University of Northern Iowa

UNI ScholarWorks

Dissertations and Theses @ UNI

Student Work

2021

Activity recognition based on thermopile imaging array sensors

Qingshan Liang
University of Northern Iowa

ACTIVITY RECOGNITION BASED ON THERMOPILE

IMAGING ARRAY SENSORS

An Abstract of a Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Industrial Technology

Approved:

Dr. Hong Nie, Chair

Dr. Jin Zhu, Co-Chair

Dr. Jennifer Waldron
Dean of the Graduate College

Qingshan Liang

University of Northern Iowa

May 2021

ABSTRACT

With aging population, the importance of caring for elderly people is getting more and more attention. In this paper, a low resolution thermopile array sensor is used to develop an activity recognition system for elderly people. The sensor is composed of a 32x32 thermopile array with the corresponding 33° × 33° field of view. The outputs of the sensor are sequential images in which each pixel contains a temperature value. According to the thermopile images, the activity recognition system first determines whether the target is within the tracking area; if the target is within the tracking area, the location of the target will be detected and three kinds of activities will be identified

Keywords- Activity Recognition, Raspberry Pi, Thermopile, Imaging Processing.

ACTIVITY RECOGNITION BASED ON THERMOPILE

IMAGING ARRAY SENSORS

A Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Industrial Technology

Approved:

Dr. Hong Nie, Chair

Dr. Jin Zhu, Co-Chair

Dr. Shahram VarzaVand, Committee Member

Dr. Mark Ecker, Committee Member

Dr. Bingqing Liang, Committee Member

Qingshan Liang

University of Northern Iowa

May 2021

ACKNOWLEDGEMENTS

Throughout the writing of this dissertation I have received a great deal of support and assistance.

I would first like to thank my supervisor, Dr. Hong Nie, whose expertise was invaluable in formulating the research questions and methodology. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to acknowledge my research partner Dr. Lei Yu from Nan Chang Hangkong University for your wonderful collaboration. I want to thank you for your patient support and for all of the contributions to my research.

I would also like to thank my tutors, Dr. Jin Zhu and Dr. Xuping Zhai, for their valuable guidance throughout my studies. You provided me with the tools, ideas, and codes that I needed to choose the right direction and successfully complete my dissertation.

In addition, I would like to thank my parents for their wise counsel and sympathetic ear. You are always there for me. Finally, I could not have completed this dissertation without the support of my friend Xuejiao Wang who provided happy distractions to rest my mind outside of my research.

TABLE OF CONTENTS

LIST OF TABLES

PAGE

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER I.

INTRODUCTION

A. Background

Aging society becomes a major trend in most of countries around the world (Bloom et al., 2015). According to the researches on aging society, although the increasing rate of population in the world is slowing down, the population aged 65 and over is increasing rapidly (United Nation). In addition, a large number of elderly people are living alone. To provide better care services to those elderly people, activity recognition and fall detection become more and more important for caregivers.

The idea of using sensors for activity monitoring and recognition appeared since the late 90s. This technology was initially pioneered and experimented by the work of the Neural Network house in the context of home automation and a location-based application (Mozer, 1998). The goal is to adapt systems to users' whereabouts (Leonhardt & Magee, 1998). The approach was useful and effective in the area of ubiquitous and mobile computing. Some improvement research has been done to make sensors' use more diversification on considerable work on context-awareness (Schmidt et al., 1999), smart appliances (Schmidt & Laerhoven, 2001) and activity recognition (Laerhoven et al., 2001). Most research at that time made use of wearable sensors, either dedicated sensors attached to human bodies or portable devices like mobile phones, with application to ubiquitous computing scenarios such as providing context-aware mobile devices. Most of mainly physical activities like motion, walking and running are monitored at the beginning of research. These early works provide a solid foundation for wearable computing and the method and outcomes still inspire and influence today's research.

Currently, a large number of sensors, including contact sensors, RFID, accelerometers, audio and motion detectors, are available for activity recognition. Those sensors are different in types, purposes, and output signals underpinning theoretical principles and technical infrastructure. However, these sensors can be classified into two main categories in terms of the way they are developed in activity recognition applications. They are wearable sensors and non-wearable sensors.

Generally wearable sensors are worn directly or indirectly on human body. Sensors generate signals when human body performs activities. As a consequence, the activity recognition systems can monitor features that are descriptive of the human body's physiological state or movement. This type of sensors is commonly embedded into clothes, eyeglasses, belts, shoes, watches, mobile devices or positioned directly on human body. The information such as body position, movement speed, pulse, skin temperature, and so on can be collected by sensors. Researchers have found several types of sensor information which are effective for determining different types of activities. Activity recognition is implemented by using wearable sensing devices like wearing on chest, arms, fingers and body fluids as shown in Fig. 1. Activity recognition is based on the assumption that specific body movements translate into characteristic sensor signal patterns, which can be sensed and classified using machine learning techniques. Wearable activity recognition relies on combinations of sensors, such as accelerometers, gyroscopes, and capacitive sensors. Patterns corresponding to activities are then detected within the streaming sensor data using either feature extraction on sliding windows followed by classification, template matching approaches, or hidden Markov modelling. (Ordonez & Roggen, 2016). The sensors in smart phones (Fahim et al., 2013), acceleration sensors (Khan et al., 2010), and capacitive sensors (Cheng et al., 2013) have been common used for activity recognitions. Researches show that the accuracy of the activity recognitions utilizing wearable sensing devices is high.

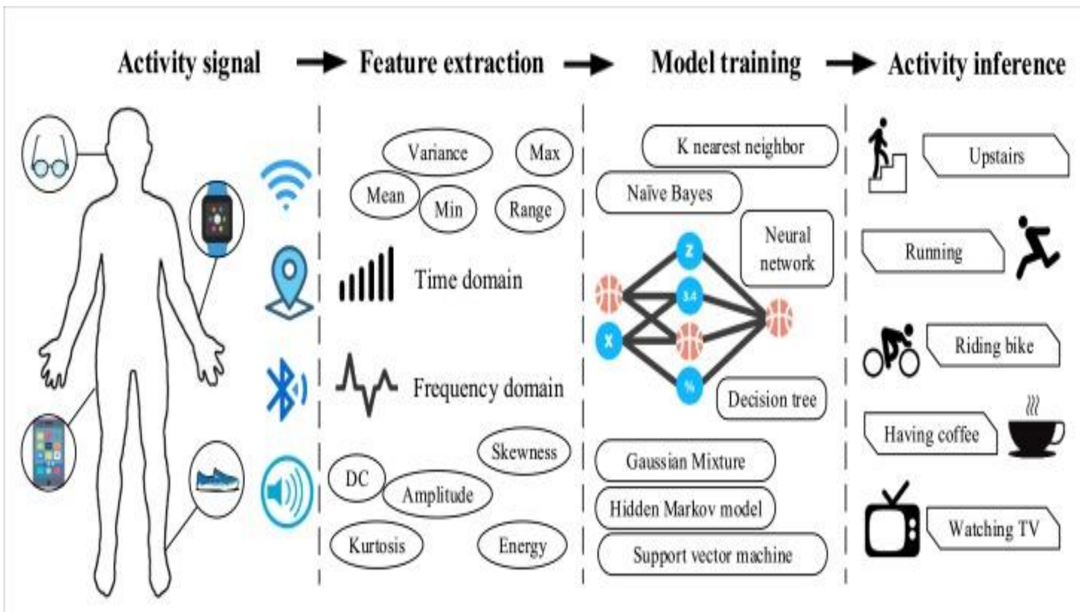

Figure 1. Wearable Activity Recognition System

However, the major limitation of wearable sensing devices is that the devices must be worn by human body, so there are stringent restrictions on the size and power consumption of the wearing detecting sensors. Using non-wearable sensing device can get rid of those restrictions and is friendlier to human body.

Typically, video cameras (Lin et al., 2008) and Doppler radar array (Hong et al., 2013) are employed to implement the non-wearable activity recognitions. The video-camera based methods is based on the designed algorithm to do image processing of continuous images captured from video-camera. The output of the algorithm classifies the activities of human target. However, by using video camera may have issues of privacy invasions, because the video camera typically directly monitors human target. In addition, the accuracy of the activity recognitions is poor in dark environments. Doppler radar array system is another method for non-wearable monitoring system. However, due to the high cost of the system, Doppler radar array is not considered in this research (Hong et al., 2013).

B. Statement of Problem

The problem of this research is to establish a non-wearable activity recognition system based on a thermopile imaging array sensor, HTPA32x32dR1L5.0/0.85F7.7eHiC, which is manufactured by Heimann Sensor GmbH (2017). The existing indoor activity recognition systems are typically based on wearable sensor like acceleration sensors (Khan et al., 2010) and capacitive sensors (Cheng et al., 2013). Those activity recognition system tracking rate is high but not convenient in real application. Video cameras (Lin et al., 2008) can be employed for non-wearable activity recognition system. However, video based system may have a privacy issue, especially when this system is used to monitor elderly people who are living alone for 24 hours a day and 7 days a week. Another type of non-wearable activity recognition system, which is based on Doppler radar array (Hong et al., 2013), is too expensive. Compared with other system, the activity recognition system based on a thermopile imaging array sensor has several advantages: lower cost, uncomplicated system, and lower risk for privacy intrusion.

C. Statement of Purpose

The purpose of this research is to create an activity recognition system, which is constituted by a thermopile imaging array sensor, a Raspberry PI microcontroller, and an algorithm implemented by MATLAB codes. The system developed will be applied in daily life to monitor elderly people who are living alone. The system will provide consecutive images collected from tracking area by a thermopile imaging array sensor and saved on a microcontroller. Based on the sequential images, a series of image processing are implemented through algorithms by MATLAB coding. Then, the target will be defined whether it is within the tracking area. Once the system has identified the human target, the activity of the target will be classified for three situations: quiescence, active, and fall.

D. Need of Study

The need of study is to address the lack of activity recognition system by using thermopile imaging array sensor. The need of study is based on the following factors:

a. Non-wearable;

   Compared with wearable sensing activity recognition systems, a non-wearable system is much friendlier to elderly people. In addition, there are no stringent restrictions on the size and power consumption for the non-wearing detecting sensors.

b. Privacy;

Compared with video based activity recognition system, the action of elderly people will not be exposed to the thermopile imaging array sensor in details. Even for the elderly people, privacy violations should be addressed.

c. Cost;

Compared with other sensing devices for wearable activity recognition systems or cameras for video based systems, the price of thermopile imaging array sensor is lower.

d. Dark environment;

A thermopile imaging array sensor is an electronic device that detects thermal energy, which is from target's body temperature. One of the advantages of this sensor is better performance even in a dark environment. Compared with visible device like video based activity recognition system, even in night without any lights, the system of thermopile imaging array sensor will still be working normally.

Based on the factors discussed above, an indoor activity recognition system based on thermopile imaging array sensor for elderly people is worth to be implemented.

E. Assumption of the Study

For this research certain assumptions are made that will serve as the basis for the calculations and measurements ensuring the performance of the design. Those assumptions are:

1. The data recorded from thermopile imaging array sensor is accurate.
2. With the increasing of distance between target and sensor, the temperature output from thermopile imaging array sensor is decreasing.

F. Research Question

The goal of this research is to develop an activity recognition based on thermopile imaging array sensor. The research questions for this study were:

1. Would the system detect the target in experiment area?
2. How far is the of the target away from sensor?
3. Would the system recognize the activity of target?
4. Would the system provide an alert when the target falls?

G. Delimitations of the Study

This study is delimited to:

1. Thermopile imaging array sensor: HTPA32x32dR1L5.0/0.85F7.7eHiC;
2. Raspberry pi 3;
3. Activity recognition of human target;

H. Limitations of the Study

The following limitations are to be applied to this study:

1. The calibration algorithm used in this study is for thermopile imaging array sensor HTPA32x32dR1L5.0/0.85F7.7eHiC.
2. The SD card for the microcontroller, Raspberry pi 3, to save the data captured from sensor is only 8 GB.
3. The system of this study is only considered for one target in the experiment area at any time.
4. The distance between the sensor and the target is within 6 meters.
5. The system designed in this research is not a real time system.

I. Definition of Terms

Activity recognition: Activity recognition aims to recognize the actions and goals of one or more targets from a series of observations on the targets' actions and the environmental conditions.

Image Processing: Image processing is the use of computer algorithms to perform image processing on digital images (Chakravoty, 2018).

CHAPTER II.

REVIEW OF RELATED LITERATURE

Human activity Recognition (HAR) aims to recognize the actions of one or more targets from a series of information observed on the targets' actions and the environmental conditions. It has been researched since 1980s. This research field has attracted lots of attention of some computer science communities due to its powerfully strength in providing personalized support for different kinds of applications. It is also connected to many other different fields of study like human-computer interaction, sociology, or artificial intelligence

Typically, there are two types of activity recognition. One of them is sensor-based and the other one is vision-based. The idea of sensor-based approach is to keep tracking target's actions by using different kinds of sensors. Through the data collected by sensors, a lot of information can be gained and the activity will be recognized. When it comes to vision-based activity recognition, it is usually implemented with cameras. The targets are observed in details and the decisions are made.

This chapter is organized as follows: Section A briefly introduces human activity recognition and its application in various contexts; Section B describes the types of sensing technologies used in HAR systems; and Section C reviews current researches in activity recognition technologies and systems.

A. Human Activity Recognition

Human activity recognition (HAR) is the ability to capture human posture through sensors and then determine human activity or action. Typically, a HAR system can be either supervised or unsupervised (Lara & Labrador, 2013). A supervised HAR system required some prior training with dedicated datasets, while an unsupervised HAR system is configured with a set of rules during development. HAR is considered as an important component in various scientific researches and applications, such as surveillance, healthcare, and human computer interaction (HCI) (Roshtkhari & Levine, 2013; Chaaraoui et al., 2013)

1. Surveillance System

   In surveillance applications, HAR was adopted in surveillance systems installed at some public places, like banks, supermarket, or airports. (Preis et al., 2012; Popoola & Wang, 2012) A new paradigm of human activity prediction is introduced by Ryoo to prevent crimes and dangerous activities from occurring at public places. (Ryoo, 2011) The findings confirm that the proposed method can identify the early stages of ongoing human interaction. Legion: AR, a system proposed by Lasecki, can offer robust and deployable activity recognition by supplementing existing recognition systems with real time activity identification through the inputs from the crowds at public places, like airport or railway station. (Lasecki et al., 2013)

2. Healthcare

   HAR is commonly used in healthcare systems, which typically are installed in residential environment, hospitals, and rehabilitation centers. Activity recognition system is widely used for monitoring the activities of elderly people who are living alone or staying in rehabilitation centers. For those elderly people who live alone, HAR systems are often combined into smart home systems for tracking the elderly people's daily activities. (Chen et al., 2012). In addition, HAR systems becomes more and more popular in rehabilitation centers to encourage physical exercises for young adults with motor disabilities (Huang, 2011), patients with dysfunction and psychomotor slowing (Gonzalez-Ortega et al., 2013), and exergaming (Alshurafa et al., 2013). Moreover, HAR systems could be applied for observing patients at home to prevent delayed care for emergency and to provide treatment and life logging. Some other behaviors, for example, stereotypical motion conditions in children with Autism Spectrum Disorders (ASD) (Paragiola & Coronato, 2013), abnormal conditions for cardiac patients (Kantoch & Augustyniak, 2012), and detection for early signs of illness (Stone & Skubic, 2012), could be monitored by HAR systems. The results provided by those systems provide clinicians with more opportunities for intervention. Another very important application related to healthcare for HAR systems is, what we proposed in this paper, fall detection and intervention for elderly people. (Nghiem et al., 2012; Vo et al., 2012; Durrant-Whyte et al., 2012)

3. Human Computer Interaction

   For this area, HAR is widely applied in gaming and exergaming, such as Kinect (Han et al., 2013), Niteno Wii (Lawrence et al., 2010) and full-body motion-based games for players (Gerling et al., 2012). By using HAR systems, the gestures of human body are recorded and recognized by machine to guide machine to complete specific tasks. Elderly people and adults with neurological injury can interact with games and exercise games conveniently by doing some simple movements. Beyond that, HAR systems also can help surgeons to have precise control of the intraoperative image monitor by using standardized free-hand movement. (Yoshimitsu et al., 2013)

B. Type of Sensing Technologies used in Human Activity Recognition Systems

As discussed in the background part of Chapter I, generally sensors play the most important role in HAR systems. In a HAR system, body gestures are set as an input, and then a human activity is recognized through a process of operations. Sensors capture all useful information about human body gesture and the recognition part, such as machine, algorithm, or interpreter, analyzes the information and determines what kind of activity has been performed by human. According to a large number of literatures about HAR systems, the sensor technologies used in HAR systems can be classified as wearable sensor and non-wearable sensor. Furthermore, non-wearable sensors also can be classified as video-based and sensor-based.

a) Wearable sensor based activity recognition

The most frequently used sensors for wearable sensor based activity recognition are accelerometer sensors. These sensors are highly effective in monitoring and recognizing actions, like walking, running, sitting, standing, and moving up or down. Bao and Intille (2004) developed a summary of research work that can recognize human body activities using acceleration data. After that, Kern et al. (2003) provide a network of 3-axis accelerometers distributed over a human body.

Each accelerometer provides information about its orientation and movement of corresponding human body part. The microphones and accelerometers were combined worn by a human body for recognizing workshop activities in the research of Lukowicz et al. (2004). Lee and Mase (2002) discussed a dead-reckoning method to determine a human body’s location and recognize activities, like walking sitting, and standing, by measuring acceleration and angular velocity (the angle of the human body’s thigh) through wearable sensors, such as accelerometers and gyroscopes. Mantyjarvi et al. (2001) provided a method to recognize human ambulation and gesture based on acceleration data collected from the hip.

Another widely used wearable sensor is a GPS sensor for locating human body and monitoring activities. This kind of sensor are pervasively used in mobile environments. For example, Patterson et al. (2003) believed that activities like boarding a bus at a special bus stop, traveling, and disembarking can be detected by a GPS sensor stream. Based on that research, Ashbrook and Starner (2003) used a GPS sensor to learn significant locations and predict movement across multiple users. With such systems, Liao et al. (2004) can learn and infer a user’s mode of transportation and their goal in addition to abnormal behaviors like taking a wrong bus based on GPS sensor data logs.

Biosensors are an emerging technology aiming to monitor activities of human body through vital signs, such as blood pressure, heart rate, electrocardiogram

(ECG), and respiratory information. Sung et al. (2004) used biosensors to check body temperature of a soldier to detect hypothermia.

Accelerometer:

Hwang et al. (2004) developed a system that was comprised of accelerometer, tilt sensor, and gyroscope. Accelerometer measured kinetic force, tilt sensor and gyroscope estimated body posture. Four case experiments of three people aged over 26 years old have been done by this system. The accuracy of fall detection is 96.7%. The article proves that accelerometer sensor based system is feasible.

The research done by Jiewen Zheng et al. (2009) proposed a system based on accelerometer sensor that worn on the waist of a human body. The device used a two-stage fall detection algorithm that sensed rapid impact and body orientation of the wearer. In this system, the device can accurately monitoring the falls of elderly people. Privacy and comfort are the advantage of this composed system.

In the above-mentioned researches, accelerometer sensor was typically fixed on the vest or waist of a human body. It may be inconvenient. In the research of Sim et al.(2011), the accelerometer sensor is put on shoes to detect the falls of elderly people. It applied energy harvesting device in this prototype system, which made the system smaller and more power efficiency.

Antenna:

Jihoon Hong et al. (2013) proposed a system that exploited an antenna array on the receiver side and decomposed received signals to improve the performance of the system. This system can classify several states and activities, such as intrusion, walking, and falling. The article shows that the array antenna sensor can be used an indoor monitoring system by communicating with wireless signals.

Smartphone:

Xing Su et al. (2014) believed that smartphones have been changing the world. The power of smartphones is not just computing, networking, but also sensing the activities of the bearer. Compared with those wearable sensors, smartphone sensor is much convenience for people to carry.

Pressure:

Gang Qian et al. (2012) presented a multimodal gesture recognition framework by using video and floor pressure sensor. Compared with typically video-based systems, the authors used additional floor pressure data to make system much more accurate and efficient. The visual results and pressure sensor results were sent to a two-stage cascaded sequential information integration scheme.

Video:

Video-based activity recognition system is another approach for automatic activity recognition. Weiyao Lin et al. (2008) described how to detect each activity by a combination of GMMs (Gaussian Mixture Models) with each GMM representing the distribution of a category feature vector (CFV), which was features with high correlations. In addition, to improve the recognition accuracy, a confident-frame-based recognizing (CFR) algorithm was proposed to recognize human activities. It means only high confidence video frames were used for activity recognition, and a specialized model is used to classify the video frames.

CHAPTER III.

METHODOLOGY

A. System Overview

The system developed by this research to implement the activity recognition system includes four subsystems:

1. Thermal image capturing subsystem;
2. Image data transmission, calibration, and saving subsystem;
3. Image processing subsystem;
4. Activity recognition subsystem.

The thermal image capturing subsystem is based on a thermopile image array sensor. The sensor used by the subsystem is HTPA32x32dR1L5.0/0.85F7.7eHiC (HTPA32x32d), which can capture thermopile images with a resolution of 32x32 pixels.

The captured thermal images are sent to a Raspberry Pi 3 through an I2C interface and are calibrated inside the Raspberry Pi 3 The calibrated images are saved into an SD card on the Raspberry Pi 3 and then are transferred to a PC for image processing. In this research, the images transferring from the Raspberry Pi 3 to a PC and the subsequent images processing are not implemented in a real-time manner, but it will not affect the validity of the research.

After the calibrated images are transferred to a PC, the image processing will be done with MATLAB codes to determine the existence and the location of the tracking target. Finally, the activities of the tracking target will be identified.

A summary about the information flow chart of the activity recognition system is shown in Fig. 2.

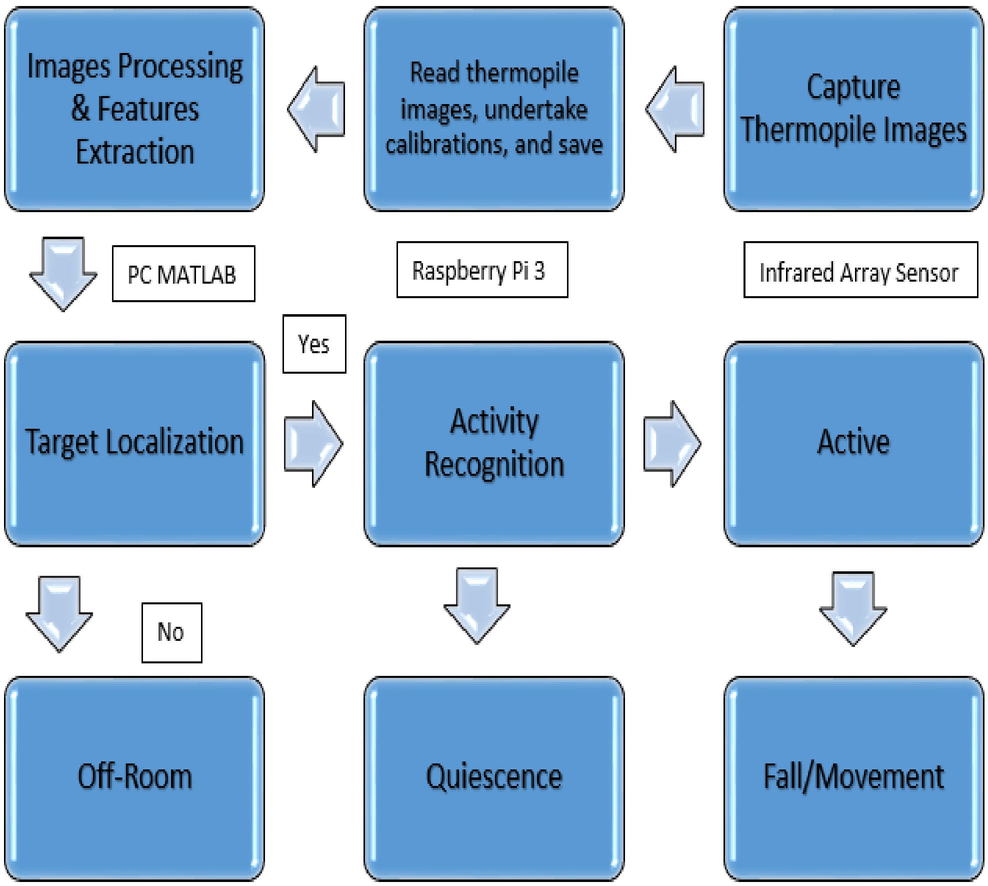

Figure 2. Information Flow Chart of the Activity Recognition System

In the following paragraphs, the details of the four subsystems of the activity recognition system will be discussed one by one.

1. Thermal image capturing subsystem

The thermal image capturing subsystem is made of a thermopile image array sensor, HTPA32x32dR1L5.0/0.85F7.7eHiC (HTPA32x32d). As we have discussed in chapter 1, there are several benefits to use thermopile image array sensors as the input of the activity recognition system, such as privacy and price. To have a better understanding about the sensor, the meaning of each parameter in the name of the sensor is shown in Table 1.

Table 1. Sensor Parameters

Types	HTPA	Heimann Thermopile Arrays
Types	32x32	32*32 elements
Output	d	Digital output
Revision	R1	Silicon revision 1
Optics	L5.0/0.85	Focal length L5.0= 5.0mm F-number 0.85
Filter	F7.7	Filter characteristics
External aperture	e	with external aperture
Sensitivity	Hi	Increased sensitivity
Version	C:	Calibrated sensor (only digital). Carries calibration constants on internal EEPROM

According to the information provided by the above table, we can find out that this thermopile array sensor can capture thermopile images with a resolution of 32×32 pixels. Compared with thermopile array sensor models, for example HTPA 8x8d, HTPA 16x4d, and HTPA 16x16d, the HTPA 32x32d model captures images with higher resolutions. It means that each image captured can provide more environment and human target data information for image processing. It may significantly increase the accuracy of this activity recognition system.

In addition, according to the datasheet of HTPA32x32d, it has a 33° field of view(FoV), which is much larger than other models. For example, the HTPA 16x16d only has a 20° FoV. That is why in the researches using low resolution thermopile array sensors, the sensors are typically installed on the center of a roof (not on a wall) to observe the experiment area from top to down. Such installation will enhance information in horizontal direction within the observe area, but lose the information in vertical direction. In the proposed activity recognition system, the HTPA 32x32d is installed on a wall with a specific height, which is good enough to cover an experiment area of six meters in length and four meter in width.

Moreover, the size of HTPA32x32d is really small, which shows in the Fig. 3 below. It is much smaller than a quarter coin.

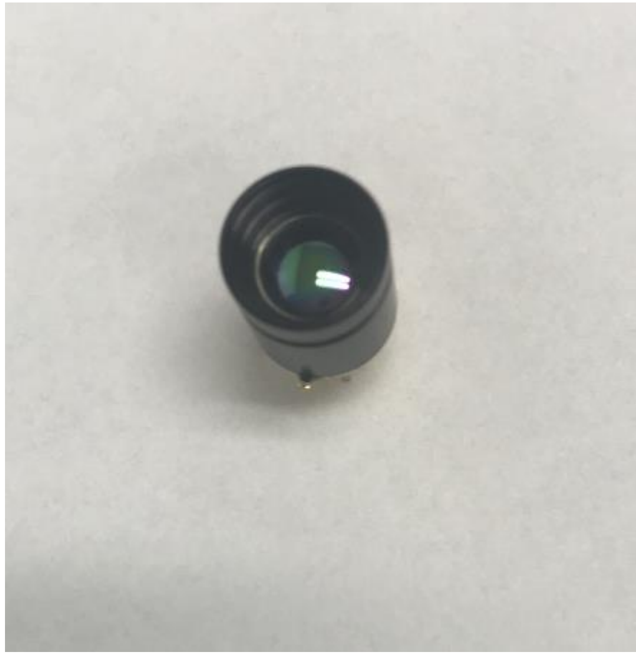

Figure 3. HTPA32x32d Sensor

The HTPA32x32d is soldered on a printed circuit board (PCB) designed by us. This PCB provides space for an HTPA32x32d sensor, a 75uf capacitor, a 100uf capacitor, and four pins to connect to the Raspberry Pi 3. A picture of the PCB is shown in Figure 4 below. The connections between the four pins on the PCB and the four pins of the HTPA32x32d are:

The first pin on the PCB (yellow wire) is connected to the pin 1 of the HTPA 32x32d, which is the digital I/O pin for serial clock (SCL).

The second pin on the PCB (orange wire) is connected to the pin 2 of the HTPA 32x32d, which is positive supply voltage (VDD).

The third pin on the PCB (black wire) is connected to the pin 3 of the HTPA 32x32d, which is negative supply voltage or ground (VSS).

The last pin on the PCB (green wire) is connected to the pin 4 of the HTPA 32x32d, which is Digital I/O pin for serial data (SDA).

All these four pins on the PCB are connected to the Raspberry Pi 3 so as to transmit the captured thermal images from the sensor to the Raspberry Pi 3.

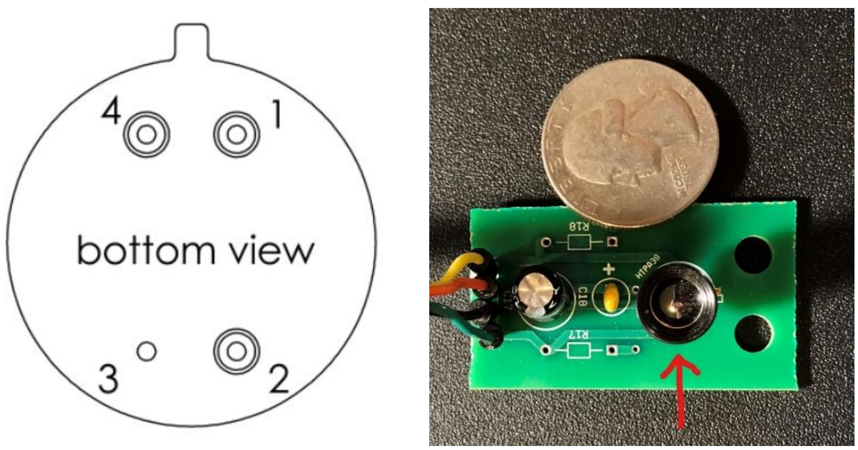

Figure 4. Sensor Bottom View & PCB Connection

The sensor with the PCB board is installed on a wall at the height of 1.8 meters and overlooks the ground at an angle of 16.5° (i.e. half of 33°) below a horizontal line. In Fig. 5 below, it shows how the sensor installed on the wall in the experiment area roughly. Such installation allows the thermopile imaging array sensor observing more area below 1.8 meters for activity recognition. We choose a height of 1.8 meters, because we assume that the height of most elderly people is lower than 1.8 meters. In addition, even the human target is higher than 1.8 meters, the activity recognition algorithm still can make a right identification on falling.

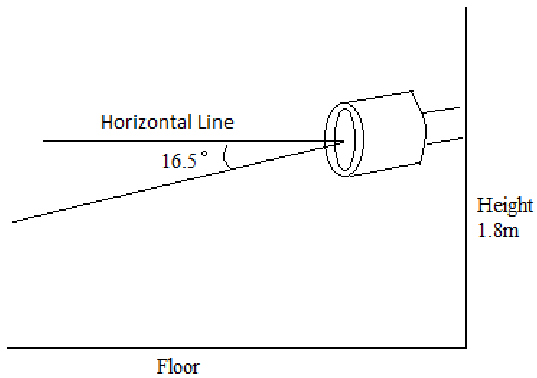

Figure 5. Sensor installation

2. Image data transmission, calibration, and saving subsystem

This subsystem has three major functions:

a. To active the thermopile array sensor and receive the thermal image data from the sensor;

b. To calibrate the received thermal images according to specific calibration algorithms;

c. To save the calibrated thermal images into a SD card for future processing.

To implement those functions, in this research, a Raspberry Pi 3 Model B is used to build the image data transmission, calibration, and saving subsystem. Raspberry Pis are a series of small single-board computers, which are developed to promote teaching of basic computer science in schools. There are several generations of Raspberry Pis. All of those models are built with a Broadcom system-on-a-chip (SOC) with an integrated ARM-compatible central processing unit (CPU) and an on-chip graphics processing unit (GPU). Raspberry Pi 3 Model B was released in February 2016 which was one of the latest model of Raspberry Pi series. The key features of this model is shown in Table 2 below.

Table 2. Features of Raspberry Pi 3

*	CPU	Quad Core 1.2GHz Broadcom BCM2837 64bit
	RAM	1GB
*	Bluetooth	Bluetooth Low Energy (BLE) on board
*	Wireless	BCM 43438 wireless LAN on board
*	GPIO	40-pin extended
*	USB 2	4 ports
*	HDMI	Full size
	CSI	camera port for connecting a Raspberry Pi camera
	DSI	DSI display port for connecting a Raspberry Pi touchscreen display
*	Micro SD	Micro SD port for loading your operating system and storing data

A general illustration about the interface of a Raspberry Pi 3 Model B is shown in the Fig. 6 below.

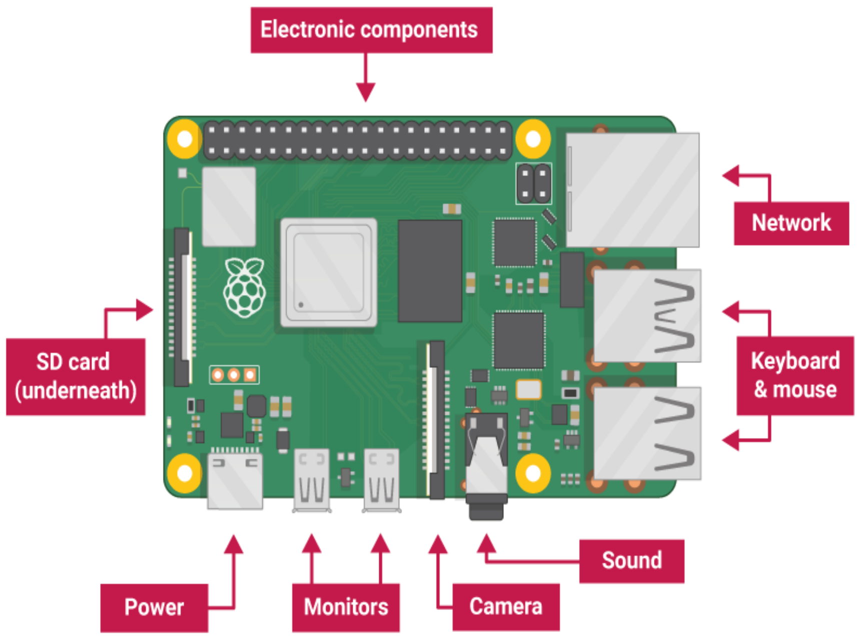

Figure 6. Raspberry Pi 3 Board

There are several reasons to build the image data transmission, calibration, and saving subsystem with a Raspberry Pi 3 Model B. First of all, the ARM cores of the Raspberry Pi 3 Model B run at 1.2GHz, making the device about 50% faster than other core (Raspberry Pi Documentation). The subsystem needs to receive image data from the thermopile array sensor and do image calibration at the same time. Thus, the speed of core becomes one of the most important reason to choose the Raspberry Pi 3 Model B. Secondly, the Raspberry Pi 3 Model B has a fully functional board, which includes USB-2 ports and HDMI ports. It means that a keyboard, mouse, and monitor can be directly connected on the Raspberry Pi 3 Model B during the experiment, which makes it convenient to operate the system and observe the output images during the experiment. The Raspberry Pi 3 Model B has an SD card port that can store all data files and the Raspberry Pi 3 Model B operating system.

In addition, the Raspberry Pi 3 Model B has an Ethernet port and a built-in WIFI function, so it can connect to Internet easily and send the captured thermal image data in real time. For the current research, this function has not yet been used, and the image data are saved to the SD card in Raspberry Pi 3 Model B and then transferred to a PC by using a USB flash disk or directly reading the SD card in a PC. That is to say, the current research is not a real-time monitoring system. However, if this activity recognition system is applied to the real life in the future, it should be a real-time monitoring system. The built-in Ethernet port and WIFI function of the Raspberry Pi 3 provides a possibility to make the system become a real-time system.

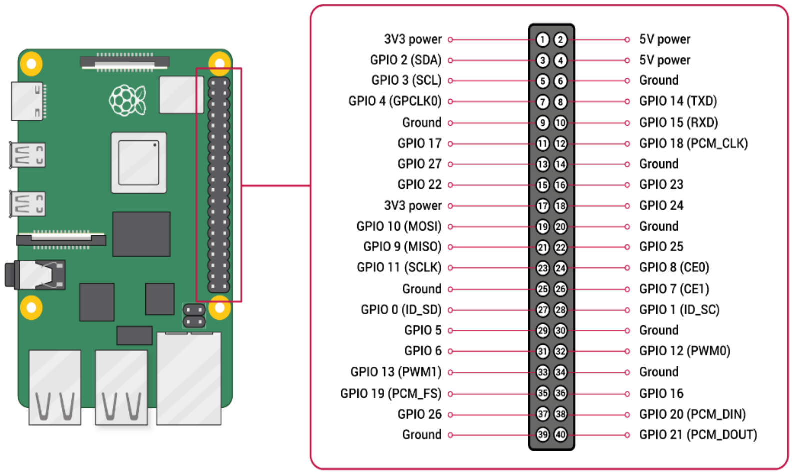

Figure 7. Raspberry PI 3 Pin Mappings

According to Fig. 7 above, a powerful feature of the Raspberry Pi 3 is the row of GPIO (general-purpose input/output) pins along the top edge of the board. As mentioned in the thermal image capturing subsystem, there are 4 pins from the PCB connected to the Raspberry Pi 3 board. The pins 1, 3, 5 and ground pin 6 on the Raspberry Pi 3 are used to control the HTPA 32x32d sensor.

At last, the price of the Raspberry Pi 3 Model B is only 35 dollars, which is much cheaper than other develop kits used in different recognition systems, like MSP 430 which is about 190 dollars.

3. Image processing subsystem

The image processing subsystem is implemented by a group of algorithms that are realized with MATLAB codes.

MATLAB is a software that provides a multi-paradigm numerical computing environment and uses a proprietary programming language developed by MathWorks. MATLAB is good at matrix manipulations, plotting of functions and data, implementation of numerical algorithms, and interfacing with the programs written in other languages. The outputs of the thermopile array sensor are 32x32 pixels of temperature, which can be treated as a 32x32 matrix for every single image. In addition, MATLAB was first adopted by researchers and practitioners in control engineering, Little's specialty, but quickly spread to many other domains. This software is the most popular software amongst scientists involved in image processing. (Moler, 2017)

The algorithms implemented with MATLAB codes will load the thermopile image data and do image processing to extract the feature point, F(n), i.e. the human target, from each frame of the thermopile image data. Here n is the index for each frame. The extracted feature points will be used by the activity recognition subsystem to identify the location of human target and identify the status of the activity of the human target.

4. Activity recognition subsystem

The activity recognition subsystem has three major functions: a. identify the location of the human target; b. determine the status of activity of the human target; c. send an alert once the system detects a fall status of the human target.

a. The location of the human target

The parameters of a feature point, P(n), include both the temperature, T(n), and the coordinates {x(n), y(n)}, of the feature point. From the value of T(n), the distance between the feature point and the thermopile array sensor, D(n), can be calculated, because T(n) = a*D(n) +b, where a and b are two known constants.

About the coordinates of the feature point, x(n) is the x^{th} pixel from the left of the thermal image and y is the y^{th} pixel from the bottom of the thermal image. With x(n), y(n), and D(n), the 3 dimensional location of the feature points in the experiment area can be identified.

b. Status of activity of the human target;

According to the information flow chart shown in Fig. 2 at the beginning of this chapter, the activity of the human target can be defined by three statuses: quiescence, movement, and fall.

Quiescence is defined that, within a certain number of frames, the coordinates of the feature point, F(n), which represent the location of human target, are not changed. It means that during these frames, the human target can be standing, sleeping, or sitting somewhere, but does not move. This status of the human target is called quiescence. Oppositely, once the coordinates of the feature point, F(n), changes, it means the human target is active. Here, movement and fall are two statuses for this situation. The movement actually means that the human target moves in a normal manner. The status of movement includes walking around, squatting down, and standing up. The fall status means that the human target has an uncommon activity.

c. An alert

Generally, the system will show the status of the human target like "Off-room", "Quiescence", and "Movement". Once the system detects a status of "Fall", the system will not only show the status of human target, and also send an alert to warn that there is a fall activity in the monitoring area.

B. Algorithm

As shown in Fig. 8, the activity recognition algorithms for this activity recognition system includes four steps: (1) image preprocessing; (2) connected domain extraction; (3) localization by the feature point; and (4) activity recognition. The details of the four steps are explained in the following paragraphs

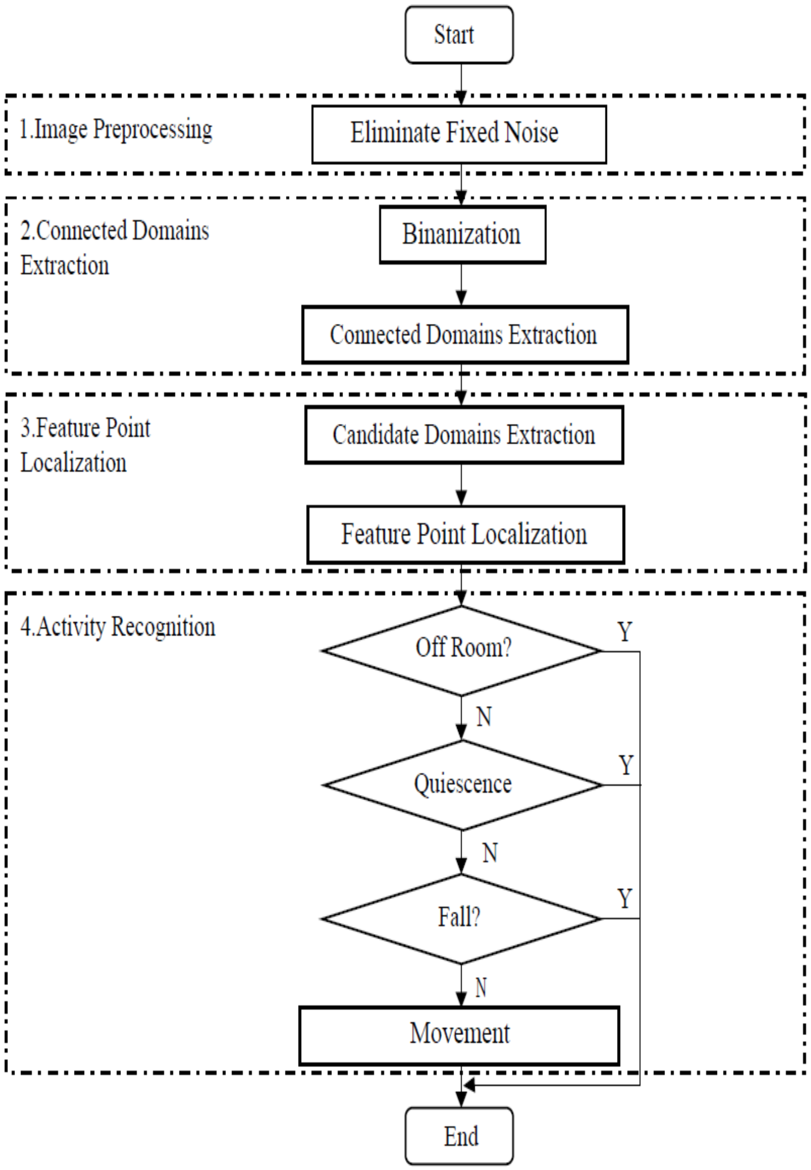

Figure 8. Activity recognition algorithm of the proposed system

- <u>Step I: Image Preprocessing;</u>

One of the original thermopile images captured by HTPA 32x32d sensor is shown in the fig. 9. In this figure, it contains two kinds of noise: fixed noises, which are circled in black, and random noises, which are circled in red. Fixed noises are some fixed temperature offsets existing in the same pixels of every single frame. Those noises are caused by some fixed high temperature objects in the monitoring area as well as some internal errors of the thermopile array senor during capturing and calibrating images. Different from the fixed noises, the random noises are temperature offset appearing randomly in different pixels of every single frame.

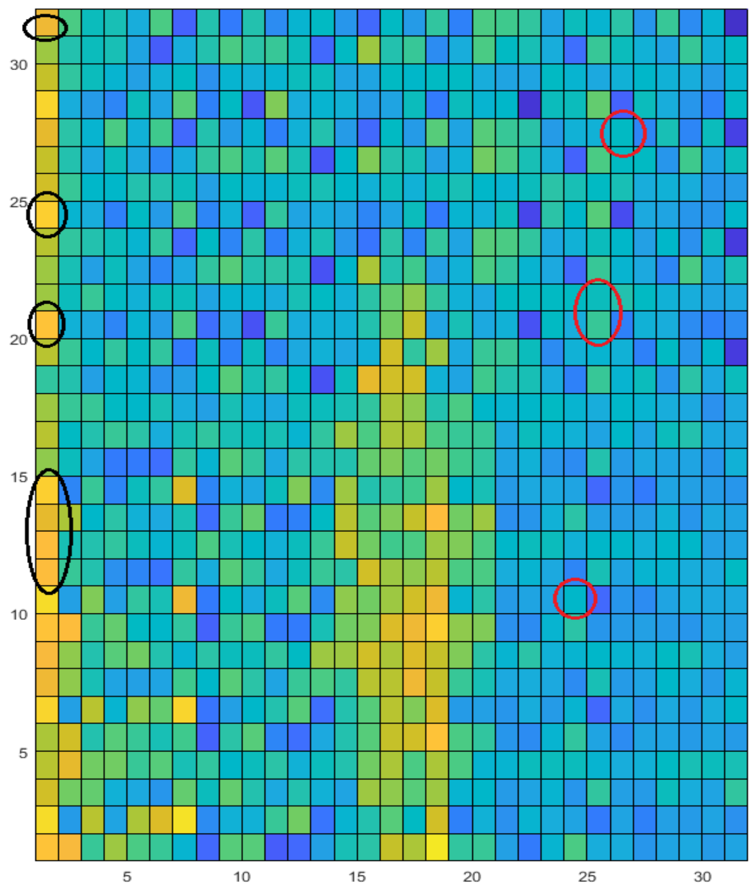

Figure 9. Noise of Original Image

Because of the existence of the noises, it is difficult to identify a human target from the original thermal images. Therefore, the first step of image processing is to mitigate the destructive effects of those noises. In this step, the algorithm is focused on eliminating the fixed noises, and the process of the algorithm is demonstrated in the following equation (1, 2)

$$Z_P(n) = X_P(n) - Y_P(n) \qquad \text{Equation 1}$$

$$Y_P(n) = \alpha Y_P(n-1) + (1-\alpha) X_P(n) \qquad \text{Equation 2}$$

where, $X_P(n)$ denotes the original temperature of the pixel P (P $\in [1, 1024]$) at the nth frame, $Y_P(n)$ represents the time-domain averaged temperature at the same pixel, and $Z_P(n)$ stands for the temperature after the image preprocessing. A large number of experiments show that when α=0.99, most background noises can be removed. An example result is shown in Fig. 10. Obviously, after the image preprocessing, compared with the original image in Fig. 9, most fixed noises have been eliminated, and there is a clearer outline of a human target in the image. However, as shown in Fig. 10, there still are some random noises circled in red in the image. Hence, the algorithms in the following step are to avoid the negative effects of the random noises on identifying the human target.

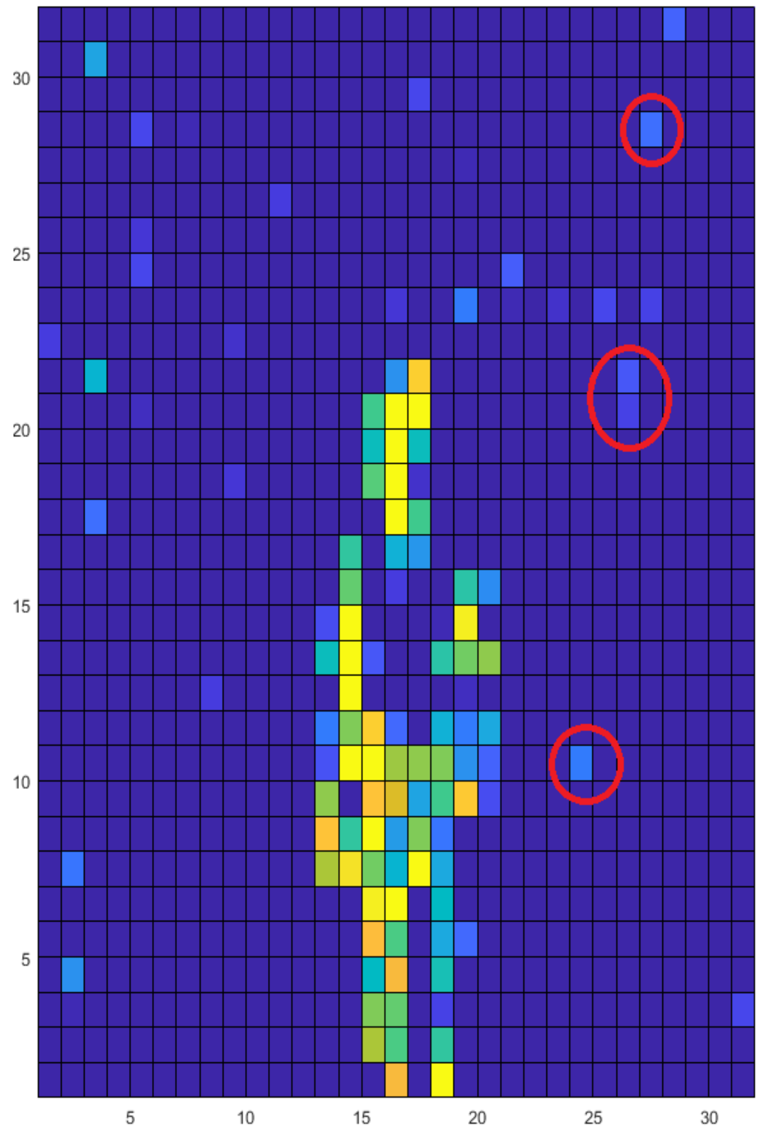

Figure 10. Image After Eliminating Noise

- <u>Step II: Connected Domains Extraction</u>

The algorithms in the step II are designed to extract the shape of the human target from the thermal image and avoid the negative effects of the random noises.

Typically, there are several methods that can achieve this goal, such as edge detection extraction, or connected domains extraction.

1. Edge detection and extraction;

   Edge detection is one of image processing methods that is to find the boundaries of objects. It works by detecting discontinuities in brightness. For the edge detection, where the gray value changes dramatically, it is defined as the edge.

   Edge definition: it is the place where the image gray level changes the most dramatically. An edge is caused by a discontinuous change in the surface normal of an image. It is generally believed that edge extraction is to identify the region where the grayscale of the image changes dramatically. From the perspective of mathematics, the most intuitive method is the differentiation operation. From the perspective of signal processing, it can also be said that a high-pass filter is used to retain the high-frequency signal. Common edge detection algorithms include Sobel, Canny, Prewitt, Roberts, and fuzzy logic methods.

   During this research, we used canny algorithm, the result is shown in Fig. 11. The edge detection results for most thermal images looks not good. As a result, the connected domain extraction algorithm becomes our priority method.

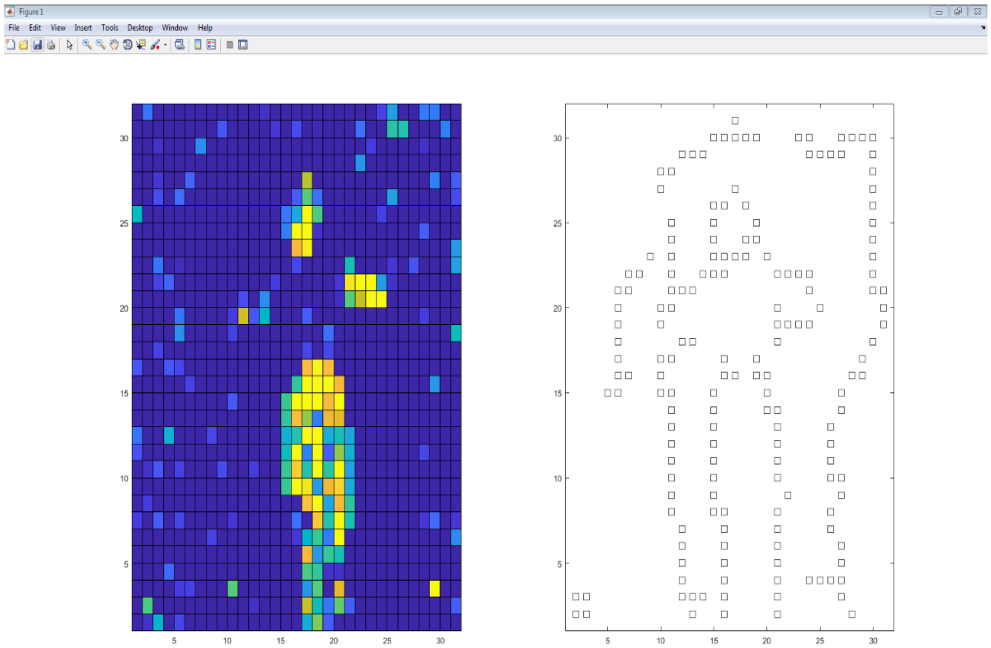

Figure 11. Results of Canny Method

2. Connected domain extraction

To find out a connected domain, the image has to be converted into a binary image. As named, the brightness value of the image has only two states: black (0) and white (255). Binary image plays a very important role in image processing and recognition because of its simple pattern and strong expressive force on spatial relationship of pixels. In practical applications, many images are converted into binary images for further image processing.

A connected domain is a region of pixels that have adjacency relation (also called pixel connectivity). In an image, the smallest unit is pixel, and there are 8 adjacent pixels around each pixel. When we define an adjacency relation (or pixel connectivity), there are two common connectivity definitions: 4-connectivity and 8-connectivity.

1. 4-connectivity

   In the 4-connectivity, pixels are defined to be connected only if they have a common edge (horizontal or vertical). As shown in Fig. 12 below, when the 4-connectivity method is used, P is connected to pixel P2, P4, P6, and P8.

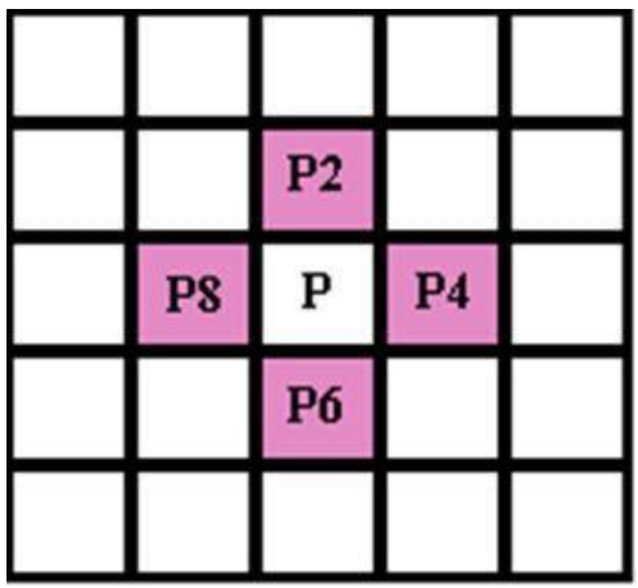

Figure 12. 4-connectivity Method

2. 8-connectivity

   In the 8-connectivity, pixels are defined to be connected if they have a common edge (horizontal or vertical) or they have a common vertex. As shown in the Fig. 13 below, P is connected to pixel P2, P4, P6, and P8, which are the same as using 4-connectivity method. In addition, P is connected to pixel P1, P3, P5, and P7.

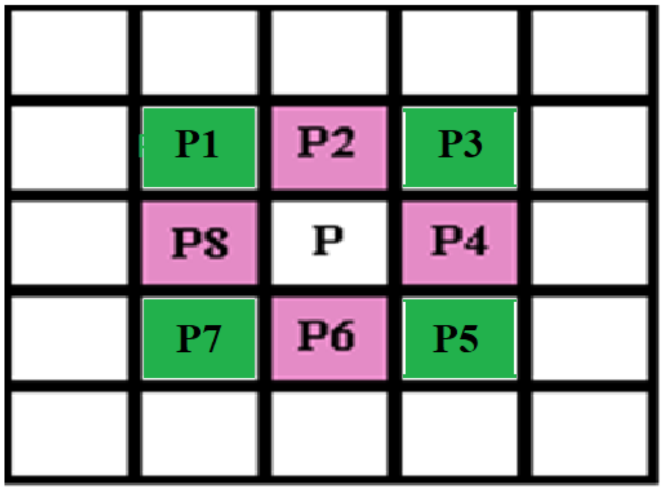

Figure 13. 8-connectivity Method

Visually, all connected pixels forms a domain, while disconnected pixels form different domains. As a result, an image can be formed by a set of domains. During this research, both 4-connectivity and 8-connectivity methods are used. However, the results of 4-connectivity method are unstable because of its low connectivity.

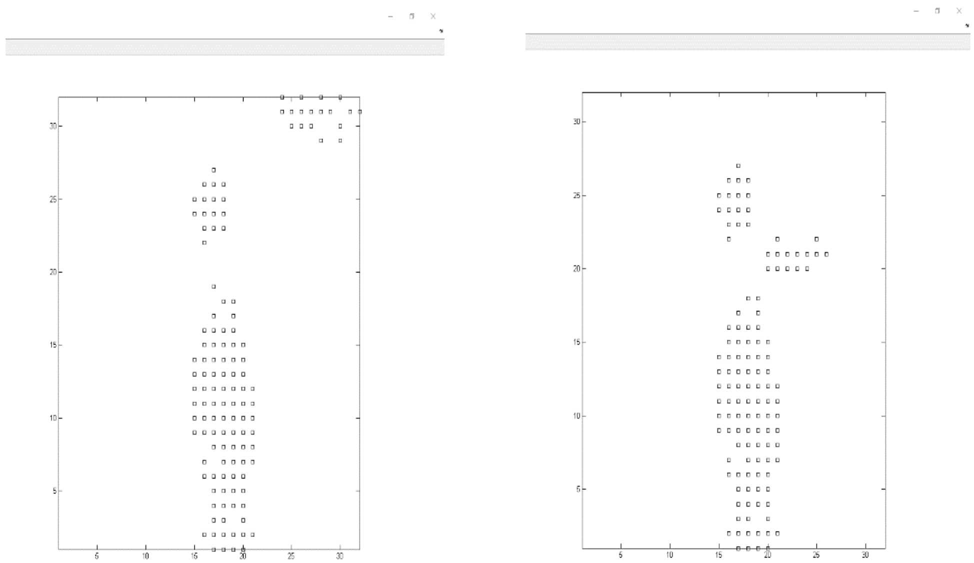

Figure 14. Comparison between 4-connectivity and 8-connectivity

As shown in Fig. 14 above, when both the 4-connectivity and the 8-connectivity methods are applied to the same image, the left one (the 4-connectivity) only shows partial of a human target as a connected domain, and the right one (the 8-connectivity) displays almost a whole human target as a domain. The performance difference of the 4-connectivity and the 8-connectivity can be justified as follows. A human body may be covered by clothes that can significantly reduce the temperature of the human body detected by the thermopile sensor. As a result, when the thermal image is converted to the binary image, some part of the human body may be marked as black. Correspondingly, the human body covered by cloths may be only represented by few pixels. Meanwhile the arms and legs of the human target are formed by a large number of pixels. Therefore, if the 4-

connectivity method is used, the arms and legs may be separated from the human body due to the low connectivity of the 4-connectivity method. If the human body is separated to several connected domains, the accuracy to identify the location of the human target will decrease significantly. That is why the 8-connectivity method is finally chosen for this research. With higher connectivity among pixels, the 8-connectivity increases the chance to find out whole human target in each frame of the thermal images.

As shown in Fig. 15, the extraction results for the connected domain include not only the human target circled in blue, but also other connected domains due to random noises. However, the connected domain of the human target contains much more pixels than the other connected domains, so it is very easy to identify the human target from the size of the connected domains.

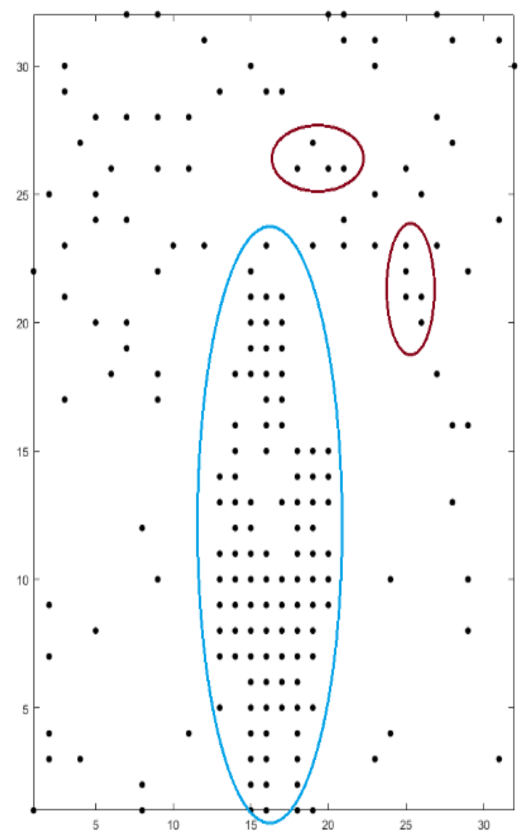

Figure 15. Connected Domains

- Step III: Feature Point Localization

Typically, the body temperature is much higher than environment temperature, so the number of pixels in the connected domain of the human target should be much larger than that in others connected domain. To further improve the accuracy on identifying the location of human target, we consider not only the number of pixels but also the temperature of each pixel in a connected domain. More specifically, the sum of the temperature of all pixels in the i^{th} connected domain, denoted as $T_i(n)$, is calculated, and the connected domains with the three highest temperature values are picked as the candidate domains, which are the circled domains in Fig. 15. After this step, most connected domains caused by random noises have been eliminated.

In the activity recognition algorithm developed by this research, in order to easily determine the status of activity of the human target, the head of the human target is selected as the feature point P(n). One reason for such selection is that other parts of the human target may be blocked by clothes, which may affect the temperature captured by the thermopile array sensor. Another reason is that the head generally is at the top of the human target, so it is easy to be identified. To identify the feature point P(n) from the three candidate domains, all the pixels of the candidate domains are scanned from top to down, starting from the candidate domain with the highest temperature value. According to the result obtained from many experiments, we found that the pixels representing the head of the human target are typically three or more consecutive pixels in horizontal direction. Therefore, if such consecutive pixels in horizontal are found in the candidate domain, the feature point P(n) is identified, which is the pixel in the middle of the consecutive pixels. The candidate domain that includes P(n) is called the feature domain F(n). As shown in Fig. 16, in the candidate domain with the highest temperature value, from top to down the first three consecutive pixels in horizontal direction are found, and the pixel (point A) with the red circle is the feature point P(n) that represents the head of the human target.

If the feature point P(n) is not found in the candidate domain with the highest temperature value, the search will move to the candidate domains with the second highest temperature values, and then the candidate domains with third highest temperature values. If the feature point P(n) is not found in all three candidate domains, the thermal image is identified as no feature point.

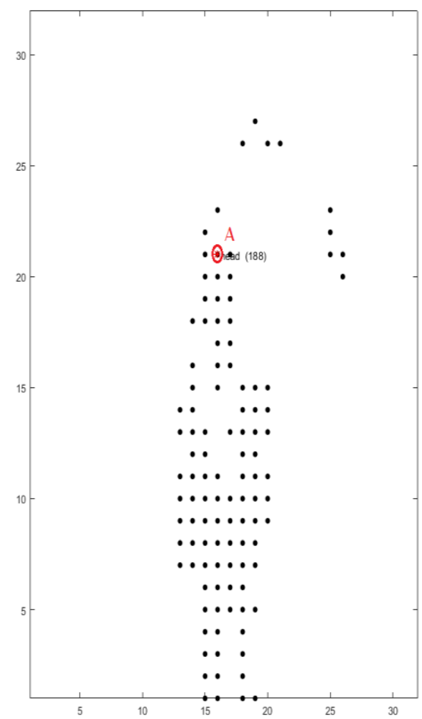

Figure 16. Feature Point P(n)

- <u>Step IV: Activity Recognition</u>

The activity of the human target will be recognized once after it is identified to be located in the room. Otherwise, the system will show “off-room”. There are two states of activities: “quiescence” and “active”. In “active” state, two kinds of activities will be recognized, namely “fall” and “movement”. The determination of the four status are discussed as follows.

1. Off-room

   The activity recognition starts from determining the location of the human target. If the feature point P(n) cannot be found in 10 consecutive frames, it is determined that the human target is not in the room, and the system will show “off room”. As an example in Fig. 17 below, there is no three consecutive pixels or more in horizontal direction in all candidate domains. After 10 consecutive frames like that, the system will show “off room”.

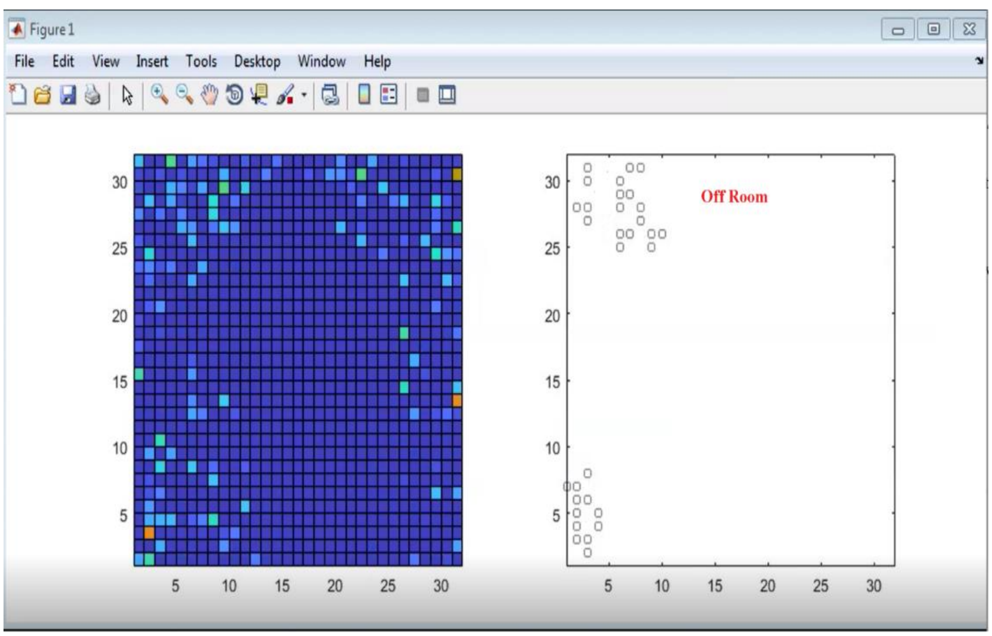

Figure 17. Output of Off Room Status

2. Quiescence

The basic rule to determine the human target is on "quiescence" is that the feature point P(n) does not significant change in both horizontal direction and vertical direction in 10 consecutive frames. In this research, the significant change is defined as 5 pixels in both directions. For example, in Fig. 18, from 100^{th} to 109^{th} frame, the P(n) is changing, but the change in both directions are less than 5 pixels. So after 109^{th} frame, the system will show "quiescence" until the target moves.

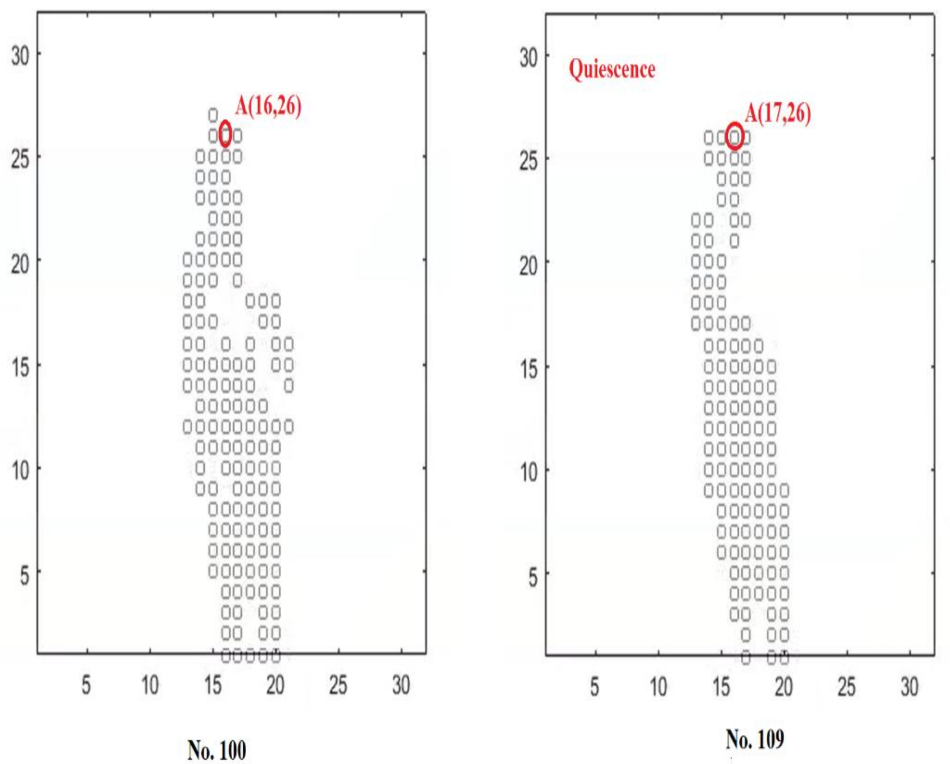

Figure 18. Quiescence Status

3. Active

a. Movement

The condition to determine the status of "movement" is set as: comparing the $(n-6)^{th}$ frame and the n^{th} frame, the feature point P(n) changes more than 5 pixels in horizontal direction and less than 15 pixels in vertical direction, i.e. no significant displacement in vertical.

Typically, when the human target walks in a room, the coordinate of feature point will not have sharp change in vertical direction. There should be only horizontal change. As shown in Fig. 19, the human target is moving from right to left. Since the feature point (the head of the human target) keeps almost in the same height, there is no significant displacement in vertical direction. Meanwhile when the human target moves to left, the feature point has a big change in horizontal direction, from 19 to 13. As a result, the activity is identified as "movement" by the system.

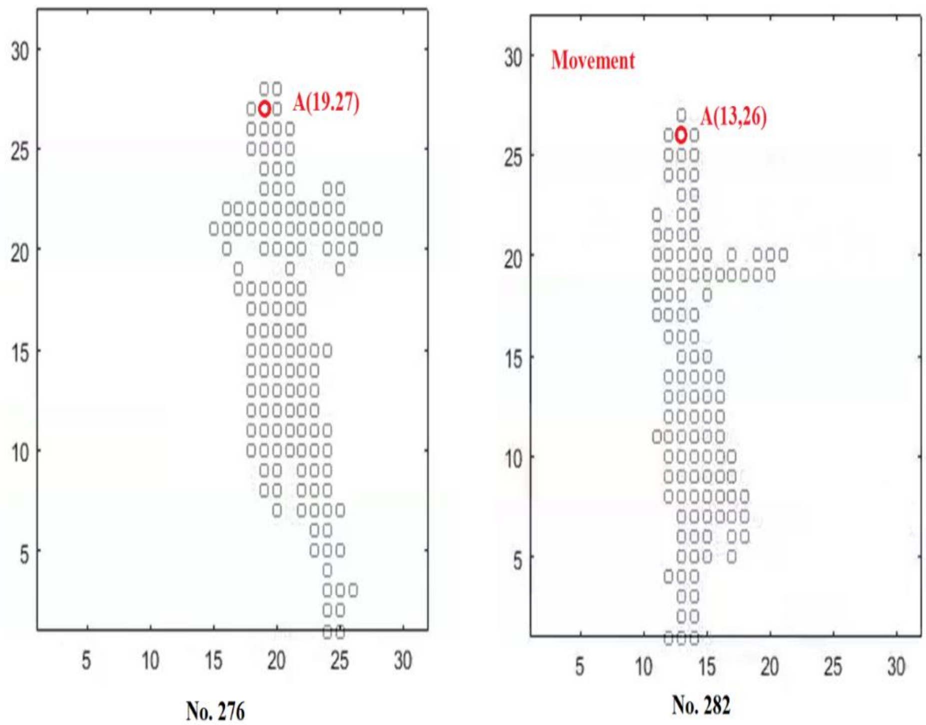

Figure 19. Movement Status

b. Fall

The condition to determine the status of "fall" is set as: comparing the $(n-6)^{th}$ frame and the n^{th} frame, the displacement of the feature point P(n) in vertical direction is more than 15 pixels; meanwhile the displacement of P(n) in horizontal direction is smaller than 10 pixels, the activity will be recognized as "fall".

It is easy to understand that when the human target is falling down, the vertical coordinate of the feature point is going down in rapidly. After undertaking a lot of experiments, we set the threshold for falling to be 15 or more pixel change in vertical direction. In addition, the activity of falling down happens in a sudden, typically with in a second. Therefore, the pixel change in vertical direction needs to be achieved within a few frame durations. In this system, the framerate of the thermopile array senor is 8 frames per second. Thus, we set the frame duration for a fall to be 7 frames, which is less than 1 second. As shown in Fig. 20, the human target is falling down during 282th and 288th frame. The vertical coordinate changes sharply from 26 to 5, meanwhile the horizontal displacement is less than 10 pixels.

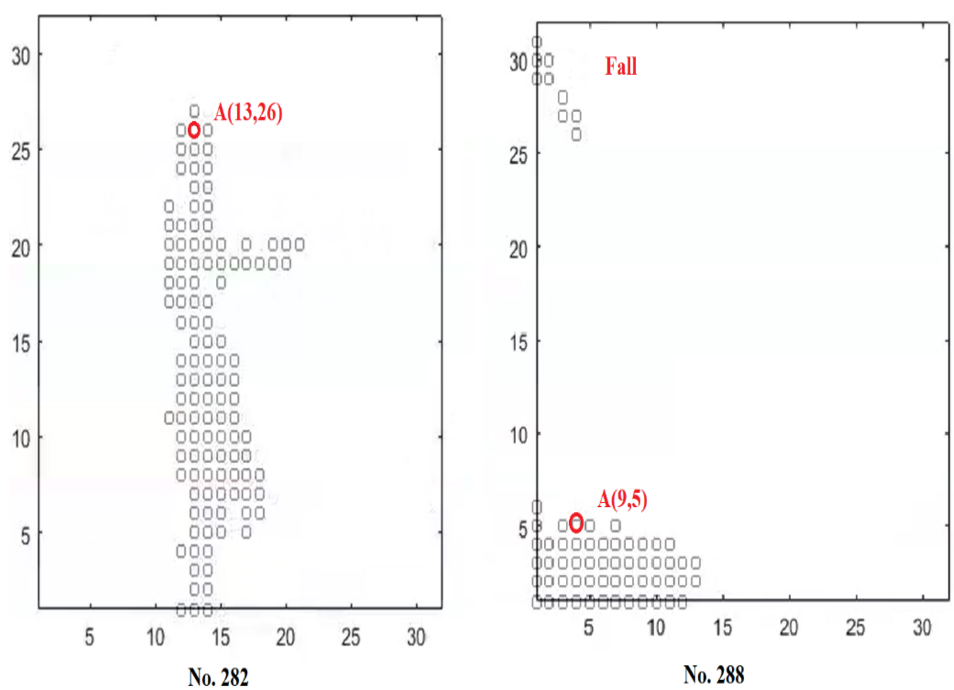

Figure 20. Fall Status

CHAPTER IV.

RESULT AND DISCUSSION

A. Experiment Results

More than 300 laboratory experiments were carried out to test the proposed algorithm in the experiment area. The experiments include 50 times for off-rooms (OFF) status tests, 50 times for quiescence (Q) status tests, 100 times for movements (M) status tests, and 100 times for falls (F) status tests. Every status test was set up at different positions. The detail information about each status test description are discussed as follows.

For the off-rooms status test, the human target stays near the boundary of the experiments area for a while, and then leaves the experiments area. The system should show “quiescence” status at the beginning when the human target is within the experiments area and then change to “off-room” status after the human target leaves the experiments area. The human target is in and out of experiment area for 50 times. The experiments results are recorded in the table below.

Table 3. Off-Room Test Results

Test result	50
Correct	49
Error	1

From the test results, we can see that the system can identify that the human target has left the experiment area with high accuracy.

For the quiescence status test, the human target is sitting, standing or laying in the experiments area. The distances between the human target and the thermopile array sensor is increasing from 1.5 meters to 6 meters with a step size of 0.5 meters. For every distance, 5 test are recorded. All the results at different positions are shown in the table below.

Table 4. Quiescence Test Results

Distance	1.5m	2.0m	2.5m	3.0m	3.5m	4.0m	4.5m	5.0m	5.5m	6.0m
Correct	5	5	5	5	5	5	5	5	5	5
Error	0	0	0	0	0	0	0	0	0	0

For the movements status test, the human target is walking around, squatting down and standing up in different positions, and the distances between the human target and the thermopile array sensor is increased from 1.5 meters to 6 meters with a step size of 0.5 meter. For example, the human target will walk from 1.5 meters to 2 meters and squat down and stand up at the position of 2 meters. After that, the human target will walk back from 2 meters to 1.5 meters and squat down and stand up again. The system should show “movement” status during all these process. Otherwise, the test result will be recorded as fail. The results for the movement status test are recorded below. The table also shows the number of misrecognition

Table 5. Movement Test Results

Distance	1.5m	2.0m	2.5m	3.0m	3.5m	4.0m	4.5m	5.0m	5.5m	6.0m
Correct	9	10	8	8	10	8	9	10	9	10
Error	1	0	2	2	0	2	1	0	1	0

For the fall status test, similarly, at different positions that are from 1.5 meters away from the thermopile array sensor to 6 meters away from the sensor with a step size of 0.5 meter, the human target will fall down and stay on the ground for a while. For each distance, the results for 10 tests are recorded and listed in the table below.

Table 6. Fall Test Results

Distance	1.5m	2.0m	2.5m	3.0m	3.5m	4.0m	4.5m	5.0m	5.5m	6.0m
Correct	10	10	10	10	10	10	9	10	9	9
Error	0	0	0	0	0	0	1	0	1	1

B. Discussion

During the experiments, the thermopile array sensor was fixed on the middle of a wall with a height of 1.8 meters. The sensor observed the experiment area with an angle 16.5˚ below the horizontal line (see Fig. 5). In the experiment area the human target was monitored for four different kinds of activities: off-room, quiescence, movement, and fall. The experiments were done for more than three hundred times. The overall results of the experiments are recorded in the table below.

Table 7. Summary of Test Results

	Off-room	Quiescence	Movement	Fall	Total tests	Correct tests	Accuracy
Off-room	49	1	0	0	50	49	98%
Quiescence	0	50	0	0	50	50	100%
Movement	0	0	91	9	100	91	91%
Fall	0	0	3	97	100	97	97%
					300	287	95.66%

According to this table, the accuracy obtained by the proposed algorithm for detecting the off-room, quiescence, movement, and fall statuses are 98%, 100%, 91%, and 97%, respectively. The overall accuracy of all detections is 95.66%. What impresses us most is that the accuracy of fall detection is 97%, which means almost falls can be detected by using the proposed algorithm.

During the off-room status tests, there is only one result wrong, which shows "quiescence" instead of "off-room". The reason for this misrecognition may be that an object with higher than environment temperature is close to the location of the human target. After the human target left the experiment area, the algorithm mistakenly detected the object and recognized it as the human target.

The accuracy for the quiescence status tests is 100%. That is to say, no matter where the human target was, however far away it is from the sensor, the algorithm can always find out the human target and determine its status.

The movement status tests got the lowest accuracy among four activity statuses, which is only 91%. As shown in Table 6 and Table 7, the misrecognitions for the movement status happened at different distances, and all misrecognitions show "fall" instead of "movement". During the movement status test, the human target was walking between different positons, squatting down, and standing up. It is worth to note that the action of squatting down is quite similar to falling down. This is the reason that squatting down was chosen to be performed during the tests. The movement status is defined that when comparing the $(n\text{-}6)^{th}$ frame and the n^{th} frame, the feature point P(n) changed significantly in horizontal direction but no significant displacement in vertical direction. The fall status is defined that when comparing the $(n\text{-}6)^{th}$ frame and the n^{th} frame, the feature point P(n) changed significantly in vertical direction. Therefore, if the human target performs squatting down, but does not standing up quickly, the activity of the human target will be determined to be the fall status instead of the movement status.

During the fall status tests, three misrecognitions are happened when the distances are more than 4.5 meters. The reason for the misrecognitions is that the threshold to identify the fall status is difficult to be set. The threshold in the proposed algorithm is that the feature point P(n) has 15 pixels or more displacement in vertical direction when comparing the $(n-6)^{th}$ frame and the n^{th} frame. However, when the human target is far away from the thermopile array sensor, the size of the human target becomes small, and the temperature of the human target detected by the sensor becomes low and close to the environment temperature. As a result, the feature domain F(n) becomes smaller and smaller, and even the human target has fallen on the ground, the vertical coordinate of the feature point P(n) may not change more than 15 pixels, and misrecognitions happened.

C. Summary

Aging society becomes a major trend in most of countries around the world. A large number of elderly people are living alone. To provide better care services to those elderly people, activity recognition and fall detection become more and more important for caregivers.

Based on a thermopile imaging array sensor, HTPA32x32dR1L5.0/0.85F7.7eHiC, the proposed activity recognition system is expected to help caregivers to determine the activities of elderly people in the experiment area, which is 6-meters long and 4-meters wide.

The outputs of the sensor are thermal images with a resolution of 32×32 pixels. The key advantages to use the thermopile imaging array sensor are: i) because the resolution of the sensor is much lower than that of video cameras, it almost has no issue of privacy invasion; and ii) because the sensor is designed to detect the infrared radiations, its performance is stable no matter in dark or bright environments.

The activity recognition system developed by this research recognizes the activities of the human target according to the following procedure: first, the system determines whether the human target is within the experiment area; if the human target is within the experiment area, the location of the target will be identified, and three kinds of activities will be determined.

This proposed system has achieved following functions:

1. Capture data output from the thermopile image sensor.

   The HTPA32x32d thermopile array sensor captures thermopile images with a resolution of 32×32 pixels and sends the images to a Raspberry Pi 3 through an I2C interface. Then the Raspberry Pi 3 calibrates the thermopile images and saves the calibrated images into an SD card. Finally, the images stored in the SD card are transferred to a PC and processed with MATLAB codes to determine the existence and the location of the human target.

2. Locate the feature point (the head of the human target) in the experiment area and track its change or show that it is “off -room”.

   The image processing operations implemented by the MATLAB codes include (1) image preprocessing; (2) connected domain extraction; (3) feature point localization; and (4) activity recognition.

   By doing the image preprocessing, the fixed noises will be eliminated. In addition, through the connected domain extraction and the feature point localization, the head of the human target, defined as the feature point P(n), is located in feature domain F(n). If the feature point cannot be found in 10 consecutive frames, the system shows “off–room” to indicate that the human target is not in the experiment area.

3. Determine three different activity statuses for the human target, “quiescence”, “movement”, and “fall”, and send out an alert when the fall status is determined. After the location of the feature point is identified in the monitoring area. The coordinates of P(n) will be tracked continuously. Based on the change of the coordinates of P(n), the proposed system can identify the three activity status of the human target. According to the test results, the proposed system can achieve 100% accuracy to determine the quiescence status. If the human target moves, the accuracy to determine the movement status is 91%, which is lower than the accuracy for the quiescence status but still acceptable. All the misrecognitions are recognizing the movement status as the fall status. The main reason for the misrecognition is that the action of squatting down is quite similar to falling down, which increases the challenge to accurately determine the movement status. Fall detection is the key function of the proposed activity recognition system. According to the test results, 97% of fall activities are recognized correctly. In addition, if the human target is within 4 meters from the thermopile array sensor, the accuracy rate for the fall detection can reach to 100%. The misrecognitions only happened when the human target is more than 4 meters from the thermopile array sensor. As discussed in the previous section, the main reason is that once the

human target is far away from the sensor, the size of the human target becomes small, and the temperature of the human target detected by the sensor becomes low and close to the environment temperature. As a result, the human target is representing by less and less pixels. During the fall down, the feature point may not have a significant change in vertical direction. In another word, the threshold of 15 pixels for the vertical displacement is no longer the optimal value, and it is better to find an alternative threshold to determine the displacement. However, an accuracy of 97% for fall detection is still a very good performance. Therefore, the system we proposed is a promising technology to be used to serve elderly people.

D. Future Work

Although the overall accuracy of proposed activity recognition system reaches more than 95%, there is still a large space to improve the performance of the system. The majority improvements can be done in three areas.

1. Improving accuracy;

   How to improve the accuracy on distinguishing the “movement” status and the “fall” status will be the major work in future work.

   An idea to avoid misrecognizing the movement status as the fall status is to check more frames after detecting a significant vertical displacement. If there is a negative vertical displacement, the activity is more likely to be squatting down and standing up, not a falling down.

   To avoid the misrecognitions for the fall status, as discussed before, the key point is to find out an adaptive threshold to determine the significant vertical displacement. In other words, when the human target is close to the thermopile array sensor, the threshold value should be large, and when the human target is far away from the sensor, the threshold should be small. In addition, it is necessary to distinguish fall activity and lay down activity, which is very similar in the captured thermal images.

2. Real-time implement.

   If the proposed activity recognition system is going to be used in real life, to make the system having real-time image processing ability is very important. Once the human target is falling down, the system should be able to detect the fall activity and send out an alert immediately. As discussed the previous chapter, in this research, all the image data captured from the thermopile array sensor are saved in a SD card, and then are transferred to a PC for analysis in a non-real time manner. In the future work, through the Ethernet or WIFI interfaces of Raspberry Pi 3, the data captured from the thermopile array sensor can be transferred to a server for image processing in a real time manner. The detection results can be immediately sent to caregivers for better service for elderly people.

E. Conclusions

In this research, an indoor human body tracking and activity recognition system with a low resolution thermopile array sensor is proposed. The sensor is installed on a wall to observe the human target within a 6 meters long and 4-meter-wide experiment area. The images captured from the thermopile array sensor are calibrated by a Raspberry Pi 3, saved into a SD card, and then transferred to a PC for image processing. In the PC, all images are processed by a group of newly developed algorithms to detect off-room, quiescent, movement or fall. Experiments show that the proposed algorithms can be used to determine whether the target is active in the experiment area and identify a fall. The overall accuracy for the detection results can reach to 95%. Therefore, the system we proposed is a promising technology to be used to serve elderly people.

REFERENCES

Alshurafa, N., Xu, W., Liu, J., Huang, M.-C., Mortazavi, B., Roberts, C. K., & Sarrafzadeh, M. (2013). Designing a robust activity recognition framework for health and exergaming using wearable sensors. *IEEE Journal of Biomedical and Health Informatics, 18*(5), 1636 - 1646. doi:10.1109/JBHI.2013.2287504

Ashbrook, D., & Starner, T. (2003). Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing, 7*(5), 275-286. doi:10.1007/s00779-003-0240-0

Bao, L., & Intille, S. (2004). Activity recognition from user annotated acceleration data. In A. Ferscha, & F. Mattern (Eds.), *Lecture notes in computer science: Vol. 3001. Pervasive computing* (pp. 1-17). Springer. doi:https://doi.org/10.1007/978-3-540-24646-6_1

Bloom, D., Canning, D., & Lubet, A. (2015). Global population aging: Facts, challenges, solutions & perspectives. *Daedalus, 144*(2), 80-92. doi:https://doi.org/10.1162/DAED_a_00332

Chaaraoui, A., Padilla-Lopez, J. R., Climent-Perez, P., & Florez-Revuelta, F. (2013). Evolutionary joint selection to improve human action recognition with RGB-D devices. *Expert Systems with Applications, 41*(3), 786-794. doi:https://doi.org/10.1016/j.eswa.2013.08.009

Chakravoty, P. (2018). What is a signal? *IEEE Signal Processing Magazine, 35*(5), 175-177. doi:10.1109/MSP.2018.2832195

Chen, L., Nugent, C., & Wang, H. (2012). A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering, 24*(6), 961-974. doi:10.1109/TKDE.2011.51

Chen, W., & Ma, H. (2015). A fall detection system based on infrared array sensors with tracking capability for the elderly at home. *2015 17th International Conference on E-health Networking, Application & Services (HealthCom), Boston, USA*, 428-434. doi:10.1109/HealthCom.2015.7454538

Cheng, J., Amft, O., Bahle, G., & Lukowicz, P. (2013). Designing sensitive wearable capacitive sensor for activity recognition. *IEEE Sensors Journal, 13*(10), 3935-3947. doi:10.1109/JSEN.2013.2259693

Durrant-Whyte, H., Roy, N., & Abbeel, P. (2012). *Robotics: science and systems VII-----Lying pose recognition for elderly fall detection.* MIT Press.

Fahim, M., Fatima, I., Lee, S., & Park, Y.-T. (2013). EFM: Evolutionary fuzzy model for dynamic activities recognition using a smartphone accelerometer. *Applied Intelligence, 39*, 475-488. doi:https://doi.org/10.1007/s10489-013-0427-7

Gerling, K., Livingston, I., Nacke, L., & Mandryk, R. L. (2012). *Full-body motion-based game interaction for older adults.* Austin, US. doi:10.1145/2207676.2208324

Gonzalez-Ortega, D., Diaz-Pernas, F. J., Martinez-Zarzuela, M., & Anton-Rodriguez, M. (2013). A kinect-based system for cognitive rehabilitation exercises monitoring. *Computer Methods and Programs in Biomedicine, 113*, 620-631. doi:https://doi.org/10.1016/j.cmpb.2013.10.014

Han, J., Shao, L., Xu, D., & Shotton, J. (2013). Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Transactions on Cybernetics*, 1318-1334. doi:10.1109/TCYB.2013.2265378

Harms, H., Amft, O., Roggen, D., & Troster, G. (2007). SMASH: A distributed sensing and processing garment for the classification of upper body posters. *The Third International Conference on Body Area Networks*. doi:10.4108/ICST.BODYNETS2008.2955

Heilman, K., & Porges, S. (2007). Accuracy of the LifeShirt®(Vivometrics) in the detection of cardiac rhythms. *Biological Psychology*, *75*(3), 300-305. doi:10.1016/j.biopsycho.2007.04.001

Heimann. (2017). *HEIMANN*. Retrieved from http://www.heimannsensor.com/Datasheets/Overview-HTPA32x32d_Rev3.pdf

Hong, J., Tomii, S., & Ohtsuki, T. (2013). Cooperative fall detection using Doppler radar and array sensor. *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC).*doi:10.1109/PIMRC.2013.6666753

Huang, J.-D. (2011). Kinerehab: a kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. *The 13th International ACM SIGACCESS Conference on Computers and Accessibility*. doi:https://doi.org/10.1145/2049536.2049627

Hwang, J., Kang, J., Jang, Y., & Kim, H. C. (2004). Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly. *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA.* doi:10.1109/IEMBS.2004.1403643

Kantoch, E., & Augustyniak, P. (2012). Human activity surveillance based on wearable body sensor network. *Computing in Cardiology (CinC).* https://www.researchgate.net/ publication/261211861_Human_activity_surveillance_based_on_wearable_body_sensor_network

Kern, N., Schiele, B., Junker, H., Lukowisz, P., & Troster, G. (2003). Wearable sensing to annotate meeting recordings. *Personal and Uboquitous Computing*, 263-274. doi:10.1109/ISWC.2002.1167247

Khan, A. M., Lee, Y.-K., Lee, S., & Kim, T.-S. (2010). A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE Transactions on Information Technology in Biomedicine, 14*(5), 1166-1172. doi:10.1109/TITB.2010.2051955

Laerhoven, K. V., Aidoo, K., & Lowette, S. (2001). Real-time analysis of data from many sensors with neural networks. *Proceedings of the Fifth International Symposium on Wearable Computers*. doi:10.1109/ISWC.2001.962112

Lara, O. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. *Commun. Sury. Tutorials., 15*(3), 1192-1209. doi:10.1109/SURV.2012.110112.00192

Lasecki, W., Song, Y., Kautz, H., & Bigham, J. P. (2013). Real-time crowd labeling for deployable activity. *Computer supported cooperative work*, (pp. 1203-1212). doi:https://doi.org/10.1145/2441776.2441912

Lawrence, E., Sax, C., Navarro, K. F., & Qiao, M. (2010). Interactive games to improve quality of life for the elderly: Towards integration into a WSN monitoring system. *2010 Second International Conference on eHealth, Telemedicine, and Social Medicine, St. Maarten, Netherlands Antilles.* doi:10.1109/eTELEMED.2010.21

Lee, S.-W., & Mase, K. (2002). Activity and location recognition using wearable sensors. *IEEE Pervasive Comuting, 1*(3), 24-32. doi:10.1109/MPRV.2002.1037719

Leonhardt, U., & Magee, J. (1998). Multi-sensor location tracking. *The 4th ACM/IEEE Int'l Conf. Mobile Comp. Net.* doi:https://doi.org/10.1145/288235.288291

Li, Y., Zeng, Z., Popescu, M., & Ho, K. (2010). Acoustic fall detection using a circular microphone array. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina*. doi:10.1109/IEMBS.2010.5627368

Liao, L., Patterson, D., Fox, D., & Kautz, H. (2004). Learning and inferring transporation routines. *Artificial Intelligence, 171*(5-6), 311-331. doi:https://doi.org/10.1016/j.artint.2007.01.006

Lin, W., Sun, M.-T., Poovandran, R., & Zhang, Z. (2008). Human activity recognition for video surveillance. *2008 IEEE International Symposium on Circuits and Systems, Seattle, WA, USA*. doi:10.1109/ISCAS.2008.4542023

Liu, T., Guo, X., & Wang, G. (2012). Elderly-falling detection using distributed direction-sensitive pyroelectric infrared sensor arrays. *Multidimensional Systems and Signal Processing*, 451-467. doi:https://doi.org/10.1007/s11045-011-0161-4

Lukowicz, P., Ward, J., Junker, H., Stäger, M., Tröster, G., Atrash, A., & Starner, T. (2004). Recognizing workshop activity using body worn microphones and accelerometers. In A. Ferscha, & F. Mattern (Eds.), *Lecture notes in computer science*: Vol. 3001. *Pervasive computing* (pp. 18-32). Springer. doi:https://doi.org/10.1007/978-3-540-24646-6_2

Mantyjarvi, J., Himberg, J., & Seppanen, T. (2001). Recognizing human motion with multiple acceleration sensors. *2001 IEEE International Conference Systems, Man. and Cybernetics*. doi:10.1109/ICSMC.2001.973004

Mashiyama, S., Hong, J., & Ohtsuki, T. (June 2015). Activity recognition using low resolution infrared array sensor. *2015 IEEE International Conference on Communications (ICC), London, UK*. doi:10.1109/ICC.2015.7248370

Moler, C. (2017). *The origins of MATLAB.* Retrieved from https://www.mathworks.com/company/newsletters/articles/the-origins-of-matlab.html

Mozer, M. (1998). The neural network house: An environment that adapts to its inhabitants [Presentation]. *The AAAI Spring Symposium on Intelligent Environment, Menlo Park, CA.*

Najafi, B., Aminian, K., François, L., Blanc, Y., & Robert, P. A. (2002). Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *IEEE Transactions on Biomedical Engineering, 49*(8), 843-851. doi:10.1109/TBME.2002.800763

Nghiem, A., Auvinet, E., & Meunier, J. (2012). Head detection using Kinect camera and its application to fall detection. *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA).* IEEE. doi:10.1109/ISSPA.2012.6310538

Ordonez, F. J., & Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors, 16*(1), 115. doi:https://doi.org/10.3390/s16010115

Paragiola, G., & Coronato, A. (2013). Intelligent monitoring of stereotyped motion disorders in case of children with autism. *2013 9th International Conference on Intelligent Environments.* doi:10.1109/IE.2013.12

Patterson, D., Liao, L., Fox, D., & Kautz, H. (2003). Inferring high-level behavior from low-level sensors. *The 5th Conference on Ubiquitous Computing (UbiComp).* doi:10.1007/978-3-540-39653-6_6

Popoola, O., & Wang, K. (2012). Video-Based abnormal human behavior recognition—A review. *IEEE Transactions on Systems, Man, and Cybernetics, 42*(6), 865-878. doi:10.1109/TSMCC.2011.2178594

Preis, J., Kessel, M., Werner, M., & Linnhoff-Popien, C. (2012). Gait recognition with Kinect. *First Workshop on Kinect in Pervasive Computing.*

Qian, G., Peng, B., & Zhang, J. (2012). Gesture recognition using video and floor pressure data. *2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA.* doi:10.1109/ICIP.2012.6466823

Roshtkhari, M., & Levine, M. (2013). Human activity recognition in videos using a single example. *Image and Vision Computing, Montreal, Canada.* doi:10.1016/j.imavis.2013.08.005

Ryoo, M. S. (2011). Human activity prediction: Early recognition of ongoing activities from streaming videos. *Proceedings / IEEE International Conference on Computer Vision. IEEE International Conference on Computer Vision.* IEEE. doi:10.1109/ICCV.2011.6126349

Schmidt, A., Beigl, M., & Gellersen, H. (1999). There is more to context than location. *Computers and Graphic*, *23*(6), 893-901. doi:https://doi.org/10.1016/S0097-8493(99)00120-X

Schmidt, A., & Laerhoven, K. v. (2001). How to build smart appliances? *IEEE Personal Communications, Special Issue on Pervasive Computing*, *8*(4), 66-71. doi:10.1109/98.944006

Sim, S., Jeon, H. S., Chung, G. S., Kim, S. K., Kwon, S., Lee, W. K., & Park, K. (2011). Fall detection algorithm for the elderly using acceleration sensors on the shoes. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA*. doi:10.1109/IEMBS.2011.6091223

Stone, E., & Skubic, M. (2012, May). Passive, in-home gait measurement using an inexpensive depth camera: Initial results. *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops.* doi:10.4108/icst.pervasivehealth.2012.248731

Su, X., Tong, H., & Ji, P. (2014). Activity recognition with smartphone sensors. *Tsinghua Sience and Technology, 19*(3), 235-249. doi:10.1109/TST.2014.6838194

Sung, M., DeVaul, R., Jimenez, S., Gips, J., & Pentland, A. (2004). A shiver motion and core body temperature classification for wearable soldier health monitoring systems. *IEEE International Symposium on Wearable Computers*. doi: 10.1109/ISWC.2004.39

Tao, S., Kudo, M., Nonaka, H., & Toyama, J. (2011). Recording the activities of daily living based on person localization using an infrared ceiling sensor network. *2011 IEEE International Conference on Granular Computing*. doi:10.1109/GRC.2011.6122673

United Nations. (2017). World population aging (ST/ESA/SER.A/408). Department of Economic and Social Affairs Population Division. http://www.un.org/en/development/desa/ population/publications/pdf/ageing/WPA2017_Report.pdf

Vo, Q., Lee, G., & Choi, D. (2012). Fall detection based on movement and smart phone technology. *2012 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future.* doi:10.1109/rivf.2012.6169847

Yoshimitsu, K., Muragaki, Y., Maruyama, T., Yamato, M., & Iseki, H. (2013). Development and initial clinical testing of "OPECT": An innovative device for fully intangible control of the intraoperative image-displaying monitor by the Surgeon. *Neurosurgery, 10*(1). doi:10.1227/NEU.0000000000000214

Zheng, J., Zhang, G., & Wu, T. (2009). Design of automatic fall detector for elderly based on triaxial accelerometer. *2009 3rd International Conference on Bioinformatics and Biomedical Engineering, Beijing, China*. doi:10.1109/ICBBE.2009.5162596

APPENDIX A

MATLAB CODE

```
clear all;
close all;
%% load data
load 'DaisyTwoMeters.txt';
OriginalData=DaisyTwoMeters;
%OriginalData(:,33)=[];
%% data analysis
[m,~]=size(OriginalData);
FrameNum=m/32;
AverageData=zeros(32,32); %findout average data for numbers of Frames;
for i=1:FrameNum
    %% Save No.i frame data into Data{i,1}
    temp=OriginalData((i-1)*32+1:i*32,1:32);
    %A=flipud(temp)';
    Data{i,1}=flipud(temp)';
    AnalysisResult(i,1)=min(min(Data{i,1}));
    AnalysisResult(i,2)=mean(mean(Data{i,1}));
    AnalysisResult(i,3)=max(max(Data{i,1}));
    AverageData=AverageData+Data{i,1};
end
AverageData1=AverageData/FrameNum;

for i=1:FrameNum
    %% Every Frame data - AverageData
```

```
    % need revise
    Data_step1{i,1}=Data{i,1}-AverageData1;
    Data_step2{i,1}=(Data_step1{i,1}+abs(Data_step1{i,1}))/2;
end

%% Show figure Result
[a,b]=meshgrid(linspace(1,32,32));
figure();set(gcf,'color','w');
set(gcf, 'position', [0 0 850 400]);
for i=10:FrameNum

    subplot(2,2,1); view([32 32 40]);%
    pcolor(a,b,Data_step1{i}');caxis([3 20]);
    %[120 220]
    %mesh(Data_step1{i}');
    %view([-27,65]);
    title('Original Figure')
    xlabel('x');ylabel('y');zlabel('z');
    %axis([1 32 1 32 -50 100]);

    subplot(2,2,2);

    Threshold=mean(Data_step2{i}(Data_step2{i}~=0));  % Setup Threshold Value, need revise
    BW=double(Data_step2{i}>Threshold);  % Binaryzation Number£¬Save to BW
    % BW=Data_step1{i}; %without binaryzation;
    [L,NUM]=bwlabel(BW,8);   % connceted components calculation,8  bwlabel(BW,4) 4
    for j=1:NUM
```

```
    %% Count numbers of Connected Domain
    num(j,1:2)=[j,length(find(L==j))];
end
for k=1:NUM  %show top three largest connceted area edge
    [x,y]=find(L==num(k,1));
    ConnectedArea=0;
    for h=1:num(k,2);
        ConnectedArea(h,1)=Data{i}(y(h),x(h));
        SumTemp(k,1)=sum(ConnectedArea);
    end
end
for j=1:NUM
    %% Count numbers of Connected Domain
    num(j,1:3)=[j,length(find(L==j)),SumTemp(j,1)];

end

num=sortrows(num,-3);   %Sort from the largest to the smallest second column

head=[0 0];%
for k=1:min(NUM,3)  %show top three largest connceted area edge
    [x,y]=find(L==num(k,1));
    plot(x,y,'sk','MarkerSize',5);
    axis([1 32 1 32]);drawnow
    hold on;

    %%
```

```
    break_=0;%·
    temp_y=unique(y);
    temp_y=sort(temp_y,'descend');

    for l=1:length(temp_y)
        if break_==1
            break;
        end
        %%
        temp_x=x(y==temp_y(l));%
        if length(find(diff(temp_x)==1))<2
            %
            continue;
        end

        for m=1:length(temp_x)-2
            if length(find(diff(temp_x(m:m+2))==1))==2
                break_=1;
                %%
                head(k,:)=[temp_x(m+1),temp_y(l)];
                break;
            end
        end
    end
end
%%
```

```
    head=sortrows(head,-2);
    AnalysisResult(i,6)=head(1,1);
    AnalysisResult(i,7)=head(1,2);
    AnalysisResult(i,8)=Data{i}(head(1,1),head(1,2));
    plot(head(1,1),head(1,2),'r*');%
    text(head(1,1),head(1,2),[' head  (' num2str(head(1,1)) ','  num2str(head(1,2)) ')'])
    title(['Frame No. ', num2str(i),]);
    hold off
    pause(0.5);

   subplot(2,2,3);

   AnalysisResult(i,9)=(220-AnalysisResult(i,8))/10;
   AnalysisResult(i,10)=AnalysisResult(i,7)*AnalysisResult(i,9)/(sqrt(3)*32)+(1.8*sqrt(3)-
AnalysisResult(i,9))/sqrt(3);

End
```