CHALMERS

Pingu: Another User-friendly Interface Device

Multi-sensor Based Gestural Interaction for Smart Home Environments

Master of Science Thesis in the Master Degree Program Communication Engineering

AYDIN GORKEM ISITAN

Department of Signal and Systems
Communication Systems Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2012
Report No. EX070/2012

REPORT NO. EX070/2012

Pingu: Another User-friendly Interface Device

Multi-sensor Based Gestural Interaction for Smart Home Environments

AYDIN GORKEM ISITAN

Examiner: Fredrik Brännström

Department of Signal and Systems
Communication Systems Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2012

Pingu: Another User-friendly Interface Device
Multi-sensor Based Gestural Interaction for Smart Home Environments

AYDIN GORKEM ISITAN

© AYDIN GORKEM ISITAN, 2012

Technical report no: EX070/2012
Department of Signal and Systems
Communication Systems Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2012

SE-41296 Göteborg
Sweden
Telephone: +46 (0) 31 – 772 1000

ABSTRACT

Smart homes improve the quality of life through integration of different electronic systems enabling simpler control of the house. The number of electronic devices in our homes is huge and it is still increasing rapidly. However, control systems are not that “smart” and still consist of traditional remote controllers. Under these circumstances, the user needs more efficient, more intuitive and more user-friendly interface to control these devices in smart home environment. The trials to substitute usual remote controllers do not yield results as expected. Proposed substitute systems either come across with technical problems or limit the user in his freedom, in other words, lead to social problems. Therefore, the need for a user-friendly interaction method for home environments directs us to seek another interaction method which is reliable and natural in a way that it can be integrated into the user’s daily life.

In this thesis, we have developed a multi-sensor based interaction system with a small multi-sensor based device called SHAKE, Sensing Hardware Accessory for Kinaesthetic Expression, to interact with other devices in smart home environment. Carrying SHAKE device and using it to perform gestures may be cumbersome and socially awkward. Consequently, we were looking for a way to render the device more attractive, socially acceptable and make its use subtle. Thus, we have adapted SHAKE into a ring-shaped device what we call as “Pingu”. Although a variety of distinct forms of wearable devices is available such as watch form, glove form and cloth form, none of them is as socially acceptable and practical as a ring. Pingu is capable of interacting even when the device is not in the vicinity of the user. To explore about the usability and practicability, we performed a user study. We have not fully analyzed the data yet but participants’ feedback show that our technique is applicable, practical and not awkward. In our design, the user can achieve natural interaction without caring the view of camera or suffering from lighting conditions or external noise.

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my examiner Assist. Prof. Fredrik Brännström from Chalmers University of Technology for giving me the opportunity to work in this thesis as well as his valuable advice and understanding throughout the project.

A very special thanks goes to Dr. Hamed Ketabdar from Quality and Usability Lab of Deutsche Telekom T-labs for his guidance in every technical and theoretical aspect and his faith in me completing this work successfully.

It is a pleasure to thank to Peyman Moghadam, Aarti Munjal and Wanlu Sun for their valuable feedbacks and contributions to overcome problems encountered during this process.

Further thanks goes to all my colleagues in T-Lab, Heng-Wei Chang, Mohammad Mohebbi and Seyed Sadra Amirkhalili for providing such a friendly and cheerful research atmosphere.

I would like to thank my brother Kemal Kagan Isitan for proofreading initial thesis draft.

I also want to express my appreciation to Elif Odabas for her persistent confidence in me and support while writing this thesis. Finally, I thank to my parents, Ali and Hedise Isitan for their support and giving me the chance to achieve a higher level of education.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF ABBREVIATIONS

MEMS	Micro Electro Mechanical Sensors
IR	Infrared
IR-LED	Infrared Light Emitting Diode
DTW	Dynamic Time Warping
EU	European Union
OS	Operating System
HCI	Human Computer Interaction
SHAKE	Sensing Hardware Accessory for Kinaesthetic Expression
USB	Universal Serial Bus
UI	User Interface
PDA	Personal Digital Assistant
QTJava	Quicktime for Java
JDK	Java Development Kit

1 INTRODUCTION

People today have more stuff, less time than ever and even less time to spend with each other. People work full time, sometimes working overtime and trying to spare time for activities after work. Coming home tired after a long day, people demand comfort at their house; in case there are house-keeping activities, they want to handle those activities in short time with minimal effort as possible. Apart from this, life expectancy has been growing steadily and birth rates have been decreasing in Europe leading to an ageing society. “It is a common understanding that population ageing, along with the increasing survival rates from disabling accidents and illnesses, will lead to an increase in the proportion of the population with impairments, disabilities or chronic illnesses.” [1] “Improving the quality of life for disabled and the increasing proportion of elderly people is becoming a more and more essential task for today’s European societies” [2], where the percentage of people over 65 years of age is due to rise to 20% by the year 2020 [3]. Smart homes improve the quality of life through integration of different electronic systems enabling simpler control of the house. For instance, appliances such as heating or lighting systems in home can be controlled easily from both inside and outside according to chosen criteria or given commands. Smart homes are one of the most important steps to overcome these comfort issues and improve the quality of life.

1.1 Smart Home Systems

Technology is being used more to make peoples life easier. Coffee machines, food processors, televisions, multimedia systems, DVD players, refrigerators, washing machines, dish washers, cell phones, adjustable lamps, air conditioners and several other devices has become part and parcel of life. In parallel with these developments, not only the quality of life improves, but also the concept of comfort changes. Remote controls for televisions and DVD players, timers for washing and coffee machines have evolved with the developing technology. Just like us, our homes also get their share from rapid technological developments. Once what we think is a fiction turns out to be a reality and the concept of “smart home” becomes no longer impossible. The next step of this process provides the ability to control the home and devices from a single point. Smart Home is the term commonly used to define a residence that responds to the needs of its residents, ease their life by providing more secure, more comfortable and economical life, processes an order in accordance with the given commands. Initial implementations were targeting comfort issues instead of focusing on assisted living systems to support older or disabled people living independently at home [4]. In the present day, systems have been focusing on assisting elderly and disabled people along with comfort. For instance, current systems provide control mechanisms for lighting, curtains, and multimedia, as well as emergency assistance to keep disabled or elderly people under constant surveillance.

1.2 Taxonomy of Smart Homes

The concept "smart" is being used enormously and there are a number of different types of smart homes due to the increase in technological developments. In an attempt to understand the smart homes in a better way, they can be classified under technological and functional perspectives.

1.2.1 Technological Perspective

Boguslaw Pilich [5] has developed the following classification in his work "Engineering Smart Houses" in order to establish a common basis for studying smart home technology. This perspective takes into account the technologic requirements of each kind of houses with the improvement that each type of houses brings to people's lives.

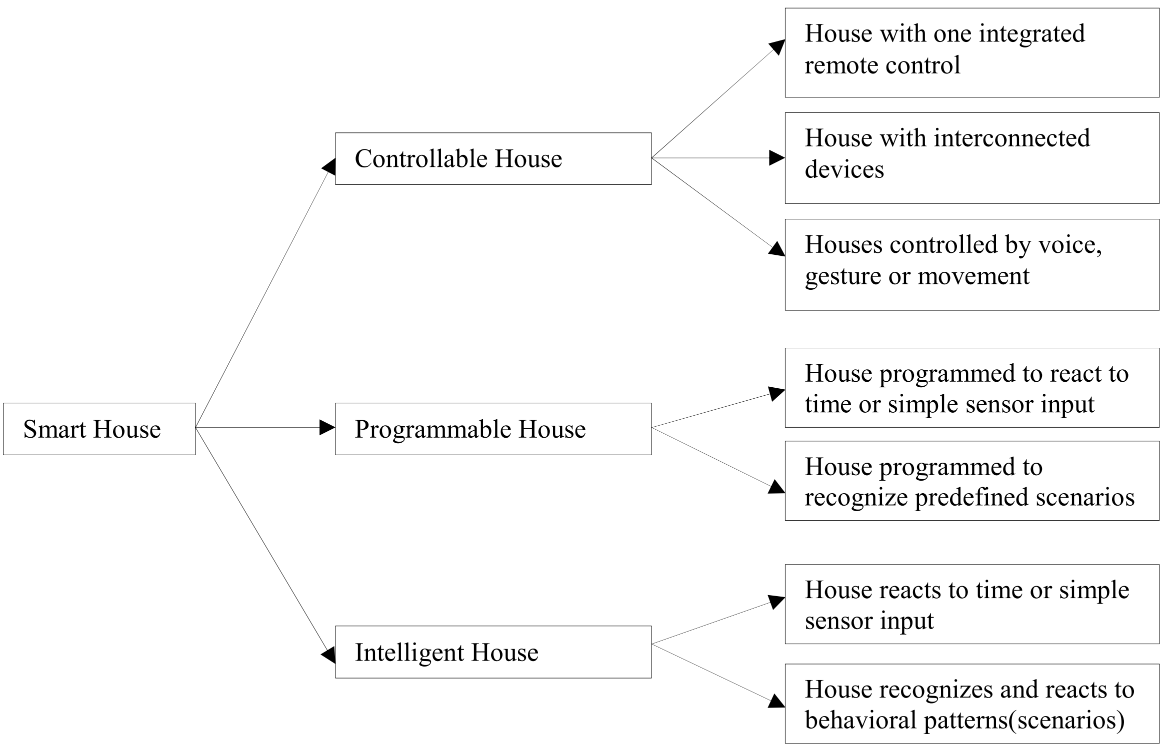


Figure 1.1. Smart home classification [5]

As shown in Figure 1.1, smart homes can be divided into three main categories:

1. **Controllable Houses**

   In controllable houses, residents can control existing devices and systems more efficiently and easier than in normal contemporary houses with various control systems. Such houses consist of three subclasses:

   i. **Houses with one integrated remote control**

      In such houses, all electronic devices are controlled from a single remote control. The remote control can be a piece of jewellery, or something larger or more functional. An example of this technology is an integrated remote controller for the VCR and TV.

   ii. **Houses with interconnected devices**

      Different electronic devices like TV sets, component stereos, video recorders, computers, additional speakers or cameras are connected with each other. In this way, data transfer and synchronization between the devices can be achieved.

   iii. **Houses controlled by voice, gesture or movement**

      Such houses are similar to the first subgroup but they are equipped with an invisible control system which is sensitive to voice, movement or gesture through integrated sensors. Rather than hardware, software is the key element in this design since the voice or gesture recognition performance needs to be really reliable. For example, clapping hands to turn the television on or the voice dial functionality of modern phones.

2. **Programmable Houses**

   Programmable houses are more advanced than controllable houses and these houses consist of two subclasses:

   i. **Programmable Houses reacting to time and simple sensor input**

      In such houses, the system and devices can be programmed and react accordingly. As an example, systems or devices can be automatically turned on and off at a predefined time or change their behaviors after predefined intervals. Similarly, temperature sensors or light sensors can be used to switch the state of the devices. Highly reliable sensors are already available on the market and used in several home environments.

   ii. **Programmable Houses assessing and recognizing situations**

      Imagine that the resident returns home being tired after long hard work, hoping to take a nap. Automatically house could turn off the lights and play some soft music on the background until it fades out after some time. As in this example, these types of houses are capable of recognizing simultaneous input from several sensors as a particular scenario. Various scenarios have to be defined and programmed in advance. However, the system has to be reprogrammed and the scenarios have to be redefined every time some changes occur. At the moment, these types of houses are the most advanced ones in existing technology.

3. **Intelligent Houses**

   These types of houses are similar to the programmable houses but they are more advanced. In programmable houses, scenarios are prepared by residents, whereas in intelligent houses such programming does not exist since these houses are capable of learning. Therefore, these houses examine the residents in their daily life, searching for repeated actions and evaluate the movements of the residents. After a pattern has been identified, houses create their own scenarios so that the next time the scenario is recognized, the house automatically switches on or off certain equipment. However, the process ignores complexity of human psychology where humans do not always react the same way.

1.2.2 Functional Perspective

Aldrich focuses on functionality available to the user since smart home technology is changing rapidly and proposes five classes of smart homes [6]:

1. **Homes which contain intelligent objects**

   These homes contain single standalone applications and objects which function in an intelligent manner.

2. **Homes which contain intelligent, communicating objects**

   These homes contain appliances and objects which function intelligently in their own right and which also exchange information between one another to increase functionality.

3. **Connected homes**

   These homes have internal and external networks, allowing interactive and remote control of systems, as well as access to services and information, both within and beyond the home.

4. **Learning homes**

   In these homes, patterns of activity are recorded and the accumulated data are used to anticipate users' needs and to control the technology accordingly.

5. **Attentive homes**

   In these homes, the activity and location of people and objects are constantly registered, and this information is used to control technology in anticipation of the occupants' needs.

1.3 Motivation

Most of us are surrounded by electronic devices in our homes such as computers, televisions, dvd players, tablets, cell phones, air conditioners and many others. The number of electronic devices in our homes is huge and it is still increasing rapidly. Under these circumstances, the user needs more efficient, more intuitive and more user-friendly interface to control these devices in smart home environment. One of the existing systems is based on speech recognition resulting in voice controlled smart home environment. Such systems provide the user the ability to be able to communicate with electronic devices in a natural way. Such a system was developed in a European Union (EU) - funded Information Society Technologies project INSPIRE [7] which allows users to control appliances like TV, lamps or video recorder via speech. Nevertheless, these systems that are on the market are insufficient because these systems are lacking reliability and require certain patterns like saying a command very loudly. Besides speech, gesture is another natural way to express ourselves and that is how we interact in the real world. The uses of gestures to control appliances in home environment are also investigated in several studies such as Starner et al [8]. They have introduced a camera-based wearable device with infrared (IR) transmitter called Gesture Pendant. The main idea behind this project is to allow users to control appliances in the house via hand motions. Hand motions are recognized with a camera and interpreted as gestures then the corresponding commands are executed to control the operation of various household devices such as room lighting, opening and closing window blinds or controlling audio and video equipment. This approach sounds promising but the camera is sensitive to ambient IR light and does not work in sunlight hence the system needs to be improved.

In this work, we used a small multi-sensor based device called "SHAKE", Sensing Hardware Accessory for Kinaesthetic Expression, to interact with other devices in smart home environment. Later, we have developed the project with a ring-shaped wearable SHAKE called "Pingu". In our design, the user can achieve natural interaction without caring the view of camera or suffering from lightning conditions or external noise.

1.4 SHAKE and Pingu

We implemented our design using a multi-sensor device called SHAKE, equipped with 3-axis accelerometer, 3-axis magnetometer, 3- axis angular rate sensor (gyro sensor), a vibrating motor, a navigation bar and a LED light providing visual feedback of the state of the device for the user. Alongside of USB connection, SHAKE also provides Bluetooth connection to interact with various devices that are equipped with Bluetooth.

The result of the designed system was satisfying enough to push us to make further improvements. We decided to transform our control device into a ring shaped SHAKE, what we call "Pingu" which is more attractive and socially accepted. We conducted a user study to see the acceptability and usability of this technique before developing smart home prototype applications. Pingu prototype looks like an electronic device rather than a normal ring. Therefore, the ultimate goal is to end up with a normal ring which is socially acceptable.

1.5 Organization of Thesis

Chapter 1 introduces the concept "Smart Home" and its classifications, and the motivation behind this research work.

Chapter 2 contains a brief analysis of the need for a more user friendly system solution and gives the detailed problem description.

Chapter 3 presents the literature review regarding gesture taxonomy and gesture recognition techniques. We focus on concepts and definitions related to gesture and gestural interaction. Section 3.1 introduces various definitions and classifications of gesture. In section 3.2, several gesture recognition techniques are discussed. Section 3.3 is devoted to the theory of recognition algorithm used in our project.

Chapter 4 presents the smart-home system design which consists of hardware and software components and gives an overview of the proposed system. We first present a brief summary of the device used in Section 4.1. Section 4.2 discusses the motivation to improve the device for better usability. We continue to present the improved device called “Pingu” in Section 4.3. Lastly, the process of gesture recognition is described in Section 4.4.

Chapter 5 shows how the gestural interaction by Pingu can be applied to control home electronics such as television, lights and ventilation system. Section 5.1 gives insight into the design and the platform. Section 5.2 demonstrates the basic functions or commands supported in the system. Section 5.3 provides information about different demos developed by us.

Chapter 6 summarizes the outcome of this thesis work and discusses potential usage areas and future steps to improve the system.

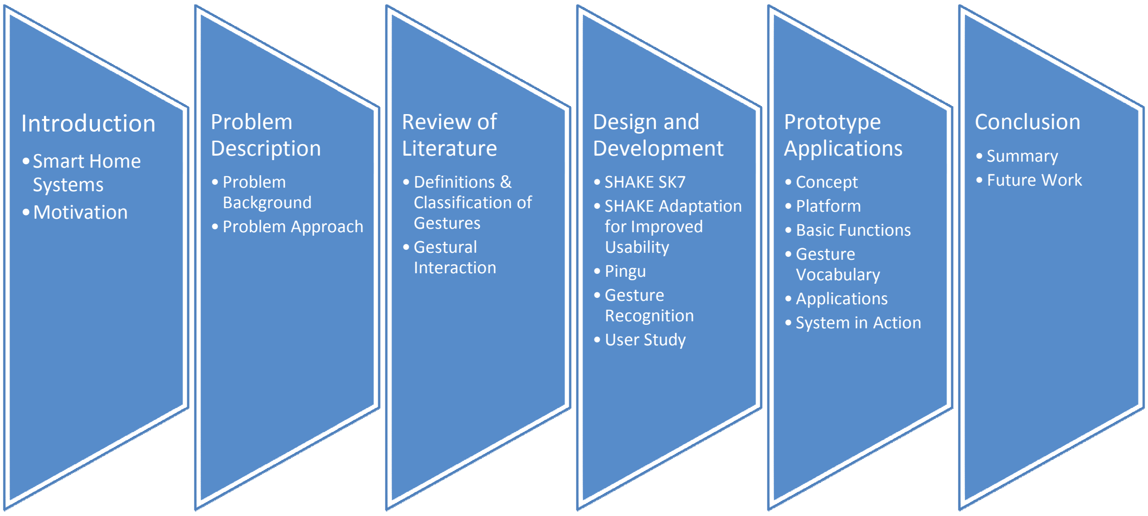


Figure 2.1. Outline of the Thesis

2 PROBLEM DESCRIPTION

2.1 Problem Background

Today, all electronic devices come with their own remote control making it difficult to keep track of which one to control which device. The user has to be in line-of-sight with the controllable appliance and use multiple remote controls, learn and remember how to use every one of them. This simply causes clutter in life. A solution could be to design a multi-device remote control by extending the number of buttons to add more functionality. However, with more keys user friendliness dramatically decreases especially for elderly people and number of malfunctions increases.

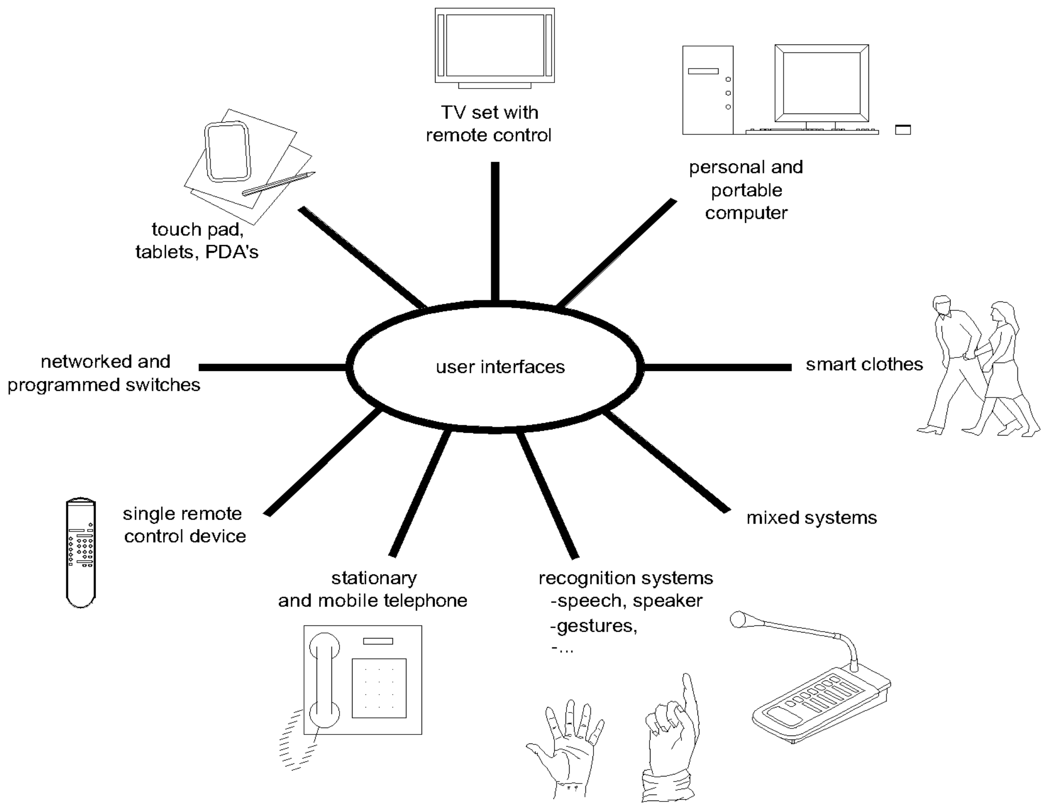


Figure 2.1. User interfaces

Numerous approaches such as the systems in Figure 2.1 are being investigated for controlling appliances in home environment as a substitute for the traditional infrared remote control system. Thanks to the success of iPhone, touchscreens became popular but they exhibit same problems as remote controller and also it is harder to learn how to use the interface especially for elderly. Besides touchscreen technology, wall panels are being used. One major drawback of this system is that the user is required to go to panel's location to interact with system.

Among all these interfaces above, speech and gesture recognition provides more intuitive and natural interaction. Speech recognition technique was used in the project INSPIRE. However, evaluation done by 24 test subjects (aged, 19 to 29) shows that the user acceptance is inadequate and the overall quality of the system was calculated as approximately 3.3 (mean) where 5 referred to excellent and 1 to bad. Speech recognition system suffers from noise sensitivity and awkward training procedures and problems arise with processing of the data from microphones. Still a successful implementation of speech based recognition system does not imply the usability and practicability of such technology in the houses. Consider that the user is not alone in the house and the residents are watching a movie or listening to music and chatting. In this case, the user may have to speak very loudly to give command to the house. Additionally, there is a problem when the residents are talking to each other, it could be very difficult to determine if a command was demanded by using a key word or if it was just an ordinary word in a conversation. Even the ambient noise may result in misrecognition and cause errors. In short, speech recognition systems have some success but it must reach to a sufficient level of reliability and usability. In noisy situations people rely more on gestural cues and this is the point where gestures come in to play.

In summary, the trials to substitute usual remote controllers do not yield results as expected. Proposed substitute systems either come across with technical problems or limit the user in his freedom, in other words, lead to social problems. Therefore, the need for a user-friendly interaction method for home environments directs us to seek another interaction method which is reliable and natural in a way that it can be integrated into the user’s daily life.

2.2 Problem Approach

We believe that a smart home will become “smarter” if the integrated interface is capable of controlling the appliances and devices in a way that we interact in real world. As discussed above, a natural, convenient and efficient way of interaction to control home appliances is missing. Therefore, our goal is to fill this gap and control the services in a smart home environment in an intuitive and natural manner. Like speech, gestures are commonly used in daily life and they are powerful means of communication among humans. Even we speak on the telephone; we gesticulate believing that the message is conveyed better. According to psychologist Albert Mehrabian [9], information conveyed in a message is mostly communicated nonverbally. His findings show that 93% of our communication (words account for 7%, tone of voice accounts for 38%, and body language accounts for 55%) is nonverbal and emotions are displayed mostly through body gestures and facial expressions. Therefore, we decided to use gestures as a tool for interacting different applications and electronic devices in a home environment.

We undertook literature review to gain insight into existing systems and gesture recognition protocols. After discussions, we came to an agreement on how the system should appear and satisfy the following conditions:

- The established interaction system must not impose restrictions on the users throughout the interaction process that is to say that the user must be able to preserve his personal freedom to do things by hand.
- The system must provide powerful human-computer interaction with high detection rates and performance as well as comfort. We agreed that the user must be able to interact with the system regardless of his position in certain area. Hence, the system should not be controlled from a fixed location.
- The use of Bluetooth infrastructure is preferred since this technology is becoming available in most devices and also solves the problem of mobility.

3 REVIEW OF LITERATURE

3.1 Definition and Classification of Gestures

Before going deeper to analyze human computer interaction (HCI) through gestures and how the gestures are recognized and processed, it is necessary to understand what gesture is and its definitions.

Gestures can be defined in several ways. For example, definitions provided by a few psychologists are as follows:

1. According to McNeill [10], gesture can be defined as movements of the arms and hands which are closely synchronized with the flow of speech.
2. Adam Kendon [11] defines gesture: "The word 'gesture' serves as a label for that domain of visible action that participants routinely separate out and treat as governed by an openly acknowledged communicative intent."

Besides these, another definition comes from biologists [12], stating, "the notion of gesture is to embrace all kinds of instances where an individual engages in movements whose communicative intent is paramount, manifest, and openly acknowledged."

Based on the different definitions of gestures and corresponding psychological criteria, we can classify gestures into three categories as suggested by Cadoz [13]:

1. Semiotic gestures: To communicate meaningful information.
2. Ergotic gestures: To manipulate the physical world and create artifacts.
3. Epistemic gestures: To learn from the environment through tactile or haptic exploration.

Alternatively Rime and Schiaratura [14] classified gestures as symbolic, deictic, iconic and pantomimic:

1. Symbolic gestures: Gestures that have single meaning within each culture such as an emblem as "OK" gesture.
2. Deictic gestures: They are most generally seen in HCI such as pointing or directing the Addressee's attention to specific events or objects.
3. Iconic gestures: They convey information about the size, shape or orientation of

the object of discourse.

4. Pantomimic gestures: These gestures are used in showing the use of movement of some invisible tool or object in the speakers' hand.

Supporting the different categorizations of gestures, Kendon specifies the types of gestures based on their speech/gesture dependency. As described in Kendon's Continuum [11], there are five types of gestures: Gesticulation, language-like, pantomimes, emblems and sign language. The necessity of accompanying speech to understand the gesture is maximum in gesticulation and presence of gestures declines in the gesture types list from left to right.

3.2 Gestural Interaction

The approaches mentioned below are the techniques used in recognizing gestures. Understanding the techniques and their drawbacks and advantages is vital to get a new perspective in implementing our design.

3.2.1 Camera-Based

Direct use of hands as an input device provides a natural way of human-computer interface without using any other extra device. For this reason, this interaction style gains researchers' attention. As a result, cameras are being widely used to capture movements and shapes in order to interpret human activities.

Many researchers endorse the use of cameras as they all agree on the disadvantages of using wearables which may require wearing a cumbersome device and carrying a load of cables to interact. As well as Wang et al [15], many other researchers also defend the usage of cameras with a similar reason [16] [17] [18].

An example demonstration of a system to capture movements is seen in Figure 3.1. Wang and Popovic [15] introduced a tracking system uses a single camera to track a hand wearing an ordinary cloth glove that is imprinted with a custom pattern.

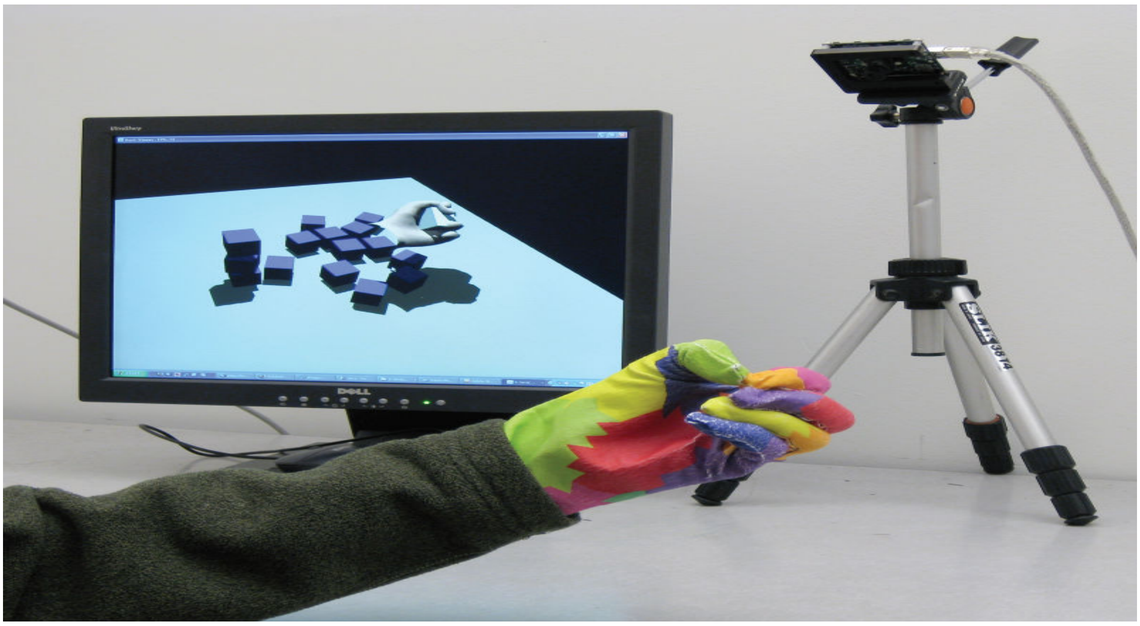


Figure 3.1. Camera-based gesture recognition using a multi-colored glove [15]

Malik and Laszlo [19] developed Visual Touchpad a low-cost vision-based input device that allows fluid two-handed interactions with desktop PCs, laptops, public kiosks or large wall displays. Lenman et al. [20] implemented a prototype based on cameras to control home appliances such as Cd-player and TV.

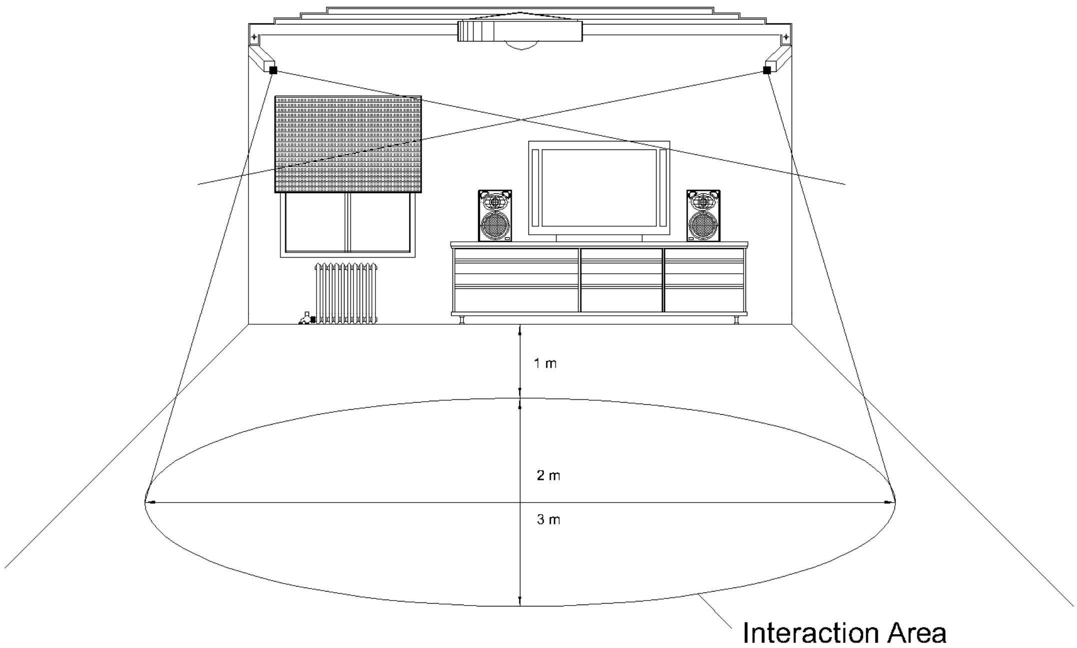


Figure 3.2. Camera-based smart home example showing interacting area

However, this type of interaction has several limitations that make things difficult for them to be used in applications for smart environments. Firstly, it restricts mobility by requiring users to stay in the range of camera as shown in Figure 3.2. This approach decreases the usability due to its complex calibration process. In addition, it is unclear how the system would work in different conditions such as under poor lighting or variant background.

3.2.2 Glove-Based

Another common technique is to instrument the hand with a glove which is equipped with sensors that can sense bending of fingers. Typically it is made of lycra with optical fibers attached along the backs of the fingers. Datagloves use sensor devices for digitizing hand and finger motions into multi-parametric data. VPL Research, Inc inspired by Zimmerman [21] developed fiber optic version of the glove in 1987. Datagloves are used in many projects to interpret hand and finger movements [22] [23] [24] [25]. Figure 3.3 shows an example of data gloves.

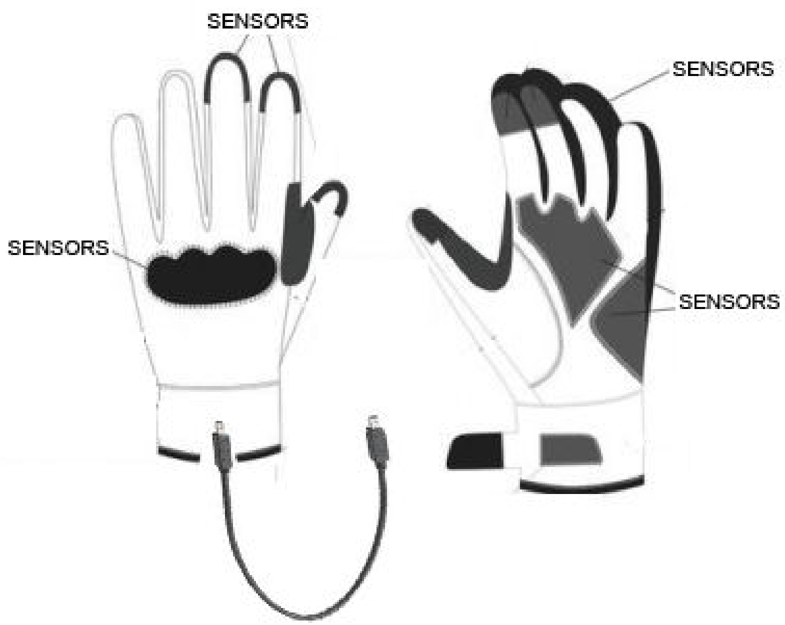


Figure 3.3. Data Glove Example

Though glove-based systems are very accurate and offer a relatively large range motion, they are very expensive, unwieldy, not portable and bring much cumbersome experience to the users. Another limitation is that as stated in Baudel and Beaudouin-lafon [22] it links the user to the computer, it is uncomfortable. Also wearing a glove limits the user's haptic sense and naturalness of movement. One advantage over camera-based interaction is that the user does not have to stand inside the camera's area of capture.

3.2.3 Magnetometer-Based

A magnetometer is a measuring instrument used to measure the strength and/or direction of the magnetic field and used in a variety of applications. Harrison et al [26] presented Abracadabra which is an input technique based on magnetic field that allows wireless, unpowered finger input for any mobile device that has a very small screen. The technique was proposed to solve difficulties in interacting with small electronic devices. The system consists of a wristwatch, with a magnetometer inside, to receive a signal from a conventional desktop computer and a magnet placed on the finger.

Figure 3.4. Abracadabra [26]

An input device in the shape of a finger ringer was presented by Ashbrook et al [27] called Nenya. Similar to Abracadabra, the system consists of a wrist worn device with sensors and an ordinary looking ring. Selection of an item from a menu such as "Pause, Radio, Next, Previous, Volume" can be done by sliding the ring along the finger.

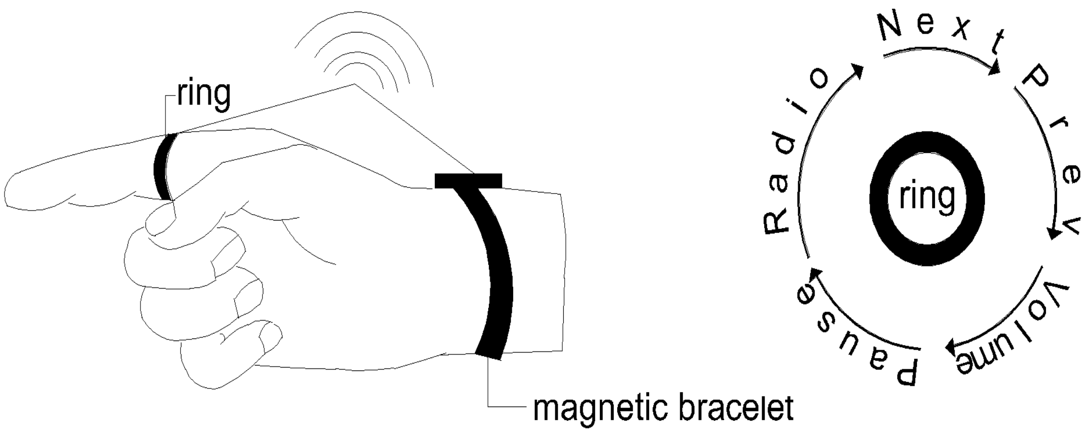


Figure 3.5. Nenya Ring Demonstration [27]

In another magnetometer scenario, Ketabdar et al [28] used a magnet to interact with a mobile using space around the phone based on change in magnetic field in their project called “MagiThings”. The project uses the magnetic sensors already integrated into the newest generation mobile devices such as iPhone 3GS, iPhone 4 and Google Nexus One. User performs movements in the 3D space surrounding the device with a properly shaped (rod, ring, pen, etc.) magnet in his hand and interact with the device. In the light of this concept, Ketabdar and his group implemented MagiSign, MagiWrite, MagiMusic which are aiming user authentication, text-digit entry and interaction with music performance applications.

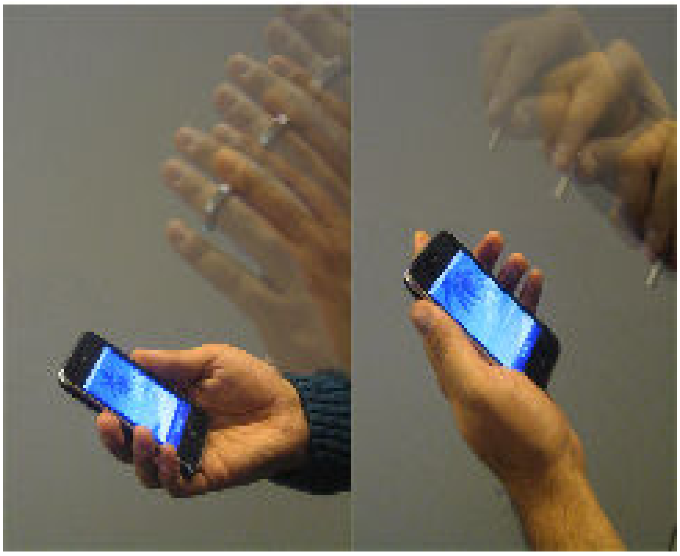


Figure 3.6. Project MagiThings. [29]

This type of interaction offers several benefits. First of all, magnets require no battery, eliminating an obstacle and minimizing the risk of unexpected service break. Line of sight is not a necessity since magnetic field can propagate through materials. Ring shaped systems look like ordinary-looking finger rings hence socially acceptable. However, as in Abracadabra and Nenya, the user may be required to wear a wrist-worn device. Another major limitation is that users have to be aware of magnetism and must be careful not to damage objects sensitive to magnetism.

3.2.4 Accelerometer-Based

Another promising technique is the use of accelerometer, a device that measures the change of speed of anything that is mounted on which is used as an input to control systems. After rapid development of the micro-electro-mechanical-sensors (MEMS) technology, most accelerometers are MEMS and using accelerometers to recognize gestures became an emerging technique.

Mäntyjärvi et al. [30][31] discussed the use of accelerometers in gesture recognition in his publications and proposed a accelerometer-based remote controller recognizing eight gestures to control DVD player. Kim et al. [32] designed a pen-like wireless device based on accelerometer to recognize single Roman and Hangul characters written in the air. Wearables, accelerometers embedded inside, were also designed. Amft et al. [33] proposed a watch called "eWatch" to control the watch and Hein et al. [34] used bracelet for gesture recognition.

Furthermore, most of the new generation electronic devices such as Apple iPhone, Nintendo Wiimote are equipped with accelerometers arousing attention to work with it. Accelerometer axes of Nintendo Wiimote is shown in Figure 3.7 which provides new interaction possibilities in a broad variety of applications such like video games and smart home systems.

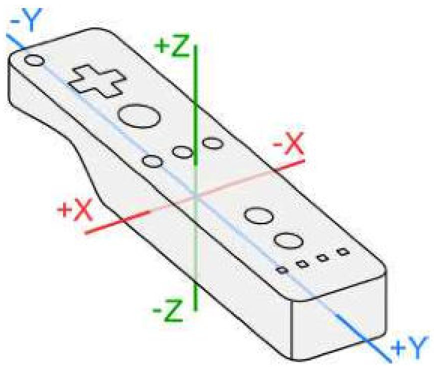


Figure 3.7. The axes of Nintendo Wiimote

4 DESIGN AND DEVELOPMENT

4.1 SHAKE SK7

The SHAKE (Sensing Hardware Accessory for Kinaesthetic Expression) is a versatile Bluetooth inertial sensing device developed by SAMH Engineering Services. SHAKE SK7 is a rectangular cuboid (sub-matchbox sized) device equipped with several sensors and a vibration motor. The dimension of SHAKE SK7 is just 43 by 32 by 18 mm and its weight is 22 grams. As shown in Figure 4.1, the largest surface of the device is just a little bit larger than a 2 Euro coin. The size and weight of the device also played an important role on our decision to choose this device because besides the technical requirements, it has to meet social ones such as portability.

Figure 4.1. SHAKE device compared to a 2 Euro coin

SHAKE SK7 senses linear and rotational movements, absolute orientation / direction (attitude and azimuth) and human body proximity and it is possible to connect to any computing device such as a mobile phone, PDA or Laptop computer that has BluetoothTM wireless or USB connectivity with SHAKE SK7. This way the information will be transferred to the connected device. In addition, Bluetooth connectivity supports up to seven SHAKE's to be connected to the same host meaning that multi-user interaction is also possible. Alternatively, SHAKE can also run in standalone mode if offline data acquisition is required and all the sensor data will be logged to the inbuilt 64Mbit FLASH memory. It also provides both tactile and visual feedback - a necessity of a successful UI design- with the help of vibration motor and LEDs.

SHAKE SK7 SENSORS

The accelerometer available in SHAKE senses the linear acceleration along three axes X, Y and Z as shown in Figure 4.2. The accelerometer can operate in two different ranges: +/-2g range and +/-6g range, where 1g is the acceleration due to gravity. The output data resolution is 1 mg. Gyro sensors complement the accelerometer to provide a six degree of freedom inertial sensing capability and it can be used to capture complex gestures. Gyro sensors measure the angular rate in reference to three axes: Pitch (rotation about the Y axis), Roll (rotation about the X axis) and Yaw (rotation about the Z axis). The measurement range is configurable between +/- 300 or +/- 900 degrees of rotation per second and the output data resolution is 0.1 degree / second. A triple axis magnetometer senses the magnetic field strength with a minimum range of +/- 0.2 mT and the output data resolution is 0.1 uT. Again there are three channels of sense data, one for each of X, Y and Z axes as indicated in Figure 4.2. There are also capacitive (proximity) sensors consist of 12 small metal pads directly under the top surface and arranged in a keypad style array of 3 * 4 as indicated in the Figure 4.2. The capacitance increases when a grounded (AC) object such as human body comes into proximity of these sense pads. On the right hand side of the SK7, there is a three way navigation switch which can be used as an input controller. One use of this might be as a “Push and Talk” switch to indicate to the host that user starts his gesture by pushing and finishes it by releasing as we did in our design.

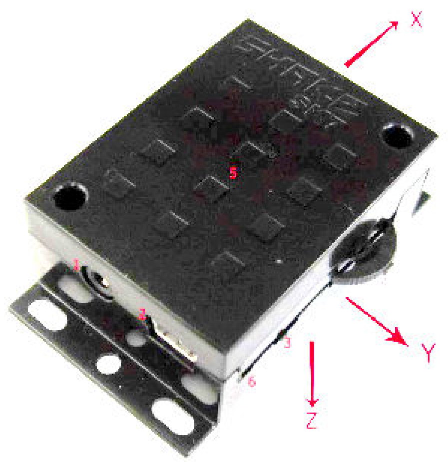


Figure 4.2. The axes of SHAKE SK7

4.2 SHAKE Adaptation for Improved Usability

Carrying a SHAKE device and using it to perform gestures may be cumbersome and socially awkward. It might be considered inappropriate or awkward to pull out the SHAKE device and perform gestures considering the people around might not be familiar with the object. The size must be minimized like recent trends in technology and a proper way to carry the device must be chosen so that in the end the device becomes more convenient to be used in daily life. Consequently, we were looking for a way to render the device more attractive, socially acceptable and make its use subtle. We came up with an idea so that SHAKE is adapted into a ring-shaped device what we call as “Pingu”.

Our decision of developing a “ring-shaped” device is based on several reasons. Our plan is to build systems to control devices through sending commands by gestural interaction. We want it to be accessible to the user for gestural input at anytime in daily life without preventing normal social activities such as hand shaking or using it to control navigation system while he is driving his car, change the music while he is jogging or change the TV channel or turn the lamps on while he is at home. Miyamae et al. [35] point out three distinctive features of wearable computing from other technologies which are: hands free, always on (accessible at anytime) and supporting daily life. In line with these reasons and facts, we reached to an agreement on wearable computing and leading to a new question – what kind of wearable?

A variety of distinct forms of wearables is available such as watch form, glove form and cloth form. Glove form can be troublesome for the user to wear it every day. We decided to have it in shape of a finger ring which is a commonly worn item making it compact and lightweight. Another reason is that users do not have to wear neither a glove nor a heavy bracelet, instead just a small sized unobtrusive ring. Unobtrusiveness is important since it should not look unusual or too “hi-tech”, on the contrary, it must be inconspicuous and uncomplicated. Since the ring is worn, it will always be available to access. It is instantly interruptible which means when the user stops using it, both hands will be free. Furthermore, it is possible to interact even with one hand.

4.3 Pingu

Figure 4.3 demonstrates the prototype of Pingu. It may look like a small machine rather than a regular ring for now, but our further goal is to develop it as in the exact shape of a regular ring in the future. It has almost the same hardware as SHAKE. The dimension without battery and battery charging connector is 10 by 22 by 4.2 mm. The battery is located under the sensory platform, but in the final version it will be at the ring band. It is also possible to connect to any computing device that has Bluetooth connectivity as in SHAKE.

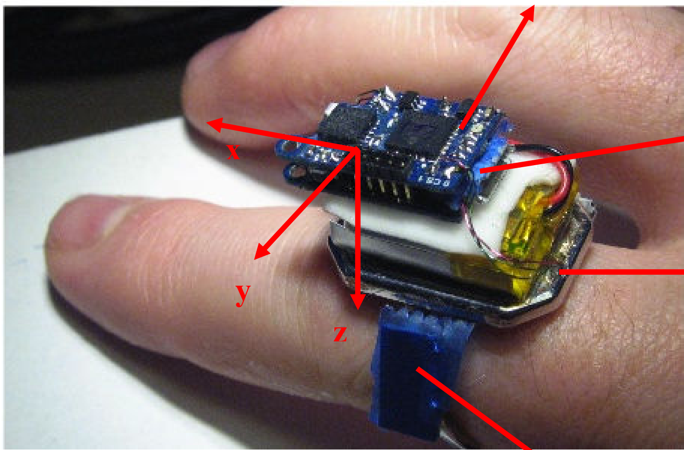


Figure 4.3. The prototype of Pingu

PINGU SENSORS

The accelerometer senses the linear acceleration along three axes X, Y and Z as shown in Figure 4.3. The accelerometer can operate in three different ranges: +/-2g, +/-4g and +/-6g range, where 1g is the acceleration due to gravity. The output data resolution is 1 mg. As in SHAKE, gyro sensors complement the accelerometer to provide a full six degree of freedom inertial sensing capability and it can be used to capture complex gestures. Gyro sensors measure the angular rate in reference to three axes: Pitch Roll and Yaw. The measurement range is configurable between +/- 250, +/- 500 or +/- 900 degrees of rotation per second and the output data resolution is 0.1 degree / second. A triple axis magnetometer senses the magnetic field strength with a minimum range of +/- 0.2 mT and the output data resolution is 0.1 uT. Again there are three channels of sense data, one for each of X, Y and Z axes as indicated in Figure 4.3. The capacitive (proximity) sensors consist of 2 metal pads attached on both sides of the ring making it easier to touch by another finger and takes place of navigation switch in Pingu. The capacitance increases when a grounded (AC) object such as human body comes into proximity of these sense pads.

Figure 4.4. The ultimate form of the Pingu

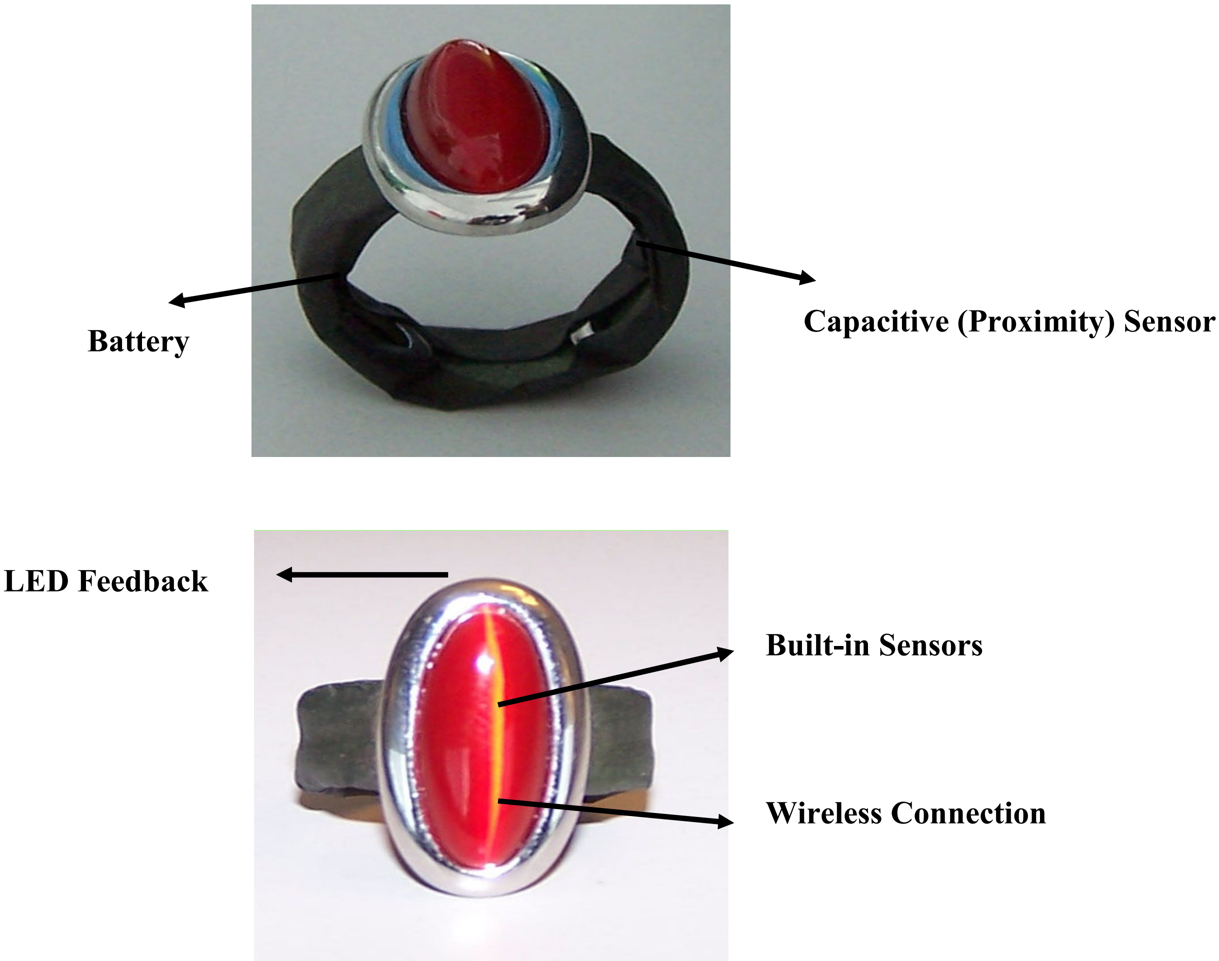


Figure 4.5. The design of the ring

Pingu prototype still looks like an electronic device rather than a normal ring. As previously discussed, the ultimate goal is to end up with a normal ring which is socially acceptable providing all these sensors, feedbacks and wireless connection. The design we had in mind is shown in Figures 4.4 and 4.5.

4.4 Gesture Recognition

The following section illustrates the process of recognition: how gestural data is captured and how it is matched to the stored gesture.

4.4.1 Data Gathering

As a first step, it is important to define the start and end points of the gesture. Although we worked on automatic detection, the probability of performing involuntary gestures seemed too high. Therefore, the system is implemented so that gesture recognition is triggered as the user touches the capacitive sensor and releases it.

Figure 4.6 illustrates the smart home system behavior controlled by gestures. Pingu sends data from various sensors to the host (Mac OSX) over the serial connection via Bluetooth. These data packets are sent periodically with a configurable frequency. One data packet contains in total 11 values: accelerometer, gyro sensor and magnetometer data for all three axis (x, y, z) and also two more values for capacitive sensing. This is what we call as "vector" in our data structure and gestures are made of several vectors. The recognition is enabled by touching the capacitive sensor plate of the ring and incoming accelerometer and gyroscope sensor readings are written into an array. Stopping the contact with capacitive sensor stops recording. The length of the gesture will be checked to eliminate the threat for an involuntary gestures resulted by an accidental touch to the capacitive sensor from being recognized as a normal gesture. If the length of a gesture is less than a certain threshold, then it is assumed that capacitive sensor plate is touched accidentally and the gesture is treated as invalid, in other words the reading is not used, hence nothing happens. Otherwise, the gesture is valid and the next step is to compare the recorded data with the gestures already stored in the system and match it to the closest one. As the gesture is identified, the command assigned for that specific gesture is sent to the smart-home system.

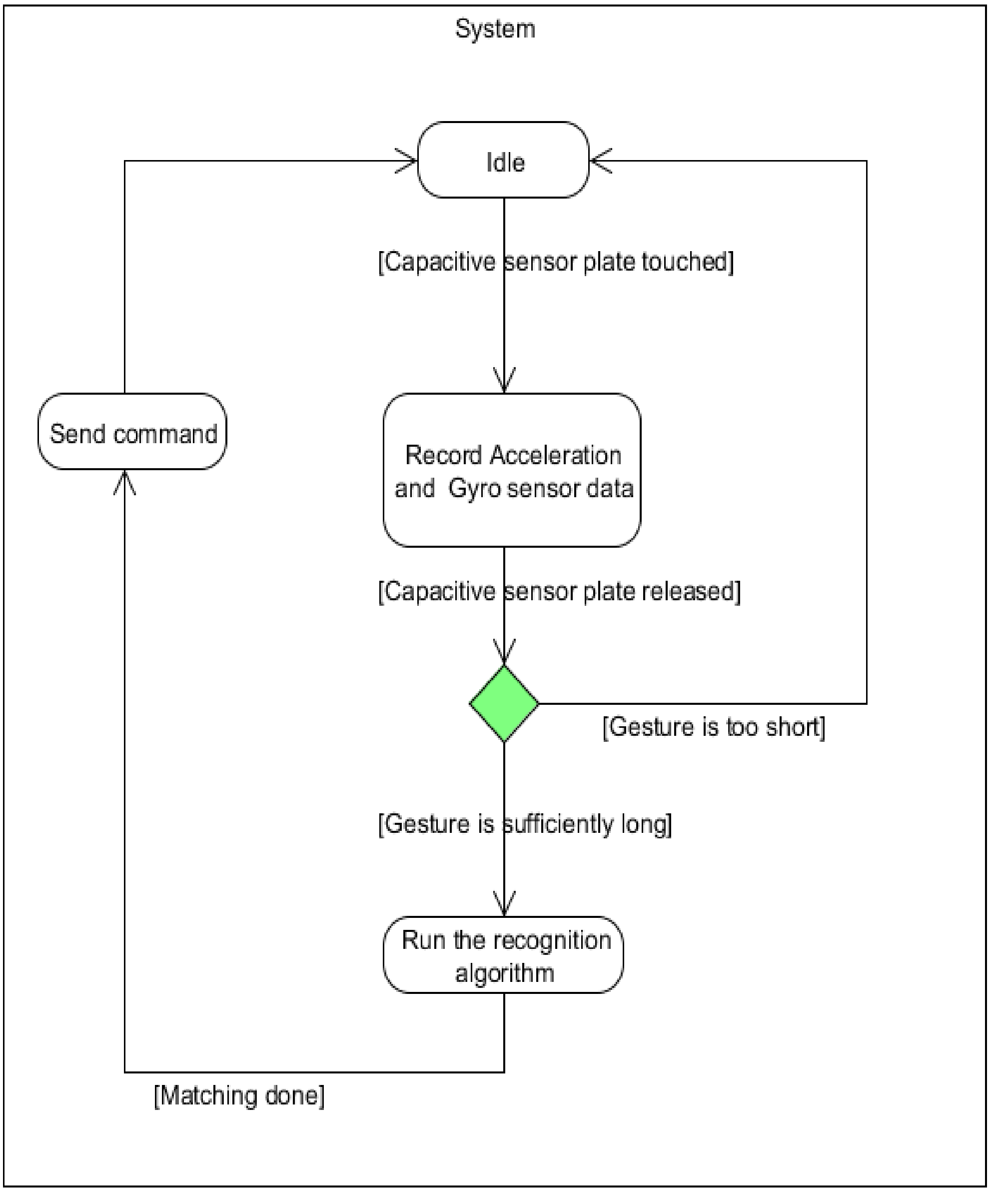


Figure 4.6. State diagram of the system

4.4.2 The Process of Recognition

Using Dynamic Time Warping (DTW) method described in next section, we find an optimal patch between two different gestures composed of vectors. Performed gesture and stored gestures are "warped" non-linearly in the time dimension. We compare performed gesture with stored gestures and compute the distance to each of them through DTW. Finally, the gestural data is matched to one of the stored gestures with minimal distance. Appendix A code segment contains the DTW algorithm implementation used in this process.

RECOGNITION ALGORITHM

One of the most popular degree of similarity measures is Euclidean distance where the cost between n-dimensional sequences $\boldsymbol{A}=(A_1, A_2, \ldots, A_i, \ldots, A_n)$ and $\boldsymbol{B}=(B_1, B_2, \ldots, B_i, \ldots, B_n)$ is computed as:

$$d_{Euclidean}(\boldsymbol{A}, \boldsymbol{B}) = \sqrt{\sum_{i=1}^{n} (A_i - B_i)^2}$$

However measuring the similarity with using Euclidean distance is very brittle. DTW is more robust and gives better results even if the two sequences are out of phase in the time axis.

Consider two sequences that have similar shape but not aligned in time axis as in below:

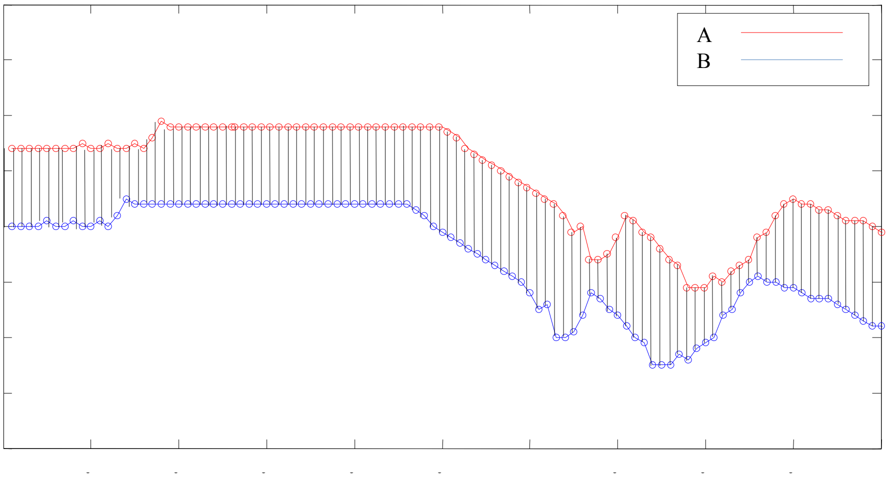


Figure 4.7. Euclidean alignment

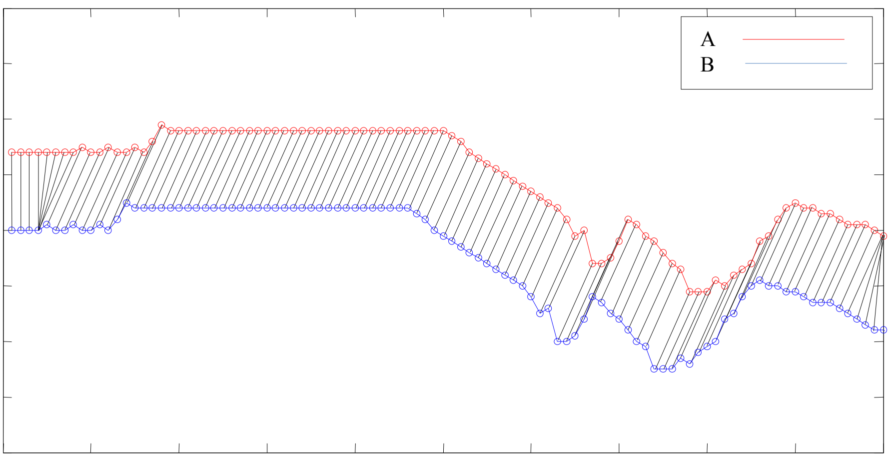


Figure 4.8. Dynamic Time Warping alignment

In Euclidean distance method, i-th points in each sequence will be aligned to each other and it will produce improper similarity measure. Dynamic time warped alignment will produce more intuitive result.

Dynamic time warping is an algorithm which is used to find optimal alignment between two (time-dependent sequences) under certain constraints. Besides, it is easy to develop, very accurate and computationally efficient. Figure 4.9 shows two sequences, $\boldsymbol{A}=(A_1, A_2, \ldots, A_i, \ldots, A_n)$ and $\boldsymbol{B}=(B_1, B_2, \ldots, B_j, \ldots, B_m)$ $n, m \in Z\text{-}+$ *(Positive Integers)*.

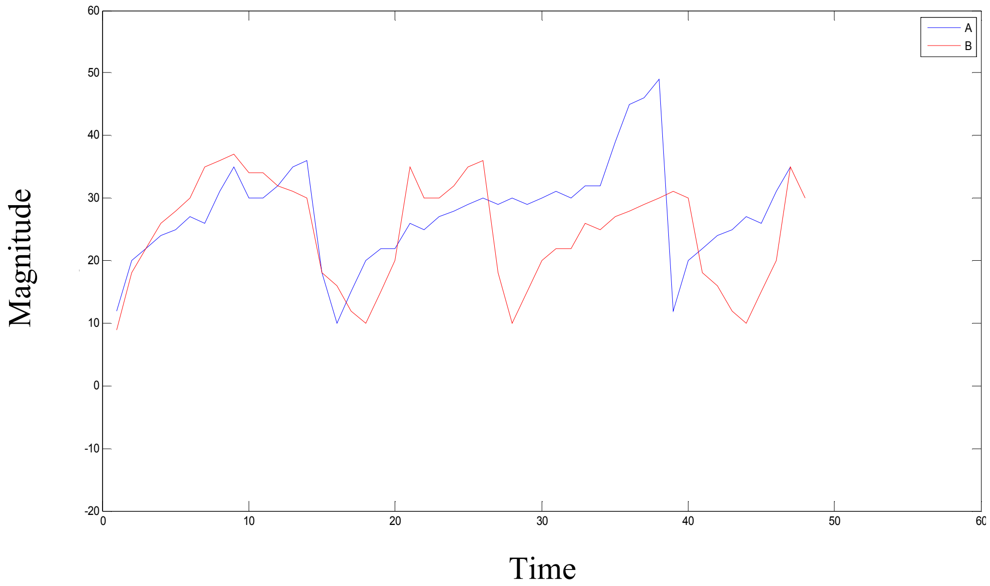


Figure 4.9. Two sequences that have similar shape

As a first step, $n \times m$ matrix D is constructed with distances $D_{ij} = d(A_i,B_j)$ which is calculated as $(A_i - B_j)^2$. A warping path defines a mapping between A and B. W is a contiguous set of matrix elements and it is defined as:

$$W_k = D[i, j]_k$$

hence $\boldsymbol{W} = (D[i,j]_1, D[i,j]_2, \ldots, D[i,j]_k, \ldots, D[i,j]_K)$ and $\max(m,n) \leq K < m+n-1$. The warping path must satisfy following conditions:

Boundary Condition: The boundary condition forces a match between the first points of the curves and a match between the last points of the curves.

$$W_1=D[1,1] \text{ and } W_K=D[n,m]$$

An example boundary condition to match starting point D*[1,1]* of the curve:

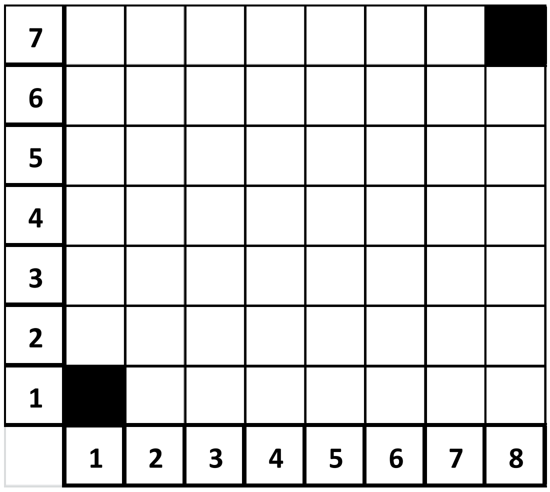


Figure 4.10. Boundary condition for $n=7$ and $m=8$

Continuity (Step Size) Condition:

Given that $W_k = D[i, j]$ and $W_{k-1} = D[i', j']$, then $i' \leq i \leq i'+1$ and $j' \leq j \leq j'+1$

This restricts the allowable steps in the warping path to adjacent cells (including diagonally adjacent cells).

Monotonicity Condition: The path can only move forward through the distance matrix and this condition prevents the matching from "going back in time".

Given that $W_k = D[i, j]$ and $W_{k-1} = D[i', j']$, where $i - i' \geq 0$ and $j - j' \geq 0$, forces the point in warping path monotonically increasing. Figure 4.11 shows an example of monotonicity condition:

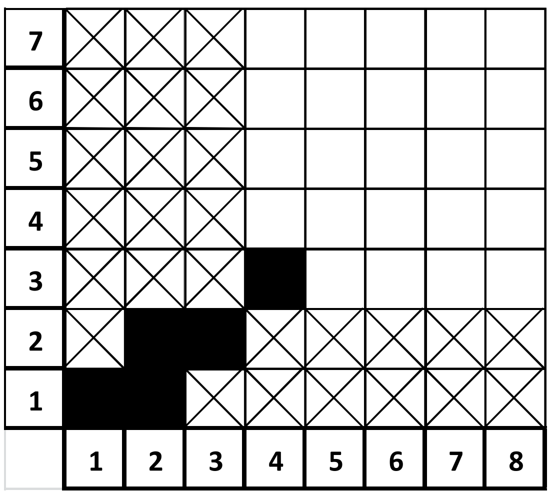


Figure 4.11. Monotonicity condition. Crossed squares are prohibited by this condition as we matched 5 shown points.

Illustration of paths of index pairs for some sequence A of length $L_A=7$ and some sequence B of length $L_B=8$. Figure 4.12 shows a violation of boundary condition because there is no match between first and last points of the curve. Figure 4.13 shows a violation of continuity condition because of the jump in row 3. Finally, Figure 4.14 is an example of a violation of monotonicity condition since the path goes backward from the cell D*[4,5]* to D*[5,4]*.

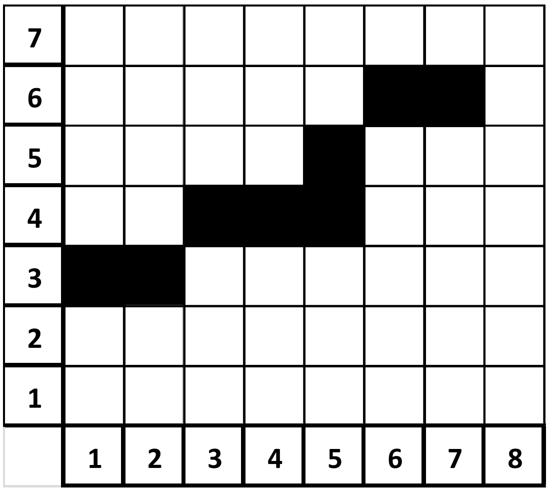


Figure 4.12. Boundary condition is violated

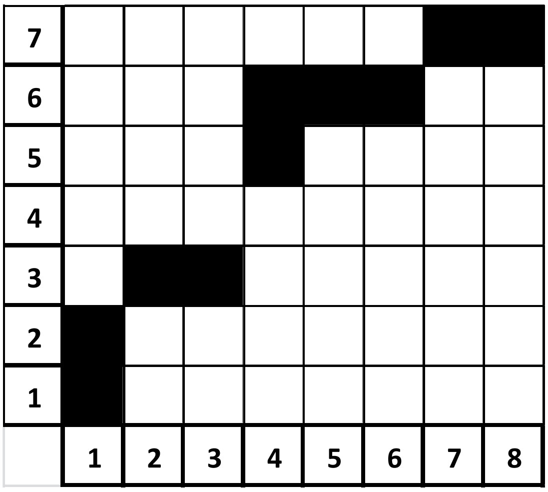


Figure 4.13. Continuity (Step size) condition is violated

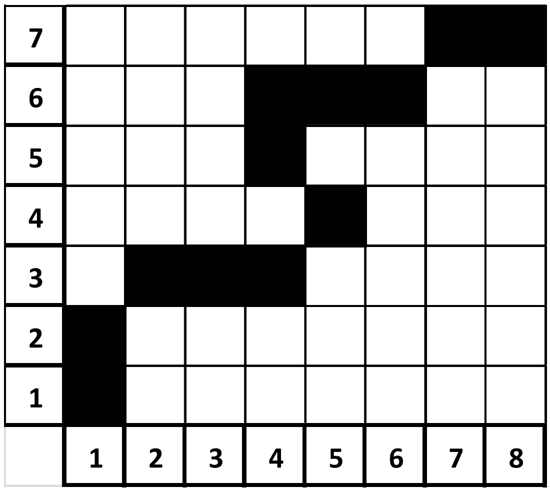


Figure 4.14. Monotonicity condition is violated

These conditions above can be satisfied by many different warping paths. However, the optimal warp path is the one with minimum cost among all possible warping paths.

$$Cost_{Optimal\ Warp\ Path} = min\begin{pmatrix} First\ Path\ Cost \\ Second\ Path\ Cost \\ \dots.. \\ Last\ Path\ Cost \end{pmatrix}, where\ Cost_{Path} = \sqrt{\sum_{k=1}^{K} W_k}$$

This warping path can be found by using dynamic programming to evaluate the following recurrence, which defines the cumulative distance *γ(i, j)* as the distance *d(i, j)* found in the current cell and the minimum of the cumulative distances of the adjacent elements:

$$\gamma(i, j) = d(A_i, B_j) + min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\}.$$

The resulting matrix satisfies the conditions above and shows the optimum alignment (warping path) as shown below.

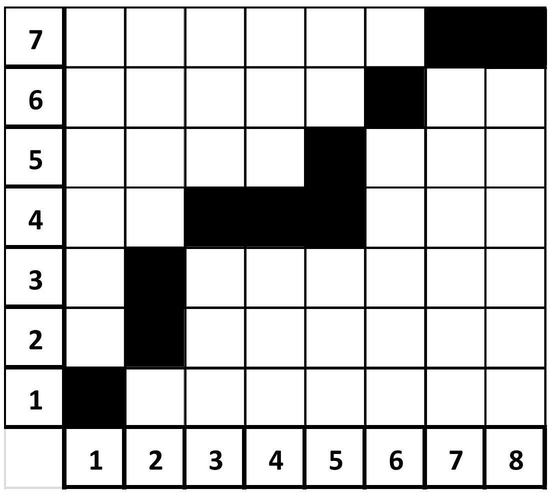


Figure 4.15. Two Sequences aligned with the optimum warping path

In our context, gestures are recognized by comparing an input gesture to a set of stored gestures in our system and measuring any similarity between them using DTW. After DTW similarity measurements are done between the input gesture and stored gestures, the input can be either matched to the template which is most similar, or rejected (no match) as belonging none of the possible stored gesture in the system if the similarity measurement is lower than pre-defined similarity threshold.

4.5 User Study

We conducted a user study with 24 participants to investigate the usability and reliability of Pingu. We used Pingu in six gesture categories (general gestures, activities, intra-hand gestures, intra-finger gestures, digit entries and signatures) and four methods of performing gestures. The four methods we described are as follows: writing in the air, on a table, on the palm and interaction with a magnet. For the last method, users wore a magnet ring on their finger. "Gesture Recognition User Study Test" questions and the guidelines for the study can be found in Appendix B and C respectively. The participants are asked to perform gestures according to the categories and methods. The gestures are recorded with a SHAKE and Pingu toolkit developed by us. We have not analyzed all the data yet but according to the participants' feedbacks, the technique was found to be very interesting, usable and reliable.

4.5.1 Gesture Categories

We divided our user study into categories:

- **General Gestures**

  In this category, we want to evaluate Pingu for general gestural interaction. This category includes natural gestures in daily life that can be used to control smart environments.

- **Activities**

  This category consists of typing on Mac keyboard and iPad screen, relaxing, running, walking and handshaking to evaluate usability of Pingu in activity monitoring

- **Intra-hand & Intra-finger Gestures**

  Interaction with magnet method is used in these two gesture categories. Participants are asked to perform gestures with their hands and fingers while wearing a ring-shaped magnet.

- **Digit Entries**

  In this category, we evaluate the usability of Pingu in character and text entry. In this section of the user study, the participants are asked to perform digit gestures in a predefined way. It is crucial to perform the digits similar to how it is defined.

- **Signatures**

  We want to evaluate Pingu for its dynamic usability in secure authentication. We asked participants to perform their own signatures to examine and analyze different signature examples with the purpose of probable signature recognition feature of Pingu.

4.5.2 Participants & Data Collection

We collected the dataset to evaluate Pingu from 24 participants. The participants were 41.7% male (10 participants) and 58.3% female (14 participants). Age ranged from 20 to 40 years old. 4 participants (16.6%) were left-handed. 20 (83.4%) were in non-technical field or areas of study. 4 (16.6%) were in technical fields. 14 participants (58.3%) own a touch screen mobile device, while 9 participants (37.5%) have a keypad device and one participant (4.2%) left this question unanswered. Of the participants, 21 (87.5%) are familiar with touch screen. Each gesture in the user study is performed 15 times per participant. The sensor readings specific to each gesture are captured via a desktop application. We made sure that the participants performed the gestures correctly so that the dataset is reliable and ready to analyze.

5 PROTOTYPE APPLICATIONS

5.1 Concept

We have developed several prototype applications considering the compliance with problem definition. The ultimate idea is to control electronic devices within Bluetooth range in a home environment via gestures by means of Pingu ring. Gestural data is sent to a host for the purpose of controlling a certain device. After the process of data by the host, the command associated with the performed gesture is sent to the relevant household appliance. This concept is illustrated in Figure 5.1.

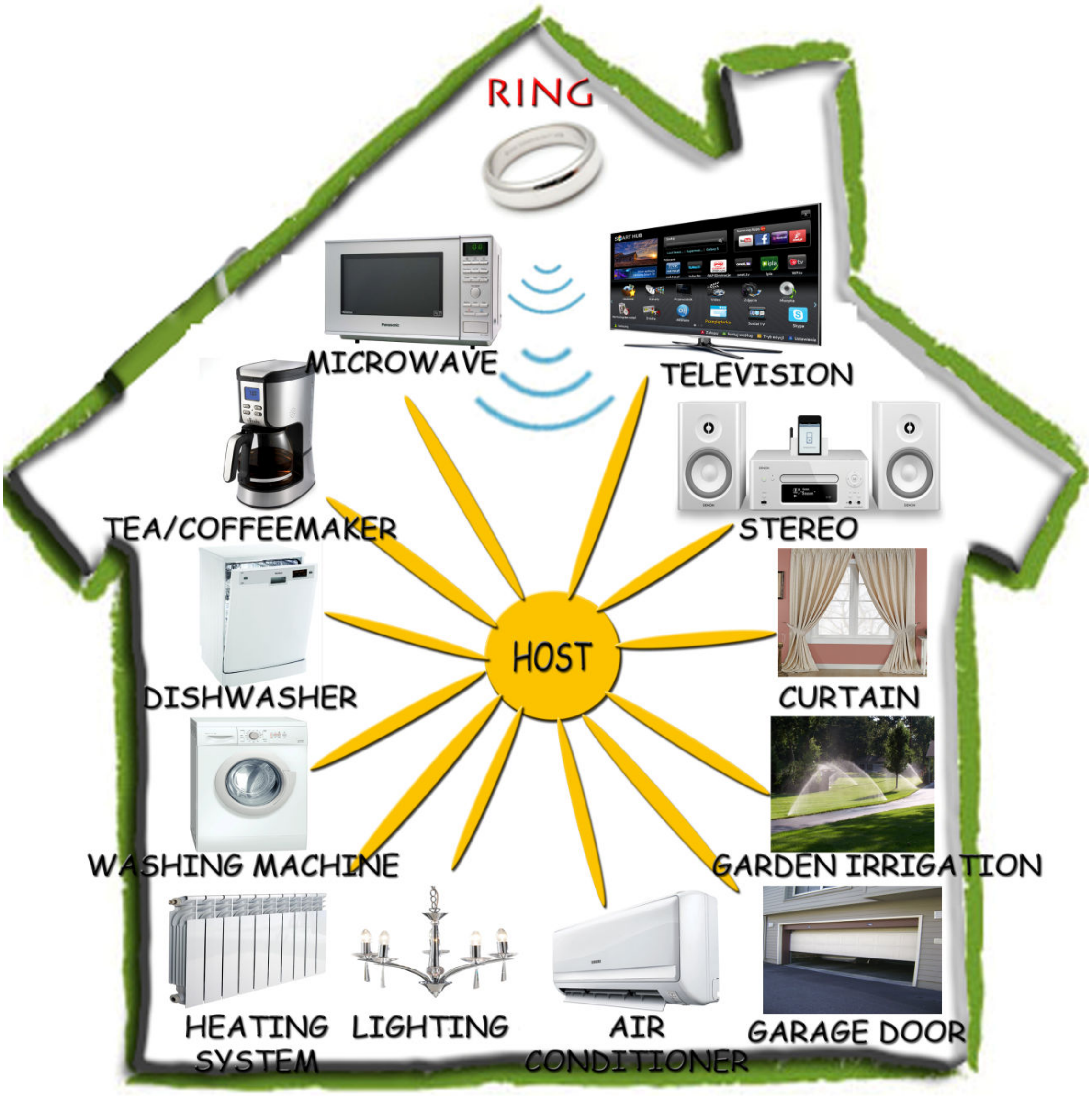


Figure 5.1. Smart Home Concept with Pingu Ring

5.2 Platform

We implement a smart-home micro-prototype system as proof of concept focusing on controlling a television, air conditioner and lighting system. The implementation is developed under Java Development Kit (JDK) 6 environment using Java programming language, on a computer running Apple MAC OS X 10.6.8. To represent a television, Quicktime for Java (QTJava) library which allows developers to use QuickTime features to play and control different types of media supported by QuickTime is used. However, QTJava is not supported on 64-bit MAC OSX thus, the application runs in 32-bit mode.

The general architecture of smart-home micro-prototype system is shown in Figure 5.2. The user communicates with the host via Bluetooth and performs gestures to send commands to the system. In addition to the television menu, a fan and two lamps are also included and connected to the host through one-channel USB relay.

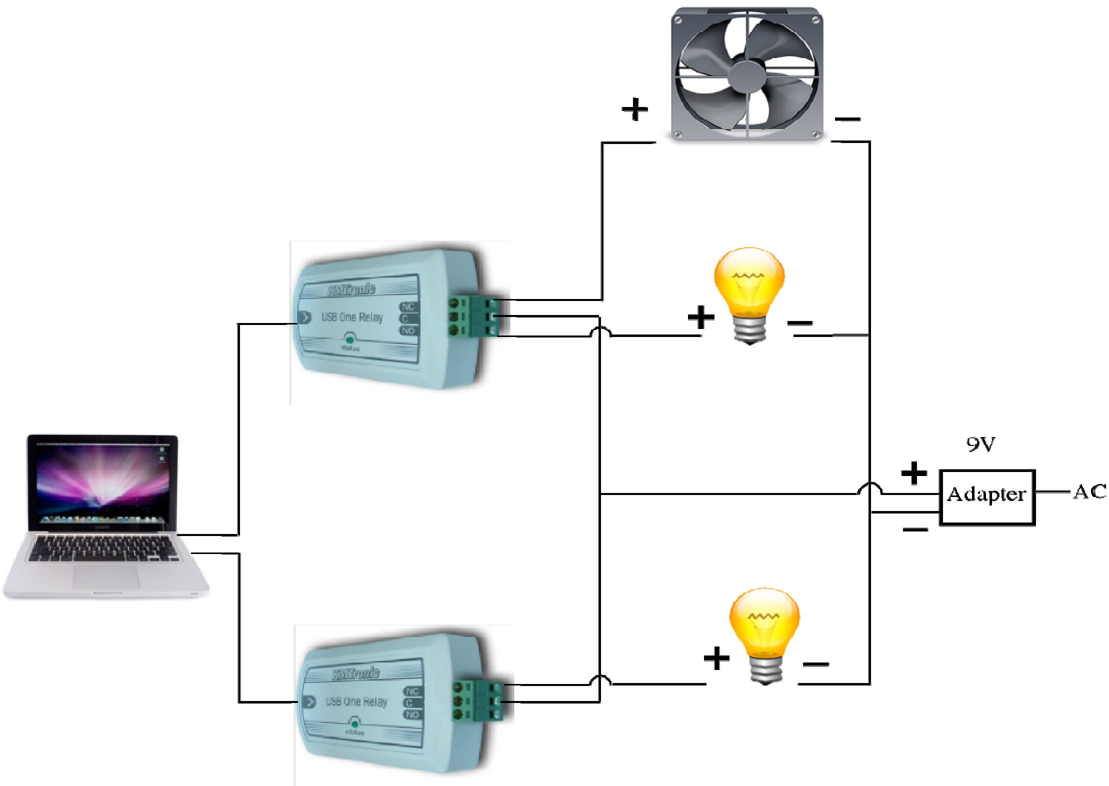


Figure 5.2. The general architecture of smart-home system behavior

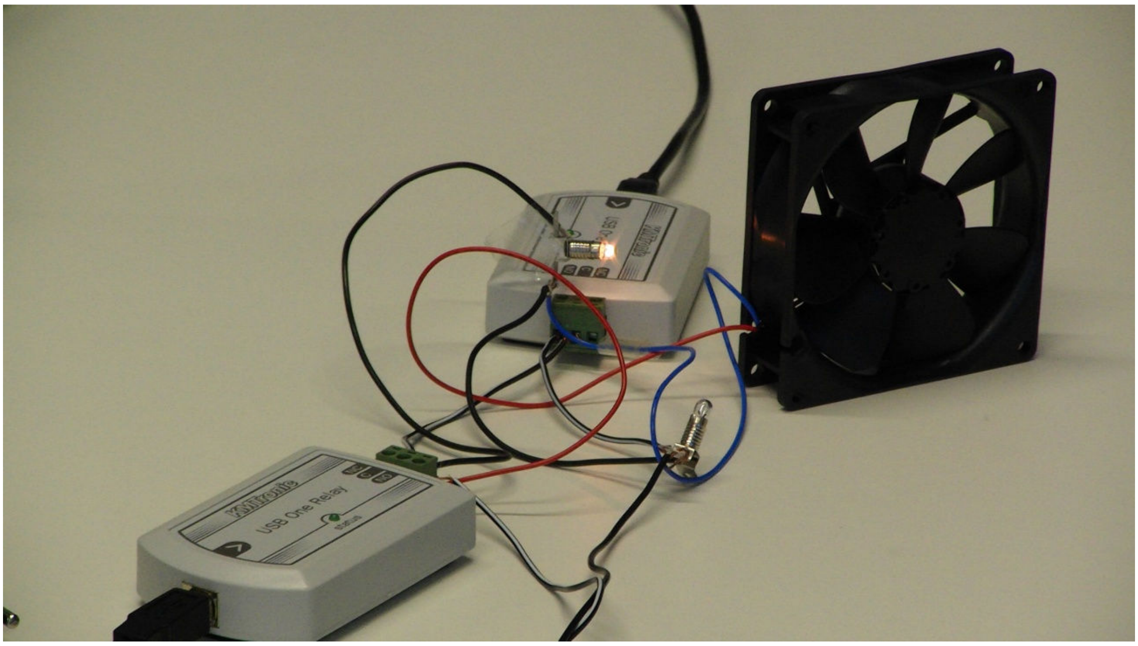


Figure 5.3. Two lamps and a fan connected through relays

5.3 Basic Functions

Our system supports the following basic functions:

Selecting a Channel

The current system supports channels 1 to 19. To select a channel, perform a gesture of the channel number that you want to select. The selected channel number appears on the screen for 5 seconds. The ways to select:

- Single digit channel - “8”
- Double digit channel - “15”
- Character entry - “F” for “FOX”
- Text entry - “BBC”

Playing Next/Previous Channel

To go to the next higher channel perform gesture “Right” shown in gesture vocabulary and to watch the next lower channel perform “Left”.

Volume Control

Perform gesture “Up” to increase the volume and “Down” to lower the volume. The level bar at the bottom of the screen indicates the relative volume level.

Turn On/Off Fan

While watching TV, perform text “FAN” gesture to turn it on and off.

Turn On/Off Lamp

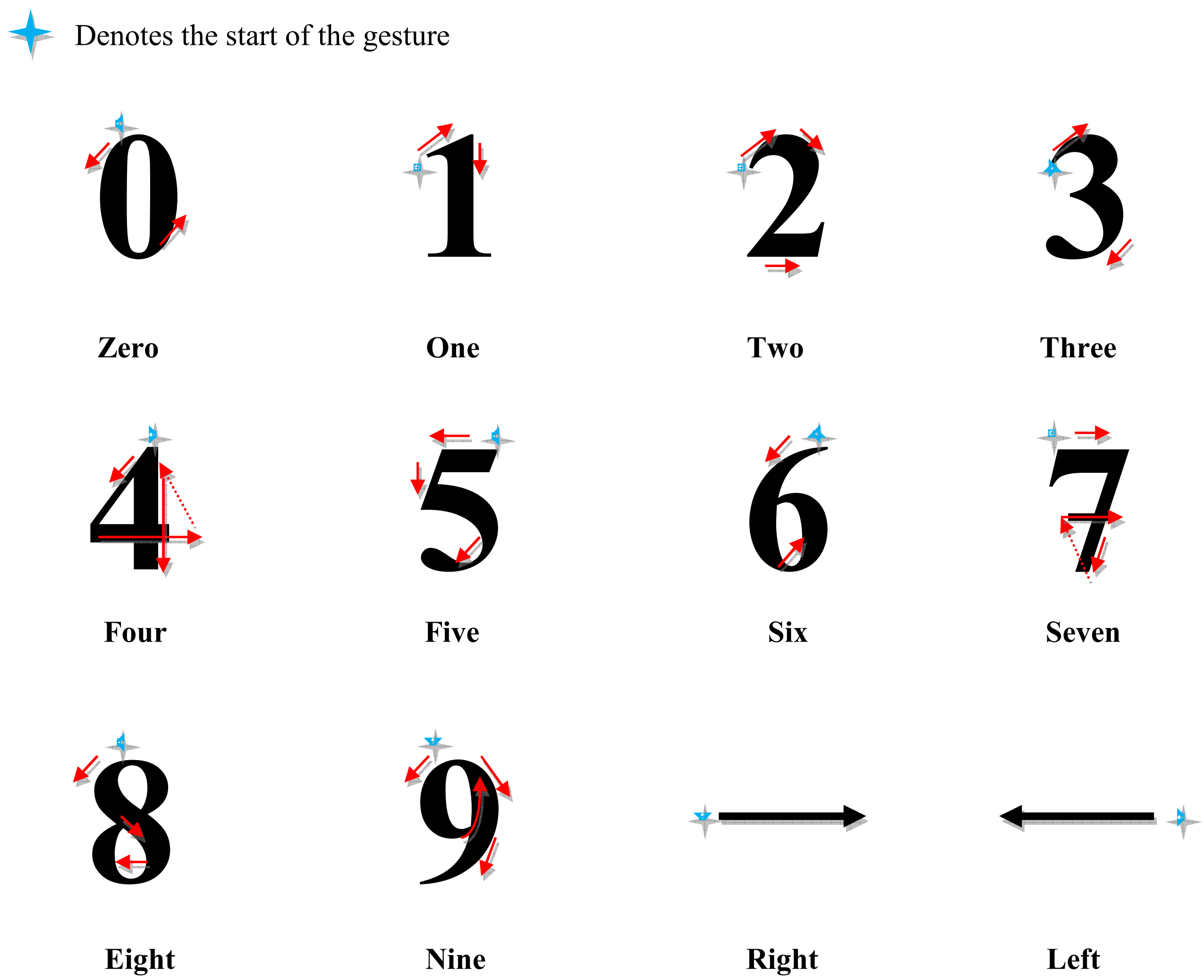
While watching TV, perform text “LAMP” gesture to turn it on and off.

Using the Menu

- The main menu will be displayed when you perform the “Check” gesture. To close, the main menu, perform “X” gesture.
- If you are at the TV menu, “X” gesture is used to deselect TV menu.
- Perform “Right” and “Left” gestures to make desired selection between TV Menu, Setup Menu, Radio Menu or Internet and select by performing “Check” gesture.
- To deselect Setup Menu and Radio Menu perform “Left” gesture. If any of them selected, now you can perform “Up”, “Down”, “Right”, “Left” to go to desired section and if the section is an end point and if you want to select it, you need to perform “Check” gesture.

5.4 Gesture Vocabulary

Currently our system supports, 26 gestures, shown in Figure 5.4. Please note that the system works for the gestures predefined in the system. In other words, the user must perform a gesture similar to shown in the figure.



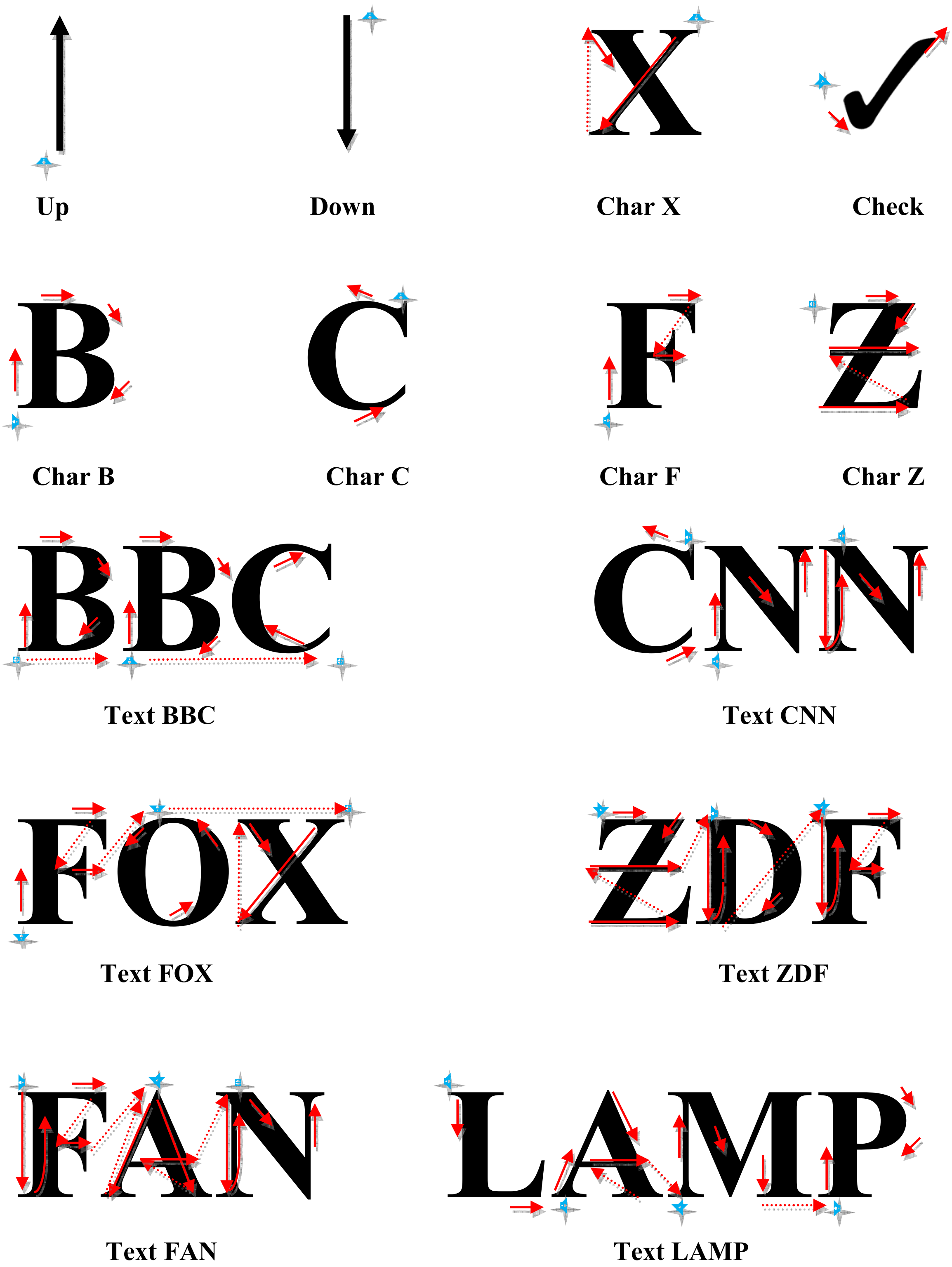


Figure 5.4. Gesture Vocabulary

5.5 Applications

As a sample based on the concept of smart home, we implement a Smart TV system with several demonstrations.

5.5.1 Basic TV System

Basic demo allows the user to perform all the gestures shown in Figure 5.4 except characters and texts. In other words, the user can perform number gestures to select a channel, “Right”, “Left” to go to the next/previous channel respectively, “Up” and “Down” to control the volume on the main frame. As soon as the user runs this application, a predefined channel appears on the screen.

- For e.g., to select a single digit channel “8” perform “8” shown in Figure 5.4 and wait. The selected channel number appears on the screen for 5 seconds.
- To select a double digit channel “12” perform gesture “1” and then perform gesture “2”.

5.5.2 Character & Text-Entry TV System

Character and Text-Entry TV System is an extension to the basic TV system and provides character and text entries. All supported gestures for the basic demo are also supported by this demo. In addition, the user can select a channel either by performing the first character of the channel or entering the name of the channel. For instance, if the user performs “B”, “BBC” will be selected or to select CNN, he/she just performs “CNN” gesture.

5.5.3 Train TV System

Different than the previous demo applications, the purpose of this demo is to teach the user how to perform a gesture and increase the system’s recognition rate. As the user runs this application, he/she is asked to repeat the basic gestures (Up, Down, Right, Left, Close, and Select) which are used to control television menu. Once this short process is finished, then the user will be more familiar with the system so that he/she can interact with the menu and repeat the same procedure for more advanced gestures (in our case which are numbers).

5.5.4 Smart Home Micro-Prototype

This application covers all the functionalities of the demos described above. In addition to the TV remote controller, the ventilation system and room lighting are embedded into this prototype application.

The user can choose channels from 1 to 19; thus, nothing happens if “0” (zero) is performed. If user enters a double digit channel number larger than 19, channel 19 will be played in all cases.

Until the date of publish of this thesis, the television menu is not fully implemented. Thus, the user may not be able to reach all the menu functions. For now, the channel menu is implemented with minor problems.

It is important how the user holds the device. For Pingu, wear it so that the yellow part is far away from your fingertip as shown in Figure 5.5.

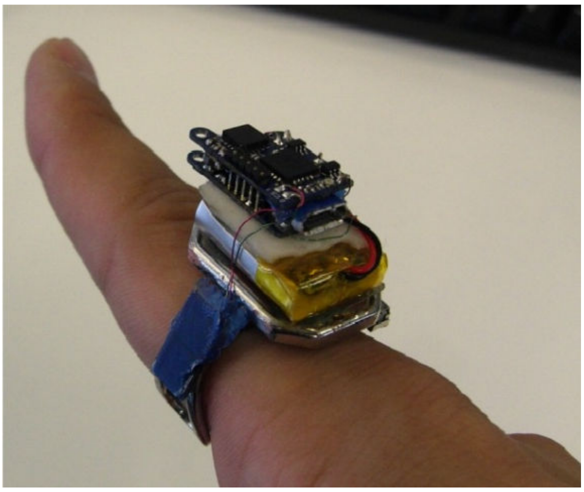


Figure 5.5. The way user must wear the Pingu

5.6 System in Action

This section gives an example on how to use the system. The following examples were made using the demo called “Smart Home Micro-Prototype”. In the beginning, user is asked to repeat the high priority gestures (Up, Down, Right, Left, Close, Select) to become familiar with the system as mentioned in Section 5.5.3. In the screen, “Please train your gesture as shown in the video” prompt appears for related gesture. Although this visual feedback is not implemented, it is a necessity and the idea is to show the user how to interact with the system and how a gesture is defined. Figure 5.6 shows the familiarization process for “Right” gesture where the user performed the “Right” gesture in an incorrect way. Therefore, at the bottom of the screen another prompt appears “Training Failed Right: Please train carefully”. After the user gets familiar with the system by performing the gestures correctly, the user is now ready to interact with the television menu as in Figure 5.7.

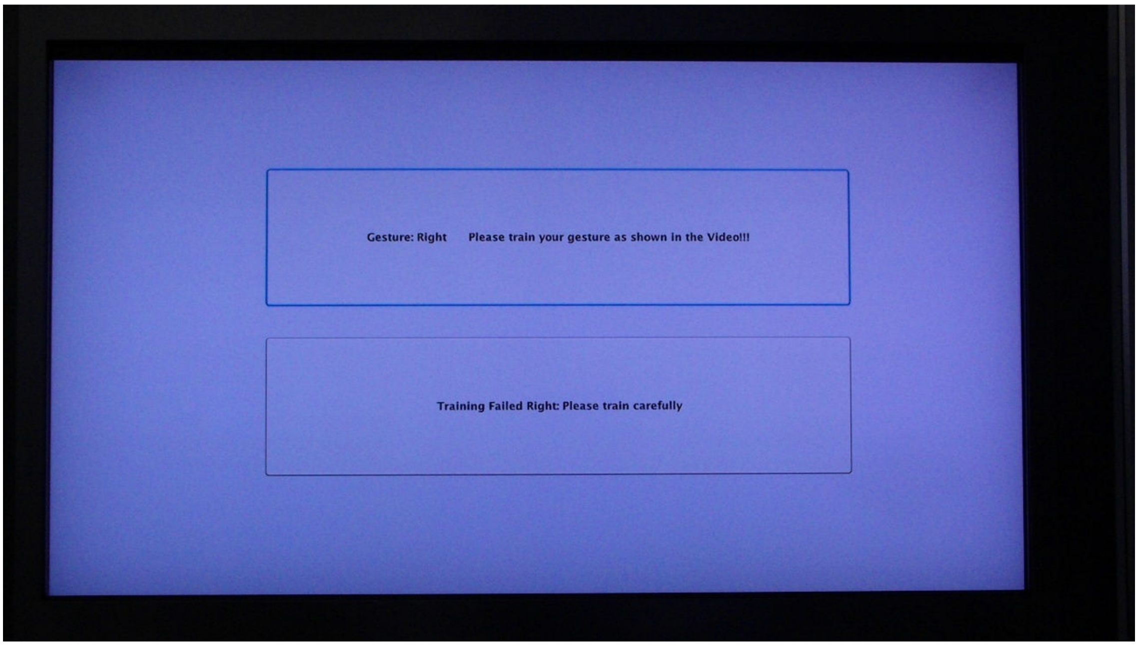


Figure 5.6 Getting familiar with predefined gestures

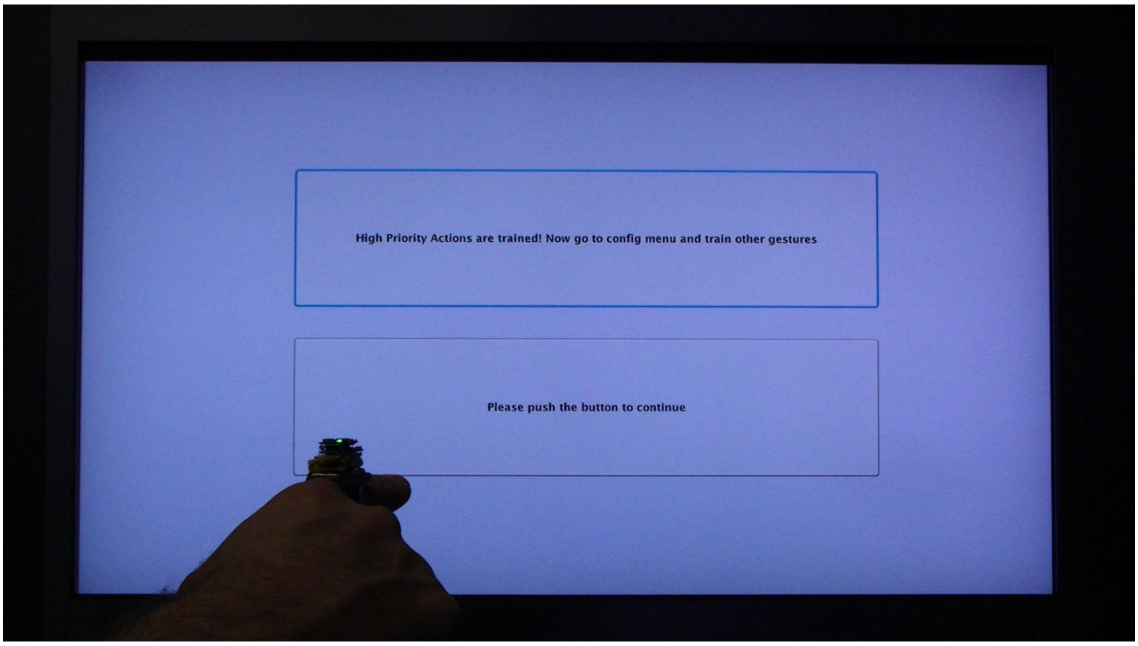


Figure 5.7 After familiarization process, ready for interaction

When the user performs "Check" gesture, it will be compared with the stored gestures and matched to the closest one. As a result, the main menu showing predefined channels appears on the screen as shown in Figure 5.8. User can select either TV menu, setup menu or radio menu by performing "Right" and "Left" gestures. For e.g., to move from TV menu to the setup menu perform "Right" shown in Figure 5.9.

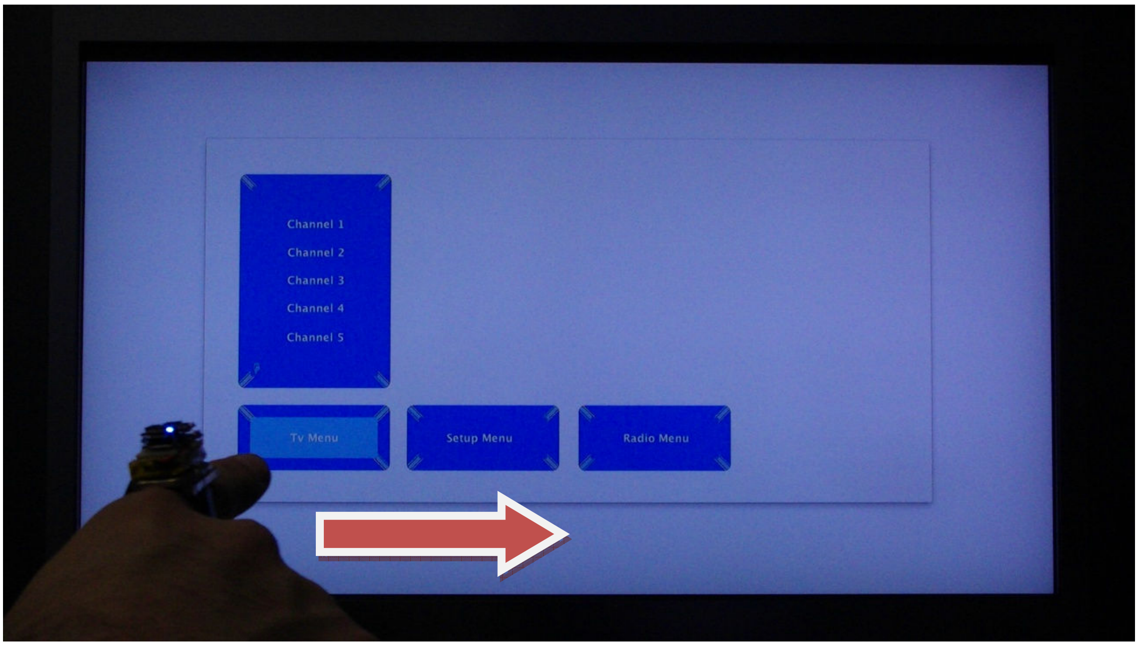


Figure 5.8 TV Menu is ON (Before performing gesture 'Right')

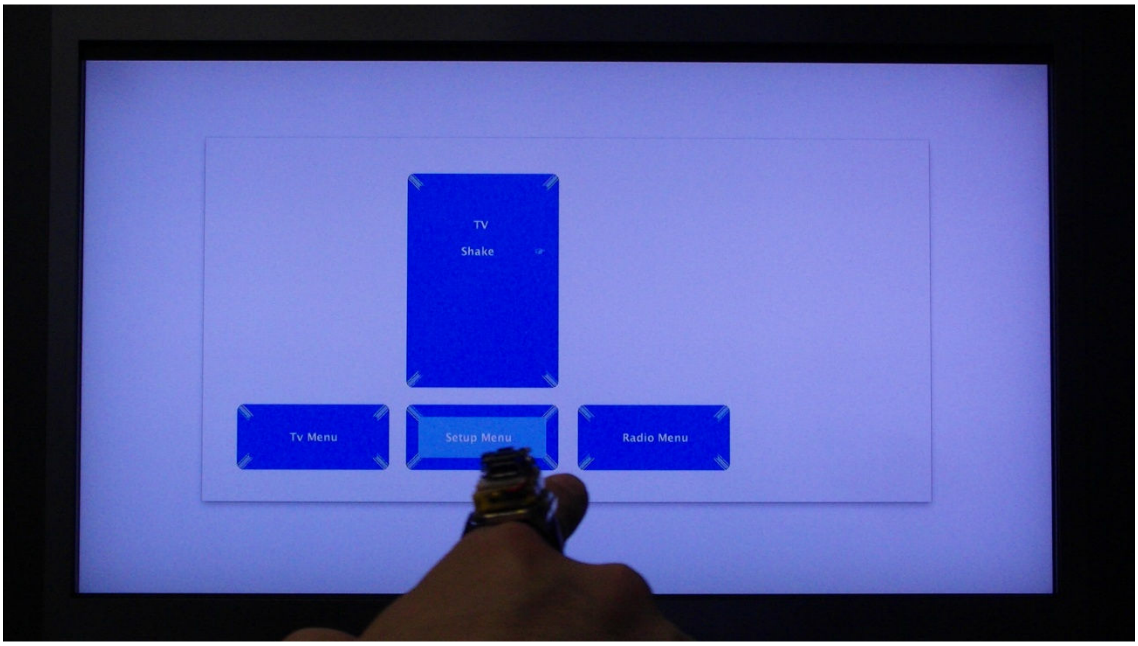


Figure 5.9 Setup Menu is enabled (After "Right" gesture was performed)

The transparent menu allows users to interact with it while watching a channel as shown in Figure 5.10. In this scenario, the user is willing to get familiar with basic gestures which are the numbers from 0 (zero) to 9 (nine). Since the user learnt how to perform basic gestures, he/she can easily interact with the television menu. First, the user performs "Right" gesture to go to setup menu and perform "Check" gesture to select. Then, "Down" gesture to go to Shake section. Here, there is a sign "☞" (hand sign) right next to Shake section as shown in Figure 5.12 which means that this section can be extended.

Figure 5.10 Transparent Menu

Figure 5.11 Setup Menu is enabled

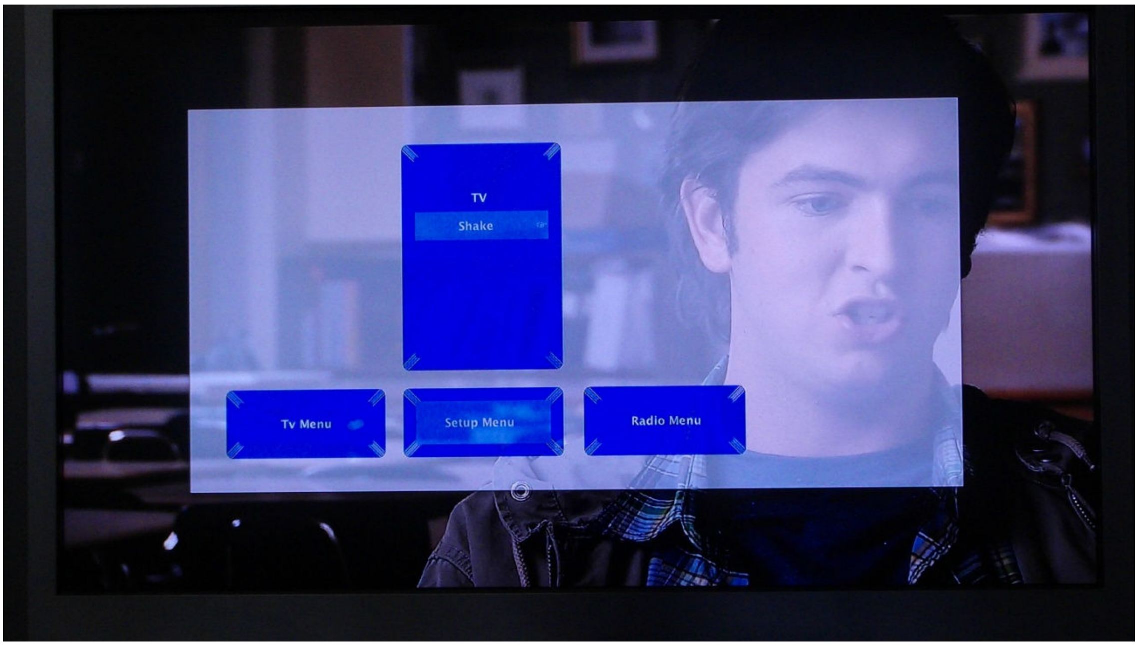


Figure 5.12 Setup Menu is selected (After “Check” gesture was performed)

The user performs two "Right" gestures consecutively and now he/she is ready to select the section which is an end point as shown in Figure 5.14. If the “Basic Gestures” section is selected the user will get familiar with the numbers (from 0 to 9) as in the initial familiarization process.

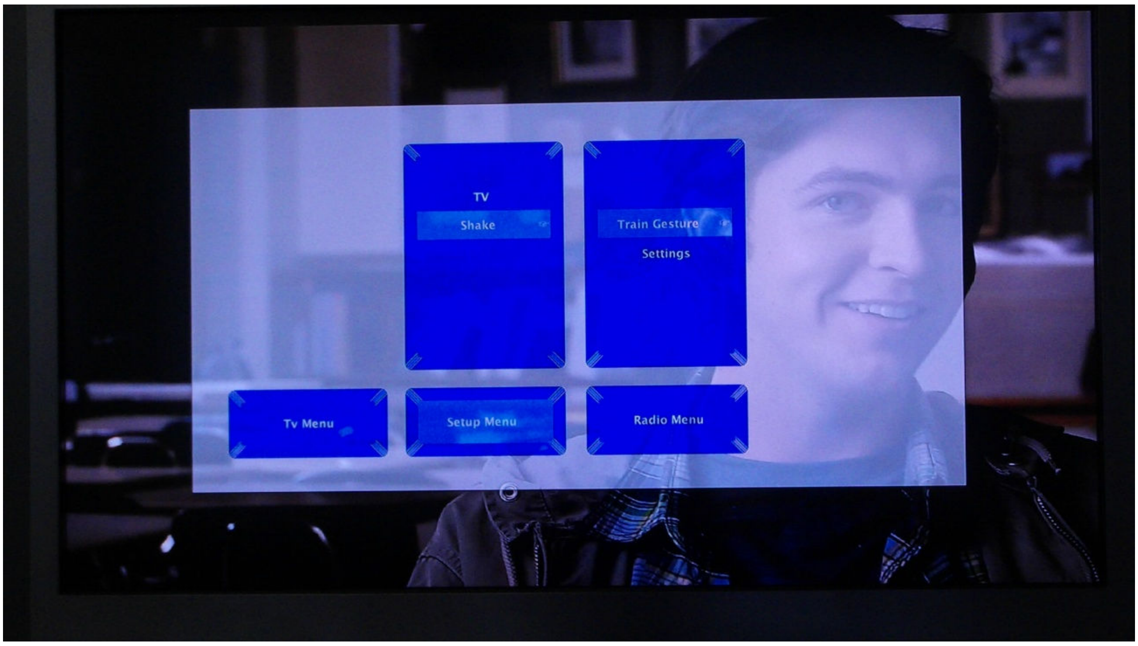


Figure 5.13 Extension of Submenus - I

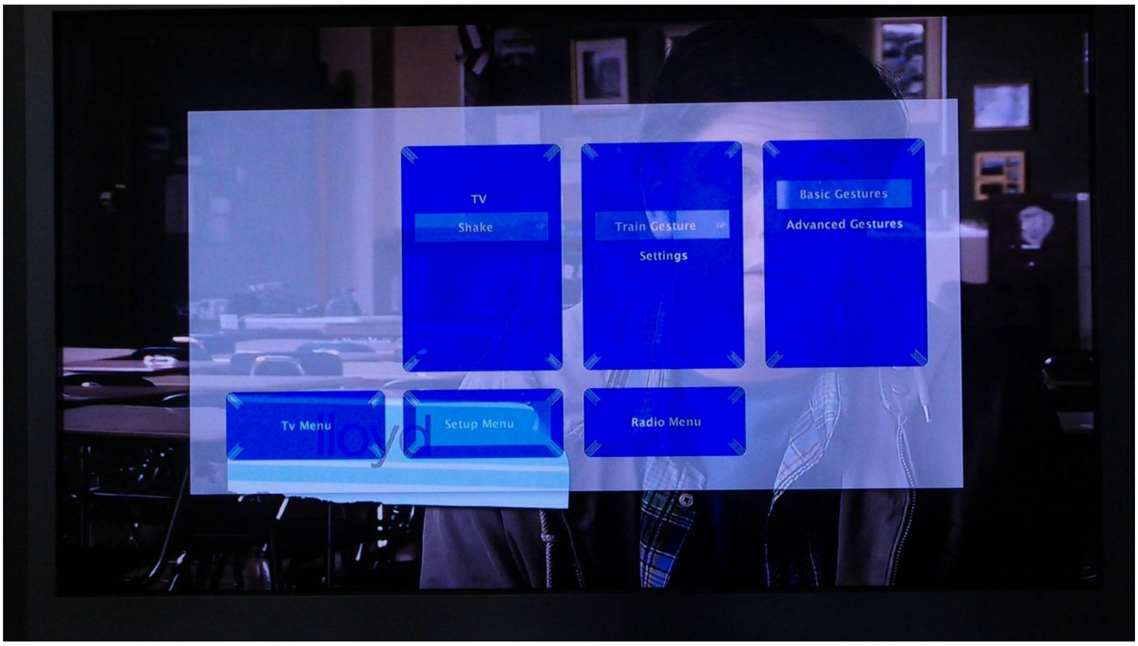


Figure 5.14 Extension of Submenus - II

The following is another scenario where the user wants to see channel lists and choose a channel while watching television. After selecting TV menu by performing “Check” gesture, the user performs “Up” and “Down” gestures to browse channels. To move from channel 1 to 2, perform a “Down” gesture, channel 5 to 4, and perform an “Up” gesture. Figure 5.15 and 5.16, show the states when “Down” gesture performed.

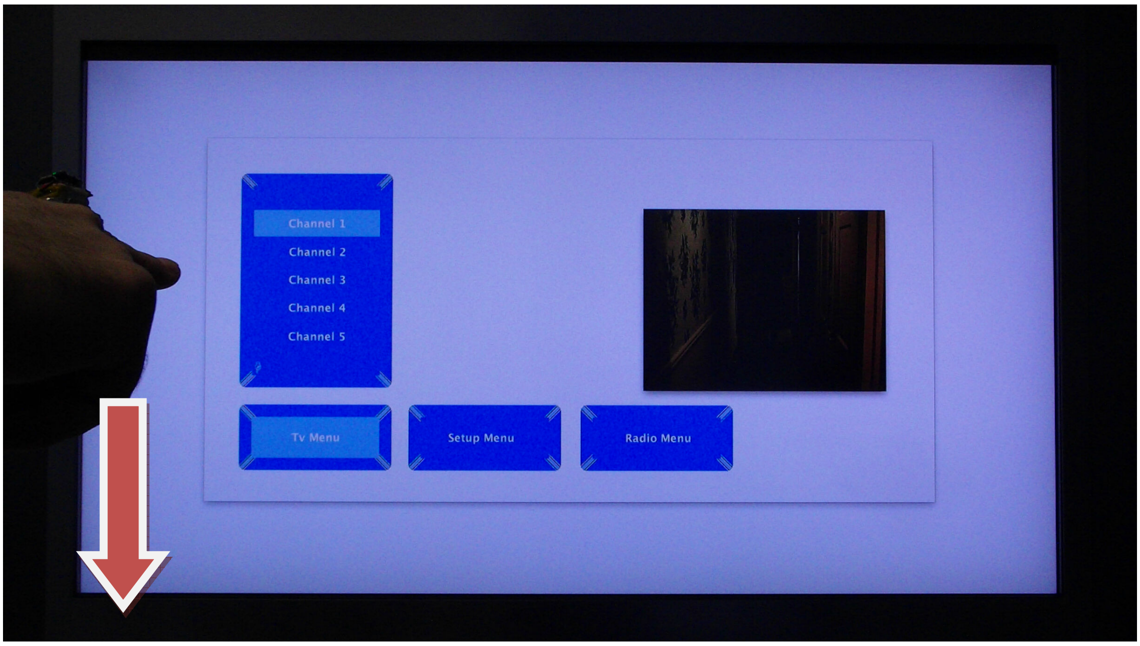


Figure 5.15 Browse channels while watching their previews

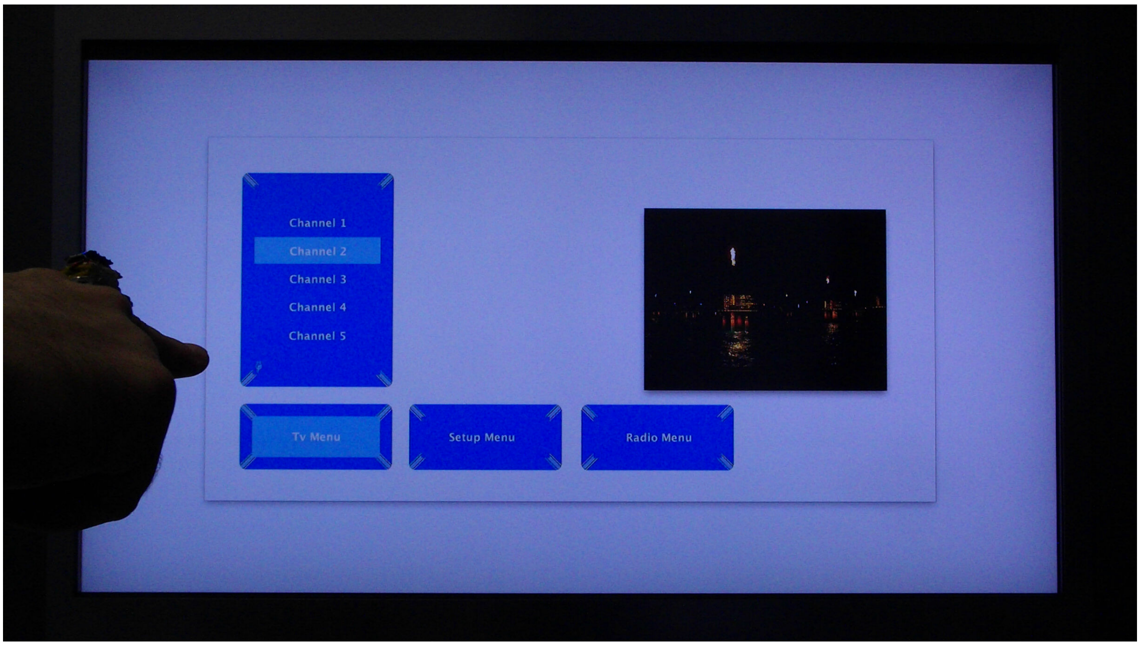


Figure 5.16 Channel Two Preview (After "Down" gesture was performed)

The users performs consecutive "Down" gestures to browse e.g. "channel 6" shown in Figure 5.17 and finally performs a "Check" gesture on the TV menu to select a channel as in the Figure 5.18.

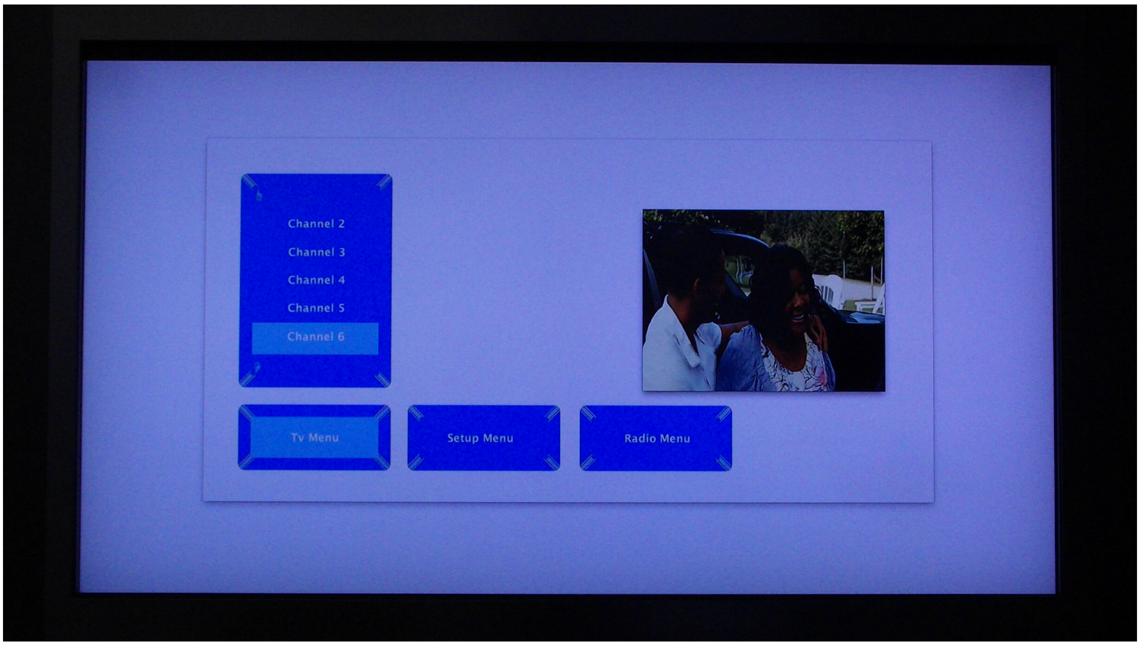


Figure 5.17 Channel Six Preview (Before "Check" gesture was performed)

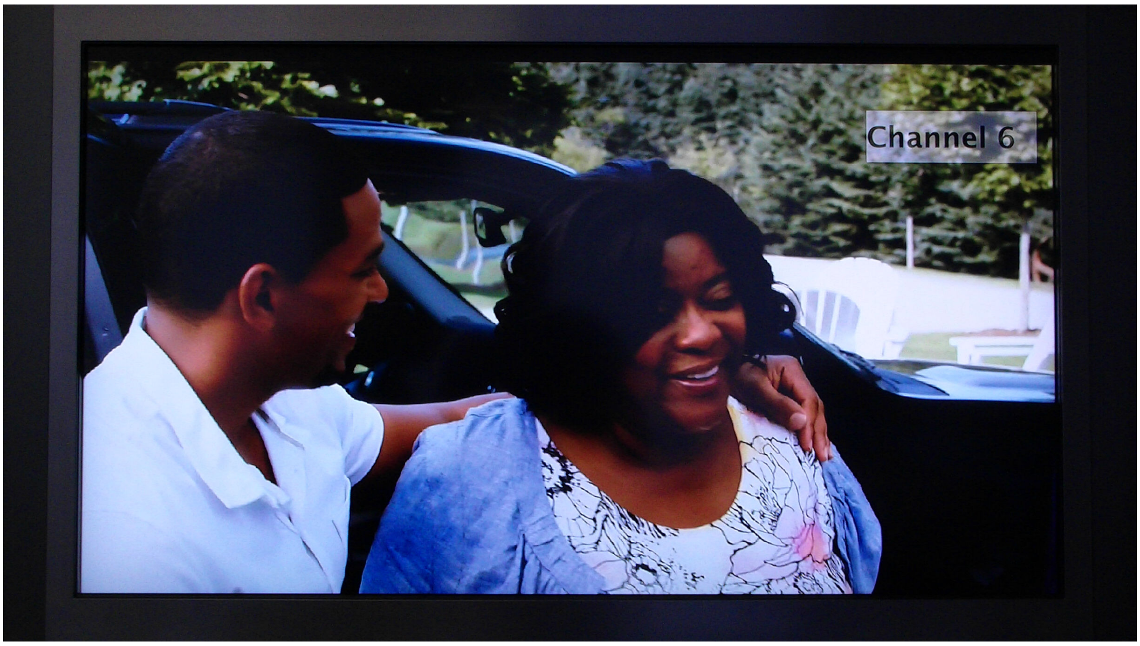


Figure 5.18 Channel Six (After "Check" gesture was performed)

The user wants to watch channel 12. He performs two consecutive gestures, initially gesture "1" followed by gesture "2". The process is shown in Figures 5.19 and 5.20.

Figure 5.19 Selection of a double digit channel - I

Figure 5.20 Selection of a double digit channel - II

As mentioned above, character and text entry is supported in our system. The user can select a channel by performing first character of the channel's name as a gesture or writing the whole name of the channel. Figure 5.21 shows the state when "F" gesture for FOX channel is performed, whereas Figure 5.22 shows when "BBC" gesture is performed.

Figure 5.21 Character Entry (Gesture "F" to watch FOX)

Figure 5.22 Text Entry (Gesture "BBC" to watch BBC)

Another possibility is to control lamps and a fan in our smart-home environment. The fan and one of the lamps are connected to the same relay meaning that the lamp which is connected to the same relay with the fan is on when the fan is off, or vice versa. This lamp and fan is controlled by the gesture "FAN" and the other lamp which is connected to the other relay is controlled by the gesture "LAMP". In Figure 5.23, both lamps are on and fan is off. User performs "FAN" gesture to turn it on as in Figure 5.24. Notice that the lamp is turned off since fan is on.

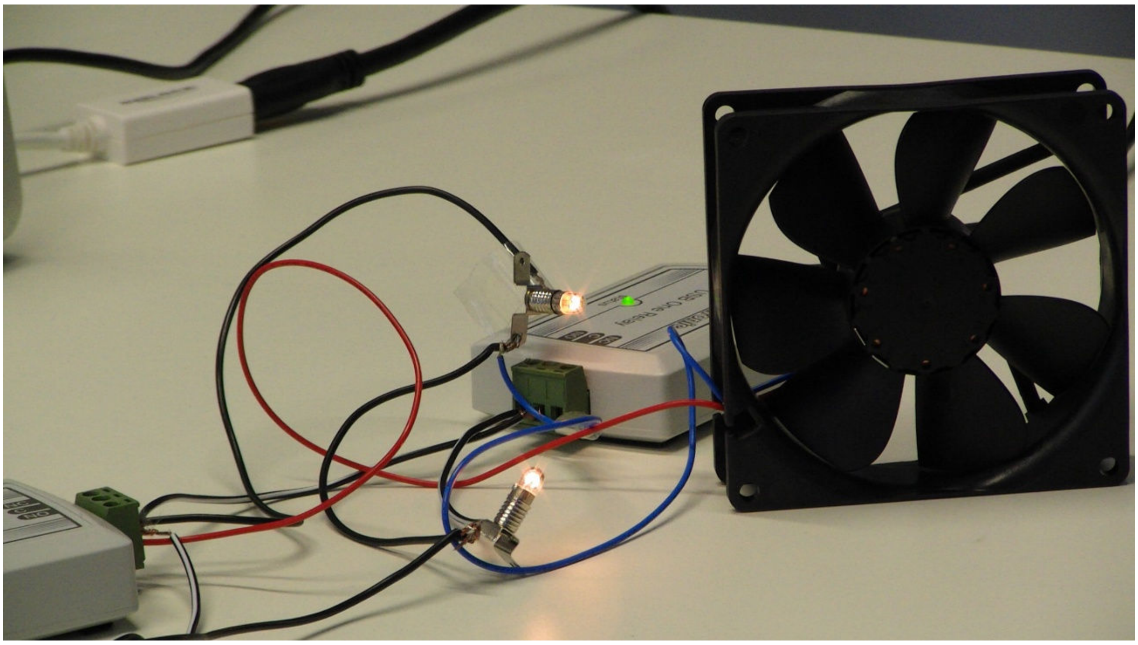


Figure 5.23 Fan is OFF (Before "FAN" gesture was performed)

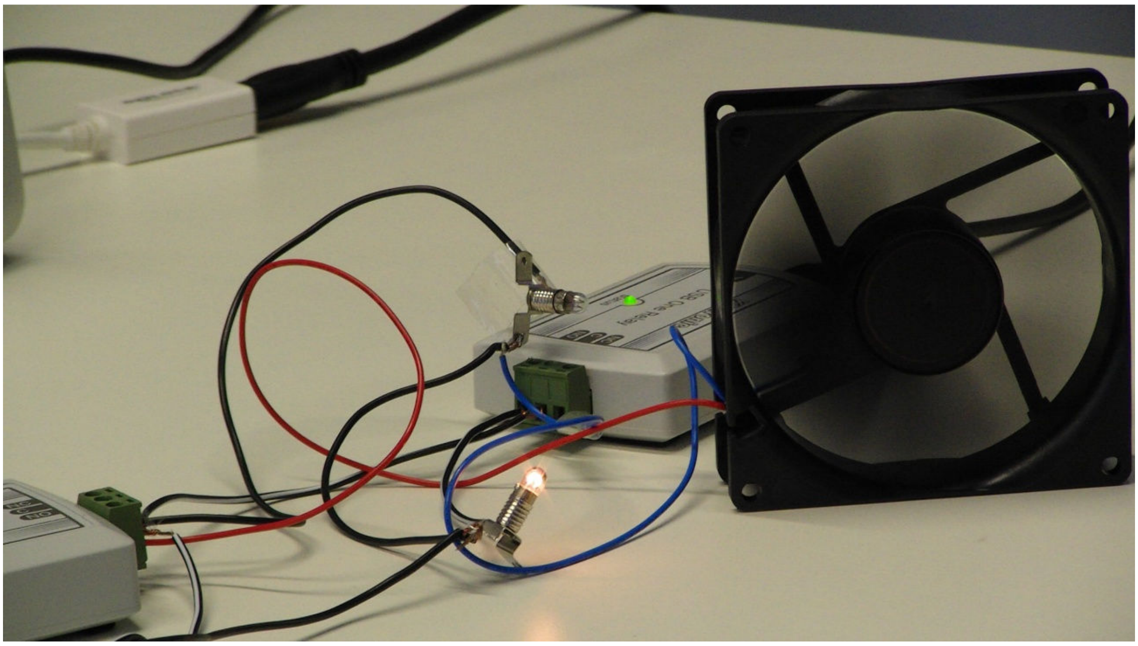


Figure 5.24 Fan is ON (After "FAN" gesture was performed)

Now, imagine that the user wants to turn the light off. He/she performs "LAMP" gesture and the light is turned off as in Figure 5.25.

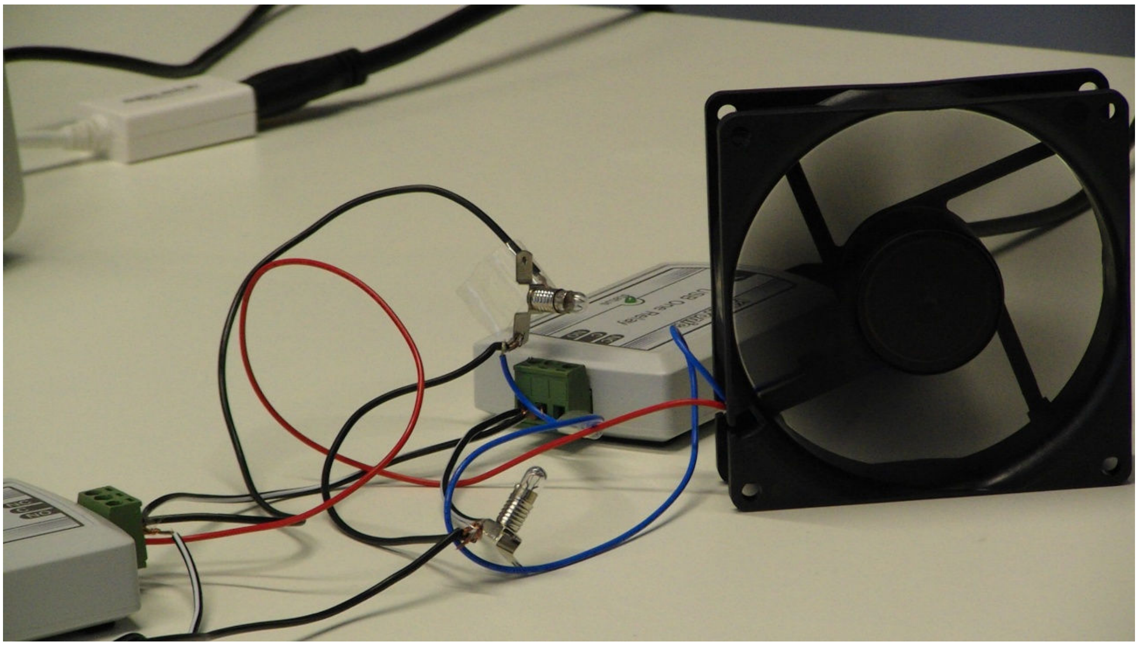


Figure 5.25 Lamp is OFF (After "LAMP" gesture was performed)

6 SUMMARY AND FUTURE WORK

6.1 Summary

People gesticulate often to communicate, yet the idea of controlling objects via gestures has not been sufficiently used in the real world. In this thesis, we have proposed a novel interaction style that allows users to control devices using gestures in a smart home environment.

We reviewed the literature related to gesture recognition to acquire understanding on the current maturity of gesture recognition, existing techniques and projects. After the examination of existing systems and techniques, we decided on how we will proceed with our case and then used a device called SHAKE and afterwards ring-shaped SHAKE called Pingu to manipulate television and interact with a lamp and a fan in a home environment.

SHAKE is a multi-sensor based device equipped with accelerometer, magnetometer, gyro sensor and proximity sensor. A vibration motor and a LED light provide haptic and visual feedback to the user. Navigation switch can be used as an input controller, in our case it indicates the beginning and end of the performed gesture, as it is pushed and released. Later we modified SHAKE due to the obtrusiveness and adapted it into a ring-shape device called Pingu for improved usability and acceptance.

We present Pingu, a new miniature wearable and tangible device in the form of a fingering. Having mostly the same hardware with SHAKE, Pingu has no navigation switch. Other than that, the number of proximity channels is two instead of twelve as in SHAKE. Pingu is accessible to the user for gestural input while he/she continues normal social activities and allows human interaction routines like handshaking.

We implemented prototype applications demonstrating the potential of Pingu as a replacement for traditional user interfaces to control devices in a smart home environment. Interaction with Pingu does not require “line-of sight”. The user can send commands as he or she stays inside the boundaries of the interaction area of the Bluetooth. For instance, the user stays in bed and ready to sleep but just realized the TV and the lamp is on. He or she can easily turn them off without leaving the bed. All he needs to do is to perform related gestures.

6.2 Future Work

6.2.1 System Improvement

- **Hardware Improvement**

  The prototype of Pingu appears to be "hi-tech" and unusual decreasing the user acceptance in daily life. Hence, a better design considering male and female preferences is essential.

  Apart from the design issue, there is another point to ponder over. The cost of the hardware must be reduced to provide people the power to purchase the product.

- **Selection of Target Device**

  The number of target devices in a smart home environment increases, hence the system must provide simple and sensuous control of multiple devices. Before using a device, authentication is required to access the target device. Devices can have predefined authentication gestures. The system can be improved in a way that the user first performs authentication hand gesture to grant permission to use the device and then control it. For instance, if the user wants to use air conditioner, he/she then needs to perform a specific gesture for selecting the device and then another gesture is required for other functionalities.

- **Scenarios**

  Predefined scenarios can be added to the system by the user with the intention of initiating series of actions. As illustrated in Figure 6.1, the user intends to turn the television, air conditioner and the heating system off and draw the curtain closed while lying on the bed. For that purpose, he/she performs predefined gesture to initiate the command.

Figure 6.1 Illustration of a Smart home scenario

- **Provide Feedback & Undo Facilities**

  A good interactive system needs complete feedback. Providing feedback not only enhances the user acceptance, but also it is a necessity for perfect two way interaction. The right "mixture" of audio and visual feedback is indispensable to remind the user about final state of the device and the functionality he is controlling.

  A gesture can sometimes be misrecognized and commands can be issued unintentionally. The command set therefore must include an undo command that gives the user the opportunity to cancel the involuntary commands easily.

6.2.2 Possible Application Areas

Using SHAKE and Pingu as a user interface is a novel technique and we believe it can find applications in several domains. Besides gesture based interaction with smart environments, Pingu can be proposed for human-vehicle interface to decrease the risk of driving by eliminating the problem of taking the eyes off the road to interact with some functionality. Figure 6.2 demonstrates the usage of Pingu in a vehicle while driving to control music in car, interact with gprs and safety switches, check emails, make calls, etc.

Figure 6.2 The idea of using Pingu in vehicles

Pingu can also be used to monitor health conditions of elderly people by the analysis of physical activities, gesture-dialing in the air, playing musical instruments, painting, dancing, browsing photos or even using Pingu as a video game console sounds quite interesting and seems possible. SHAKE and Pingu offer great potential to replace current user interfaces.

REFERENCES

[1] European Commission, (2004). The social situation in the European Union 2004. Office for Official Publications of the European Communities, Luxembourg.

[2] Steg H, Strese H, Loroff C, Hull J, & Schmidt S, (2006). Europe is facing a demographic challenge. Ambient assisted living offers solutions. Report compiled within the Specific Support Action “Ambient Assisted Living – preparation of an article 169-initiative”, Berlin.

[3] OECD Factbook 2005, (2005). Economic, Environmental and Social Statistics, OECD, Paris.

[4] Stefanov DH, Bien Z. and Chul Bang W. (2004). The smart house for older persons and persons with physical disabilities, IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 12(2), 228-250

[5] Pilich B. Engineering Smart Houses [MSc Thesis] Lyngby: Technical University of Denmark; 2004.

[6] Aldrich FK, (2003). Smart Homes: Past Present and Future. In R. Harper (Ed.) Inside the smart home (pp. 17- 39). London: Springer- Verlag UK.

[7] Moeller S, Krebber J, Raake A, Smeele P, Rajman M, Melichar M, Pallotta V, Tsakou G, Kladis B, Vovos A, Hoonhout J, Schuchardt D, Fakotakis N, Ganchev T, and Potamitis I. Inspire: evaluation of a smart-home system for information management and device. In proceedings of the LREC 2004 international conference; Lisbon, May 26-28 2004. P. 1603 - 1606

[8] Starner, T. and Auxier, J. and Ashbrook, D. and Gandy, M, The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring, International Symposium on Wearable Computing, 2000, pp. 87-94.

[9] Mehrabian A. Silent messages. Oxford, England: Wadsworth; 1971. viii 152 pp.

[10] McNeill D. Hand and mind: What gestures reveal about thought. Chicago, USA: University Of Chicago Press; January 15, 1996.

[11] Kendon, Adam. “Current issues in the study of gesture.” In: The Biological Foundations of Gestures: Motor and Semiotic Aspects. Jean-Luc Nespoulous, Paul Perron, and Andre Roch Lecours, eds. Hillside, NJ: Lawrence Erlbaum Associates, Inc. 1986, 23-47.

[12] Jean-Luc Nespoulous, Paul Perron, and Andre Roch Lecours. The Biological Foundations of Gestures: Motor and Semiotic Aspects. Lawrence Erlbaum Associates, Hillsdale, MJ, 1986.

[13] Claude Cadoz. Les realites virtuelles. Paris, France: Flammarion; 1994.

[14] Rime B, and Schiaratura L. Gesture and speech: Fundamentals of nonverbal behavior. New York, USA: Cambridge University Press; 1991.

[15] Robert Y. Wang and Jovan Popovic, Real-Time Hand-Tracking with a Color Glove, ACM Transaction on Graphics (SIGGRAPH 2009), 28(3), August 2009.

[16] Lenman S, Bretzner L, Thuresson B. Using marking menus to develop command sets for computer vision based hang gesture interfaces. In proceedings of the second Nordic conference on Human-computer interaction, NordCHI '02; Denmark, 10/02/2002. P. 239-242.

[17] Morency LP and Darrel T. Hand gesture recognition in intelligent interfaces: The role of context in improving recognition. In proceedings of the 11th international conference on Intelligent user interfaces, IUI '06; Sydney, Australia; January 2006. P. 32-38.

[18] Morris M, Huang A, Paepcke A, and Winograd T. Cooperative gestures: multi-user gestural interactions for co-located groupware. CHI 2006: Proceedings of SIGCHI conference on Human Factors in computing systems, New York, NY, USA, ACM Press, pp. 1201 - 1210.

[19] Malik S and Laszlo J. Visual touchpad: a two-handed gestural input device. In R. Sharma, T. Darrell, M.P. Harper, G. Lazzari, and M. Turk, editors, ICMI, pages 289–296. ACM, 2004.

[20] Lenman S, Bretzner L and Thuresson B. Computer Vision Based Recognition of Hand Gestures for Human-Computer Interaction. Technical report TRITANA-D0209, CID-report, June 2002

[21] Zimmerman TG, Lanier J, Blanchard C, Bryson S, and Harvill Y. A hand gesture interface device. In Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface - CHI '87, pages 189–192, New York, New York, USA, 1987. ACM Press. 7, 12

[22] Baudel T and Beaudouin-Lafon M. Charade: remote control of objects using free-hand gestures. In Communications of the ACM, volume 36, pages 28–35, July 1993.

[23] 5DT. 5DT data glove. http://www.5dt.com/products/pdataglovemri.html, 2005. (Accessed June 2011)

[24] Park I, Kim J, and Hong K. An Implementation of an FPGA-Based Embedded Gesture Recognizer Using a Data Glove. In Proceedings of the 2nd international conference on Ubiquitous information management and communication, pages 496–500. ACM, 2008a. 7, 8

[25] Fakespace Systems Inc. Pinch glove. http://www.fakespacelabs.com/tools.html, 2007. (Accessed June 2011)

[26] Harrison C and Hudson SE. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. In Proceedings of the 22nd

Annual ACM Symposium on User interface Software and Technology (Victoria, British Columbia, Canada, October 4 - 7, 2009). UIST '09. ACM, New York, NY. P. 121-124.

[27] Ashbrook D, Baudisch P, White S: Nenya: subtle and eyes-free mobile input with a magnetically-tracked finger ring. CHI 2011: 2043-2046

[28] Ketabdar H, MagiThings: Interaction with Mobile Devices via Magnetic Sensor, available online as:
http://www.deutsche-telekom-laboratories.de/~ketabdar.hamed/MagiThings/ (Accessed June 2011)

[29] Ketabdar H, Yüksel KA, and Roshandel M, "MagiTact: interaction with mobile devices based on compass (magnetic) sensor" Proceeding of the 14th international conference on Intelligent user interfaces, Hong Kong, China: ACM, 2010, P. 413-414.

[30] Mäntyjärvi J et. Al. (2004), Enabling fast and effortless customization in accelerometer based gesture interaction, Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia

[31] Mäntyjäarvi J, Kela J, Korpipää P, Kallio S, Savino G, Jozzo L and Marca D. Accelerometer-based gesture control for a design environment. Personal Ubiquitous Computing, Springer London (2006), P. 285–299.

[32] Kim D, Choi H, and Kim J. 3d space handwriting recognition with ligature model. In Ubiquitous Computing Systems, volume 4239/2006 of Lecture Notes in Computer Science, pages 41–56. Springer Berlin / Heidelberg, 2006.

[33] Amft O, Amstutz R, Smailagic A, Siewiorek D, and Tröster G. Gesture-controlled user input to complete questionnaires on wrist-worn watches. In Human-Computer Interaction. Novel Interaction Methods and Techniques, volume 5611 of Lecture Notes in Computer Science, pages 131–140. Springer Berlin / Heidelberg, 2009.

[34] Hein A, Hoffmeyer A, and Kirste T. Utilizing an accelerometric bracelet for ubiquitous gesture-based interaction. In Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments, volume 5615 of Lecture Notes in Computer Science, pages 519–527. Springer Berlin / Heidelberg, 2009.

[35] Miyamae M, Terada T, Tsukamoto M and Nishio S. Design and Implementation of an Extensible Rule Processing System for Wearable Computing. In Proceedings of MobiQuitous. 2004, P. 392-400.

Appendix A

```
double computeDTW(ArrayList test, ArrayList trained) {
    int xSize = test.size();
    int ySize = trained.size();
    double[][] dtwArray = new double[xSize][ySize];

    // initialization of the DTW array
    dtwArray[0][0] = 0;
    for (int i = 1; i < xSize; i++) {
        dtwArray[i][0] = 999999999;
    }
    for (int i = 1; i < ySize; i++) {
        dtwArray[0][i] = 999999999;
    }


    // compute the dtw
    for (int i = 1; i < xSize; i++) {
        for (int j = 1; j < ySize; j++) {
            dtwArray[i][j] = (((Vector) test.get(i - 1)).subtractVector((Vector)
(trained.get(j - 1)))).length
                + Math.min(Math.min(dtwArray[i - 1][j], dtwArray[i][j - 1]), dtwArray[i
- 1][j - 1]);


        }
    }
    //System.out.println(dtwArray[xSize-1][ySize-1]);
    return dtwArray[xSize - 1][ySize - 1];

}
```

Appendix B

Gesture Recognition User Study Test
11-22July 2011

Personal Details form subject No.
Date:

Note: You can ignore any question if you do not want to answer.

1. Age:

0-20 years old
20-40 years old
40-60 years old
60+ years old

2. Gender:

Male
Female

3. Which hand do you write with?

Left hand
Right hand

4. Your most recent job:

5. Your level of education

Elementary School
Secondary School
College
Undergrad degrees
Postgrad degrees

6. What type of mobile devices do you use?

TouchScreen
Regular with Keypad
Others:

Have you use any touch screen device before?

7. If you have such technology as a product where you would like to use it for making gestures?

Air
Palm
Table

8. Your impression about the technology under the test:

9. What specific application you suggest for this technology? Please briefly describe! We will give a prize to the best idea!

Appendix C

Gesture Datasets

- A = Air
- T = Table
- P = Palm
- Im = Interaction with Magnet

1

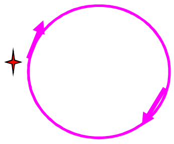
Order of the test

- General gesture
- Genetal gesture with magnet
- Activity set
- Intra Hand
- Intra Finger
- Digit Entry
- Signature

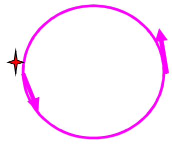
2

General Gesture (A-T-P)

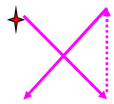
1) Circle-clockwise



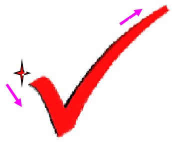
2) Circle-counter clockwise



3) X



4) Check



5) Right one hand-2 times

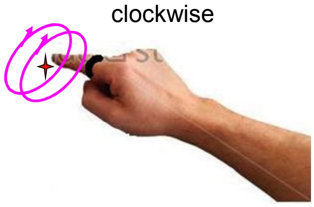
6) Left one hand-2 times

3

General Gesture (A-T-P)

7) Double click on the screen

8) Turn the finger around itself-clockwise

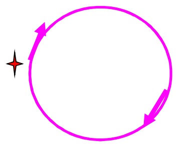


9) Turn the finger around itself-counter clockwise

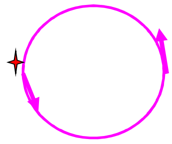
4

General Gesture with magnet (Im)

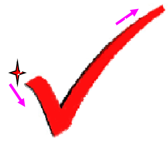
1) Circle-clockwise



2) Circle-counter clockwise



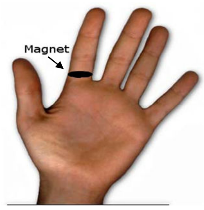
3) Check



4) Right one hand-2 times

5) Left one hand-2 times

Writing on the palm of a hand with magnet facing to palm



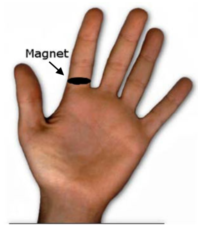
5

Activity Set

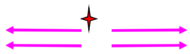
1. Typing on the Mac Keyboard(keyboard)
2. Typing on the ipad Screen (ipad)
3. Relax and rest for 1-2 min. (relax)
4. Running on the spot (running)
5. Walking (walking)
6. Shaking hand (handshaking)

6

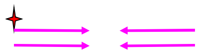
Intra Hand



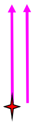
1) Zoom in- 2 times



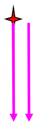
2) Zoom out -2 times



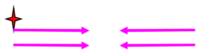
3) Upward 2 times (palm down)



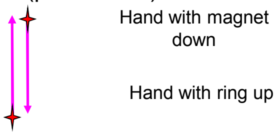
4) Downward 2 times (palm down)



5) Hand clapping



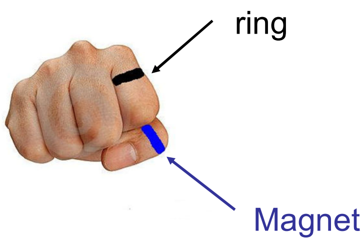
6) Reverse movement 2 times (palm down)



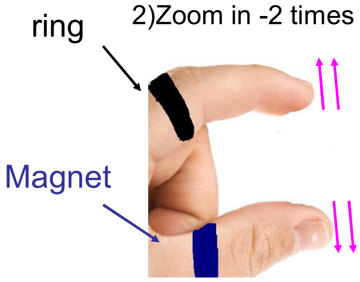
7

Intra Finger

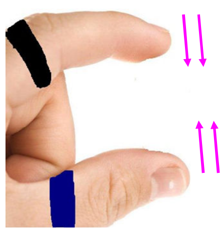
1)Punching 2 times



2)Zoom in -2 times

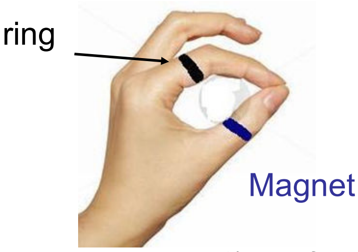


3)Zoom out -2 times



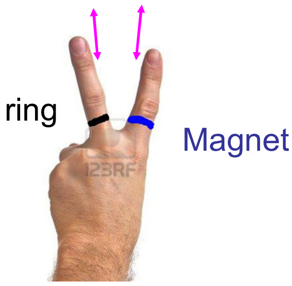
Intra Finger

4)hitting-2 times



5)pinching-no magnet-2 times

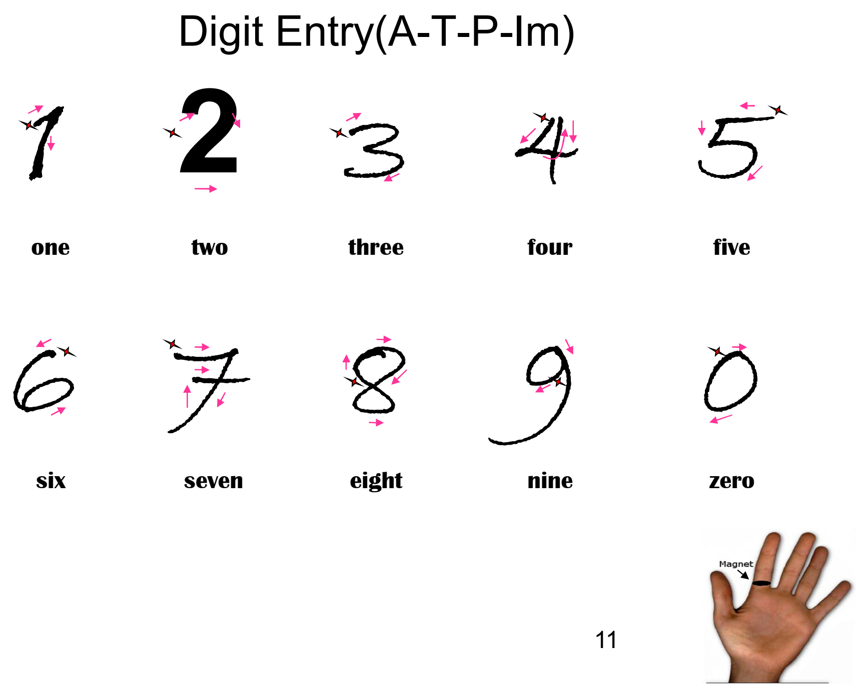
6) Two fingers go up and down – 2 times



9

Signature(A-T-P)

10


Digit Entry(A-T-P-Im)
one
two
three
four
five
six
seven
eight
nine
zero
Magnet
11