-1}$ [10]. On the other hand, the detection of temperature changes at the contact point depends on the camera's sensitivity. The changes must be higher than the camera's temperature sensitivity to be visible by the camera.

To detect heat traces, it is necessary to consider its decay time. Based on the Netwon's law of cooling, the rate of heat loss of a body ($Rate_{cooling}$) is proportional to the temperature difference between the body and its surrounding. The higher the difference is, the lower is the cooling rate, thus, the trace lasts longer. If the cooling rate ($Rate_{cooling}$) is smaller than the time one frame lasts ($1/FrameRate_{camera}$), the camera can not sample the trace before it decays. It should be mentioned that the cooling rate depends on other additional factors such as the surface area of the heat being transferred and the heat transfer coefficient between surfaces.

We calculated T_c for the surfaces used in the study by measuring the T_{skin}, the $T_{surface}$, and obtaining the $b_{surface}$ from [10] (Table 1). The sensitivity of the camera used is 0.08$^\circ C$ and it's frame rate is 120Hz. Based on our camera property, the change in temperature at the contact point should be bigger than 0.08$^\circ C$ and should last at least 8.3 milliseconds. The result unveiled that the difference between T_c and $T_{surface}$ is more than the threshold for all surfaces except for aluminum. Thus, the temperature changes on aluminum surfaces is invisible to our camera. In [12], it is empirically tested and reported that no trace can be detected on aluminum using this thermal camera. A camera with higher sensitivity and/or a higher frame rate may reveal other results.

Mid-air gestures through thermal reflectivity

For mid-air interaction specular thermal reflectivity is required which depends on the roughness of a surface. It is reported that the roughness (R_a) of the surface should not exceed 1/8 of the human body radiation, i.e., 1.18 micrometer, to recognize interaction behind the camera's direct FOV [12]. The results of our study reveals that the lower the roughness the sharper the rendered image from the reflection, respectively, higher is the recognition performance. Otherwise, the interaction in extended space is too blurry and cannot be used.

Material Space for Interactive Surfaces

Our investigation reveals that knowing certain information about a surface enables us to determine whether it can support on surface interaction and/or mid-air interaction using thermal reflectivity. Based on our study and prior work we derived a material space for an Optris PI160 thermal camera (Figure 2). As previously described, for on surface interaction the temperature at the contact point (T_c) and its decay

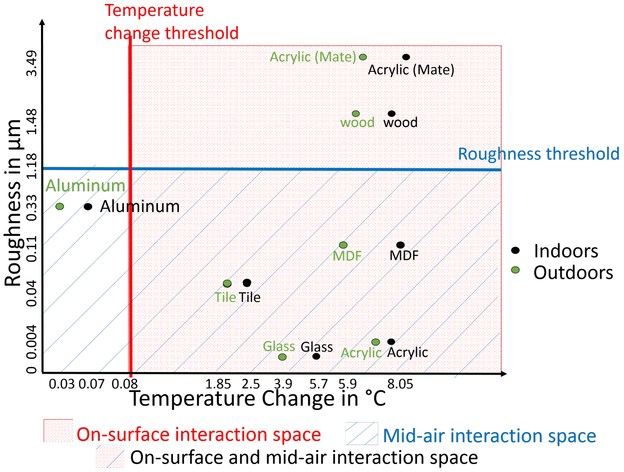

Figure 2. Material space for the Optris PI160 thermal camera: it depicts which materials support on-surface and/or mid-air interaction.

rate should be taken into account (T_c decay_rate 8ms). For For mid-air interaction using thermal reflectivity the surface's roughness should be smaller than 1.18 micrometer.

Based on our setup and our thermal camera, surfaces such as glass, MDF, and tile can be used for both types of interaction. Whereas, aluminum can be used only for mid-air interaction and wood only for on surface interaction.

Constrains

Thermal imaging has shown promising potentials for interactive systems. However, there are still limitations and constrains concerning deploying thermal cameras as sensing technologies. Sensitivity of the camera and its frame rate are one of the main constraints. Furthermore, it effects the sharpness of reflected images rendered. The resolution of the camera should be also considered. We report the material space (Figure 2) for a specific thermal camera (Optris PI160). Using the approach described above, it can be easily derived for cameras with improved properties.

CONCLUSION

In this paper we assessed the recognition accuracy of mid-air interactions sensed through thermal reflectivity using surfaces with different reflection characteristics in indoor and outdoor setups. We further derived the material space and the constraints of interaction enabled through thermal imaging. The guideline allows identifying whether a surface supports on-surface interactions and/or mid-air gestures using a thermal camera as the sensing technology. Knowing the thermal penetration coefficient of a surface, it is possible to determine if heat traces last long enough at the contact point on the surface to be detected by the thermal camera. Further, it is possible to find out if the mid-air interaction space can be extended beyond the camera's direct FOV through thermal reflectivity by knowing the roughness of the surface.

Our work as well as previous work that uses thermal imaging for interactive systems uses standard computer vision techniques originally developed for the visual spectrum. This approach already provides reasonable performance. However, significant improvements can be expected when using algorithms specifically designed to exploit the characteristics of thermal imaging. To sense users through thermal reflection techniques that were initially designed to remove reflections from recorded thermal images [13] could be applied. By separating the scene sensed through reflection and the directly observed, both scenes could be analyzed independently. A further improvement could be achieved by using more advanced thermal cameras. While such cameras were only available for the military, they currently become also available for normal use.

ACKNOWLEDGMENTS

This work is funded by the German Research Foundation within the SimTech Cluster of Excellence (EXC 310/1).

REFERENCES

1. Anthony, L & Wobbrock, J. A lightweight multistroke recognizer for user interface prototypes. In *Proc. GI* (2010).
2. Cooper, M., Mikic, B., and Yovanovich, M. Thermal contact conductance. *International Journal of heat and mass transfer 12*, 3 (1969).
3. Hilliges, O., Izadi, S., Wilson, A. D., Hodges, S., and et al. Interactions in the air: adding further depth to interactive tabletops. In *Proc. UIST* (2009).
4. Iwai, D., and Sato, K. Heat sensation in image creation with thermal vision. In *Proc. ACET* (2005).
5. Iwai, D., and Sato, K. Limpid desk: see-through access to disorderly desktop in projection-based mixed reality. In *Proc. VRST* (2006).
6. Kane, S. K., Avrahami, D., Wobbrock, J. O., Harrison, B., and et al. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In *Proc. UIST* (2009).
7. Larson, E., Cohn, G., Gupta, S., Ren, X., Harrison, B., Fox, D., and Patel, S. Heatwave: thermal imaging for surface user interaction. In *Proc. CHI* (2011).
8. Manresa, C., Varona, J., Mas, R., and Perales, F. Hand tracking and gesture recognition for human-computer interaction. *ELCVIA 5*, 3 (2005).
9. Murugappan, S., Vinayak, and et al. Extended multitouch: recovering touch posture and differentiating users using a depth camera. In *Proc. UIST* (2012).
10. Parsons, K. Contact between human skin & hot surfaces equivalent contact temperature. In *Proc. ICEE* (1992).
11. Ray, R. The theory and practice of safe handling temperatures. *Applied ergonomics 15*, 1 (1984).
12. Sahami Shirazi, A., Abdelrahman, Y., Henze, N., and et al. Exploiting thermal reflection for interactive systems. In *Proc. CHI* (2014).
13. Vollmer, M., Henke, S., Karstädt, D., and et al. Identification and suppression of thermal reflections in infrared thermal imaging. In *Proc. Inframation* (2004).
14. Wilson, A. D., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. UIST* (2010).
15. Wobbrock, J. O., Wilson, A. D., and Li, Y. Gestures without libraries, toolkits or training: a recognizer for user interface prototypes. In *Proc. UIST* (2007).