

# ThermalRing: Gesture and Tag Inputs Enabled by a Thermal Imaging Smart Ring

Tengxiang Zhang<sup>12</sup>, Xin Zeng<sup>12</sup>, Yinshuai Zhang<sup>3</sup>, Ke Sun<sup>4</sup>, Yuntao Wang<sup>5</sup>, Yiqiang Chen<sup>12</sup>

<sup>1</sup>Beijing Key Laboratory of Mobile Computing and Pervasive Device, ICT, CAS

<sup>2</sup>UCAS

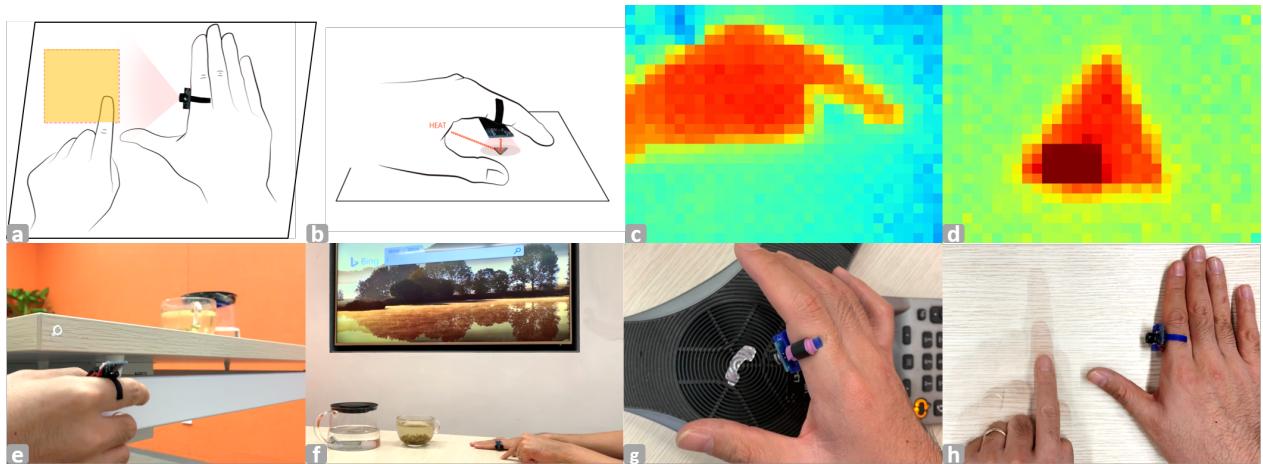
<sup>3</sup>Lenovo Research

<sup>4</sup>Huawei

<sup>5</sup>Tsinghua University

{zhangtengxiang,zengxin18z,yqchen}@ict.ac.cn

zhangys21@lenovo.com sunke36@huawei.com yuntaowang@tsinghua.edu.cn



**Figure 1.** a) Bimanual gesture input on flat surfaces; b) Imaging passive ThermalTags that reflect heat from the hand; c) The thermal image for the interacting hand; d) The thermal image of a triangular ThermalTag; e) Scan a Search ThermalTag to search on a TV; f) Handwrite letters to input search keywords; g) Scan a tag on a landline phone to pair with it; h) Finger slide to adjust call volume.

## ABSTRACT

The heterogeneous and ubiquitous input demands in smart spaces call for an input device that can enable rich and spontaneous interactions. We propose ThermalRing, a thermal imaging smart ring using low-resolution thermal camera for identity-anonymous, illumination-invariant, and power-efficient sensing of both dynamic and static gestures. We also design ThermalTag, thin and passive thermal imageable tags that reflect the heat from the human hand. ThermalTag can be easily made and applied onto everyday objects by users. We develop sensing techniques for three typical input demands: drawing gestures for device pairing, click and slide gestures for device control, and tag scan gestures for quick access. The

study results show that ThermalRing can recognize nine drawing gestures with an overall accuracy of 90.9%, detect click gestures with an accuracy of 94.9%, and identify among six ThermalTags with an overall accuracy of 95.0%. Finally, we show the versatility and potential of ThermalRing through various applications.

## Author Keywords

Thermal imaging;smart ring;gesture recognition;interactive tags

## CCS Concepts

•Human-centered computing → Graphics input devices; Gestural input; Usability testing;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-6708-0/20/04 ...\$15.00.

<http://dx.doi.org/10.1145/3313831.3376323>

## INTRODUCTION

More and more everyday objects are becoming interactive with the trend of Internet of Things (IoT). Such smart objects use different input interfaces for their unique interaction demands. Also, they are often scattered across various locations. The heterogeneous demands and varied locations call for a unified input device that can provide a rich, spontaneous, and consistent interaction experience. Currently, smart phone is widely used to interact with different smart devices. However, the size and weight of a smart phone can also introduce inconveniences under certain circumstances. The process is also complicated since users need to find the target device in an APP first. Vocal commands can also be used for remote control. However, the recognition can be difficult due to user accent variances and background noises. Also, vocal control systems can raise privacy concerns.

Many previous studies propose to use smart wearables as an alternative to provide a more spontaneous interaction experience. Among them, smart rings can enable rich interactions thanks to the dexterity of human fingers. Many of the previously proposed smart rings use Inertial Measurement Unit (IMU) to recognize dynamic gestures. The IMU-based rings could detect finger movements and recognize gestures [13, 14, 31, 37]. TypingRing [25] uses an IMU, two proximity sensors, and a displacement sensor for on-surface typing purpose. Even though IMUs are small and power efficient, they are not able to detect static gestures. Smart rings with RGB cameras [24, 4] are used to recognize gestures and current context. However, the RGB camera can lead to severe privacy issues. Its performance can vary greatly under different illumination conditions. The power consumption of a regular RGB camera is also usually too high for a small battery that can be integrated into a ring.

We propose ThermalRing, a low-resolution thermal imaging smart ring to enable a wide variety of spontaneous interactions in smart spaces. As shown in Figure 1a, ThermalRing can sense gestures performed by one hand (interacting hand) when worn and aligned on the other hand (auxiliary hand). It can also sense temperature differences, thus able to detect heat radiative or reflective objects in the surrounding area. We then propose ThermalTag, thermal-visible passive tags with different shapes made by high heat reflective materials (e.g. aluminum foil). When scanned by ThermalRing, ThermalTags will be ‘illuminated’ by the heat radiated from the human hand and become visible to thermal cameras (Figure 1b). ThermalTags can be easily made by end users in a DIY manner using widely available materials and conveniently deployed on everyday objects and surfaces. Users can then use ThermalRing to scan and identify ThermalTags for quick access and pairing purposes.

Compared with IMU-based wearables, ThermalRing captures image streams to recognize both dynamic and static gestures. Compared with smart wearables using RGB cameras, ThermalRing is identity-anonymous, illumination-invariant, and more power-efficient. Compared with other wearables like smart watches enabled on-skin input techniques [9, 8], the ring form factor makes it easy to adjust camera orientations

to enable more applications. Combined with easily made and deployed ThermalTags, ThermalRing can satisfy various types of input demands in smart spaces.

We built a proof-of-concept ThermalRing using MLX90640 thermal sensor array with a 32x24 low resolution<sup>1</sup>. With this prototype, we focus our discussion on smart device remote control tasks and demonstrate three example application domains: 1) Planar surface-based drawing gesture sensing for smart device pairing; 2) Virtual button and slider input sensing for smart device control; 3) ThermalTag identification for instant smart device access.

In the first example domain, we propose a BoW (Bag-of-Words) algorithm based six-step sensing pipeline for drawing gesture recognition. Nine drawing gestures are selected to represent commonly used devices in smart spaces. The results of the user study show that our system achieves an overall classification accuracy of 90.9%. In the second example domain, we implement a virtual user interface with two buttons and one slider. The study results show that our system can detect click events with an accuracy of 94.9%. It takes 7.57 seconds on average for a participant to complete a sliding task correctly. In the third example domain, we propose inexpensive, passive, and thin thermal-visible tags that can be easily made and applied by end users. We optimize the sizes of the tags and design a tag scan procedure for a quick scan experience. We use Hu’s Moments [20] as learning features to differentiate among six different ThermalTags and achieve an overall accuracy of 95.0%. The results of subjective interviews show that users are willing to wear ThermalRing for a more spontaneous and consistent input experience. At last, we show more possible interactions to demonstrate ThermalRing’s potential as a universal input device.

Our contributions are three-fold:

1. For the first time, we propose to use thermal imaging smart ring as a spontaneous input device in smart spaces. Our solution is identity-anonymous, illumination-invariant, and power-efficient compared with rings equipped with regular RGB cameras.
2. We implement a ThermalRing prototype, develop sensing techniques for three example domains that achieve high recognition accuracy, and conduct user study to validate our system for each application domain.
3. We discuss more potential applications to illustrate the versatility of ThermalRing as an input device in smart spaces.

## BACKGROUND AND RELATED WORK

Thermal cameras capture the object *emitted* Long-wavelength Infrared (LWIR, 8 to 15  $\mu\text{m}$  wavelength) energy. RGB cameras and Near Infrared (NIR) sensors, on the other hand, capture the object *reflected* visible light (400nm-700nm wavelength) and NIR energy (750nm-1.4  $\mu\text{m}$  wavelength), respectively<sup>2</sup>. In this section, we first compare thermal cameras

<sup>1</sup><https://www.melexis.com/en/product/MLX90640/Far-Infrared-Thermal-Sensor-Array>

<sup>2</sup><https://en.wikipedia.org/wiki/Infrared>

with RGB cameras and NIR sensors when integrated on smart wearables, then review works on thermal imaging based interactions, and at last look at related works on tag-based interactions.

### Smart Wearables with RGB Cameras and NIR Sensors

Small and low-cost RGB cameras can be integrated into smart wearables for interaction purposes. The advanced Computer Vision techniques enable RGB cameras to recognize gestures, objects, and contexts. CyclopsRing [4] and EyeRing [24] integrate RGB camera into a ring to recognize input gestures and objects in the surrounding area. MagicFinger [36] uses RGB camera to sense different textures for context detection. One major drawback of RGB camera based wearables are that they can severely infringe user privacy. Also, RGB cameras' sensing performance can vary significantly under different illumination conditions since they only detect reflected energy.

The sizes of NIR transmitters and receivers are also small, which makes them suitable for wearables integration. SensIR [19] and Novest [11] emit NIR signal from a wristband and detect the reflected signal to recognize hand gestures and localize fingertips on the hand back respectively. WaveSense [34] detects in-air gestures by IR arrays on the back of HMD devices. iRing uses an IR transceiver looking onto the finger skin to detect single hand gestures like finger rotations [26]. Combining NIR sensors with RGB cameras can improve recognition performance. For example, PalmGesture [32] uses a wrist wearable composed of an NIR transceiver and a micro-camera for stroking gesture recognition on the palm. The generation of NIR signal increases the power consumption though. The number of NIR transmitters and receivers are thus limited due to the power and size constrains of smart wearables. The low resolution leads to a limited gesture set that can be recognized by NIR-based smart wearables.

The 32x24 thermal camera resolution enables ThermalRing to detect more gestures while ensuring user anonymity. The thermal camera is also illumination-invariant, which can be used in more scenarios. Thermal cameras capture the hand emitted LWIR energy for gesture sensing, which is more power efficient than NIR sensing systems which require extra signal emitters. The temperature differences between human hand and background objects also leads to an easier separation of the recognition target and the background.

### Thermal Imaging Enabled Interactions

Researchers have explored high-resolution thermal cameras in the environment for interaction purposes. For example, ThermoTablet [12] senses object-based inputs on a surface by monitoring its temperature distribution using a bottom-view thermal camera. HeatWave uses a top-view high resolution thermal camera to sense the heat residues after finger touches for touch positions and pressures detection [16]. The set-up thermal cameras can only support interactions on specific surfaces. Sahami et al. expands the thermal camera's field of view by leveraging the reflected heat from smooth surfaces [29], and demonstrates the technique's application for interactions using hand-held projectors. The high resolution thermal cam-

eras are usually large and expensive, which are difficult to integrate into wearable devices.

The passive infrared (PIR) motion sensor is widely used to detect human presence. In essence, PIR sensors can be viewed as thermal cameras with only one pixel. Pyro [5] integrate PIR sensors onto a smart glass and a smart watch to recognize thumb-index finger micro-gestures. PIR sensor arrays are shown to be able to detect simple gestures [35, 3]. The sensing ability of PIR sensors are limited though due to their passive nature.

The ring-form of ThermalRing support gesture input on any surfaces, which enables a more spontaneous interaction experience in smart spaces. The dexterity of fingers also enables more interaction possibilities than sensing devices worn on the wrists or heads. ThermalRing can also recognize a much larger gesture set than that of PIR sensors, which is important for an input device in smart spaces.

### Tag-based Interactions

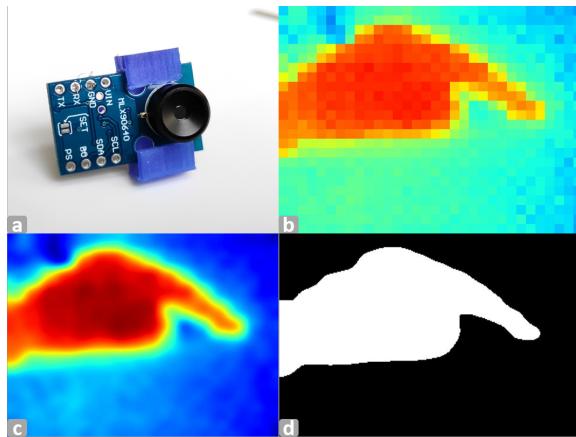
Various types of interactive tags have been used as ubiquitous interactive interfaces to bridge the gap between the physical and virtual world [33]. One widely used optical tag is barcode. Users can scan a barcode using smart phone camera to retrieve its information [28, 21]. Bokode is a small optical tag that can be detected by phone cameras from a distance of up to 2m [22]. Instead of power-heavy cameras, acoustic tags only require high frequency IMU or microphones for sensing purposes. ViBand [15] shows that information can be extracted by touching a vibrating transducer with fingertips. The transducer needs to be powered up by batteries though. Acoustic Barcode [10] enables information extraction by sliding fingernails on the barcode surface. RFID are popular electromagnetic tags due to their passive nature, thin form factor, and extreme low cost. When applied on everyday objects, RFID tags can be used to sense a variety of input gestures [30, 18, 17, 27]. Users can also customize RFID tags to DIY binary sensors [38]. Large and expensive RFID readers need to be installed inside the room to interrogate the passive tags though, which limits their applications.

ThermalTags can be easily made from widely available materials. They are also passive and thin so that users can deploy them on everyday surfaces without any further maintenance. Instead of advanced computing devices like RFID readers, only a low-resolution wearable thermal camera is required for interaction. ThermalTags also enable intuitive interactions since their shapes can be visually interpreted by users (e.g. a heart shape means 'Favorite' operation), which creates a more certain interaction experience.

## THERMALRING

### Hardware Implementation

Our current implementation of ThermalRing uses a MLX90640 thermal camera (40 USD) module for thermal imaging. The module consists of a thermal camera and a STM32F103 microcontroller. The thermal camera is chosen because of its wide Field of View ( $110^\circ \times 75^\circ$ ) for gesture capture, low resolution ( $32 \times 24$ ) for user privacy, as well as small



**Figure 2.** a) Current prototype; b) Raw temperature data; c) Scaled and filtered image; d) After Otsu thresholding and contour filtering.

size (6mm height with a diameter of 8mm) and low power consumption (20mA typical) for ring integration. An AirPod earphone battery (93mWh, 4mm diameter<sup>3</sup>) will last more than one day assuming one-hour usage time. We set the frame rate to 8Hz, which does not change under different illuminations. The microcontroller reads frame data from the camera via I2C and calculates temperature on-chip. The data is then sent to a PC (Core i7 8700@3.2GHz CPU, 16GB RAM) via a serial port. A 3D printed case is used to hold the module and worn on the user's index finger (Figure 2a). The current ring size is 25mm x 15mm x 5mm with a camera height of 8mm.

This prototype is designed for concept validation only. We imagine future implementation integrated with an IMU for purposes like activation, mode switch, and single ring based bimanual gesture recognition. The ring can either classify gestures on-chip or stream data to smartphones for recognition. A higher frame-rate and resolution could benefit recognition accuracy and enable more applications (e.g. tracking heat traces on a table [16]), albeit require more computing resources and power.

### Thermal Image Preprocessing

The temperature data from the camera module needs to be preprocessed. First, the temperature data is assembled into a 32x24 image array (Figure 2b). To reduce the impact of nearby heat sources and reflected heat from the desktop for a cleaner hand silhouette, all pixels with temperatures below 0.8 of the image's maximal temperature is set to 0. The image is then max-min scaled to 0-255 (Figure 2c). Otsu thresholding is then employed to generate a binarized image. At last, only the contour with the largest connected area is kept, which is assumed to be the interacting hand (Figure 2d).

ThermalRing can enable a wide variety of interactions thanks to the dexterity of human fingers. In the following three sections, we describe ThermalRing as an input device in three distinct example domains. For each example domain, we first

describe the input method, then explain the sensing algorithm, and at last conduct a user study to validate the input technique.

### EXAMPLE DOMAIN 1: DRAWING GESTURE SENSING

#### Interaction Design

ThermalRing can recognize a user's drawing gestures on a flat surface in an asymmetrical bimanual interaction setup (Figure 1a). This allows the user to pair and interact with different smart devices remotely via gestures. Specifically, the user places the auxiliary hand on any surface, palm down and stretching the index finger and thumb in orthogonal directions. The index finger and thumb together indicate the interaction area (similar to Imaginary Interfaces [7]). During interaction, the user first uses the index finger of the interacting hand to perform drawing gestures. A gesture starts when the finger touches down on the surface and ends when the finger is lifted up.

Such asymmetrical bimanual interaction is used for three reasons. 1) Asymmetrical bimanual interaction is in line with many common user activities and feels natural. For example, a user usually need to anchor a paper using the auxiliary hand first, then write on it using the interacting hand; 2) By allowing the user to naturally place the entire interacting hand on the surface, such interaction requires less physical efforts compared with mid-air interactions and capacitive touch panels; 3) Users can better locate imaginary UI elements and interact within the recommended area by referring the auxiliary hand. As Guiard [6] stated in his *Right-to-Left Spatial Reference* principle, "motion of the right hand typically finds its spatial references in the results of the motion of the left".

#### Sensing Algorithm

The recognition of drawing gestures consists of six steps (Algorithm 1):

##### Fingertip Extraction

Based on the binarized side-view thermal image, the convex hull of the interacting hand is calculated and the point with the largest horizontal pixel coordinate is designated as the fingertip (Figure 3b, Point F).

##### Finger Lift Detection

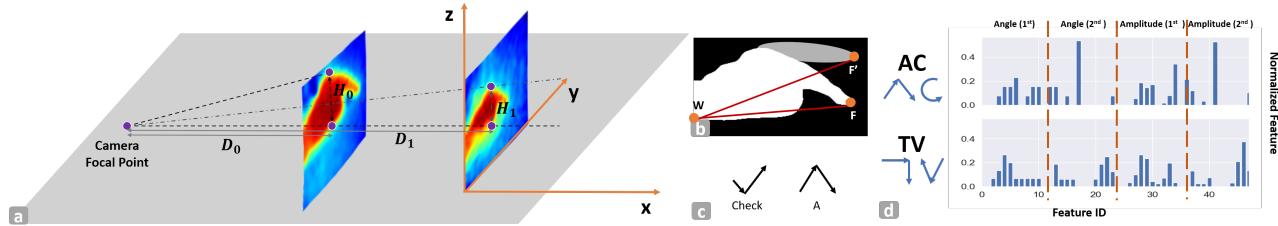
The finger lift status is detected based on the slope of Line WF (Point W being the lower intercept point of wrist and a image boundary, Figure 3b). If the slope is smaller than a preset threshold, the finger is recognized as touched down. Otherwise, it is recognized as lifted up.

##### X/Y Coordinates Estimation

As shown in Figure 3a, the fingertip moves on the x-y plane while the imaged captured by ThermalRing is only hand projection on the y-z plane. The y coordinate can be easily calculated by scaling the y-axis pixel position of Point F (Figure 3b). For x coordinate estimation, we calculate the distance D between the camera focal point and the hand using triangular similarity principle. The perceived focal length FL can be calculated by

$$FL = (H_{real} \times D) / H \quad (1)$$

<sup>3</sup><https://en.wikipedia.org/wiki/AirPods>



**Figure 3.** a) x coordinate estimation; b) Finger lift status detection; c) The check and Graffiti 'A' gestures have the same stroke angles; d) Two example feature vectors for gesture 'AC' and 'TV';;

in which  $H$  is the pixel height of the hand in the thermal image, and  $H_{real}$  is the actual height of the hand. Since  $FL$  and  $H_{real}$  are fixed, we get  $D \propto 1/H$ . Since we only need to estimate the fingertip's position on the x-axis when it is touched down,  $H$  is calculated as the largest distance between a point on the hand contour to Line FW (aligned with the hand bottom when the fingertip touched down).

#### Kalman Filtering

The estimated x and y coordinates is not accurate since the thermal image is distorted and can be noisy. So we implemented a Kalman filter ( $P = 1000$ ,  $R = 5$ ,  $Q$  is modeled as gaussian white noises with a variance of 0.5, empirically decided) to smooth the estimated fingertip position on the x-y plane.

#### BoW based Feature Extraction

Note that the absolute position on the x-axis is unknown using the above estimation method. So we only leverage the finger movement angles and amplitudes to recognize gestures. We divide 0 to 360 degrees equally into 12 angel ranges, each represented by a word in the BoW model [1]. The fingertip movement direction between consecutive frames can be calculated by  $\theta = \arctan(\frac{\Delta y}{\Delta x})$ . The word count increases by 1 if  $\theta$  falls into the angle range of the word. To capture the travel distance along different angels, the movement amplitude is also summed and normalized for each angle range. Some gestures might have the same overall movement angles (e.g. a check and a Graffiti 'A' shown in Figure 3c). To differentiate between such gestures, we extract angle and amplitude histogram features separately for the first half and the second half of the gesture drawing period. The size of each feature vector is then 48. Two example gestures are shown in Figure 3d.

#### SVM Prediction

The extracted feature is fed into a SVM model (rbf kernel, C=1) for training or prediction.

#### User Study 1

##### Study Design and Procedure

We conducted a user study to validate our drawing gesture sensing algorithm. We recruited 8 participants (4 males) with ages ranging from 23 to 30 (MEAN=23.9, SD=2.31) from the local institution. The study lasts for about 1 hour and each participant was compensated with 20 USD. We chose nine drawing gestures that could be used in a smart home pairing scenario: 1) 'AC' for Air-Conditioner; 2) 'TV' for Television;

---

#### Algorithm 1 Drawing Gesture Classification Algorithm

---

```

while image do
     $Pos_{ft} = findFingertip(image)$ 
    if TouchDown( $Pos_{ft}$ ) then                                ▷ 1st touch down
         $kalmanFilter = initKF()$ 
         $hist_{ft} = initHist()$ 
    end if
    if isTouchDown( $Pos_{ft}$ ) then                      ▷ touched down
         $x_{ft}, y_{ft} = coordCal(Pos_{ft})$ 
         $x_{ft}, y_{ft} = kalmanSmooth(kalmanFilter, x_{ft}, y_{ft})$ 
         $\Delta x_{ft} = x_{ft} - x_{Prev_{ft}}$ ,  $\Delta y_{ft} = y_{ft} - y_{Prev_{ft}}$ 
         $\theta = \arctan(\Delta y / \Delta x)$ 
         $r = \sqrt{\Delta x^2 + \Delta y^2}$ 
         $hist_{ft} = BOW(hist_{ft}, \theta, r)$ 
         $x_{Prev_{ft}} = x_{ft}$ ,  $y_{Prev_{ft}} = y_{ft}$ 
    end if
    if LiftUp( $Pos_{ft}$ ) then                         ▷ lift up after touch down
        gesture = svmPredict( $hist_{ft}$ )
    end if
end while

```

---

3) 'A' for Alexa; 4) 'T' for Telephone; 5) 'L1' for Light1 ; 6) 'L2' for Light2; 7) 'P' for Printer; 8) 'K' for Kettle; 9) 'Backspace' for disconnection. We asked the participants to follow the Graffiti gestures [2] and write each letter on top of each other for multi-letter gestures like 'AC' (Figure 4b).

After the participant arrived, we first explained the sensing principles of ThermalRing and the interaction paradigm. Due to the limited frame rate of the thermal camera, we reminded each participant to spend at least one second writing each character. Then the participant warmed up for 5 minutes. The study contained three sessions, each with nine blocks. Each block corresponded to one drawing gesture. Within each block, the participant performed 20 trials of the gestures, then rested for 30 seconds. The gesture sequence was randomized within each session. Between each session, we asked the participant to take down the ring, stand up, and walk around for one minute. We measured the hand temperature of the participant using Non-contact Infrared thermometer at the start and end of each session. At last, we interviewed the participant to understand his/her subjective feelings for the technique and using smart rings as an input device in general (5 point Likert scale, the higher the better).

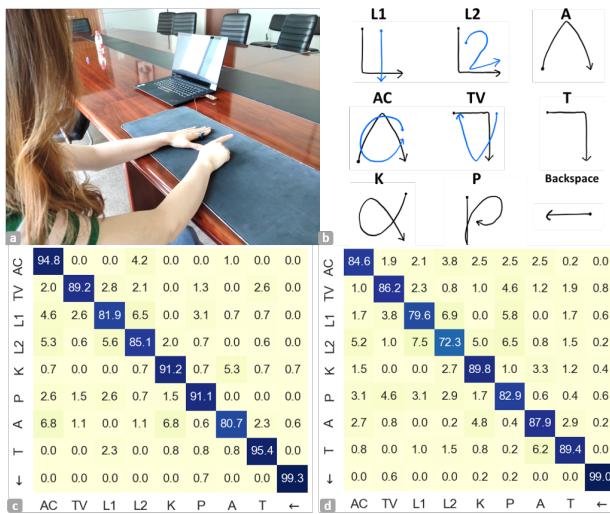


Figure 4. a) Experiment setup; b) Gesture set used in the study; c) Averaged within-user confusion matrix; d) Averaged between-user confusion matrix.

### Study Results

A total of  $9 \text{ gestures} \times 6 \text{ participants} \times 3 \text{ sessions} \times 20 \text{ trials} = 3240$  drawing gesture trials are collected, 360 for each gesture type. The hand temperatures of participants range from  $35.9^{\circ}\text{C}$  to  $36.7^{\circ}\text{C}$  (MEAN =  $36.3$ , SD =  $0.23$ ). The maximal hand temperature variations within a session is  $0.2^{\circ}\text{C}$  and  $0.5^{\circ}\text{C}$  during the whole study for all participants. The overall classification accuracy is 90.9% with 3-fold cross validation. Within-user classification accuracy uses the data collected from the same participant for both training and testing. For each participant, 70% of the trails are randomly selected for training and the rest for testing. We average the within-user accuracies for all participants, and show the results in Figure 4c. The overall within-user accuracy is 89.2% (SD = 0.04). Between-user classification accuracy is calculated by using data collected from five participants for training and the rest one for testing. The overall between-user accuracy averaged across all participants is 85.7% (SD = 0.06). The lower between-user confusion is expected since different users may draw the same gesture with different rhythms and timings. Another possible reason is that the participant's hand can drift outside of the suggested interaction area. The drawing area was determined by each participant based on the thumb and index finger position. This may cause variations in calculated angles for the same gesture, which leads to deteriorated recognition performance among gestures with similar drawing strokes like 'L1', 'L2' and 'P'. We believe a thermal camera with higher frame rate can capture more details of the drawing gestures, thus able to better differentiate such gestures. The averaged processing time between two consecutive frames is 30.6ms, which is fast enough to capture the gesture movements in the study.

All participants are willing to wear one smart ring for smart home remote control (MEDIAN=5, MODE=5) and smart device quick access (MEDIAN=4, MODE=4) purposes. If they are already wearing a ring, all participants are willing to use it

for interaction purposes (MEDIAN=5, MODE=5) for a convenient and spontaneous input experience. All participants do not mind rotating the ring for different interactions since it is quite easy (MEDIAN=5, MODE=5). For the drawing gesture interaction technique, all participants rate the interaction as comfortable (MEDIAN=4, MODE=4) and convenient (MEDIAN=4.5, MODE=4). They are neutral on the input speed though (MEDIAN=3, MODE=3). The input speed can be improved by using a camera with higher frame rate.

### EXAMPLE DOMAIN 2: CLICK AND SLIDE GESTURE SENSING

#### Interaction Design

After paired with the target device, it is more convenient to continue control the devices with the same bimanual setup by manipulating virtual buttons and sliders in the interaction area via click and slide gestures. The user can refer to the locations and visual features of the auxiliary hand's thumb and index finger to localize virtual UI elements. Figure 5a shows an virtual interface with two buttons and one slider.

#### Sensing Algorithm

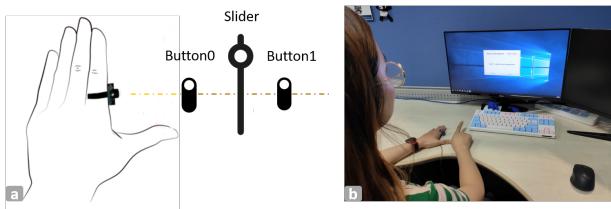
As mentioned above, the finger lift status can be detected by monitoring the slope between index fingertip and the wrist. So a click gesture can be detected when the fingertip briefly touches down and then lifts up within a threshold duration ( $TH_t = 1.5\text{second}$ ). A slide gesture is recognized if the fingertip touches down for longer than  $TH_t$ . UI elements at different locations in the interaction area can be separated by the estimated x and y coordinates. A calibration process is required when there are multiple UI elements at different x positions. Before each control session, users perform calibration by a simple click gesture with interacting hand placed right next to the stretched thumb of the auxiliary hand. The pixel height of the hand is then recorded. UI elements at different x positions can then be differentiated by the ratio of the current pixel hand height to the calibrate pixel hand height value.

### User Study 2

#### Study Design and Procedure

We conducted this study to validate the usability of the virtual UI interface. More specifically, we want to know whether our technique can robustly detect finger click gestures and slide gestures. We evaluated a UI interface with two virtual buttons (one close to the auxiliary hand, the other is farther away) and one slider with 5 scales. The layout of UI elements was invisible to users. Only recognition results were shown on the screen (Figure 5b). We recruited 8 participants (4 males) with ages ranging from 23 to 30 (MEAN =  $24.3$ , SD =  $2.38$ ) from the local institution. The study lasted for about 20 minutes, and each participant was compensated with 15 USD.

We first explained the input method and sensing principles to the participant. Then the participant warmed up for 2 minutes. The study contained three sessions. At the beginning of each session, the participant calibrated the hand height by click the index finger once. The participant then clicked each button 16 times and adjusted the slider 8 times (2 times for each level). The clicked button and slider level was displayed on a monitor



**Figure 5.** a) The virtual sliders and buttons; b) Experiment interface

for reference. A 20 seconds timeout was enforced between tasks. The task sequence was randomized for each session. The participant rested for one minute between sessions. We measured the hand temperature of the participant using Non-contact Infrared thermometer at the start and end of each session. At last, the participant finished a questionnaire on a 5-point Likert Scale (the higher the better).

#### Study Results

A total of 8 participants  $\times$  2 buttons  $\times$  16 clicks  $\times$  3 sessions = 768 click gestures and 8 participants  $\times$  1 slider  $\times$  8 slides  $\times$  3 sessions = 192 slide gestures were collected. The hand temperatures of participants ranged from 36.1°C to 36.7°C (MEAN = 36.3, SD = 0.18). The maximal variations of hand temperatures within the study is 0.2 °C with data from all participants. The overall button click detection accuracy is 94.9% (SD=0.02). Some participants mentioned that sometimes they did not follow the instruction correctly, which account for some of the errors. All but one slider task were successfully finished before timeout. On average, it took the participants 7.57 seconds to finish slider adjustment tasks. Our technique requires a touch down duration of 1.5 second to avoid false positives of sliding gestures. One-way ANOVA does not show significant differences of button click recognition accuracies ( $F_{1,7} = 2.35, p = .12$ ) or slide time ( $F_{1,7} = 0.50, p = .61$ ) between sessions.

On average, the participants felt they can reliably locate 4 buttons (SD = 1) and 2 sliders (SD = 0.71) at most within the recommended interaction area. All participants felt it easy to locate the two buttons in this study (MEDIAN = 5, MODE = 5), the interactions were precise (MEDIAN = 5, MODE = 5) and not tiring (MEDIAN = 5, MODE = 5).

#### EXAMPLE DOMAIN 3: THERMALTAG IDENTIFICATION

##### Principle and Interaction Design

Thermal cameras can identify materials with different heat reflectivity [29]. We leverage this nature and propose ThermalTag-thin and passive tags made of high heat reflection materials and applied on surfaces with lower heat reflectivity. When close to human hand, the tag reflects the heat from the hand and form contrast on the thermal image between the tag and the surface. The thermal image can then be processed to recognize the tag shape.

In this application, the tags with different shapes (e.g. triangle for 'Play', rectangle for 'Stop') are applied on surfaces of various objects. For example, users can apply tags on a desk for TV control, on a wall for AC control, or even on a cup

handle for lights control. One advantage of ThermalRing is that such tags can even be embedded inside a surface since ThermalRing is robust under different illumination conditions. Users can easily make such tags by cutting widely available materials like aluminum foil or tapes (Figure 6a).

Users can scan tags on various surfaces through a two step process(Figure 6b). The first step is to align ThermalRing with the tag to capture a complete tag image. However, this could be challenging when the ring camera points downward and the tag is blocked by the hand. The second step is to hold ThermalRing at a distance from the tag to capture a complete tag image. To enable an easier and faster scan experience, we designed a two-step tag scan procedure: 1) Touch the camera on the tag to ensure precise alignment; 2) Lift the hand and keep it at a distance from the tag to take a complete picture.

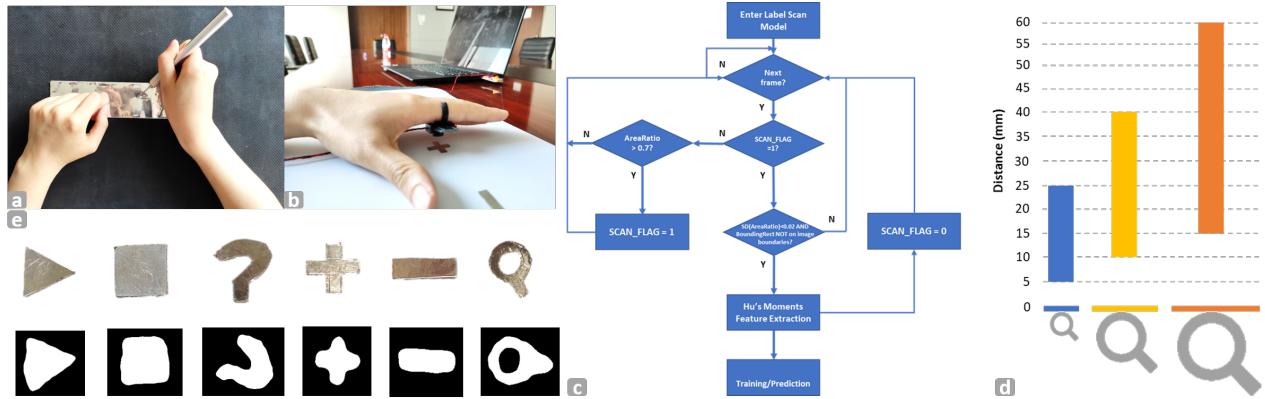
The size of the tag needs to be carefully designed. For smaller tags, a close distance between the camera and the tag is needed for a complete and clear image due to the camera's low resolution. This requires precise tag alignment using the finger-worn camera, which can be difficult for users. It is easier to align with bigger tags since the imaging distance can be larger. However, there will be less heat radiated by the hand reflected by the tag as the camera (thus the hand) moves farther away, thus parts or all of the tag may not be 'illuminated'. We conducted a pilot study to find the tag size that enables better scan experience.

##### Sensing Algorithm

For a more natural scan experience, users should be able to approach the tag from different angles and keep the camera at different distances above the tag, as long as a clean and complete image can be captured. The Hu's Moments [20] are used as learning features since they are translation, scale, and rotation invariant shape descriptors. When the camera touches on the tag, most pixels of the image should be high temperature due to camera's self-heat. So the start of the scan can be detected when the bounding rectangle of the largest contour fills more than 70% (area ratio) of the image. When the hand moves away from the tag, the area ratio decreases until the user holds the hand still for image capture. The end of the scan is then detected when the standard deviation of the area ratio is less than 0.02 (decided in the pilot study). To ensure a complete image is captured, the edges of the contour's bounding rectangle should not reside on any image boundaries. The Hu's Moments of the bounded image area is then calculated and log-transformed for a comparable magnitude. The features are then fed into a RandomForest classifier (1500 estimators) for prediction. The sensing flow chart is shown in Figure 6c.

##### Pilot Study

We conducted a pilot study to determine the tag sizes used for the later study. To understand the tag-camera distance ranges for tags with different sizes, we made the Search tag with three different sizes-8mm, 15mm, 20mm height. The 'Search' tag is chosen for its rich shape features. The minimal distance is set when a complete image fills 90% of the image, and the maximal distance is set when the search tag cannot



**Figure 6.** a) ThermalTags can be fabricated by cutting aluminum tape; b) Study 3 Experiment Setup; c) ThermalTag identification algorithm flow chart; d) Camera-tag distances for the 'Search' tag with 8mm, 15mm and 20mm heights; e) ThermalTags used in the study and their corresponding example scan image.

be distinguished from a circle anymore. In general, larger tags require a greater camera-tag distance to form a complete tag image. They also work within a larger range of distances, which can better account for variances among different scans (Figure 6d). During the study, we found that the heat from one hand is enough to 'illuminate' the entire tag for small tags. Due to the larger distance between the hand and the tag, the heat from the hand is not enough to 'illuminate' the entire 20mm tag though. The incomplete image can impact tag recognition accuracies. For an easier and faster scan experience, we choose a tag size of 20mm and add one step to the scan process. When the camera touches the tag, put both hands together and then lift hands. The heat from two hands ensures a complete tag imaging.

### User Study 3

#### Study Design and Procedure

To validate our sensing algorithm, we conducted another user study. We made six ThermalTags from aluminum tapes with different shapes: 1) Play, a triangle; 2) Stop, a square; 3) Search, a magnifier; 4) Help, a question mark; 5) Up, a plus sign; 6) Down, a minus sign. Their shapes and the corresponding binarized thermal image are shown in Figure 6e. They were placed on a white PVC plastic sheet with 70mm separation between each other. We recruited 8 participants (4 females) with ages ranging from 23 to 30 (MEAN = 24.3, SD = 2.38). The study lasted for about 30 minutes and each participant was compensated for 12 USD.

The participant sat on a chair with the tags placed around 20cm (within arm reach) on the desk. We first explained the sensing principles to the participant. Then the participant warmed up for 5 minutes to get familiar with the scan procedure. The study contained two sessions, each with six blocks. In each block, the participant scanned one tag for 20 times, then rested for 30 seconds. The order of the block sequence was randomized within each session. We also recorded the scanning time for each trial. Similar to Study 1, we asked the participant to take down the ring, stand up, and walk around for 1 minute between sessions. After the study, we asked the participants

	Up	Down	Play	Stop	Search	Help	Up	Down	Play	Stop	Search	Help
Up	93.1	0.0	1.6	3.8	0.8	0.7	93.2	0.0	0.3	0.9	1.5	4.1
Down	0.0	95.5	4.5	0.0	0.0	0.0	0.9	95.0	3.1	0.3	0.3	0.3
Play	1.2	0.0	97.9	0.8	0.0	0.0	2.2	0.0	96.6	0.9	0.3	0.0
Stop	0.0	0.0	0.0	100.0	0.0	0.0	4.7	5.0	1.2	89.1	0.0	0.0
Search	0.0	0.0	0.0	0.0	98.2	1.8	1.2	0.0	0.9	0.6	87.3	9.9
Help	0.8	0.0	0.0	0.0	4.7	94.5	2.8	1.6	0.6	0.0	15.4	79.6

**Figure 7.** a) Within-user classification confusion matrix; b) Between-user classification confusion matrix.

to rate the physical, mental, and temporal effort of the scan operation using a 5-point Likert scale (higher score indicates lower efforts).

#### Study Results

A total of 8 participants  $\times$  2 sessions  $\times$  6 tags  $\times$  20 trials = 1920 scans were collected, 320 for each tag. The hand temperatures of participants range from 36.1°C to 36.4°C (MEAN = 36.1, SD = 0.15). The maximal variations of hand temperatures within the study is 0.5 °C across all participants. The detection accuracy using a 3-fold cross validation is 95%. The averaged within-in user sensing accuracy is 95.0% (SD = 0.04, Figure 7a) while the averaged between-user sensing accuracy is 90.1% (SD = 0.08, Figure 7b). The difference is due to that different participants scanned the tag by approaching from different angles at slightly different tag-camera distances. The confusion rate is high between the 'Help' and 'Search' tag for between-user validation (13.2%). One reason could be that when the distance is large (e.g. close to 60mm), the shapes of the two tags become small and similar, thus are difficult to be separated using Hu's Moments. The sensing performance can be improved if user scan the tag with a recommended distance (e.g. 30mm) or a camera with higher resolution is used. Advanced shape descriptors (e.g. shape context [23]) can be used to identify more complex shapes. The averaged scan time is 3.5 seconds, which shows that ThermalTag can



**Figure 8.** a) Drop a curtain on a desktop; b) Start music on a chair arm; c) Navigate slides from a whiteboard; d) Turn on lights while pushing a door open;

provide a free and instant scan experience. In general, the participants felt that the scan interaction did not require high physical (MEDIAN = 4, MODE = 4) and mental (MEDIAN = 4, MODE = 4) efforts. They also felt the scan is fast enough for quick access tasks (MEDIAN = 4, MODE = 4).

## APPLICATION SCENARIOS

### Smart Curtain Control on a Table

ThermalRing enables easy control of a smart curtain from any surface in a natural, spontaneous, and intuitive way (Figure 8a). For example, the user can write 'CT' to pair with the curtain, then perform a down sliding gesture to drop the curtain. An up sliding gesture would rise the curtain.

### Smart Speaker Control on a Chair

ThermalRing works even on a small surface like a chair armrest (Figure 8b). While sitting on a chair, the user can write a 'K' to pair with a nearby music player, click a button to turn on the speaker, then slide to adjust volume. Users can keep interacting with the music player. For example, users can draw a heart shape to add the current song to the favorite list.

### Slides Navigation on Whiteboard

ThermalTag enables users to add control functions to existing objects in an easy and spontaneous way (Figure 8c). For example, the user can make and apply ThermalTags on a whiteboard to navigate the projected presentation slides.

### Smart Light Control on a Door

The scanning of ThermalTag can be combined with the existing interaction gestures for an implicit interaction experience (Figure 8d). For example, a ThermalTag can be applied on a door. The user can scan the ThermalTag while pushing the door open to turn on the lights inside the door.

## DISCUSSION AND LIMITATIONS

### Robustness

In this paper, we filtered out pixels with low temperatures to mitigate impacts from nearby heat sources and reflections. An absolute temperature upper bound can also be set to further filter other heat sources in the environment. Background removal techniques can be leveraged to further improve robustness to interferences.

ThermalRing can robustly localize the index fingertip between 4cm to 32cm away from the ring, which is sufficient for applications discussed in this paper. Thermal imaging also works

as long as enough hand heat radiates through the occlusion (e.g. a thin cotton glove).

Heat can be reflected by specular surfaces (e.g. glass) and blocked by curved surfaces (e.g. lap). For specular surfaces, we can leverage the formed symmetric image due to heat reflection for hand segmentation, and only keep the upper part of the image. The surface curvatures can block the imaging of index fingertip, which leads to inaccurate fingertip localization. One possible solution is using hand models to reconstruct the blocked image and extrapolate to estimate the blocked fingertip position.

Since ThermalRing is mainly intended for interactions in smart spaces, we only conducted indoor lab controlled evaluations in this paper. In-the-wild and outdoor studies are still valuable to fully evaluate ThermalRing's robustness as a final product.

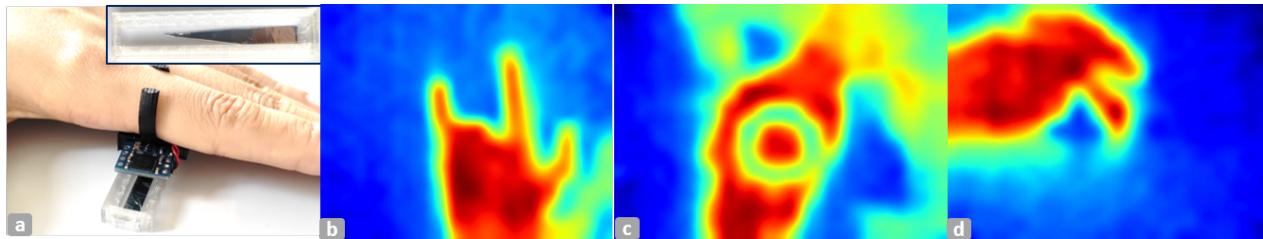
### Mode Switch and Feedback

Thanks to the flexibility of fingers, ThermalRing can support various applications by rotating the ring or wearing the ring on different fingers. A mode switch mechanism is thus necessary for such a multi-functional input device, since different tasks may require different sensing pipelines. We imagine shaking (with IMU) and pressing (with buttons) the ring as two practical ways to iterate through modes. ThermalRing can also switch modes by recognizing current gestures. For example, double taps on the camera can iterate through different modes; closing the thumb to cover the camera for several seconds can put the camera into sleep mode to save power; a three seconds spreading the interacting hand could start the handwriting input mode; a three seconds thumb-index finger pinch could start the virtual UI mode.

Visual and haptic feedback can enable a more certain interaction experience. For example, an LED can single blink for three seconds when successfully paired with a smart device, and double blink when the smart device refuses the connection request. Different vibration patterns can indicate the current mode and remind users when the fingertip moved outside of the interaction area.

### Discreet and One-handed Interaction

Bimanual interactions discussed in our paper are less likely intended to be discreet. That being said, users can still place one hand on each leg to interact on-lap. Aside from the bi-manual gestures discussed in this paper, ThermalRing can also recognize single-hand gestures. For example, the camera can capture thumb tap, swipe, and finger pinches [4]. Thanks



**Figure 9.** a) ThermalTag enabled slider; b) Thermal image of a 'Rock&Roll' gesture; c) Thermal image of a tape on the palm; d) Thermal image when holding a pen.

to ThermalRing's illumination-invariant imaging, users can perform such gestures inside a pocket or a bag, which leads to a discreet interaction experience. ThermalRing can also scan tags at discreet places (e.g. on the bottom side of a table) with proper haptic designs that help alignments.

## FUTURE WORK

Aside from the three example domains, ThermalRing can enable more interaction possibilities. We believe the following four domains are particularly interesting for further exploration:

**ThermalTag-based UI Elements** Our technique allows users to scan thin passive tags in close proximity. However, scan distances and angles can impact the recognition accuracy. 3D printed tag cases and rails can enforce a fixed scan angle and distance, which will improve the sensing performance. Figure 9a shows a slider ThermalTag with a 3D printed rail.

**In-air Gesture Recognition** The on-surface interaction can provide tactile feedback for a more precise and certain input experience. ThermalRing can also enable less precise but more spontaneous in-air interactions. For example, a user can perform a 'Rock&Roll' gesture to start playing music (Figure 9b).

**Everyday Object Recognition** The temperature of the object can either be higher or lower than that of the background for thermal imaging based recognition. Aside from radiating or reflecting heat, an object can be imaged if it blocks the heat (e.g. from a human body). Figure 9c shows that the tape can be clearly imaged when placed on the palm.

**Context-aware Interactions** ThermalRing can also detect current contexts based on the number, shape, and relative locations of the heat sources, either by adjusting the ring's orientation or in an opportunistic manner. After the context is detected (e.g. typing, drawing with a stylus), users can perform single hand gesture to facilitate the current task. For example, when a typing scenario is detected, the users can swipe thumb on the ring to switch modes; when a stylus drawing scenario is detected (Figure 9d), users can tap on the ring to change color.

## CONCLUSION

In this paper, we introduced ThermalRing to enable rich, consistent, and spontaneous interactions in smart spaces. Our

minimal-viable prototype uses small, low-resolution, low-power thermal sensor arrays camera, which can fit into a ring for identity-anonymous imaging. We discussed three example input techniques, described their sensing algorithms, and conducted studies for validation. The first example domain is recognizing drawing gestures for smart device pairing tasks. We implemented a BoW-based six-step sensing pipeline and validated that it can recognize nine drawing gestures with a within-user accuracy of 89.2% and between-user accuracy of 85.7%. The second example domain is recognizing click and slide gestures on a virtual UI interface. We implemented a virtual UI of two buttons and one slider and validated that users can click the buttons with an accuracy of 94.9%. It took 7.57 seconds on average to finish a control task for a 5 level slider. In the third example domain, we introduced ThermalTag, thin and passive tags that can be imaged by thermal cameras. Our system can correctly classify 6 ThermalTags with an accuracy of 95.0% while allowing a free and natural scanning process. In the end, we demonstrated more applications that are supported by the proposed input techniques, as well as those that can be enabled by ThermalRing in the future.

## ACKNOWLEDGMENTS

We thank Yang Zhang and Yiqin Lu for their helpful suggestions. We also thank the anonymous reviewers for their insightful comments. This work is supported by the National Key Research and Development Plan of China (No. 2016YFB1001200), National Natural Science Foundation of China (No. 61972383), R&D Plan in Key Field of Guangdong Province (No. 2019B010109001), and Beijing Municipal Science & Technology Commission (No. Z171100000117017).

## REFERENCES

- [1] 2019a. Bag-of-words model in computer vision. (Sept. 2019). [https://en.wikipedia.org/w/index.php?title=Bag-of-words\\_model\\_in\\_computer\\_vision&oldid=915448145](https://en.wikipedia.org/w/index.php?title=Bag-of-words_model_in_computer_vision&oldid=915448145) Page Version ID: 915448145.
- [2] 2019b. Graffiti (Palm OS). (Sept. 2019). [https://en.wikipedia.org/w/index.php?title=Graffiti\\_\(Palm\\_OS\)&oldid=915451732](https://en.wikipedia.org/w/index.php?title=Graffiti_(Palm_OS)&oldid=915451732) Page Version ID: 915451732.
- [3] T. David Binnie, A. F. Armitage, and P. Wojtczuk. 2017. A passive infrared gesture recognition system. In *2017 IEEE SENSORS*. 1–3. DOI: <http://dx.doi.org/10.1109/ICSENS.2017.8234402>

- [4] Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 549–556. DOI: <http://dx.doi.org/10.1145/2807442.2807450> event-place: Charlotte, NC, USA.
- [5] Jun Gong, Yang Zhang, Xia Zhou, and Xing-Dong Yang. 2017. Pyro: Thumb-Tip Gesture Recognition Using Pyroelectric Infrared Sensing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 553–563. DOI: <http://dx.doi.org/10.1145/3126594.3126615>
- [6] Yves Guiard. 1987. Asymmetric Division of Labor in Human Skilled Bimanual Action. *Journal of Motor Behavior* 19, 4 (Dec. 1987), 486–517. DOI: <http://dx.doi.org/10.1080/00222895.1987.10735426>
- [7] Sean Gustafson, Christian Holz, and Patrick Baudisch. 2011. Imaginary phone: learning imaginary interfaces by transferring spatial memory from a familiar device. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM Press, Santa Barbara, California, USA, 283. DOI: <http://dx.doi.org/10.1145/2047196.2047233>
- [8] Chris Harrison, Julia Schwarz, and Scott E. Hudson. 2011. TapSense: enhancing finger interaction on touch surfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM Press, Santa Barbara, California, USA, 627. DOI: <http://dx.doi.org/10.1145/2047196.2047279>
- [9] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: appropriating the body as an input surface. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 453. DOI: <http://dx.doi.org/10.1145/1753326.1753394>
- [10] Chris Harrison, Robert Xiao, and Scott Hudson. 2012. Acoustic barcodes: passive, durable and inexpensive notched identification tags. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, Cambridge, Massachusetts, USA, 563. DOI: <http://dx.doi.org/10.1145/2380116.2380187>
- [11] Yu Ishikawa, Buntarou Shizuki, and Junichi Hoshino. 2017. Evaluation of finger position estimation with a small ranging sensor array. In *Proceedings of the 5th Symposium on Spatial User Interaction - SUI '17*. ACM Press, Brighton, United Kingdom, 120–127. DOI: <http://dx.doi.org/10.1145/3131277.3132176>
- [12] Daisuke Iwai and Kosuke Sato. 2005. Heat sensation in image creation with thermal vision. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '05*. ACM Press, Valencia, Spain, 213–216. DOI: <http://dx.doi.org/10.1145/1178477.1178510>
- [13] Lei Jing, Zixue Cheng, Yinghui Zhou, Junbo Wang, and Tongjun Huang. 2013. Magic Ring: a self-contained gesture input device on finger. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia - MUM '13*. ACM Press, Luleå, Sweden, 1–4. DOI: <http://dx.doi.org/10.1145/2541831.2541875>
- [14] Lei Jing, Yinghui Zhou, Zixue Cheng, and Tongjun Huang. 2012. Magic Ring: A Finger-Worn Device for Multiple Appliances Control Using Static Finger Gestures. *Sensors (Basel, Switzerland)* 12, 5 (May 2012), 5775–5790. DOI: <http://dx.doi.org/10.3390/s120505775>
- [15] Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*. ACM Press, Tokyo, Japan, 321–333. DOI: <http://dx.doi.org/10.1145/2984511.2984582>
- [16] Eric Larson, Gabe Cohn, Sidhant Gupta, Xiaofeng Ren, Beverly Harrison, Dieter Fox, and Shwetak Patel. 2011. HeatWave: thermal imaging for surface user interaction. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM Press, Vancouver, BC, Canada, 2565. DOI: <http://dx.doi.org/10.1145/1978942.1979317>
- [17] Hanchuan Li, Eric Brockmeyer, Elizabeth J. Carter, Josh Fromm, Scott E. Hudson, Shwetak N. Patel, and Alanson Sample. 2016. PaperID: A Technique for Drawing Functional Battery-Free Wireless Interfaces on Paper. ACM Press, 5885–5896. DOI: <http://dx.doi.org/10.1145/2858036.2858249>
- [18] Hanchuan Li, Can Ye, and Alanson P. Sample. 2015. IDSense: A Human Object Interaction Detection System Based on Passive UHF RFID. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2555–2564. DOI: <http://dx.doi.org/10.1145/2702123.2702178>
- [19] Jess McIntosh, Asier Marzo, and Mike Fraser. 2017. SensIR: Detecting Hand Gestures with a Wearable Bracelet using Infrared Transmission and Reflection. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*. ACM Press, Québec City, QC, Canada, 593–597. DOI: <http://dx.doi.org/10.1145/3126594.3126604>
- [20] Ming-Kuei Hu. 1962. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory* 8, 2 (Feb. 1962), 179–187. DOI: <http://dx.doi.org/10.1109/TIT.1962.1057692>

- [21] Kento Miyaoku, Anthony Tang, and Sidney Fels. 2007. C-Band: A Flexible Ring Tag System for Camera-Based User Interface. In *Virtual Reality (Lecture Notes in Computer Science)*, Randall Shumaker (Ed.). Springer Berlin Heidelberg, 320–328.
- [22] Ankit Mohan, Grace Woo, Shinsaku Hiura, Quinn Smithwick, and Ramesh Raskar. 2009. Bokode: Imperceptible Visual Tags for Camera Based Interaction from a Distance. In *ACM SIGGRAPH 2009 Papers (SIGGRAPH '09)*. ACM, New York, NY, USA, 98:1–98:8. DOI: <http://dx.doi.org/10.1145/1576246.1531404>
- [23] G. Mori, S. Belongie, and J. Malik. 2005. Efficient shape matching using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 11 (Nov. 2005), 1832–1837. DOI: <http://dx.doi.org/10.1109/TPAMI.2005.220>
- [24] Suranga Nanayakkara, Roy Shilkrot, and Pattie Maes. 2012. EyeRing: an eye on a finger. (2012), 4.
- [25] Shahriar Nirjon, Jeremy Gummesson, Dan Gelb, and Kyu-Han Kim. 2015. TypingRing: A Wearable Ring Platform for Text Input. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '15*. ACM Press, Florence, Italy, 227–239. DOI: <http://dx.doi.org/10.1145/2742647.2742665>
- [26] Masa Ogata, Yuta Sugiura, Hirotaka Osawa, and Michita Imai. 2012. iRing: intelligent ring using infrared reflection. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, Cambridge, Massachusetts, USA, 131. DOI: <http://dx.doi.org/10.1145/2380116.2380135>
- [27] Swadhin Pradhan, Eugene Chai, Karthikeyan Sundaresan, Lili Qiu, Mohammad A. Khojastepour, and Sampath Rangarajan. 2017. RIO: A Pervasive RFID-based Touch Gesture Interface. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking (MobiCom '17)*. ACM, New York, NY, USA, 261–274. DOI: <http://dx.doi.org/10.1145/3117811.3117818>
- [28] Jun Rekimoto and Yuji Ayatsuka. 2000. CyberCode: Designing Augmented Reality Environments with Visual Tags. In *Proceedings of DARE 2000 on Designing Augmented Reality Environments (DARE '00)*. ACM, New York, NY, USA, 1–10. DOI: <http://dx.doi.org/10.1145/354666.354667>
- [29] Alireza Sahami Shirazi, Yomna Abdelrahman, Niels Henze, Stefan Schneegass, Mohammadreza Khalilbeigi, and Albrecht Schmidt. 2014. Exploiting thermal reflection for interactive systems. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, Toronto, Ontario, Canada, 3483–3492. DOI: <http://dx.doi.org/10.1145/2556288.2557208>
- [30] A. P. Sample, D. J. Yeager, and J. R. Smith. 2009. A capacitive touch interface for passive RFID tags. In *2009 IEEE International Conference on RFID*. 103–109. DOI: <http://dx.doi.org/10.1109/RFID.2009.4911212>
- [31] Takayuki Iwamoto and Hiroyuki Shinoda. 2007. Finger ring device for tactile sensing and human machine interface. In *SICE Annual Conference 2007*. IEEE, Takamatsu, Japan, 2132–2136. DOI: <http://dx.doi.org/10.1109/SICE.2007.4421339>
- [32] Cheng-Yao Wang, Min-Chieh Hsiu, Po-Tsung Chiu, Chiao-Hui Chang, Liwei Chan, Bing-Yu Chen, and Mike Y. Chen. 2015. PalmGesture: Using Palms as Gesture Interfaces for Eyes-free Input. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*. ACM Press, Copenhagen, Denmark, 217–226. DOI: <http://dx.doi.org/10.1145/2785830.2785885>
- [33] Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. 1999. Bridging Physical and Virtual Worlds with Electronic Tags. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 370–377. DOI: <http://dx.doi.org/10.1145/302979.303111>
- [34] Anusha Withana, Shanaka Ransiri, Tharindu Kaluarachchi, Chanaka Singhabahu, Yilei Shi, Samitha Elvitigala, and Suranga Nanayakkara. 2016. waveSense: Ultra Low Power Gesture Sensing Based on Selective Volumetric Illumination. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16 Adjunct*. ACM Press, Tokyo, Japan, 139–140. DOI: <http://dx.doi.org/10.1145/2984751.2985734>
- [35] Piotr Wojtczuk, David Binnie, Alistair Armitage, Tim Chamberlain, and Carsten Giebel. 2013. A Touchless Passive Infrared Gesture Sensor. In *Proceedings of the Adjunct Publication of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13 Adjunct)*. ACM, New York, NY, USA, 67–68. DOI: <http://dx.doi.org/10.1145/2508468.2514713>  
event-place: St. Andrews, Scotland, United Kingdom.
- [36] Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic Finger: Always-available Input Through Finger Instrumentation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 147–156. DOI: <http://dx.doi.org/10.1145/2380116.2380137>
- [37] Boning Zhang, Yiqiang Chen, Yueliang Qian, and Xiangdong Wang. 2011. A ring-shaped interactive device for large remote display and mobile device control. In *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*. ACM Press, Beijing, China, 473. DOI: <http://dx.doi.org/10.1145/2030112.2030177>

- [38] T. Zhang, N. Becker, Y. Wang, Y. Zhou, and Y. Shi. 2017. BitID: Easily Add Battery-Free Wireless Sensors to Everyday Objects. In *2017 IEEE International*

*Conference on Smart Computing (SMARTCOMP)*. 1–8.  
DOI:<http://dx.doi.org/10.1109/SMArtComp.2017.7946990>