

Exploiting Thermal Reflection for Interactive Systems

Alireza Sahami Shirazi, Yomna Abdelrahman, Niels Henze,

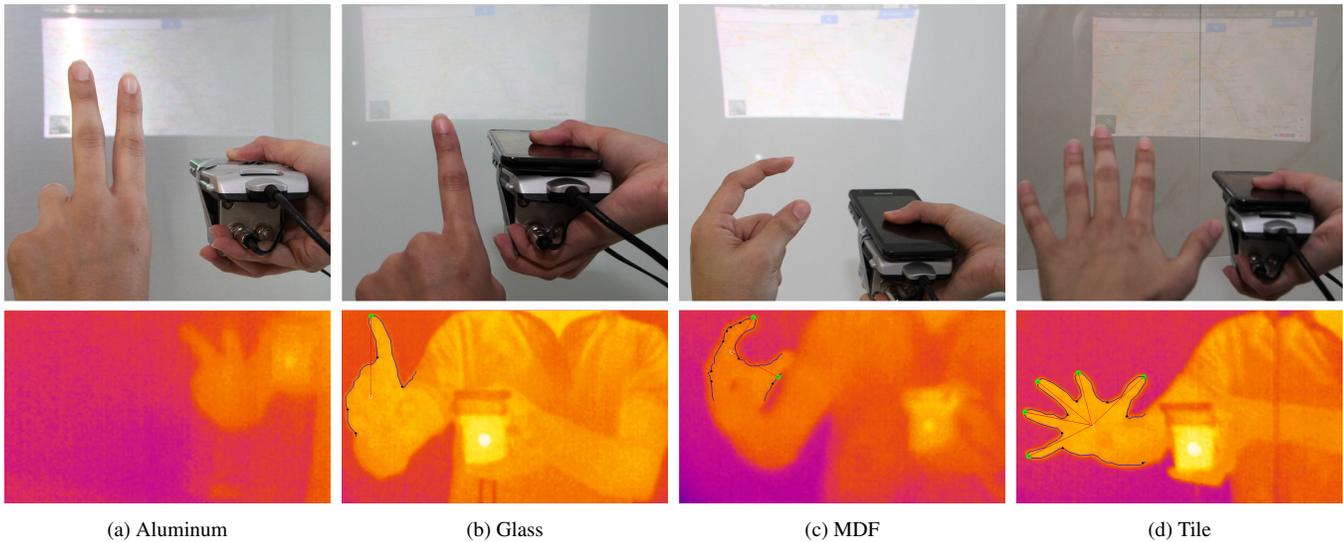
Stefan Schneegass, Mohammadreza Khalilbeigi[†], Albrecht Schmidt

VIS, University of Stuttgart
Stuttgart, Germany

{firstname.lastname}@vis.uni-stuttgart.de

[†]Technische Universität Darmstadt
Darmstadt, Germany

khalilbeigi@tk.informatik.tu-darmstadt.de



(a) Aluminum

(b) Glass

(c) MDF

(d) Tile

Figure 1: We explored the thermal reflectivity of different surfaces for interaction with projected surfaces. The thermal reflectivity allows to sense users that perform in-air gestures inside and outside the thermal camera's direct field-of-view. Four of the eight surfaces we analyzed and which can be found in normal office environments are presented above.

ABSTRACT

Thermal cameras have recently drawn the attention of HCI researchers as a new sensory system enabling novel interactive systems. They are robust to illumination changes and make it easy to separate human bodies from the image background. Far-infrared radiation, however, has another characteristic that distinguishes thermal cameras from their RGB or depth counterparts, namely *thermal reflection*. Common surfaces reflect thermal radiation differently than visual light and can be perfect thermal mirrors. In this paper, we show that through thermal reflection, thermal cameras can sense the space beyond their direct field-of-view. A thermal camera can sense areas besides and even behind its field-of-view through thermal reflection. We investigate

how thermal reflection can increase the interaction space of projected surfaces using camera-projection systems. We moreover discuss the reflection characteristics of common surfaces in our vicinity in both the visual and thermal radiation bands. Using a proof-of-concept prototype, we demonstrate the increased interaction space for hand-held camera-projection system. Furthermore, we depict a number of promising application examples that can benefit from the thermal reflection characteristics of surfaces.

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

Author Keywords

heat; thermal imaging; reflection; camera-projector system; roughness

INTRODUCTION

Thermal cameras operate in the far-infrared (F-IR) band and provide a heat map of a scene without an additional illumination source. As thermal cameras are light invariant and robust to illumination and color changes, they have been used for various applications, such as surveillance

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 the author(s) 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'14, April 26 – May 01 2014, Toronto, ON, Canada

Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-2473-1/14/04...\$15.00.

<http://dx.doi.org/10.1145/2556288.2557208>

systems [48] and medical applications [47]. In the field of human-computer interaction, thermal imaging has been recently used to determine users' body pose, emotions [30], and for face recognition [22]. Due to their unique properties, thermal cameras have shown remarkable opportunities for interactive systems [14, 15, 23]. In these systems, a thermal camera is not only used to detect the users' position and body parts but also to trace heat trails left behind by body parts touching a surface.

Another unique property of thermal cameras results from objects' varying reflectivity in different electromagnetic wavebands. Brass, for example, diffuses visible radiation whereas it is practically a mirror in the far infrared spectrum. As shown in Figure 2, a surface that seems non-reflective for a human observer and an RGB camera, is a mirror for thermal radiation. Previous work considered such thermal reflection from surfaces as noise [23] and developed approaches to filter them out [37]. In contrast, we propose to exploit *thermal reflection* as a unique characteristic of thermal radiation and use it for interactive systems. If a thermal camera faces towards a surface that reflects thermal radiation in a specular manner the surface gains properties of a thermal mirror for the camera. In addition to its direct field-of-view, a thermal camera can thereby sense the space besides and even behind the camera through the thermal mirror (see Figure 2). For user-surface interaction, this extends the interaction space beyond the space between the camera and a surface. It enables sensing a human body and hand gestures in areas that are not directly visible from the camera's position. While recent prior work has used time-of-flight imaging to also detect objects that are not in the camera's direct field-of-view [36], thermal imaging can provide a more detailed picture, is more robust and requires little computational power.

In this paper, we investigate how thermal reflectivity, which has been considered as noise in previous work [23, 37], can be exploited for creating novel interactive systems. We overview the laws of physics to explore which surfaces diffuse visual light but provide specular thermal reflectivity. Such surfaces can be used with normal visual displays such as video projectors but at the same time enable a thermal camera not only to observe its direct field-of-view but also the space besides and behind the camera. We build a handheld camera-projection prototype that uses thermal imaging. The system supports on-surface interaction and mid-air gestures in the camera's direct field-of-view. In addition, it can also trace body movements outside the camera's direct field-of-view. Using the prototype, we assess the thermal reflectivity of common surfaces which can be found in many offices and living rooms. We envision further promising use cases that can benefit from this phenomenon.

This paper makes the following contributions:

- Investigation of thermal reflection and how it can be used to build novel interactive systems extending the interaction space beyond the field of view of the camera.
- Implementation of a thermal camera-projection prototype as the proof-of-concept allowing users standing in front

of a surface to perform hand gestures outside the thermal camera's direct field of view.

- Analysis and assessment of the reflection characteristics of different surfaces that can be found in our vicinity.
- Discussion of four promising application examples that can benefit from thermal reflection.

The paper is structured as follows: after reviewing the related work, we discuss the physical properties of thermal and visual radiation. We then present an interactive prototype that exploits thermal reflection to extend the interaction space beyond the camera's direct field of view. Afterward, we use the prototype to assess the reflection characteristics of common materials in the visual and the thermal spectrum. Finally, we discuss a number of potential areas of application.

RELATED WORK

Enabling natural gestural and touch-based interaction techniques is one of the main goals of human-computer interaction research. In the following, we discuss prior work in three different areas aimed to provide natural interaction using (1) thermal imaging, and techniques used for recognizing human gestures to interact with (2) stationary as well as (3) mobile projected screens.

Thermal Imaging

Thermal cameras have been used for various purposes, including health-related applications, user recognition, and to build interactive systems. Wong et al. used thermal imaging and proposed algorithms to detect persons' head positions and parameter ratios such as the height and the width of objects [48]. Such algorithm were not only used for anti-theft surveillance systems, but also applied in nursing homes for monitoring Alzheimer patients and preventing them from leaving the nursing home without being attended to. The same approach is used for faint detection [47]. Pham et al. proposed a system for posture analysis in a dense crowd of people to detect irregular behavior using thermal imaging [29].

Face recognition remains one of the most challenging tasks in computer vision. By utilizing the light independent properties of thermal imaging, a vast enhancement in the performance of face recognition can be achieved [22]. Thermal imaging has not only been used to recognize faces but also to classify facial expressions [19] and it has also been shown that by measuring the temperature of a user's face it is possible to determine the cognitive load and the emotional state [30].

Thermal imaging has recently also been used as a sensor for enabling the interaction with arbitrary surfaces [14, 15, 23]. This is achieved by integrating thermal imaging and existing computer vision techniques to improve user surface interaction by utilizing advantages of thermal imaging to overcome common RGB and depth cameras' drawbacks. Daisuke and Koskue, for example, used short-lived heat traces that result from the heat transfer from one object to another as means for interaction and created an interactive

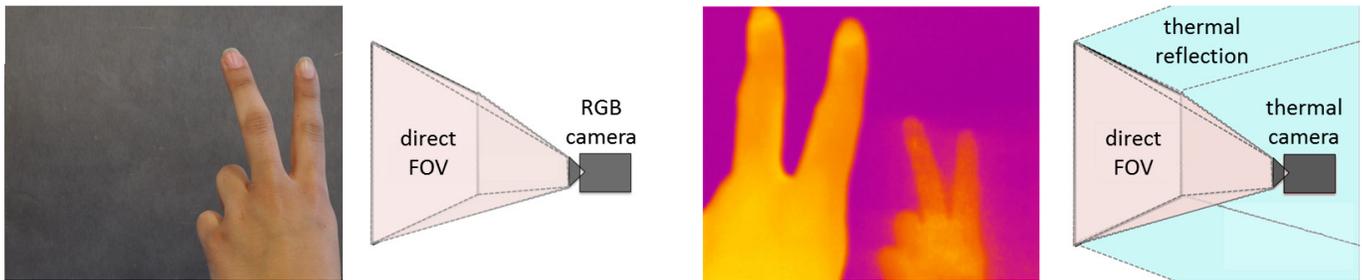


Figure 2: A standard RGB camera observes objects in the direct field of view as shown on the left. Through thermal reflection, a thermal camera facing a surface such as brass shown on the images also observes objects besides and behind the camera.

tabletop surface [14]. Larson et al. used heat traces caused by fingers touching a surface and detected pressure for interaction on surfaces [23]. They reported that the thermal reflectivity of surfaces induces noise for their system.

Interaction with Projected Surfaces

Touch and mid-air gestures are common techniques to interact with projections. These are typically detected using either RGB, infrared (IR), or depth cameras. There exists a large body of work focusing on detecting and tracking hands and fingers to enable multi-touch and mid-air gestures using RGB cameras [7, 21, 18, 25]. Such systems typically use skin color detectors [18] or template matching [21] to segment the hand and then calculate contour and convexity defects [25] to identify fingers.

Infrared imaging is a popular technique to enable multi-touch and mid-air gesture when interacting with projection screens [13, 12, 17, 40, 41]. In such systems, the space behind the screen is typically illuminated with an infrared source and all except the infrared light is blocked from the camera using an infrared-pass filter. This technique has been widely used for tabletop interaction by combining a rear-mounted IR camera and a projection unit. Using the depth map provided by depth cameras is another approach for detecting touch and hand gestures on projected screens [27, 42, 43, 44]. These systems generally utilize either a 2D view above the surface [43, 27] or a selective 2D projection of 3D sensed data [44] for processing users input on or above the surface using common 2D computer vision techniques.

Ubiquitous Mobile Interaction

Our research is also related to previous work on interaction with mobile projectors. According to the work by Rukzio et al. [32], interaction with mobile projected displays can be divided into four categories. A common approach is to separate input and output and use the touchscreen of a mobile phone [10] or a touch sensor for input [3]. Researchers have also investigated input for mobile projection by moving and gesturing with the projector itself [3, 5, 39] or by aiming with the projector at objects in the environment [33]. Although these solutions allow users to focus on the projection and perform intuitive gestures, tracking the projector's movements requires additional hardware equipment to be installed in the room or mounted on the projector unit.

Directly touching the projection screen with the fingers [46, 11] or using a stylus [4] is another approach to interact with mobile projections. However, such a setup requires users to be very close to the projection leading to a small projection area and large shadows on the projection. Another solution is to use mid-air finger pointing and hand gestures to interact with the projection [26, 8]. The SixthSense system [26], for example, offers a set of mid-air hand gestures to support interaction with the projection. The system uses a color-based approach to track fingers. ShadowPuppets [8] provides shadow gestures as input to a handheld projector system by casting hand shadows for co-located collaborative scenarios. Winkler et al. [45] have argued that it is more preferable to perform gestures besides or even behind the projector. This is not possible with current handheld projectors-camera systems as both face the same direction and the projection and sensing spaces overlap. Thus, users occlude the projection while performing mid-air gestures in front of the camera. A thermal camera, however, can detect direct interactions on a surface and mid-air gestures as well as users interaction out of the camera's direct field-of-view for instance behind the camera.

THERMAL REFLECTIVITY

Thermal radiation, as a result of energy transitions of molecules, atoms, and electrons of a substance, is continuously emitted by all matter whose temperature is above absolute zero. The spectrum and intensity of blackbody radiation depends on the object's temperature as expressed by the Planck's and Stefan-Boltzmann laws. The radiation emitted by objects at room temperature falls into the infrared region of the spectrum, which extends from 0.76 to 100 micron. The human body's net radiation is, for example, around 142 watt (W), with a skin temperature of 33°C, at an ambient temperature of 22°C, and a peak wavelength of 9.5 micrometer (μm).

When radiation strikes a surface it is partially absorbed, partially reflected, and the remaining part, if any, is transmitted. Based on the first law of thermodynamics the sum of absorbed, reflected, and transmitted radiation is equal to the incident radiation. For fully opaque surfaces the transmissivity is zero, thus, the sum of absorptivity and reflectivity is one. The absorptivity is independent of a surface's temperature. However, it depends on the temperature of the source at which the incident radiation is generated. The reflectivity of a surface depends not only

on the direction of the incident radiation but also on the direction of the reflection. Surfaces are assumed to reflect in two manners: specular and diffuse. In specular (mirror-like) reflection, the angle of reflection equals the angle of the radiation beam. For diffuse reflection the radiation is reflected equally in all directions regardless of the incident radiation's direction. The reflectance of a surface depends on its roughness and the wavelength of radiation strikes [2]. If the wavelength is smaller than the surface roughness, light is scattered diffusely. For wavelengths much larger than the roughness dimensions, the radiation is specularly reflected as from a mirror [37]. Beckmann & Spizzichino reports that reflectance is specular if the roughness (R_a) is smaller than one eighth ($1/8$) of the wavelength and otherwise diffuse [1]. The smaller the roughness, the higher the reflectivity: reflection from smooth and polished surfaces is mirror-like, whereas it is diffuse from rough surfaces [37]. Surfaces with roughness smaller than approximately 1.18 micrometer (μm) reflect a human's radiation (with a peak wavelength of $9.5 \mu\text{m}$) in specular manner.

A thermal camera produces thermograms of a surface based on the incident radiation from the surface. This radiation includes the energy the surface emits (based on its temperature) as well the reflection of objects' radiation from the surrounding. If the reflectivity of all objects is diffuse the camera only views the objects in its direct field of view. However, if a surface reflects radiation in a specular manner, it acts as a mirror for the thermal camera. Thus, the camera is additionally able to view objects which are out of its direct field of view but visible through the surface's reflection. With such surfaces it is possible to extend the camera's field of view and the space of interaction, respectively.

Objects reflect thermal radiation and visual light differently. Surfaces made of different metals or with a smooth paint can act as a mirror in the thermal spectrum and can still be used for visual projection. Other materials such as transparent glass and plastic are transparent for visual light but still a mirror for thermal radiation. In the following, we show how the reflection of thermal radiation can be exploited to build interactive systems that can sense body gesture in front, besides and even behind a thermal camera. We show that a wide range of materials exist that diffuse visual light and can thus be used for projecting visual content but still reflect thermal radiation. As the human body radiation is in the F-IR range, we are interested in surfaces that have high specular reflectivity for F-IR radiation but diffuse reflectivity in the visual spectrum.

PROTOTYPE

In the following, we describe the design of our prototype that is used as proof-of-concept to explore the concept of thermal reflectivity and investigate opportunities for interactive systems. The prototype runs a sample application that highlights the usage mid-air gestures and on surface interaction using a single portable thermal camera. It further allows to have an ad-hoc mobile setup and transforms any surface into an interactive canvas.

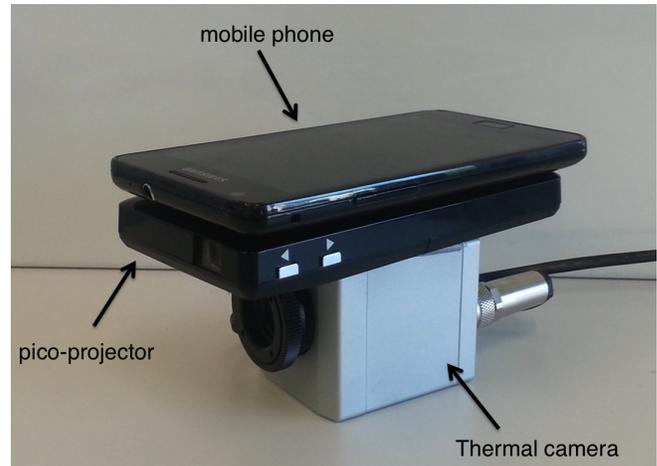


Figure 3: The prototype consisting of a thermal camera, a pico-projector, and a smartphone that has been used for the experiment.

Hardware Unit

As shown in Figure 3, our prototype consists of a smartphone augmented with a pico-projector (MicroVision) and a thermal camera (Optris PI160). The thermal camera is a contactless sensor that measures the temperature of objects in its field of view. The optical resolution is 160×120 pixels and it has a frame rate of 120Hz. The camera is able to measure temperatures between -20°C and 900°C . It operates with a thermal sensitivity of 0.08 Kelvin represented by the noise equivalent temperature difference (NETD)¹. The wavelengths captured by the camera are in the spectral range between $7.5\mu\text{m}$ and $13\mu\text{m}$. The lens we use in the following provides a $23^{\circ} \times 17^{\circ}$ field of view. The thermal camera is faced in the same direction as the pico-projector.

The camera uses USB as power source as well as to transfer data. It provides raw data in one of three formats: pre-processed energy values, temperature values, or YUV-color values. A dynamic link library (DLL) allows to access the camera through inter-process communication. In our implementation we dynamically link the DLL into our application. We utilize the 16-bit color values from which we can compute the corresponding temperature based on the specification of the thermal camera.

Algorithm

We implemented a Google Maps application which allows users to pan and zoom the map information either by directly touching the projection surface or performing mid-air hand gestures. In order to recognize touch input and hand gestures we used OpenCV library² for image processing and feature extraction. For each retrieved frame from the thermal camera pre-processing is required before extracting the features. The pre-processing includes the following steps: noise filtering, background subtraction, and thresholding.

¹NETD refers to the electronic noise that is interpreted as a temperature difference of an object

²OpenCV: <http://opencv.org/> (last access 15.01.2014)

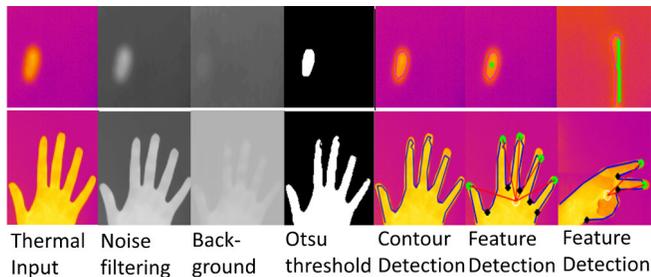


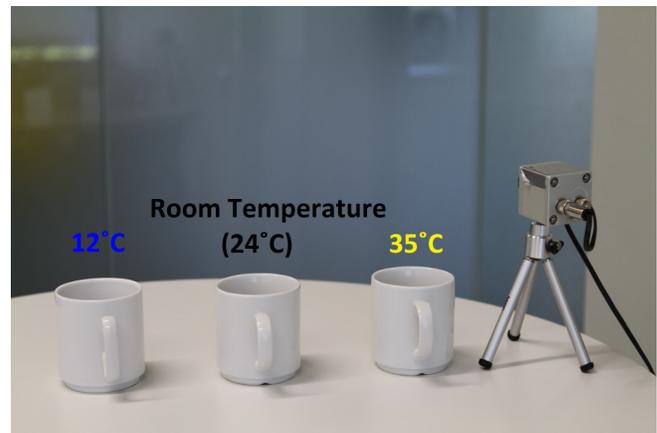
Figure 4: Feature extraction in different steps of the algorithm. The top row shows the features extracted from the heat trace on a surface, the bottom row presents features retrieved from an in-air gesture.

Noise filtering: As the first step we apply a filter to reduce the noise in the image provided by the camera. Similar to [23], we use a 5×5 pixels median filter. Then, we convert the image to gray scale and reapply the median filter for better noise reduction.

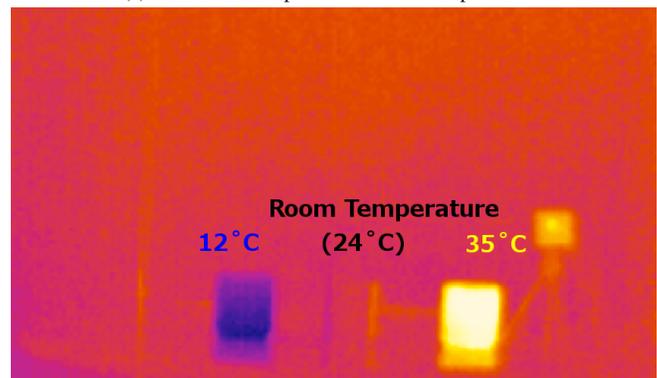
Background subtraction: To remove the background, a model is computed using an accumulated weighted model for each pixel. It is chosen to allow a dynamic adaptation of the background. The dynamic update is dependent on the learning rate parameter (α) that controls how fast the background model is updated which is essential to detect heat traces. The α value lies between zero and one and can be adjusted to either maintain heat traces until they start decaying or disappear completely. The higher α , the more sensitive the background model becomes to changes in the image sequence. We systematically tested the algorithm with different α values and $\alpha = 0.1$ showed the best result. Old heat trace lasts in the foreground for detection yet are merged to the background fast enough to enable interaction on the same spot.

Thresholding: We use Otsu's thresholding method [28] for identifying parts of the image that are relevant to detect body parts and heat traces. The thresholding algorithm separates parts of a human's body such as hands and fingers as well as heat traces from the rest of the foreground. This method assumes the existence of two classes of pixels and aims to set the threshold to have minimal class overlap. Since the interactive parts stand out in the image, it separates the image into two classes (interactive part and the remaining foreground). An additional morphological closing operation is applied to enhance the boundaries of the segmented foreground and shrink the background.

Heat trace detection: The heat transmitted from the user's hand to the surface leaves a trace that could be detected by searching for contours in the image and examining the area, shape, and temperature of the contour. The gesture mapping is implemented by matching the shape of the contour detected. In our system we consider two types of interaction on surfaces, that is, touch-points or swiping in any arbitrary direction. By fitting the contour to a circle/eclipse or a line using Hough Transform [9] the interaction type is determined. By tracking the center of the detected line-shape



(a) Three coffee cups with different temperatures.



(b) Thermal image of the coffee cups with different temperatures.

Figure 5: When an object has a temperatures that is very similar to the temperature of the background, there is no clear difference between the object and the background in the thermal image. Whereas, as higher the temperature difference is the easier the object segmentation becomes.

contour the direction of swipe gestures is recognized. The center of the contour is computed by getting the spatial moment of the extracted contour. By detecting this trace, we can differentiate between directly touching a surface and the hovering above it. Figure 4 shows step by step processing for detecting the touch-point interaction on the surface as well as mid-air gestures.

No calibration is required for detecting heat traces if the amount of pressure that is applied to a surface by the user is not important. Otherwise, the pressure can be detected through a calibration process and determining the number of frames the heat trace lasts (β). With two levels of pressure (high pressure and low pressure) the number of frames in which the trace persists is identified and stored. During the interaction the number of frames are compared to stored β values to deduce the pressure level of the interaction. This calibration is processed after noise filtering and before background subtraction due to the fact that we adjust the background subtraction learning rate (α) in such a way that we do not view old heat traces. This results in latency for

detecting the pressure level. The latency depends on the camera frame rate as well as the decay rate of the heat trace on the surface.

Hand and finger detection: Hand and finger detection is considered as one of the most challenging tasks using ordinary RGB cameras. Such a task is achieved either by using skin color filter [6, 31], relying on a color map of the human skin, or by instrumenting the hands using colored gloves [16, 38] to extract and track the hand. Since the hands have a different temperature than the room temperature, they can be robustly segmented from the image in any lighting conditions using thermal imaging. Our algorithm computes the convex hull and convexity defects of the hand contour. To determine the hand center, we moreover calculate the average depth points of the defects. The finger tips are then identified as points which have local maximum distance to the computed hand center. This approach allows detecting the hand and fingers at any orientation with all possible cases (fist, open hand, 1-5 fingers). Figure 4 presents the procedure of extracting a hand and its finger tips.

Hand gesture detection: Approaches to detecting hand gestures include model-based and view-based approaches. Model-based approaches utilize a 3D hand model for tracking, which makes it a complex and challenging, yet not very robust [34] task. However, view-based approaches use the hand and finger information extracted and matches these features to patterns for hand gesture recognition. We have utilized this approach by tracking fingertips and computing their relative distance. We have implemented for instance, the pinch and pan gestures for zooming in and out by tracking the relative distance between the finger tips over the image sequence captured. The approach can be used in real-time without any latency or delay.

Robust detection of hands and fingers through the processing pipeline explained above depends on the sharpness and contrast of the thermal images. In order to retrieve a sharp image the roughness of the surface should be as low as possible (e.g., smooth and polished surfaces). Surfaces with high roughness disturb the sharpness of the image rendered from the reflected heat waves, making the detection process very challenging. The contrast of objects in the thermal images also depends on the object's temperature and the thermal imaging sensitivity. If the object's temperature is similar to the surface (e.g., both have room temperature), the object may neither be visible through the reflection nor in the direct field of view (cf., Figure 5).

Changes in the temperature of surfaces or fingertips after a long interaction session decreases the robustness of the system. Such a limitation does not apply to mid-air gestures.

As described above, there is a latency for detecting the pressure level. This latency relays on the surface material and is directly proportional to the decay of the trace. In order to differentiate between hovering over a surface and mid-air gestures we defined a set of feature constraints. We use, for example, the heat traces for interaction on the surface and at least two fingertips the in-air.

Surface	β (High Pressure)	β (Low Pressure)	R_a Roughness
Tile	200 (1.7 sec)	60 (0.5 sec)	0.04
Glass	220 (1.8 sec)	100 (0.8 sec)	0.004
Acrylic (transparent)	170 (1.4 sec)	40 (0.3 sec)	0.005
Acrylic (white)	160 (1.3 sec)	44 (0.4 sec)	0.06
Acrylic (mate)	250 (2.1 sec)	48 (0.4 sec)	3.49
MDF	120 (1.0 sec)	50 (0.4 sec)	0.11
Polished Wood	260 (2.2 sec)	70 (0.6 sec)	1.48
Polished Wood (glossy painted)	400 (3.3 sec)	120 (1.0 sec)	0.11
Aluminum Plate	0 (0.0 sec)	0 (0.0 sec)	0.33

Table 1: The surfaces used for the experiment. The β represents the number of frames and the respective time in seconds the heat trace lasts.

EXPERIMENT

We conducted an experiment to explore the thermal radiation characteristics of various surfaces found in a normal office or home environment. We focused on surfaces that reflect thermal radiation and can serve as a projection surface. Such surfaces can be turned into interactive surfaces with a large interaction space even outside the camera's direct field of view.

Based on our pre-observation of several office and home environments, we choose frequently used surface materials for our experiment. We considered nine surfaces for the experiment (cf., Table 1). We tested shiny tiles, transparent glass, transparent acrylic, white acrylic, medium-density fiberboards (MDF), polished wood, and aluminum plates. We used the prototype and experimentally tested each surface. The α value was set to 0.1 during the whole experiment.

Observations

In general, shiny surfaces showed enough reflectivity to be used with our prototype. We examined the properties of the thermal reflectivity and whether it is sufficient to detect mid-air gestures outside of the camera's direct field of view for each of the nine surfaces. For all surfaces except the polished wood and the aluminum plate, the reflectivity is high and the camera provides images with enough sharpness and contrast to detect mid-air gestures. Figure 1 shows the images provided by the thermal camera next to a visual photo of the recorded scene for four of the nine surfaces that we tested in the experiment. While glass has the highest reflectivity, the aluminum plate is more diffuse and provides blurry reflection. The polished wood is completely diffuse and provides unclear image through reflection. Therefore, we decided to paint the normal wood board and repeat the experiment. We used a glossy white color for painting the board. The color added a layer to the wood that reflects thermal wavelengths. However, the reflectivity is not as good as the MDF, for instance.

Surface Roughness

As described above, the reflectivity of a surface depends on its roughness. We measured the roughness of the surfaces used in the experiment according to ISO 4287. The

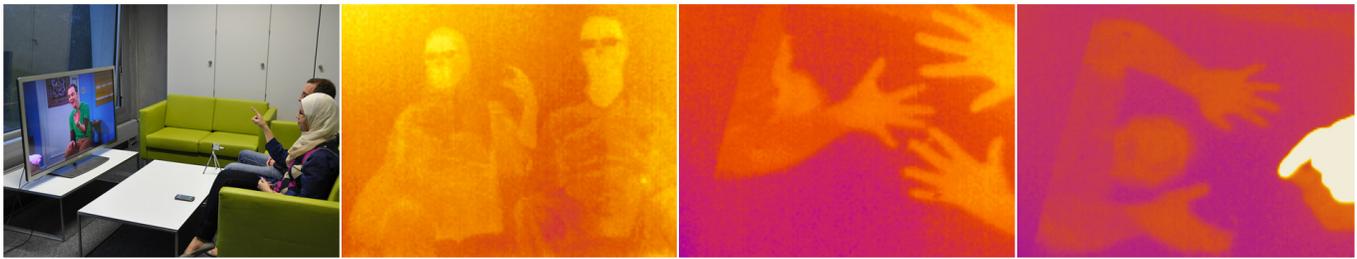


Figure 6: A thermal camera facing a TV can monitor users sitting behind the camera through thermal reflection (left images). Facing a living room table from the top an ad-hoc tabletop setup can be created. The right images show two persons interacting while the thermal camera solely observed through thermal reflection and by observing one user through reflection and other one in the direct field of view.

topography of the surfaces were measured with a spinning micro lens confocal microscope [35, 24] (50x magnification, 0.8 numerical aperture, optical resolution below 50 nm). The roughness Ra was calculated with the NIST reference software³ according to ISO 4287 from a representative slice along the measured surface. Table 1 shows the roughness of the surfaces. The glass has the smallest roughness, whereas the polished wood has the highest one. The smaller the roughness value, the more mirror-like is the surface. Painting the wood decreases its roughness by more than 90%.

Heat trace decay

The speed of heat trace decay depends on the surface's material. Therefore we measured the number of frames the heat trace last on each of the nine surfaces (β) after a finger is lifted from the surface. We measured the β values for touch with high and low pressure. Table 1 shows the values for all surfaces. While the heat trace stays very long on glass, it decays very fast on acrylic. For the aluminum plate no heat trace is left behind, hence, the β value is zero.

LIMITATIONS

It should be mentioned that the described approach has its own limitations. The object's temperature plays an important role for object detection. When objects have a similar temperature as the background there is no clear difference between objects and background in the thermal image (see Figure 5). This makes it very hard to detect such objects. Whereas, a higher temperature difference between object and background makes objects segmentation much easier. The sensitivity of the thermal camera in measuring of temperature should be also taken into account.

The spatial configuration is another important factor. The camera has a limited viewing angle. Therefore, the interaction space through thermal reflectivity is also limited. A movement of the camera and, in general, a change in its spatial configuration with the surface results in rendering a new interaction volume based on thermal reflectivity.

Furthermore, the approach cannot be used with every surface. Surfaces should be smooth and polished to have a mirror-like

reflectivity in the thermal frequency range. It is particularly critical when it is used in combination with projection. Best surfaces for thermal imaging, such as glass, cannot be used for projection and vice versa. The best combinations are surfaces that are diffuse for projection and reflective for thermal imaging.

POTENTIAL USE CASES

We envision that thermal reflection can be used for a variety of applications. In the following we describe four potential applications in which thermal reflectivity can enlarge the interaction space and consequently opens up novel possibilities for interaction in front of surfaces that reflect thermal radiation.

Interactive (Mobile) Projection

Given the recent technological advances in manufacturing pico-projectors, we already witnessed mobile phones equipped with an embedded pico-projector. The projector extends the space for visualizing information beyond the mobile phone's display. Using current commercial systems, users interact with such systems through the phone's display. Equipping the pico-projector with a camera allows users to interact in front of the camera using mid-air gestures. Using RGB cameras in such a setup forces the user to perform the gestures in front of the camera which results in occlusion. Winkler et al. [45] already indicated that enabling users to perform gestures behind the phone might result in higher performance.

We envision that future mobile phones can be equipped with a combination of a thermal camera and a pico-projector. If the projector projects on a surface that is a mirror for the thermal camera, the camera can also observe gestures performed besides or even behind the camera and the phone. Hence, the user can perform gestures without occluding the projection. This advantage of an increased interaction radius can not only be used for mobile projectors but also for stationary projectors. Equipping a stationary projector with a thermal camera and using a surface that reflects thermal radiation as projection screen enables to create a setup with large interactive space. The advantage is that projector and camera can be co-located which leads to a very simple setup. Users can interact directly on the surface in the camera's direct field of view as well as while being behind or besides

³National Institute of Standards and Technology: <http://physics.nist.gov/VSC/jsp/> (last access 15.01.2014)

the projector. However, finding a proper surface is essential to use this setup.

Interactive TVs

With the increasing success of inexpensive commercial sensing technologies such as depth cameras (e.g., Microsoft Kinect⁴), they have been used widely for sensing natural gestures at home. The typical setup consists of an RGB camera or a depth camera that is situated above or below the TV. Thereby the camera is looking at the living rooms. It can observe gestures performed in their direct field of view.

We envision that thermal cameras to be integrated into TV's remote controllers (RC) or another mobile device and leverage thermal reflectivity once the device aims at the TV (see Figure 6). This can be used in an ad-hoc scenario while holding the RC in the hand and doing bimanual interactions. For example, using the RC for activating the Electronic Program Guide (EPG) and using hand gestures for navigating through the EPG. Further, the RC can be on the living room table looking towards the TV so users can perform mid-air gestures while sitting on the sofa behind the camera. This setup may provide a larger interaction space compared to setups with depth or RGB cameras. However, it requires a specific spatial arrangement of the camera and the TV.

Interactive tabletops

A typical vision-based setup to create an interactive tabletop is orienting a projector and a camera at a table that can serve as a projection screen. Camera and projector are either located above or below the surface. Each setup has its own shortcomings [41]. In a top-projection setup the user's body parts can occlude the projected image. In a rear-projection setup it becomes challenging to capture objects that are on top of the surface or hover above. Another proposed approach is to place projector and camera at the side of the interactive surface [41]. In this setup, a difficult overhead installation of the projector is avoided. Further, the camera and projector are oblique and the occlusion problem is minimized.

Using a thermal camera as the sensing camera together with a reflective surface can enable setups that further reduce occlusion or increase the interaction radius (see Figure 6). When the thermal camera is oblique and the surface reflects thermal radiation, its field of view is extended. Interaction can mainly take place outside of the projection space to avoid occlusion. To have a surface that reflects human body thermal radiation, its roughness should be smaller than $1.18\mu\text{m}$. Typical tables provide this reflectivity and further surfaces can be easily created by, for example, painting a smooth wood surface glossy.

Wearable Computing

The miniaturization of computing devices has led to new types of devices that users can wear and use in various contexts (e.g., Google Glass). Interaction with such wearable devices is of great importance. Various sensors such

⁴Microsoft Kinect - <http://www.microsoft.com/en-us/kinectforwindows/> (last access 15.01.2014)

as infrared cameras [20] have been proposed to facilitate full-body interaction. However, these sensors might be sensitive to illuminations. Further, with such camera, the user must hold the hands in the camera's direct field of view to be observable by the camera.

Considering Google Glass, for example, its RGB camera provides the opportunity to be used for observing mid-air gesture. Substituting the RGB camera with a thermal camera in front of a reflective surface, the interaction space is enlarged. Users can perform interaction behind/beside themselves out of their sight while their gestures are still visible for the camera. Furthermore, such a camera works in any light condition which is very useful for wearable devices. However, not all surfaces can be used for interaction.

CONCLUSION

RGB and depth cameras are widely used to build interactive systems by research and industry. Users can interact with such systems while they are in the camera's direct field of view. In such systems, occlusion may occur if the interaction space overlaps with the projection volume. As such cameras are sensitive to illumination or color changes, thermal cameras have been proposed as a robust alternative. A unique property of thermal cameras results from objects' varying reflection characteristics in different electromagnetic wavebands. A surface that diffuses visual light can still reflect thermal radiation. In previous work, thermal reflection has only been considered as noise. In this paper we exploit thermal reflection to build novel interactive systems.

Surfaces that either diffuse visual light or are transparent for visual light can be used as projection screen but still reflect thermal radiation. Using such common surfaces we can extend the area observable by a thermal camera beyond its direct field of view. Using such a setup we built a camera-projector system that can monitor users in front of the camera but also besides and even behind the camera. After reviewing the laws of physics we present an interactive prototype system used to explore the characteristics of different materials. We show that a wide range of surfaces is suitable for projection and can also provide an enlarged interaction space through thermal reflection. Through four application examples that could benefit from thermal reflection we demonstrate that the concept is widely applicable.

Current thermal cameras, the algorithms that have been developed for thermal imaging, and the resulting systems are all optimized to reduce thermal reflection. We assume that by developing cameras, algorithms, and systems that are optimized to exploit thermal reflection we can further extend the range of surfaces that can be used for interaction through reflection. We further believe that using thermal cameras beyond the niche applications where they are currently used can bring down their cost dramatically.

A comprehensive assessment of further gestures and their recognition rate is an interesting direction which can be subject of future work. In particular, to develop algorithms that are optimized to exploit thermal reflection. Various use

cases can be implemented that benefit from using thermal reflectivity for interaction.

ACKNOWLEDGMENTS

We are very grateful to Prof. Dr. James Fogarty (University of Washington) for his valuable feedback on this paper. We also thank Mr. Ahmad Faridian and Mr. Wolfram Lyda (ITO, University of Stuttgart) for helping us with the measurement of surfaces roughness. This work is funded by the German Research Foundation within the SimTech Cluster of Excellence (EXC 310/1).

REFERENCES

1. Beckmann, P., and Spizzichino, A. *The scattering of electromagnetic waves from rough surfaces*. Artech House Radar Library, 1987.
2. Bennett, H., and Porteus, J. Relation between surface roughness and specular reflectance at normal incidence. *Journal of the Optical Society of America* 51 (1961).
3. Blasko, G., Feiner, S., and Coriand, F. Exploring interaction with a simulated wrist-worn projection display. In *Proceedings of the International Symposium on Wearable Computers* (2005), 2–9.
4. Cao, X., and Balakrishnan, R. Interacting with dynamically defined information spaces using a handheld projector and a pen. In *Proceedings of the Symposium on User interface software and technology* (2006), 225–234.
5. Cao, X., Forlines, C., and Balakrishnan, R. Multi-user interaction using handheld projectors. In *Proceedings of the Symposium on User interface software and technology* (2007), 43–52.
6. Ciaramello, F. M., and Hemami, S. S. Real-time face and hand detection for videoconferencing on a mobile device. In *Proceedings of the Workshop on Video Processing and Quality Metrics for Consumer Electronics* (2009).
7. Cohen, C. J., Beach, G., and Foulk, G. A basic hand gesture control system for pc applications. In *Proceedings of the Workshop on Applied Imagery Pattern Recognition* (2001), 74–79.
8. Cowan, L. G., and Li, K. A. Shadowpuppets: supporting collocated interaction with mobile projector phones using hand shadows. In *Proceedings of the Conference on Human Factors in Computing Systems* (2011).
9. Duda, R. O., and Hart, P. E. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM* 15, 1 (1972), 11–15.
10. Greaves, A., and Rukzio, E. Evaluation of picture browsing using a projector phone. In *Proceedings of the Conference on Human computer interaction with mobile devices and services* (2008), 351–354.
11. Harrison, C., Benko, H., and Wilson, A. D. Omnitouch: wearable multitouch interaction everywhere. In *Proceedings of the Symposium on User interface software and technology* (2011), 441–450.
12. Henze, N., Löcken, A., Boll, S., Hesselmann, T., and Pilot, M. Free-hand gestures for music playback: deriving gestures with a user-centred process. In *Proceedings of the Conference on Mobile and Ubiquitous Multimedia* (2010), 16:1–16:10.
13. Hilliges, O., Izadi, S., Wilson, A. D., Hodges, S., Garcia-Mendoza, A., and Butz, A. Interactions in the air: adding further depth to interactive tabletops. In *Proceedings of the Symposium on User interface software and technology* (2009), 139–148.
14. Iwai, D., and Sato, K. Heat sensation in image creation with thermal vision. In *Proceedings of the Conference on Advances in computer entertainment technology* (2005), 213–216.
15. Iwai, D., and Sato, K. Limpid desk: see-through access to disorderly desktop in projection-based mixed reality. In *Proceedings of the Symposium on Virtual reality software and technology* (2006), 112–115.
16. Iwai, Y., Watanabe, K., Yagi, Y., and Yachida, M. Gesture recognition by using colored gloves. In *Proceedings of the Conference on Systems, Man, and Cybernetics* (1996), 76–81.
17. Izadi, S., Hodges, S., Taylor, S., Rosenfeld, D., Villar, N., Butler, A., and Westhues, J. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proceedings of the Symposium on User interface software and technology* (2008), 269–278.
18. Kane, S. K., Avrahami, D., Wobbrock, J. O., Harrison, B., Rea, A. D., Philipose, M., and LaMarca, A. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology* (2009), 129–138.
19. Khan, M. M., Ingleby, M., and Ward, R. D. Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variations. *ACM Transactions on Autonomous Adaptive Systems* 1, 1 (2006), 91–113.
20. Kim, D., Hilliges, O., Izadi, S., Butler, A. D., Chen, J., Oikonomidis, I., and Olivier, P. Digits: freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In *Proceedings of the Symposium on User interface software and technology* (2012), 167–176.
21. Koike, H., Sato, Y., and Kobayashi, Y. Integrating paper and digital information on enhanceddesk: a method for realtime finger tracking on an augmented desk system. *ACM Transactions on Computer-Human Interaction* 8, 4 (2001), 307–322.
22. Kong, S. G., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B. R., Koschan, A., Yi, M., and Abidi, M. A. Multiscale fusion of visible and thermal ir images for illumination-invariant face recognition. *International Journal of Computer Vision* 71, 2 (2007), 215–233.

23. Larson, E., Cohn, G., Gupta, S., Ren, X., Harrison, B., Fox, D., and Patel, S. Heatwave: thermal imaging for surface user interaction. In *Proceedings of the Conference on Human Factors in Computing Systems* (2011), 2565–2574.
24. Lyda, W., Zimmermann, J., Burla, A., Regin, J., Osten, W., Sawodny, O., and Westkämper, E. Sensor and actuator conditioning for multiscale measurement systems on example of confocal microscopy. In *SPIE Europe Optical Metrology*, vol. 7389 (2009).
25. Manresa, C., Varona, J., Mas, R., and Perales, F. Hand tracking and gesture recognition for human-computer interaction. *Electronic letters on computer vision and image analysis* 5, 3 (2005), 96–104.
26. Mistry, P., Maes, P., and Chang, L. Wuw - wear ur world: a wearable gestural interface. In *Proceedings of the Conference on Human Factors in Computing Systems (extended abstracts)* (2009), 4111–4116.
27. Murugappan, S., Vinayak, Elmqvist, N., and Ramani, K. Extended multitouch: recovering touch posture and differentiating users using a depth camera. In *Proceedings of the Symposium on User interface software and technology* (2012), 487–496.
28. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* 11, 285–296 (1975), 23–27.
29. Pham, Q.-C., Gond, L., Begard, J., Allezard, N., and Sayd, P. Real-time posture analysis in a crowd using thermal imaging. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (2007), 1–8.
30. Puri, C., Olson, L., Pavlidis, I., Levine, J., and Starren, J. Stresscam: non-contact measurement of users' emotional states through thermal imaging. In *Proceedings of the Conference on Human Factors in Computing Systems (extended abstracts)* (2005), 1725–1728.
31. Raheja, J. L., Das, K., and Chaudhary, A. Fingertip detection: A fast method with natural hand. *arXiv preprint arXiv:1212.0134* (2012).
32. Rukzio, E., Holleis, P., and Gellersen, H. Personal projectors for pervasive computing. *IEEE Pervasive Computing* 11, 2 (2012), 30–37.
33. Schmidt, D., Molyneaux, D., and Cao, X. Picontrol: using a handheld projector for direct control of physical devices through visible light. In *Proceedings of the Symposium on User interface software and technology* (2012), 379–388.
34. Stenger, B., Thayananthan, A., Torr, P. H., and Cipolla, R. Model-based hand tracking using a hierarchical bayesian filter. *Transactions on Pattern Analysis and Machine Intelligence* 28, 9 (2006), 1372–1384.
35. Tiziani, H., Haist, T., and Reuter, S. Optical inspection and characterization of microoptics using confocal microscopy. *Optics and lasers in engineering* 36, 5 (2001), 403–415.
36. Velten, A., Willwacher, T., Gupta, O., Veeraraghavan, A., Bawendi, M. G., and Raskar, R. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications* 3 (2012).
37. Vollmer, M., Henke, S., Karstädt, D., Möllmann, K., and Pinno, F. Identification and suppression of thermal reflections in infrared thermal imaging. In *Proceedings of Inframation*, vol. 5 (2004), 287–298.
38. Wang, R. Y., and Popović, J. Real-time hand-tracking with a color glove. In *ACM Transactions on Graphics*, vol. 28 (2009), 63:1–63:8.
39. Willis, K. D., Poupyrev, I., and Shiratori, T. Motionbeam: a metaphor for character interaction with handheld projectors. In *Proceedings of the Conference on Human Factors in Computing Systems* (2011), 1031–1040.
40. Wilson, A. D. Touchlight: an imaging touch screen and display for gesture-based interaction. In *Proceedings of the Conference on Multimodal interfaces* (2004), 69–76.
41. Wilson, A. D. Playanywhere: a compact interactive tabletop projection-vision system. In *Proceedings of the Symposium on User interface software and technology* (2005), 83–92.
42. Wilson, A. D. Depth sensing video cameras for 3d tangible tabletop interaction. In *Proceedings of Workshop on Horizontal Interactive Human-Computer Systems* (2007).
43. Wilson, A. D. Using a depth camera as a touch sensor. In *Proceedings of the Conference on Interactive Tabletops and Surfaces* (2010), 69–72.
44. Wilson, A. D., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the Symposium on User interface software and technology* (2010), 273–282.
45. Winkler, C., Pfeuffer, K., and Rukzio, E. Investigating mid-air pointing interaction for projector phones. In *Proceedings of the Conference on Interactive tabletops and surfaces* (2012), 85–94.
46. Winkler, C., Reinartz, C., Nowacka, D., and Rukzio, E. Interactive phone call: synchronous remote collaboration and projected interactive surfaces. In *Proceedings of the Conference on Interactive Tabletops and Surfaces* (2011), 61–70.
47. Wong, W. K., Lim, H. L., Loo, C. K., and Lim, W. S. Home alone faint detection surveillance system using thermal camera. In *Proceedings of the Conference on Computer Research and Development* (2010), 747–751.
48. Wong, W. K., Tan, P. N., Loo, C. K., and Lim, W. S. An effective surveillance system using thermal camera. In *Proceedings of the Conference on Signal Acquisition and Processing* (2009), 13–17.