High resolution, Low cost, Privacy Preserving Human Motion Tracking System via Passive Thermal Sensing

by

Sharmeen Browarek

B.S., Electrical Engineering, M.I.T., 2008

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2010

Author .....
Department of Electrical Engineering and Computer Science
February 2, 2010

Certified by .....
Ramesh Raskar
Associate Professor, MIT Media Laboratory
Thesis Supervisor

Accepted by .....
Dr. Christopher J. Terman
Chairman, Department Committee on Graduate Theses

High resolution, Low cost, Privacy Preserving Human Motion Tracking System via Passive Thermal Sensing

by

Sharmeen Browarek

Submitted to the Department of Electrical Engineering and Computer Science
on February 2, 2010, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Thermal imaging is powerful but expensive. This thesis presents an alternative thermal sensing system capable of tracking human motion by using a novel projection mechanism from an array of inexpensive single-bit thermal sensors. Since thermal sensor performance is currently limited by lengthy refractory periods, the motion tracking system is further developed in the visible spectrum to demonstrate a proof of concept. A single array of sensors is coupled with Gray-coded binary masks to spatially divide a room, allowing for the tracking of a moving persons location. By analyzing the output over time, a motion flow map is rendered and displayed in a software graphical user interface. The software is also capable of preliminary two-person motion tracking schemes. This overall system provides a low-cost alternative for privacy preserving activity monitoring, and generating alerts based on data anomalies, such as a break in motion pattern or an unusually low level of overall activity. Future hardware and software advancements will allow a full-scale motion tracking system to be developed.

Thesis Supervisor: Ramesh Raskar
Title: Associate Professor, MIT Media Laboratory

Acknowledgements

I owe my deepest gratitude to Ramesh Raskar, for his continuous guidance and support as my thesis advisor over the course of the past year. It has been an honor to work with and learn from such a gifted researcher.

Particular thanks to Raphaële Thévenin, my partner for the thermal application portion of this project. I could not have asked for a better start to my research.

I would also like to thank Dennis Miaw, for his patience and hours of help with debugging my circuits. And Douglas Lanman for his amazing MATLAB skills, helping along the way with my time-dependent visualization software. And also to Ahmed Kirmani, for his guidance with the two-person tracking algorithm. Lastly, thanks to Rohit Pandharkar, my officemate, for letting me talk through my research problems and brainstorming new ideas with me to further my work.

I am grateful to the many people and companies that helped with the extensive thermal sensor research that went into this project:

- Professor Joe Paradiso and Mark Feldmeier of the Responsive Environments group at the MIT Media Laboratory
- Professor Neil Gershenfeld, Director of MIT's Center for Bits and Atoms
- Dr. Hideaki Nii, author of the Prakash project, which has inspired so many aspects of this thesis
- Matthias Wagner of RedShift Systems
- Stuart McMuldroch of Blue Sky Imaging Ltd.
- Stewart Hall of Tyco International
- Bob Reich, Dan Schuette, and Dr. Andrew Siegel from Lincoln Laboratory

To Joy Ebertz, Ahmed Kirmani, and Daniel Schultz, thank you for taking the time to read and edit my thesis.

Finally, to my family and close friends, thank you so much for all of your love and support over these months. Most importantly, I especially dedicate this thesis to my mother, for her endless sacrifices and dedication to making all of my success and achievements possible.

Contents

List of Figures

Chapter 1

Introduction

This thesis presents a novel technology that delivers a motion flow map of a person's movements through a room and potentially generate warnings or alerts if a break in the motion pattern is observed. This is achieved by using thermal IR sensing combined with the theory of spatial Gray coding and powerful algorithms for computational post-processing.

1.1 Overview

The technology put forth by this thesis is centered around two key points: using single pixel sensors to detect the thermal InfraRed (IR) radiation emitted from a human body, and coupling these sensor outputs with Gray-coded binary masks to spatially divide a room. One Gray-coded mask is needed per sensor to divide a room into zones of detection and non-detection. The output of the system, at any given time t, is a binary code that encodes the position of a person in the room. By storing this data in a series of consecutive time sequences, the different locations are extracted from the binary codes. This information is used to create a motion flow map of the user's movement through the room and compute the velocity at which the user is moving.

1.2 Related Work

The majority of current motion tracking technologies focus on research in software for human tracking and pattern recognition from images captured by ordinary cameras [1-4]. Instead of focusing on software, this thesis presents a new technology that uses IR sensing to capture data for pattern recognition analysis. The algorithms for the data analysis in this system are drastically different from existing systems, but the motion pattern recognition still involves techniques similar to those described in previous papers.

A variety of technologies can be used to implement motion tracking systems, including mechanical sensing, acoustic sensing, magnetic, optical, and radio frequency sensing [5]. The most common commercial methods used to track people are based on either acoustics, like the AmbienNet technology [6], radio frequency signals, such as Microsoft's Radar research project using RFID tags [7], or optical devices. The advantage of using optical technology is that these systems tend to be faster and more accurate when compared to others.

A majority of current optical human tracking systems involve the user needing to wear a device. For example, the HiBall Presence technology [8] requires its clients to wear a head-mounted sensor, which captures IR light emitted by LEDs that have been attached to the ceiling. In order to have accurate data collection, many sensors have to be installed on the ceiling beforehand, in precise locations. Another type of system that accomplishes full-body motion tracking is the Prakash technology [9]. In Prakash, the user wears clothing embedded with photosensor tags that detect emissions in the near-IR wavelengths. Projectors located around the room emit coded light signals that are picked up by the sensors on the person. The captured data is transmitted wirelessly and analyzed to output the 3D positions of the sensors. Similarly to HiBall, Prakash also entails setting up projectors around the room in very specific positions to perform 3-d motion capture. Yet another approach comes from the Active Badge System [10], which involves badges that emit a unique IR signal at 10 second intervals. At known locations throughout the building, sensors are situated that can detect the

signal from the badge. Any user wearing the badge can be tracked in all parts of the building, but at much lower resolutions than the previous two systems. Fairly accurate location information can be obtained from all of these techniques, but they have many disadvantages, such as requiring the user to wear sensors or tags as tracking devices.

A slightly different, yet popular approach, involves motion tracking with tags using high speed cameras that observe passive visible markers or active light emitters [11-15]. However, this technology is extremely expensive and still entails the user wearing the markers.

Finally, a unique type of monitoring has been installed in smart homes to track the daily activities of people. Individual sensors are placed on crucial devices and appliances that are used often daily. Reports from these sensors would generate patterns indicating whether a person is keeping their autonomy in various rooms of a house, checking on their activity throughout the home and generating alerts if rate of activity diminishes. Yet, this this technology has several drawbacks. Even though it gives information about a user's daily behavior, it cannot detail the motion pattern of a person in their home. If someone is struggling to move around, or if they have fallen and need immediate assistance, this system would not be ideal. Instant feedback is a crucial element when monitoring a patient in their home, or applying such a system as a security device.

Much work has been done in the computer vision area of research on developing software for human motion tracking. This thesis puts forward a unique method with distinct advantages over existing research in the area.

1.3 Applications

Motion capture is a relevant technology in the market for monitoring elderly people in their homes. As times progress, older people are increasingly reluctant to move into nursing homes. The factors driving this shift include both the expensive costs

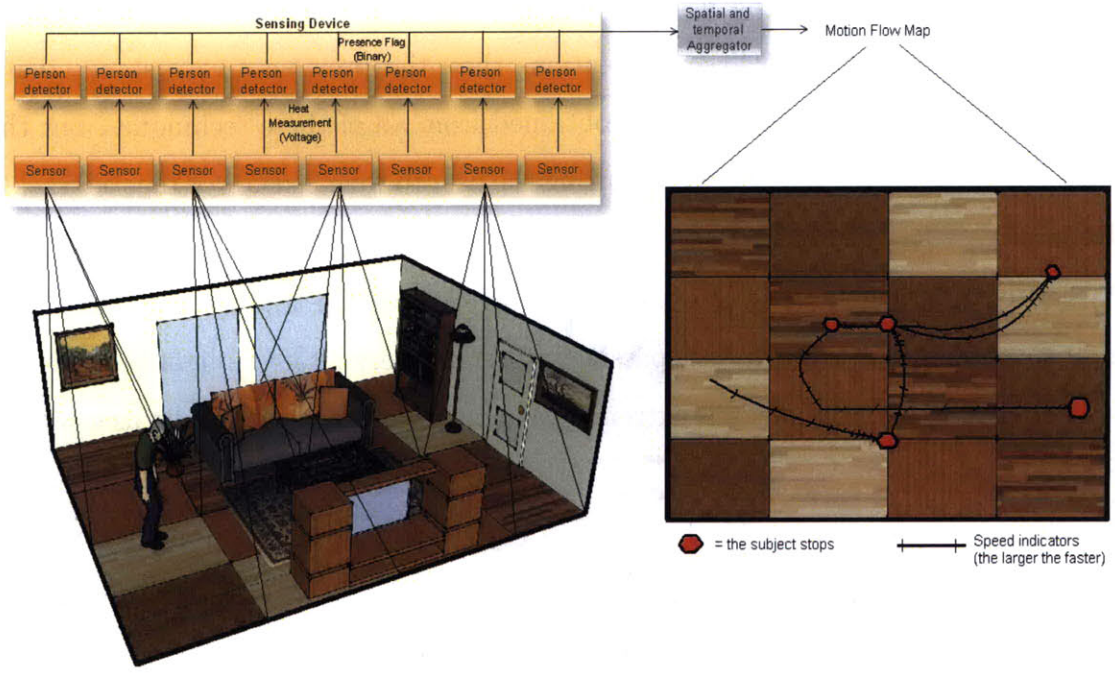

Figure 1-1: Example of a usecase scenario of an elderly user monitored by this technology. The system detects the presence of the user in voltage levels, converts the results to binary, decodes the locations over time, and renders the motion flow map. Scenario in this example: 1. Goes to TV, 2. Goes to sit on couch, 3. Goes to bookshelf, 4. Returns to sit on couch, 5. Heads towards door.

of nursing homes or hospitals and the desire to remain in the comfort and familiarity of one's home. However, close family members worry for the health of their elderly relatives, especially when there isn't anyone to constantly watch over them. This is where motion capture technologies come into play. The perfect solution to this concern would involve a non-invasive method of tracking senior citizens; something that could immediately alert families when assistance might be needed. An example of how this device would work and what a rendered output might look like is illustrated in Figure 1-1.

This device would not only be useful in personal homes, but also in hospitals as well. In situations where hospitals are short of staff, nurses and doctors would be able to monitor a patient's motion remotely and be warned of slight changes in a patient's behavior.

In addition, this system could function as a security monitor in locations where privacy

must be preserved. For example, lock box vaults in banks are not allowed to take any visual footage, so security cameras are out of the question. Yet, customers still want to feel that their valuables are fully protected from theft. In such cases, the device presented in this thesis provide an ideal monitoring technique.

1.4 Motivation

The method for motion capture described by this thesis has many advantages over existing technologies: it is privacy preserving, completely passive, independent of visible illumination in a room, low cost compared to thermal IR cameras, and the system expands on an order of $\log_2(n)$ as opposed to n^2.

Privacy Preserving

This technique for motion tracking is entirely privacy preserving, absolutely no images are captured. Typical cameras may carry more information, but are extremely intrusive. Specifically in the case of monitoring the elderly, families want the comfort of knowing their loved ones are safe, but imaging is not necessary to acquire such a goal. Furthermore, it is important that users maintain their privacy within their homes. The output of this device consists solely of a motion pattern on a screen. The detectors are sensitive purely to motion, rather than intensity. For example, thermal IR cameras are able to detect intensity levels in the thermal range, and use that information to form a thermal image of the person. There is no image capture in this device; the user is represented by a series of colored blocks and remains fully anonymous.

Passive

Another distinguishing factor of this system is that it is completely passive. Since thermal detectors sense the heat radiated from a human body, the user does not need to wear a badge to be tracked. Hence, elderly users are not inconvenienced by wearing a tracking device, and there is no concern as to whether or not the user remembered

to put on the device. Plus, if applied as a security device, intruders would not break in while wearing a marker or tag. This system would be able to catch anyone that crosses its field of view.

Environment Independent

Another benefit of this technology includes that it functions independent of the visible illumination in the room, whether it be natural sunlight or indoor lighting. Since the outputs of the sensors are differenced in processing, the system only detects thermal radiation in motion. Thus, warm and stationary objects in the background scene do not obstruct the final data. The device is not affected by the environment and operates equally well in all conditions.

Cost-Effective

Thermal IR cameras today cost approximately $10,000. They consist of arrays of thermal detector pixels with nitrogen cooling units to speed up response times. An image of the scene is taken based on thermal variations of the objects in the camera's field of view. This device will cost roughly $200 and is made entirely of off-the-shelf components. Such a low cost range makes this device a feasible purchase for families and hospitals for tracking, and for banks or museums for security purposes.

Scalable

The idea of coupling outputs of single pixel detectors with Gray-coded masks allows the system to benefit from the accuracy and lower latencies of optical technology. this technology offers a meaningful compression of motion data because the output necessitates a very low bit rate and no significant post processing is required. More information is retrieved from each coupled detector than could be attained from those that are independently utilized. Furthermore, for n number of sensors, the number of possible spatial divisions in this system is 2^n. This means that expanding the system occurs on an order of $\log_2(n)$ sensors rather than n^2 sensors. Fewer sensors are needed

to have the same resolution as other systems, making this device more efficient and scalable.

1.5 Structure of Thesis

- Chapter 1 motivates the thesis, describes the background and related work in the field, and explains the targeted applications and goals of the system.
- Chapter 2 focuses on the theory behind the thesis; it gives the details of the thermal IR and visible spectrums, and introduces Gray coding.
- Chapter 3 presents the design of the thermal spectrum system, including the mechanical device, analog signal processing, data collection, and visualization tool.
- Chapter 4 examines the visible spectrum system and its parallel components to the thermal IR system. It delves into the mechanical device, intricate electronic circuitry, and the data collection.
- Chapter 5 illustrates the software involved in the project. It shows how a user would interface with the graphical software client, describes how the data is processed, and brings insight to the two-person tracking algorithm developed for the device.
- Chapter 6 concludes the thesis and expresses areas of future work and expansion for the motion tracking system.

Chapter 2

Theory

2.1 Thermal IR Spectrum

The thermal IR spectrum lies just outside the range of light visible to the naked eye. The full IR spectrum ranges from 700 nm up to 1000 μm. This spectrum is divided into four sections: the near-IR, the short-wavelength IR, the mid-wavelength IR, the long-wavelength IR, and the far-IR spectrums. The human body has an average temperature of 37 °C, so it emits heat at a wavelength of approximately 9.5 μm, which lies in the long-wavelength IR range. A number of thermal sensors can detect radiation emitted in this range.

2.1.1 Sensor Models

Sensors in the thermal IR range come with varying strengths and drawbacks. Two key factors needed in an ideal thermal sensor for this application are a far detection range and a fast refractory period. Since the goal of this device is to detect a person's motion through a room, the ideal detection range is 10-15 ft, covering the lengths and widths of average sized rooms. The refractory period is the amount of time it takes for a sensor's output voltage to return to neutral after it has sensed a heated body. This is the largest drawback in thermal technology today. Due to their sensing material, most thermal detectors have extremely slow response times. Few solutions exist,

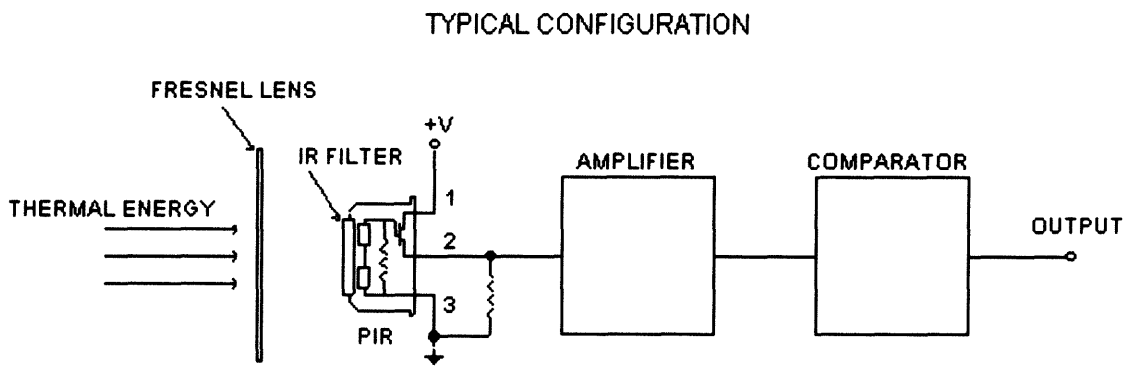

Figure 2-1: PIR module block diagram.

such as those used in thermal IR cameras or microbolometers, but these solutions are applied to arrays of hundreds of sensors and require very expensive nitrogen cooling units in tandem. This project also strives to be cost-effective and easily reproducible, so such units are not an option. Most commonly, thermal sensors are able to achieve one of these two key factors, but rarely both at the same time. The following sections analyze the in-depth research done on thermal sensing technologies and their outcomes relative to this system.

PIR modules

The first detectors used for the thermal system are PIR modules, Passive InfraRed sensors with an integrated circuit attached to do the signal processing. The block diagram for the full module, shown in Figure 2-1, consists of a sensor followed by an amplifier and comparator circuit. The final output of the integrated circuit is a binary value: either 5 V high or 0 V low. The sensor used in this module is the Pyroelectric Passive IR Sensor model RE200B made by NICERA. More specifically, the RE200B is a differential pyrosensor, which means that it has two sensing elements, as illustrated in Figure 2-2. The two sensing elements allow the differential pyrosensor to only detect thermal radiation that is in motion. These sensors have a transmission rate higher than 70% in the range of 7-14 μm. On average, the background of a room is typically at a temperature of 20 °C, which translates to an IR energy of approximately 10 μm. This wavelength is well within the detection range of the pyrosensor. Thus,

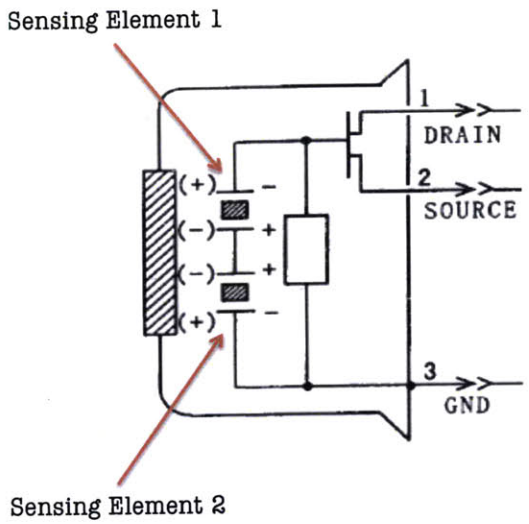

Figure 2-2: Two sensing elements of a differential pyrosensor.

the detectors do not only sense a human body's thermal radiation, but also the background scene behind a person. However, having two sensing elements prevents the background scene from affecting the output. The two sensing elements, or pixels, are oriented with opposite charges relative to each other. If an IR emitting object remains motionless in front of the sensor for a given time t, the two pixels would detect exactly the same sensor response over period t. Since one is oriented negatively and the other positively, their overall output would be zero. If the radiating body is in motion, though, the pixels no longer have the same output. The first pixel will detect the emission at time t, and the second pixel will detect the radiation at time $t + \delta t$. Consequently, only thermal bodies in motion are detected by differential pyrosensors. Figures 2-3 and 2-4 demonstrate an example of how the differential pyrosensor would function if a heated body were to walk in front of it. One advantage of the pyrosensor is its ability to detect the radiation of moving bodies from distances up to 12 ft away. This comes from its large field of view (FOV). The x-axis has a FOV of 138°, and the y-axis has a FOV of 125°. Figure 2-5 illustrates the orientation of the two sensing elements in the pyrosensor. The figure shows a view looking down on the round sensor casing from above, and the two rectangular boxes represent the two sensing elements within the detector. Figure 2-6 depicts how the FOV differs for the sensing elements

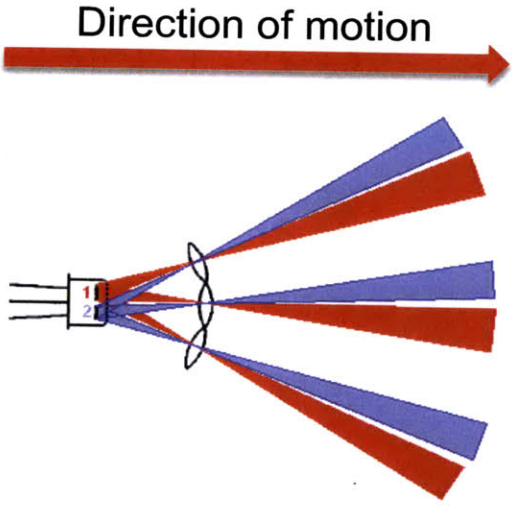

Figure 2-3: Direction of movement as a thermal body crosses in front of a pyrosensor. Red and blue strips represent the zones of detection for each sensing element.

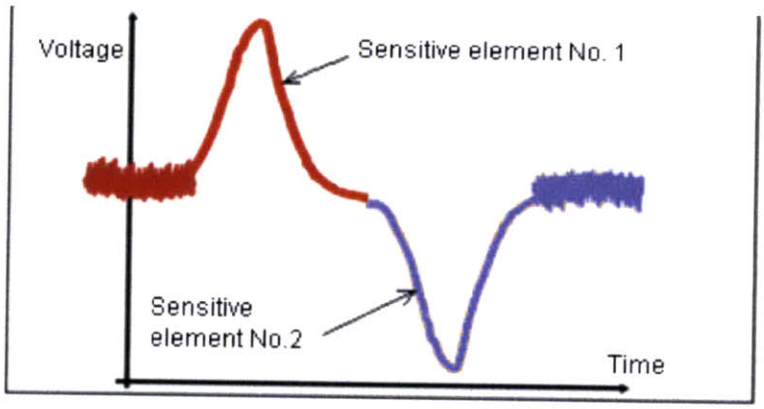

Figure 2-4: Example voltage output of a pyrosensor over time, after a thermal body has crossed in front of it, as shown in Figure 2-3.

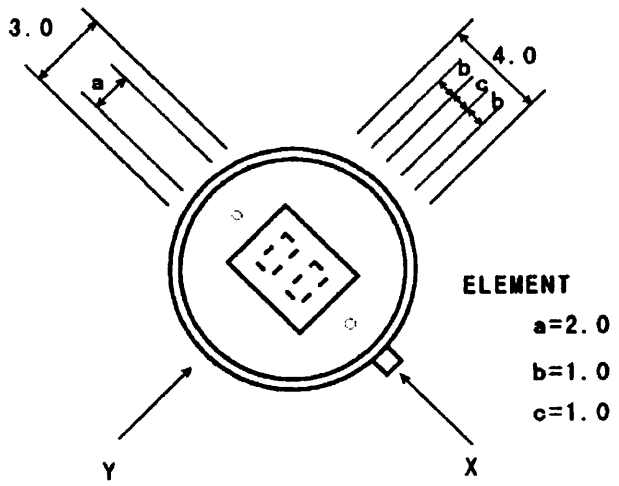

Figure 2-5: Top view of the pyrosensor, showing x and y axes fields of view.

Figure 2-6: Pyrosensor field of view for the x and y axes directions.

with respect to the x-axis and y-axis orientations.

As mentioned previously, thermal sensors do not currently possess both the key factors in the same detection device. Although the detection range of a differential pyrosensor is within acceptable limits, the refractory period is lacking. At 3.5 sec, the refractory period for the PIR module is much slower than most thermal sensing options. More details on the circuitry implemented to mitigate this issue can be found in Section 3.2.1.

Thermopiles

Thermopiles are another type of sensor tested for for this device; they are detectors that turn thermal energy into electrical energy. They consist of thermocouples, which output a voltage level relative to the temperature they sense in the environment.

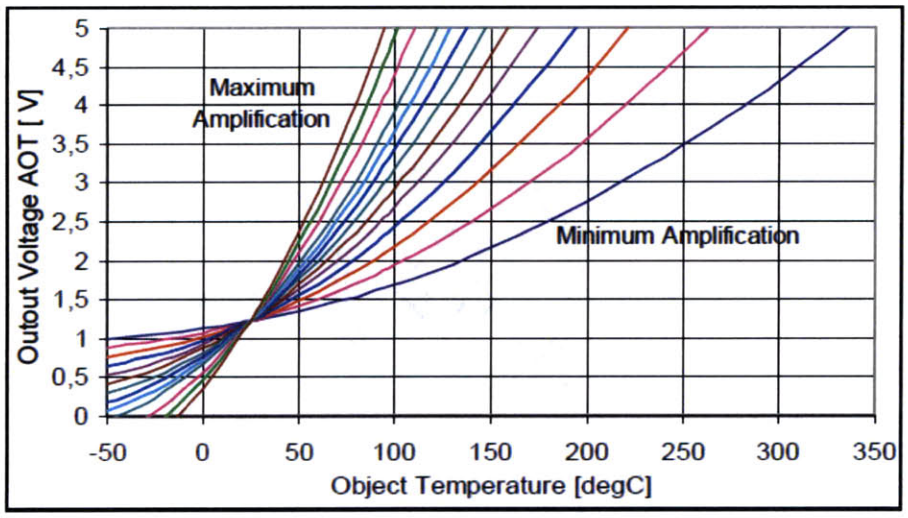

Figure 2-7: Sensing range of the Heimann Sensor HTIA series. The colored lines represent the output amplification of sensed wavelengths. The higher the temperature, the stronger the output voltage representation.

Thermopiles were investigated beacuse because of their extremely fast response times. One of the thermopiles evaluated for this device is Exergen's micro IRt/c. The micro IRt/c has a time constant of only 60 ms, which is extremely appealing for the purposes of this motion tracking system. Unfortunately, It has a very small FOV of 14° and is strongly affected by ambient temperatures. Furthermore, the thermopile is predominantly used for sensing gas leaks so it detects temperature ranges up to 625°C. This far extends the range necessary for this system and thus the voltage outputs are not sensitive enough to represent human body temperature.

Another thermopile that was analyzed is Heimann Sensor's HTIA series with optics, provided for by the Boston Electronics corporation. Detailed investigation resulted in a similar outcome to the micro IRt/c. The HTIA series is also most commonly used for gas sensing, and has a maximum sensing range of 500°C. Once again, the wide sensing range dwarfs the detector's response to the temperature of a human body. Figure 2-7 demonstrates this issue. The temperature of a human body is approximately 37°C. As one can see, this is in an area of null amplification for the HTIA sensor. A person walking through a room will not be strongly detected by this sensor. Thermopiles are not an adequate sensing device for this specific motion

tracking system.

Heimann Sensors does carry thermopile arrays with germanium lenses that have ideal characteristics for this application, but the detectors enter a price range that is beyond the scope of this project. Cost-efficiency is an equally important characteristic.

Thermal IR Cameras and Microbolometer Arrays

Similar to the thermopile arrays, thermal IR cameras and microbolometer arrays also have both characteristics needed for this motion tracking system to be successful. As mentioned above, though, cost-effectiveness is a crucial aspect of this project. Microbolometer arrays are approximately $3,000 to $20,000 depending on the resolution, and thermal IR cameras are an equally staggering $8,000 - $16,000. They are not a plausible option for this system.

MP Motion Sensors

The thermal sensors that are employed in this motion tracking system are Panasonic's MP Motion Sensor. They offer the best tradeoff between detection range and refractory period. They are passive IR type motion sensors, which also act as differential pyrosensors. They only detect changes in infrared radiation that are different from the surrounding background temperature. This sensor family has a variety of options available and can be tailored to match a specific application. The detectors chosen for this project are of the "slight motion detection" type. They are capable of detecting movements of 20 cm at a 2 m distance. Detection range has to be sacrificed in order to have a faster refractory period for the system.

The MP Motion Sensors use a quad type PIR element for its thermal sensor. Unlike the differential pyrosensor, these actually have four elements as opposed to two that are differencing the thermal radiation inputs. The quad type PIR allows for increased accuracy and faster response times. Figure 2-8 shows the block diagram of the MP Motion Sensor. The fields of view for the sensors are 91 ° along both the horizontal and vertical axes. They come coupled with a multi lens that has 104 zones of detection; Figure 2-9 illustrates the detection area. This means that the detectors are

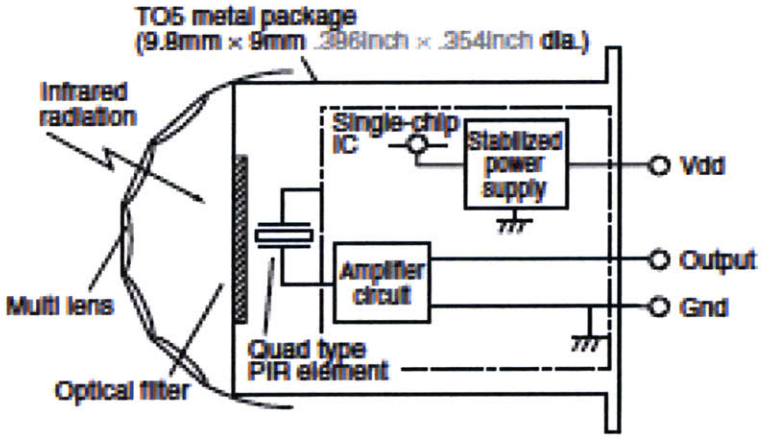

Figure 2-8: Diagram of the MP Motion Sensor casing. The on-board circuit, sensing element, optical filter, and multi lens are depicted.

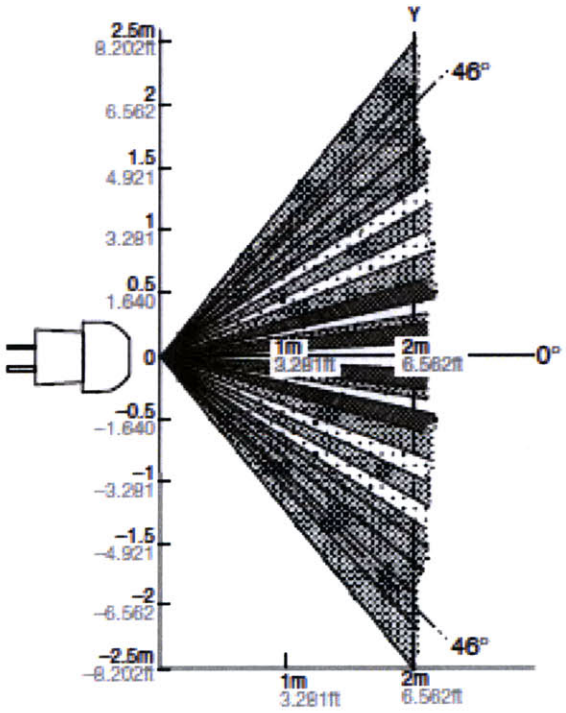

Figure 2-9: Cross-sectional diagram showing the area of zones of detection for the slight motion type sensor. The sensor is depicted on the left, and the detection zone patterns are indicative of the projections of the multi lens. The overall FOV is 91°.

more sensitive to the movement in a room. Lower-quality sensors have fewer zones of detection and larger gaps between the zones. Therefore, they are unable to capture as much information as the slight motion type sensors. Ultimately, the slight motion type sensors present an ideal balance between detection range and response time, making them the best fit for this thermal tracking system. Using these sensors, a 4-bit thermal motion tracking system is implemented with an accuracy of approximately 85%.

The table below summarizes the findings from the thermal sensor research:

Sensor	*Cost*
Nicera PIR Module	$4
MP Motion Sensor	$30
Heimann Sensor HTIA series	$19
Exergen micro IRt/c	$99
Thermal IR Camera	$8,000 - $16,000
Microbolometer Array	$3,000 to $20,000

2.2 Visible Spectrum

The visible spectrum is the alternative spectrum chosen in which to develop the proof of concept system. With the visible spectrum at wavelengths of 380-750 nm and the near-IR spectrum at 700-1400 nm, the upper range of the visible spectrum slightly overlaps with the near-IR spectrum. Using a combination of both allows for more flexibility when testing the system.

2.2.1 Sensor Models

Sensor technology in the visible spectrum has progressed much further than the thermal spectrum. The detectors are much more sensitive at farther distances, have larger fields of view, and have extremely fast frequency response times. Hence, working with visible spectrum sensors allows for more accuracy in data collection, which leads to a device with much higher resolution.

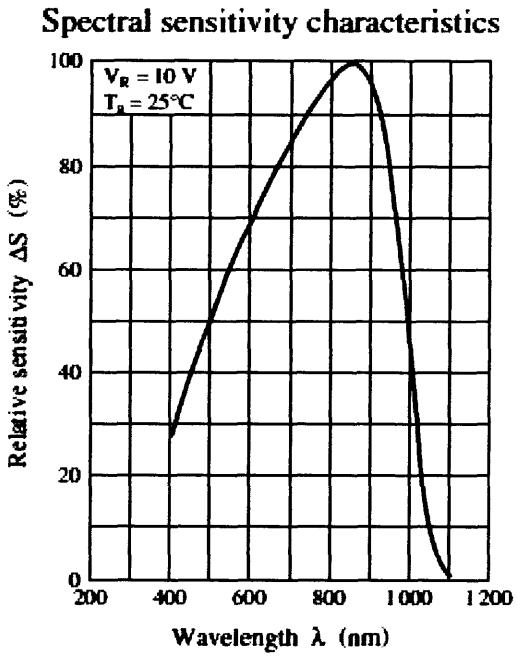

Figure 2-10: Spectral sensitivity characteristics for the PNZ334 PIN Photodiode.

PIN Photodiodes

The first sensors tested for the visible spectrum device are Panasonic's PIN Photodiodes, specifically the PNZ334 model. These sensors detect wavelengths between 400 and 1000 nm, with a peak at 850 nm; see Figure 2-10 for the specific curve. Even though these sensors output a very clean signal, they are not in fact the best fit for this specific application. Ideally, this system should work equally in all indoor environments, regardless of sunlight entering through windows or bright indoor illumination. However, due to the increased sensitivity to all visible light wavelengths, these sensors are strongly affected by the background lighting. This obstructs the valuable data in the output signal. A better fit sensor for this system would have a more concentrated wavelength spectrum.

IR Emitting Diodes

The sensors actually used in the visible spectrum device are High Speed Infrared Emitting Diodes made by Vishay Semiconductors. Their peak wavelength lies at 870

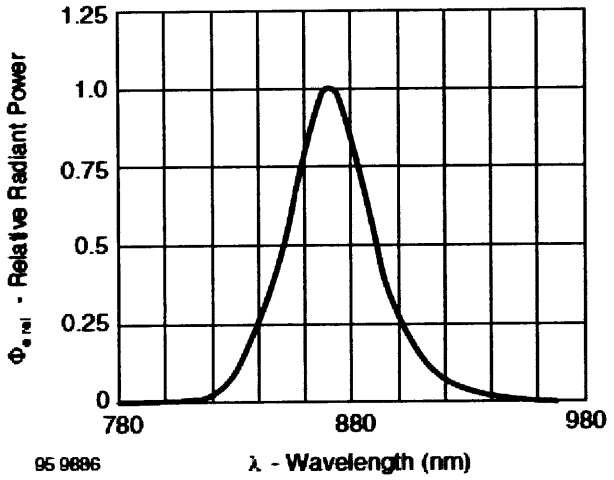

Figure 2-11: Relative radiant power vs. wavelength for the Vishay Semiconductors High Speed IR Emitting Diode.

nm, overlapping with the lower end of the near-IR spectrum. Unlike the photodiodes, the sensitivity of the IR emitting diode has a much stronger peak at 870 nm and drops off drastically for neighboring wavelengths, as illustrated in Figure 2-11. Having such a concentrated peak in sensitivity allows the sensor to detect illumination from a direct light source, without nearly as much interference from the ambient lighting. The IR emitting diode has a field of view of 27° and an angular resolution of 0.1°. At a distance of 3 m, it has an accuracy of 3 mm. The most significant characteristic of these detectors is that their rise and fall times are both 15 ns. This is many orders of magnitude smaller than the shortest refractory period achieved by the thermal sensors. Thus, the the visible spectrum system is able to achieve a much higher resolution than the thermal system, and with greater precision.

2.2.2 Implementation As Thermal System

Since sensor technology in the thermal IR range is still lacking for this application's purpose, the project scope is shifted to achieve a similar goal in the visible spectrum. The first step is to do a simple proof of concept with the visible spectrum, and then circuitry is applied to make the system behave in the same manner as the thermal

system. The visible spectrum system is set up exactly like the thermal system, only with the near-IR sensing diodes in the place of the differential pyrosensors. The key difference is that the person being tracked needs to hold a light emitting source for the diodes to sense; a simple flashlight will do. This new system still needs one more step before it functions like the thermal system. Differential pyrosensors only detect a moving heated body, not a stationary one. So in order to have the visible spectrum system perform this way, the output signals of the detectors need to be differenced. After differencing, the signals are non-zero only when there is a change in the original signal's slope. In other words, the signal detects motion of the light source, and otherwise remains zero. Now the visible system is capable of behaving exactly like the differential pyrosensors in the thermal spectrum system.

2.3 Gray Coding

Gray coding is the key tool in this system that allows for the spatial division of a room. Gray-coded binary sequences have a Hamming distance of one, meaning consecutive terms differ from one another by a single bit. A 3-bit Gray-coded sequence is shown in Figure 2-12, along with what the Gray-coded masks would like for a 3-bit system. In order to visually convert the binary sequence to a Gray-coded mask, the binary values need to be looked at horizontally. Each of the eight rows of the binary sequence translates to a vertical strip of the Gray-coded masks. The top row of three zeros represents the the black strip down the left-most side of the three masks. The masks can be divided into eight distinct vertical strips. Each strip has a unique combination of black and white values for the three masks. This code follows the pattern of ones and zeros in the binary sequence of numbers. A visualization of an 8-bit system in Gray-coded masks is depicted in Figure 2-13. Once again, the codes are found in each vertical strip of mask combinations. The 8-bit system has 256 unique codes.

This logic is applied to the binary masks in the motion tracking system. Each of the eight masks is placed in front of a detector. A one, or a high value, means that the mask allows the sensor to detect signals in that area. In contrast, a zero, or low

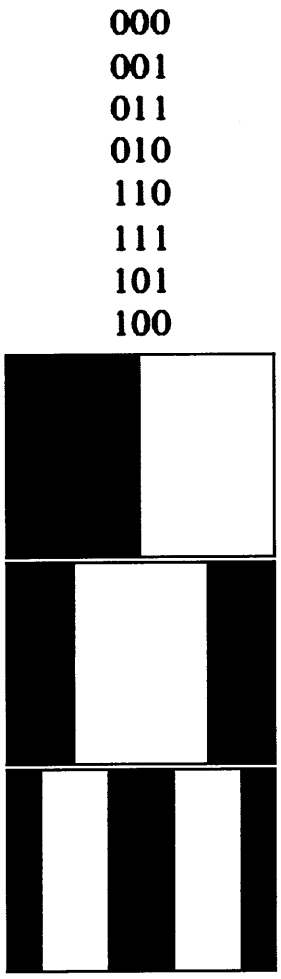

Figure 2-12: (Top) 3-bit Gray-coded sequence in binary values. (Bottom) The conversion from the binary sequence to a set of Gray-coded masks [21].

value, denotes an area where the mask blocks signals from being detected. When n masks are placed in a vertical array, the area in front of them is virtually divided into 2^n horizontal spaces. The 3-bit system in Figure 2-12 creates eight distinct spatial divisions. Along these lines, the 8-bit visible system creates 256 zones of detection and non-detection. Two of these devices positioned orthogonal to each other in a room results in a 256x256 grid. Each combination of high and low output values from the system represents a specific location in the grid, thus allowing a person's movement in the room to be pinpointed to one of these 256 spaces.

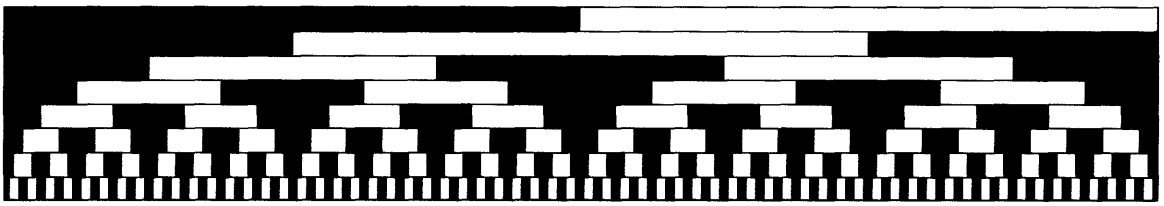

Figure 2-13: 8-bit Gray-code structured mask sequences. Masks can be divided into 256 vertical strips, each with their own unique code [21].

Chapter 3

Thermal Spectrum System Design

The thermal system consists of three main components: the mechanical device, the signal processing circuitry, and the data collection.

3.1 Mechanical Device

The mechanical device prototype for the thermal system is purposely designed to be large and functions as a prominent visual demo. The device is shown in Figure 3-1.

3.1.1 Material

One of the constraints for the mechanical device is that the masks and surrounding material need to be thermal IR opaque. Examples of such material would be cardboard, wood, or gold. It is also necessary to have a material that is both sturdy and easy to work with, but also cheap. This is why masonite was selected. Masonite is made of wooden chips that are heated and pressed into the form of a board. It is much sturdier than cardboard, but not nearly as dense or heavy as wood. The next limitation of the mechanical device, which explains the reason for such a large-scale prototype, comes from the resolution of the highest-order bit. The thermal system is a 6-bit design, and a 6th Gray-coded bit has 32 oscillations in its signal. In other words, 16 vertical strips needed to be carved into the mask for the 6th bit. A laser

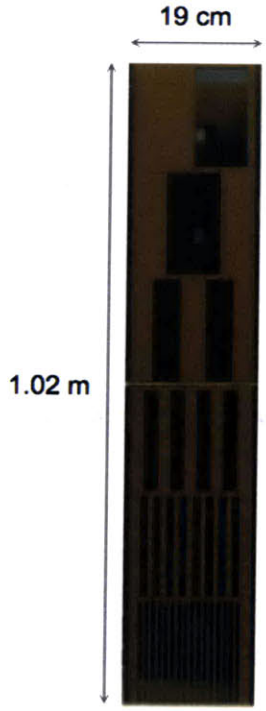

Figure 3-1: The mechanical device for thermal spectrum system.

cutter was used to carve these sections. Due to the resolution of the laser cutter, the vertical strips had to be wide enough such that the mask could be precisely carved. As a result, a mask size of 16x16 cm was chosen for the thermal mechanical device.

3.1.2 Optics

Condensing optics are used to focus the energy of the sensors in the thermal IR system. However, thermal IR optics are extremely expensive, and this project is concerned with making a cost-effective device. Thus, single-zone Fresnel lenses with a least absorption loss in the 8 μm to 14 μm range are used in this system. The lenses have a focal length of $f' = 10$ mm, and a diameter of $\phi_{sensor} = 4$ mm. The dimensions of the active zones of the sensors are 2x4 mm along the z and x axes. The dimension along the x-axis and the focal length are needed to calculate the field of view (see Figure 3-2) of the sensor and lens system. The expression used to find the

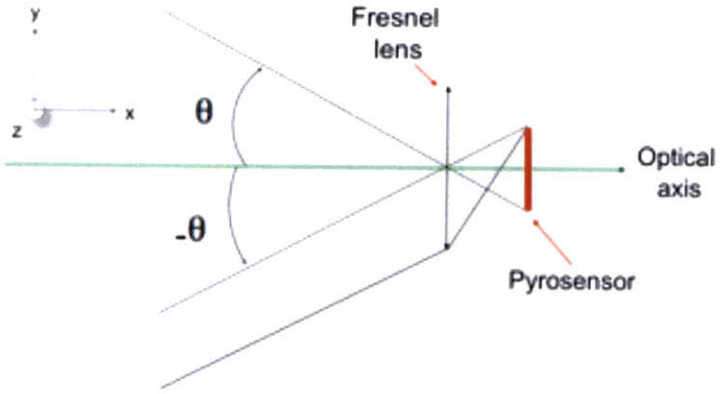

Figure 3-2: Field of view of a lens and sensor system.

field of view of the overall system is:

$$\theta = \arctan(\frac{\phi_{sensor}/2}{f'}) \tag{3.1}$$

In this case, $\theta = 11.3°$. This is a very small field of view, and in order to optimize the system, a lens with a shorter focal length is needed. To obtain a shorter focal length, two Fresnel lenses are used in series with one another. The new optical system uses two off-the-shelf Fresnel lenses glued together to divide the focal length and consequently double the field of view. Now $\theta = 21.8°$ and the sensor and lens system sees $43.6°$ of the room.

3.1.3 Calcuations

Further calculations are needed to decide the depth placement of the sensors, lenses and masks. The masks must be in the field of view of the sensor and lens system, and the border rays of the field of view need to rest on the contours of the sensor. In order to fulfill both of these conditions, the masks are placed precisely at a distance of 18.6 cm from the lenses. The zones of detection are now be calculated for each complete sensor, lens, and mask system; see Figure 3-3 for more details on the first three bits. These are crucial values for the data analysis to determine the speed and

location of a person in a room.

3.2 Analog Signal Processing

The initial detectors used for the thermal system were PIR modules, Passive InfraRed sensors with an integrated circuit (IC) attached to do the signal processing. The circuit converts the sensor's analog output into a binary 5 V high or 0 V low value. Since the attached IC resulted in output signals with extremely slow refractory periods, the sensor was removed from the IC module. The processing circuitry could then be recreated by hand to allow for more control over the signals, in an effort to reduce the refractory period. This signal conditioning is accomplished in two stages of circuits. Texas Instruments TL082 general purpose, dual operational amplifiers are used for both of these stages.

The first stage of circuitry is a low-pass filter to remove unwanted noise from the signals. Figure 3-4 illustrates the circuit diagram used to filter the signals. The equation to solve for the cutoff frequency, f_c, of the filter is:

$$f_c = \frac{1}{2\pi R_2 C_1} \tag{3.2}$$

The f_c for this filter is 1 KHz. In junction with the filtering, this circuit also amplifies the sensor output to give it a full range from 0-5 V. The expression used to set the gain in this specific circuit is:

$$gain = \frac{R_2}{R_1} \tag{3.3}$$

This results in a positive gain of 3 at the output of the circuit. Next, a comparator is employed to convert the signals from analog to digital outputs, depicted in Figure 3-5. An operational amplifier is used to function as a comparator in this case. The V_+ value is the signal input and the V_- value is set as a voltage reference. For this situation, a threshold value of 3 was desired, so a voltage divider is used to obtain 3 V from the 5 V power source. Resistors 1 and 2 in the circuit form the voltage

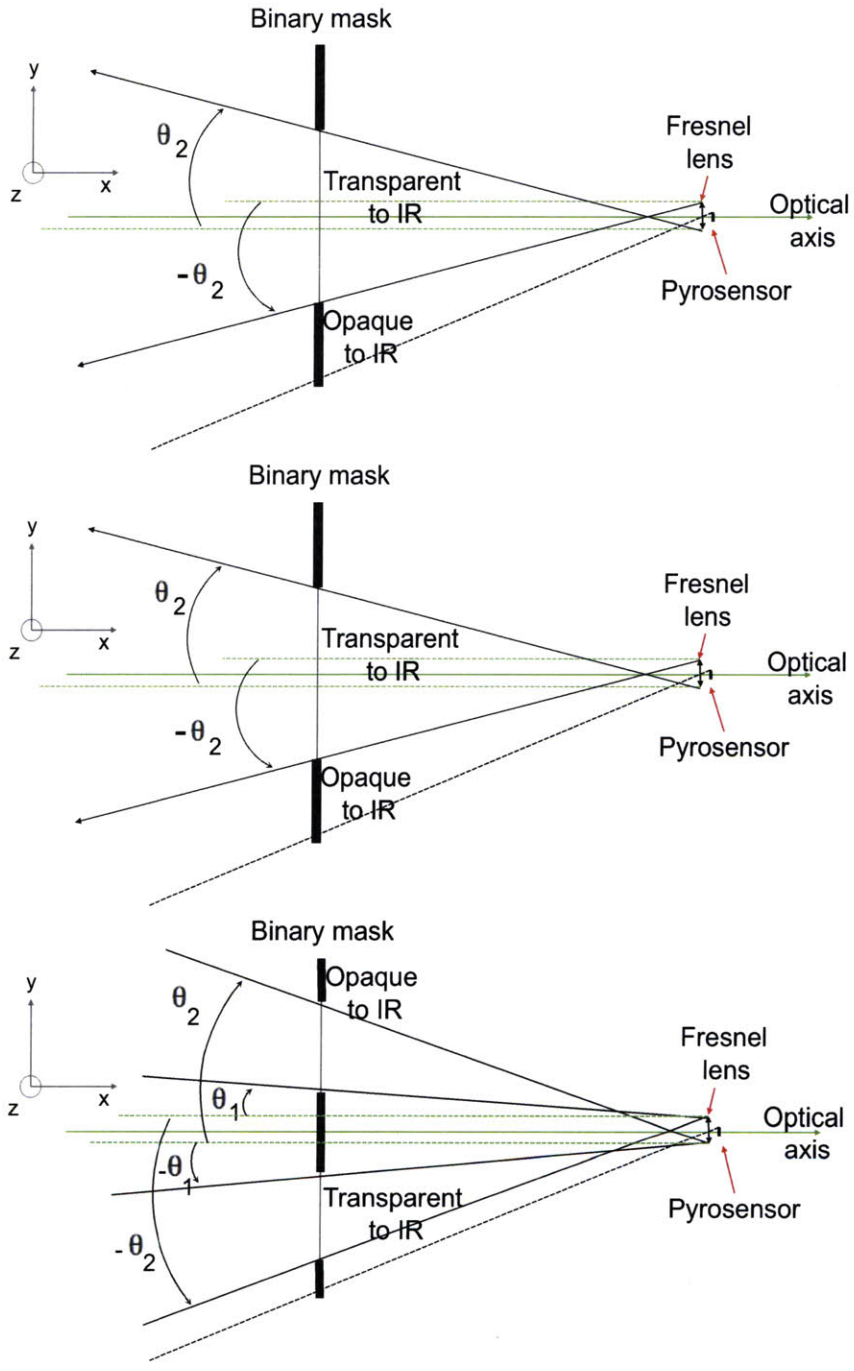

Figure 3-3: (Top) Field of detection for the first bit [mask + lens + sensor] system. $\theta_1 = 1.54°$ and $\theta_2 = -21.8°$. (Middle) Field of detection for the second bit [mask + lens + sensor] system. $\theta_1 = 13.6°$. (Bottom) Field of detection for the third bit [mask + lens + sensor] system. $\theta_1 = 4.6°$ and $\theta_2 = -19.3°$.

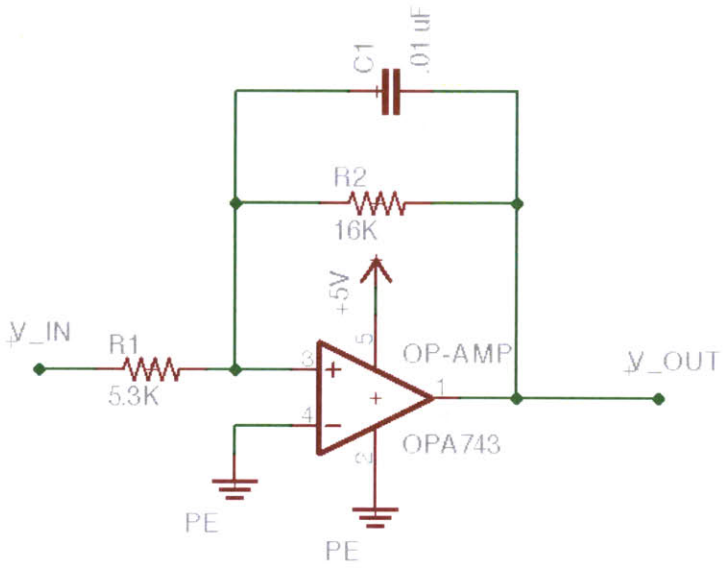

Figure 3-4: Combined low-pass filter and amplifier circuit.

divider. The equation used to set the dividing voltage is:

$$\frac{V_{out}}{V_{in}} = \frac{R_2}{R_1 + R_2} \tag{3.4}$$

A threshold value of 3 V is chosen to reduce false positives from noise. The comparator circuit outputs 5 volts when its input signal goes over the 3 V threshold, but otherwise outputs a 0 V signal. This essentially turns the analog sensor signal into a digital signal with binary high or low values. At this point, the signals are ready for the microcontroller.

Unfortunately, due to the slow frequency response of the pyrosensors, the circuitry is unable to speed up the refractory period to a feasible range for this application. The frequency response of the PIR sensor ranges from 0.3 - 3.0 Hz, which translates to 0.3 - 3.3 sec in time. Even though the refractory period is reduced drastically from 3.5 sec to 1 sec, it is still very far from the desired 50 ms. A 1 sec refractory period allows the thermal system to collect valid data for 3 bits, which spatially divides a room into only 8 locations. This is a very limited resolution for the system, and does not suffice.

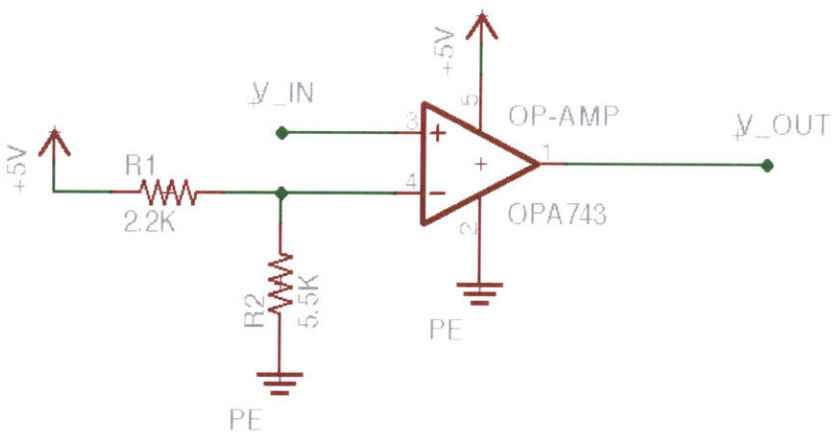

Figure 3-5: An operational amplifier used as a comparator circuit.

3.3 Data Collection

3.3.1 Microchip PIC

The binary outputs of the processed sensor signals are connected to a Microchip PIC18F25K20 microcontroller. The PIC (Programmable Interface Controller) is programmed to interpret the data and apply the logic algorithms to process it. A serial-to-USB connection port is used to transfer the data from the microcontroller to a computer and store it in a CSV file for post-processing.

3.3.2 System Visualization

Given that the thermal system operates as a demo, it is crucial to have a setup that provides real-time visual validation. This is done through eight LEDs, one to represent each location division of the room. An LED will turn on when a person is moving through that respective location. As a person walks from left to right in front of the thermal device, they step through these eight locations in order, and the LEDs will flash in that same order. See Figure 3-6 for a visual. The same Microchip PIC used for data collection is also programmed to control these LEDs. Eight of the PIC output bits are designated to go high in accordance with the logic outputs of the sensors, and the LEDs are directly connected to these bits. This allows the LEDs to

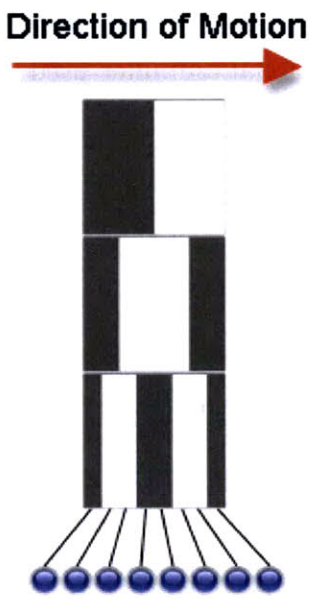

Figure 3-6: A 3-bit Gray-coded mask system is displayed; the masks divide the room vertically into 8 locations. As a person walks in front of the system from left to right, the LEDs corresponding to each location will light up.

display a person's location as he or she walks in front of the device.

Chapter 4

Visible Spectrum System Design

The visible spectrum system has analogous components to the thermal IR system. A small mechanical device holds the detectors, analog circuitry processes the outputs of the detectors, and the processed data is collected with an Arduino microcontroller.

4.1 Mechanical Device

The mechanical device for the visible system comes from the Prakash project, depicted in Figure 4-1. This device is on a much smaller scale compared to the thermal spectrum prototype. As a result, each component has to be precisely positioned to get accurate results. The various layers are illustrated by Figure 4-2. The detectors,

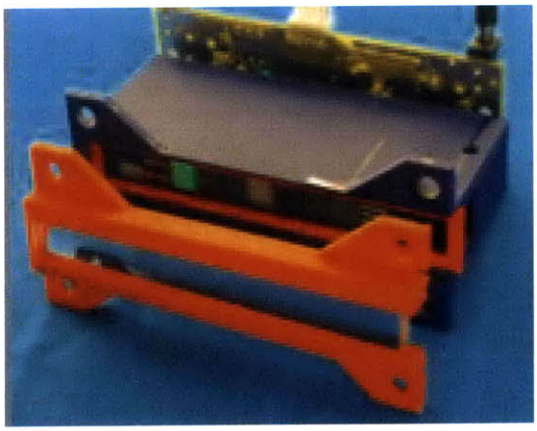

Figure 4-1: The Prakash projector unit.

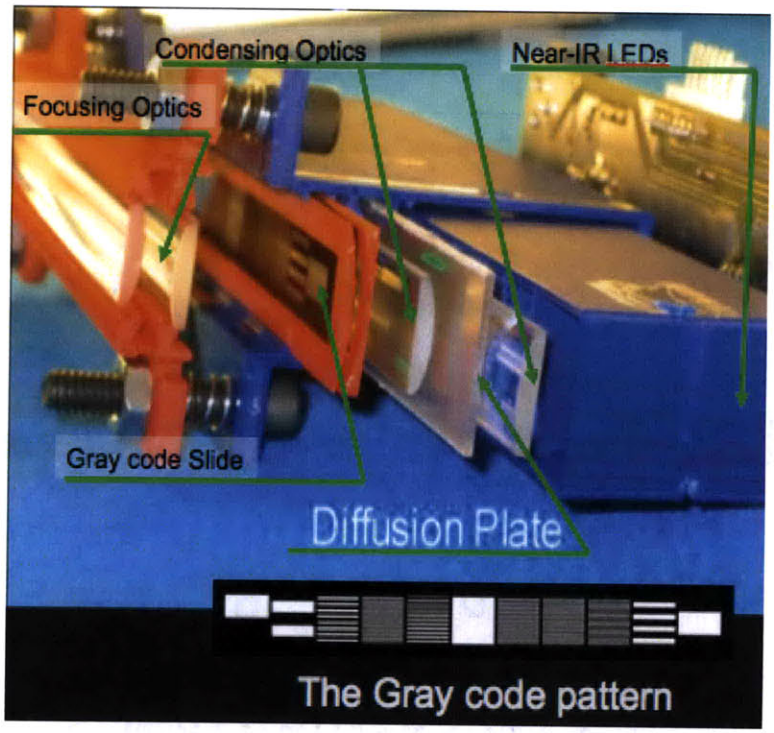

Figure 4-2: A layer-by-layer breakdown of the Prakash projector unit. The binary Gray-coded mask is displayed at the bottom.

near-IR LEDs, are inserted at the back layer of the device, forming a vertical array of single-bit sensors. There are eleven possible openings for detectors, but only eight are used in this system. Beyond eight bits, the resolution of the system surpasses the need of applications. Division of a room into 512 locations is not necessary. The condensing optics are the following layer in the device. These optics ensure the signal sent to the detectors is properly conditioned. A diffuser mask is placed between two cylindrical lenses to ensure the signal strength is equal across all eight vertical detector bits. The next layer is the binary coded mask. The mask is printed on a transparency and its pattern is also shown at the bottom of Figure 4-2. The eight lowest-order bits are utilized, which correspond to the upper four and lower four bits of the mask, ie: all but the center three bits. In order for the device to spatially divide a room, the incoming signals need to equally span the entire vertical length of the eight bits, and they cannot be wider than the highest bits mask width. This is accomplished by the front layer of the device, the focusing optics. It consists of two cylindrical lenses

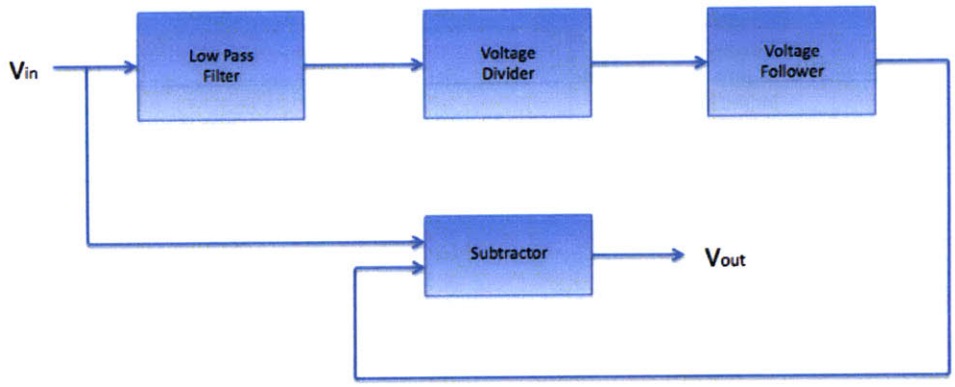

Figure 4-3: Block diagram of how a makeshift AGC circuit is achieved in this system.

with the convex sides facing each other. This placement of lenses focuses all entering signal wavelengths into a thin, vertical strip, allowing each detector in the array to sense the signal.

4.2 Analog Signal Processing

The detector outputs are fed through various stages of circuits to process the raw analog signals. In order for the visible spectrum to behave like the thermal spectrum, the method of an automatic gain control (AGC) circuit is employed. Figure 4-3 illustrates the block diagram of the necessary stages to obtain the data. An AGC circuit is an adaptive system used in many technologies - most commonly in radio, radar, and cellular phone audio signals - to maintain a constant DC level while preserving the AC values of a signal. A typical block diagram of an AGC circuit, illustrated by Figure 4-4, intelligently employs a variable gain amplifier (VGA). The control voltage to the VGA is set by taking the original signal, processing it with filters, and taking its difference. This control voltage varies the amplifying gain on the original signal itself, thus creating a negative feedback loop. The circuitry designed for the visible system does not use an exact AGC circuit, but rather leverages the same methodology and applies it to cater to the needs of the system. The following subsections will describe the circuitry designed specifically for the visible system.

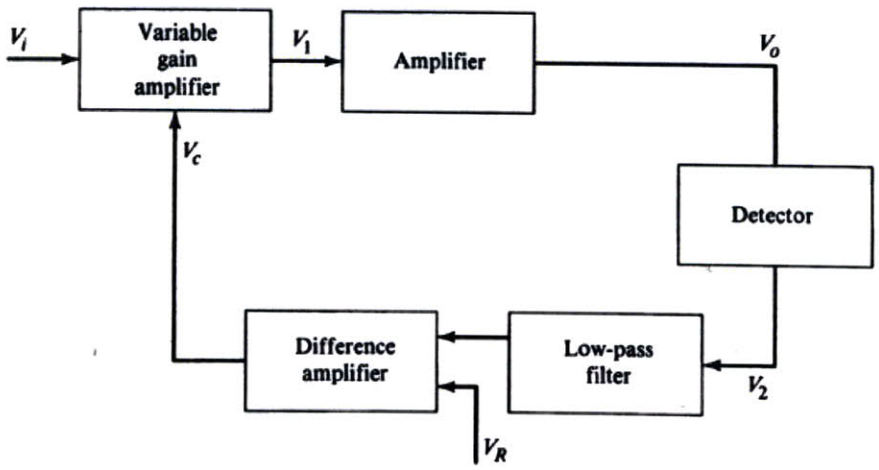

Figure 4-4: Block diagram of a standard AGC circuit [18].

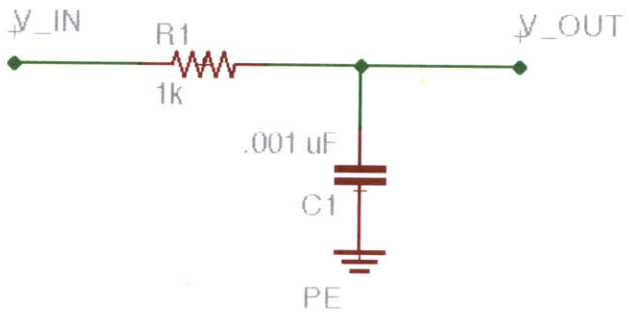

Figure 4-5: Noise filter circuit.

4.2.1 Noise Filter

The first stage is a simple RC circuit to filter out initial noise and to buffer the signal before it is processed. If the outputs of the sensors are connected directly to an integrated circuit, the information in their signals will be lost. The sensor outputs are extremely sensitive, and the high frequency information is dwarfed by the DC value of the integrated circuit. A voltage follower can be used to rectify this problem. The output of the voltage follower is strong enough to withstand further conditioning, and the AC values of the signal are not lost. The noise filter and voltage buffer circuit are shown in Figure 4-5. A 1 KΩ resistor is used with a .001 μF capacitor to remove

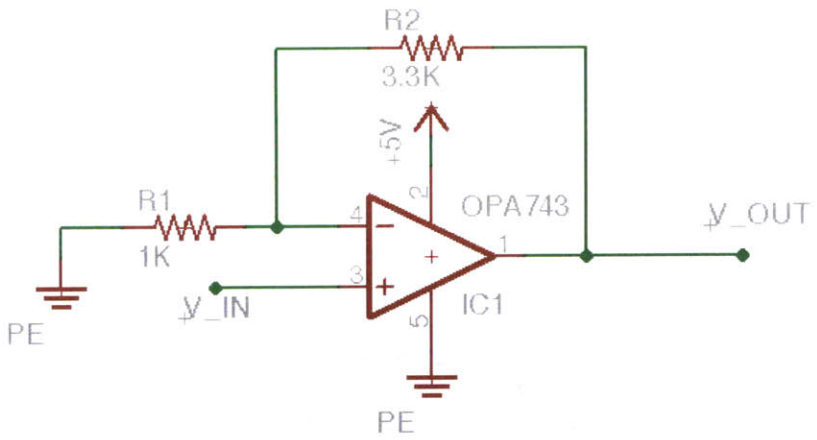

Figure 4-6: Amplifier 1 circuit.

initial noise without adding a long decay to the signal. This also provides enough of a buffer that the signal is processed without losing the raw data.

4.2.2 Amplifier 1

The next step is to do an initial amplification of the signal. The detectors output a range of 0-.75 V, but a full 5 V range is available for data capture. A non-inverting amplifier circuit is used, as depicted in Figure 4-6. So as not to risk exceeding a 5 V value, a gain of 3.2 is used.

4.2.3 Notch Filter

The output signals also pick up the electric hum, an oscillating current resulting from the main electricity line, at a 60 Hz frequency. An example of the sensors' output with this oscillation is illustrated by Figure 4-7. The electric hum may not affect data capture for the lower-order bits, but it does interfere with the accuracy of the system in the higher-order bits. Therefore, it becomes necessary to remove the oscillation to acquire precise data. A notch filter is the ideal technique employed to eliminate the electric hum from the signal. Similar to a band-stop filter, the notch filter allows all frequencies to pass except those immediately surrounding a set center frequency, f_c. The twin active notch filter design is used in this circuit, and is depicted in Figure

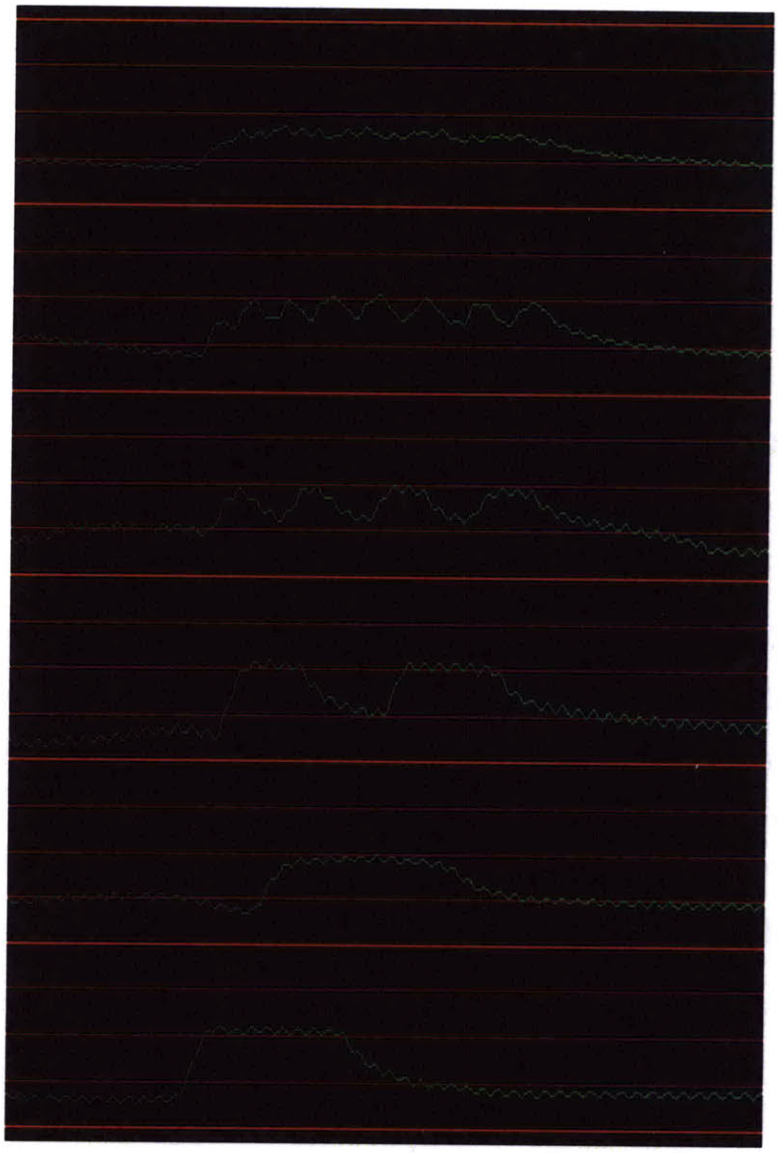

Figure 4-7: Example of 6 bits of data collection with the 60 Hz hum.

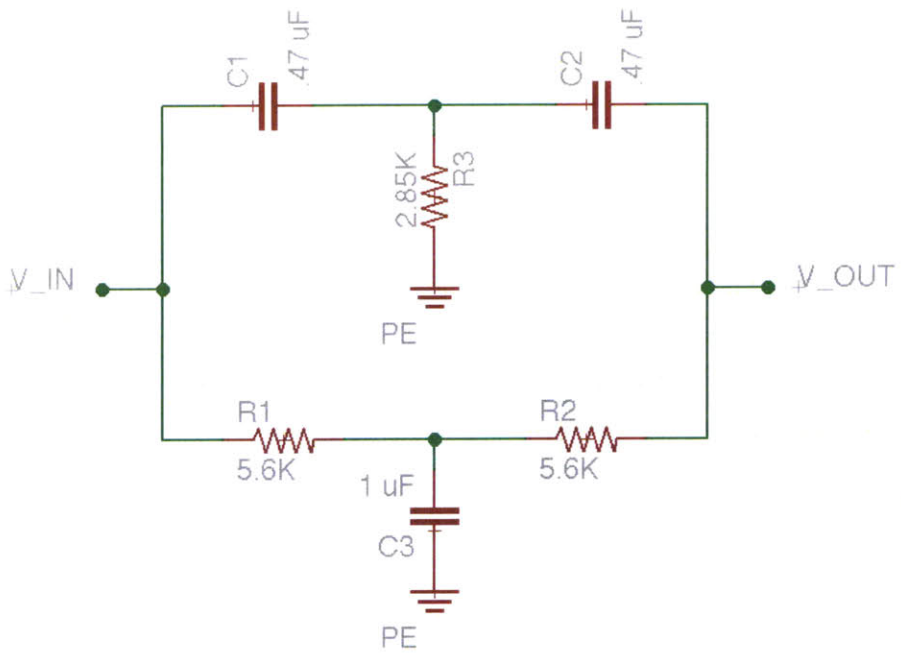

Figure 4-8: Twin active notch filter circuit.

4-8. The resistor and capacitor values are chosen to give an f_c of 60 Hz, and the filter successfully removes all frequencies in the 50-85 Hz range from the signal. Figure 4-9 demonstrates the notch filter at work; the signals no longer have the 60 Hz oscillation obstructing data collection.

4.2.4 Subtractor Circuit

The next stage of the circuitry is the makeshift implementation of the AGC circuit. This is done by subtracting a low-pass filtered version of the signal from the original signal. Thus, the final output signal only deviates from a constant signal when there is a change in slope of the original signal, as portrayed by Figure 4-10. Essentially, the subtraction pairing creates a differential version of the sensor output. Figures 4-11 to 4-14 show the four circuits needed to execute this stage.

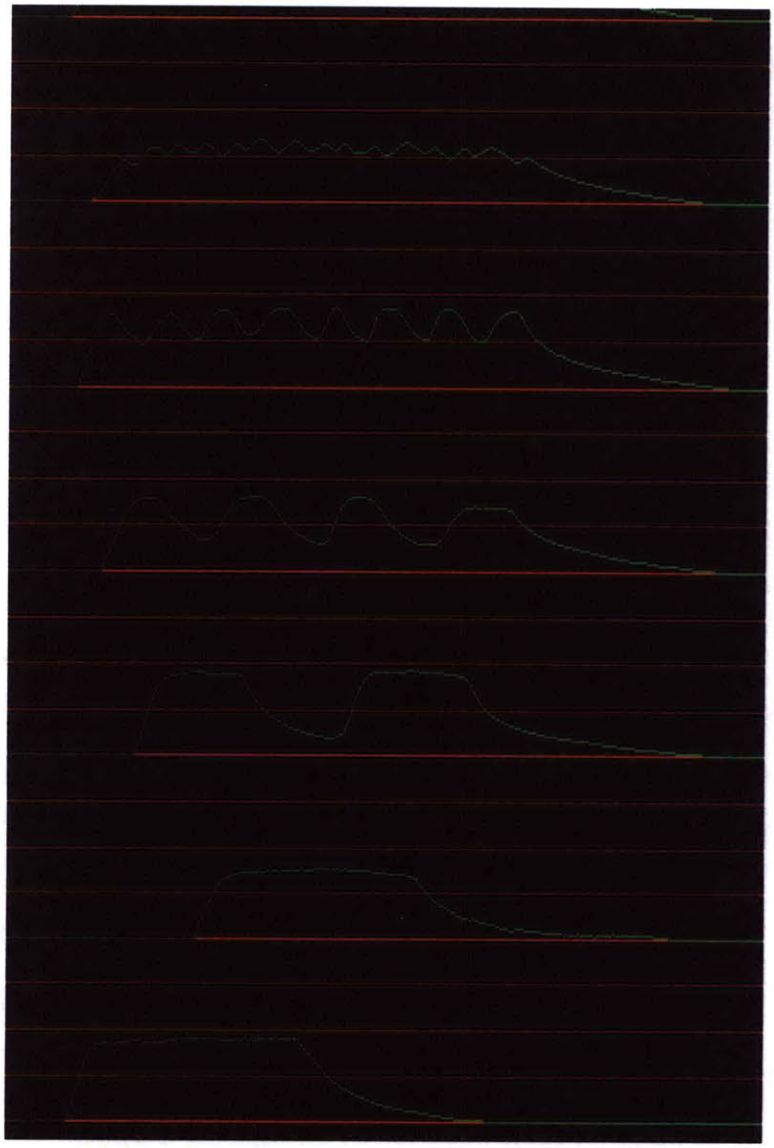

Figure 4-9: Example of 6 bits of clean data collection after filtering out the 60 Hz hum.

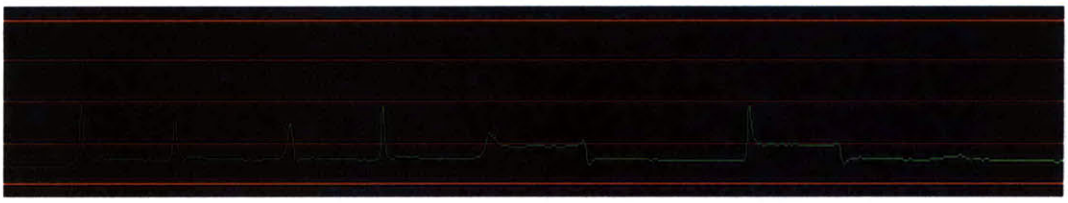

Figure 4-10: Example of an output signal in differential form. The first four spikes represent a light source quickly passing in front of sensor; this results in four impulses. The next two deviations represent a light source entering FOV of sensor and pausing before the light source exits sensor FOV. Result is a smaller positive spike, signal remaining neutral, followed by a negative spike as light source exits.

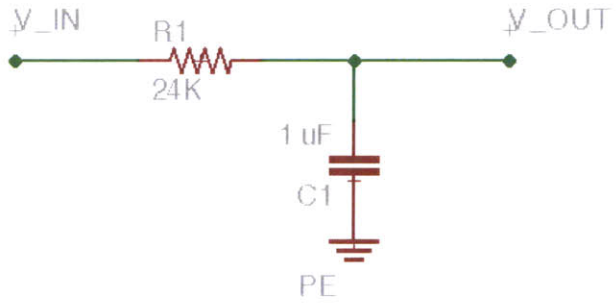

Figure 4-11: Low-pass filter circuit.

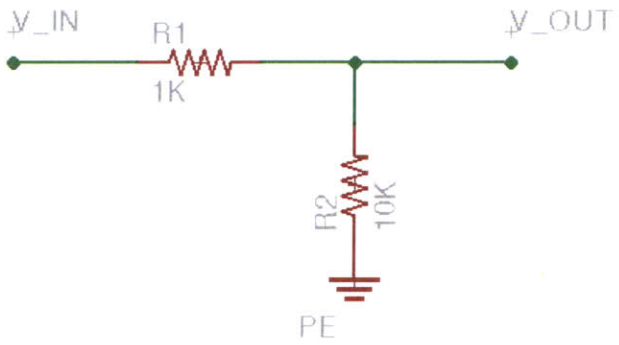

Figure 4-12: Voltage divider circuit.

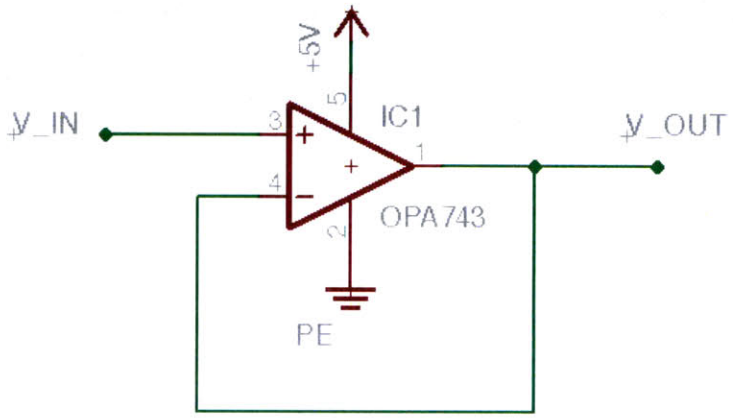

Figure 4-13: Voltage follower circuit.

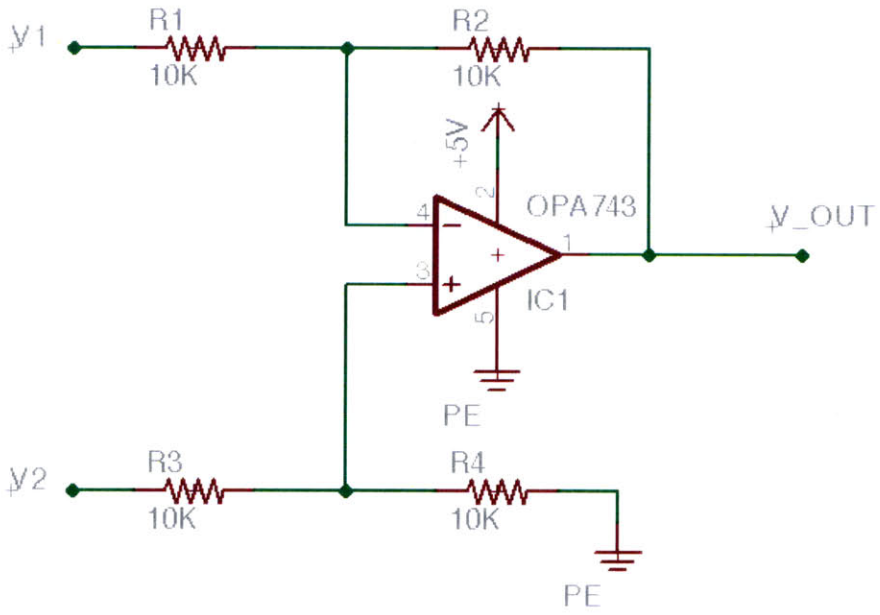

Figure 4-14: Subtractor circuit.

Low-Pass Filter

The first circuit is the low-pass filter. A simple RC circuit is used again, but this time, the capacitor value is crucial. The RC constant needs to be slow enough such that there is a difference between the signals in the lower-order bits, but not so slow that the signal is entirely lost in the higher-order bits. A capacitor value of 1 μF coupled with a 24 KΩ resistor was finally chosen for the lower-order bits, but this value was decreased for the higher-orders. The 6th bit uses a 0.1 μF capacitor, and the 7th and 8th bits use a 0.01 μF capacitor.

Voltage Divider

Next, a voltage divider is needed to reduce the magnitude low-pass filtered signal. Taking the difference of a signal results in a null signal unless there is a change in slope. There is a positive impulse when the original signal has a positive slope, and a negative impulse when the original signal has a negative slope. The circuitry for this project operates on a 0-5 voltage range with no negative voltage. Therefore, the difference signal needs to be raised above zero in order to see the negative impulses. Slightly decreasing the magnitude of the low-pass filtered signal accomplishes this. The voltage divider has an output voltage ratio of 91% of the input voltage.

Voltage Follower

Following the voltage divider, a voltage follower is required to bfer the signal before sending it to the subtractor. The subtractor circuit consists of a network of resistors that would alter the performance of the previous circuit if the two were to be put in series. Thus, the two individual stages need to be separated by a buffer to prevent their circuits from interfering with each other.

Subtractor

Finally, the actual subtraction is executed. The input to this circuit is the clean, notch-filtered detector output. The voltage subtracted from it is the low-pass filtered

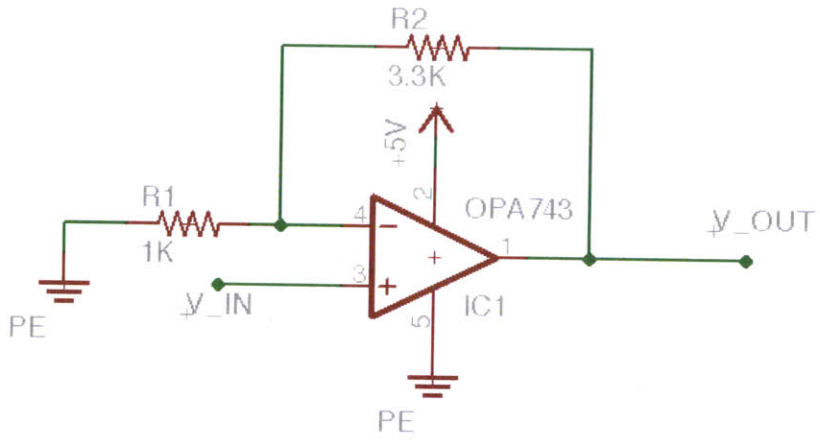

Figure 4-15: Amplifier 2 circuit for first 6 bits of system.

version of the same signal. The result is a differential signal representing the slopes of the detectors. V_2 is the input signal to the subtractor, and V_1 is the low-pass filtered version of the input signal. This circuit subtracts V_1 from V_2. The output of this stage finally resembles the output of the original thermal system.

4.2.5 Amplifier 2

The last stage of the circuitry is to perform a final amplification of the signals. To optimize the full 5 V available in the analog range, the differential signals are amplified by a gain of eleven. Resistor values of $R_1 = 1$ KΩ and $R_2 = 10$ KΩ are used for the first 6 bits of the system. As expected, the two highest-order bits receive less signal amplitude, and are accordingly amplified with a gain of 16. For them, $R_1 = 1$ KΩ and $R_2 = 15$ KΩ. The circuit for the amplifier is shown in Figure 4-15.

4.2.6 Higher-Order Bits Processing

The detailed circuitry above gives precise outputs for the first 6 bits of the visible spectrum system. However, an additional step is required to obtain the same accuracy for the 7th and 8th bits. These bits are put through a repeated step of subtraction and amplification, depicted in Figures 4-16 and 4-17, to capture the valuable information contained in their signals.

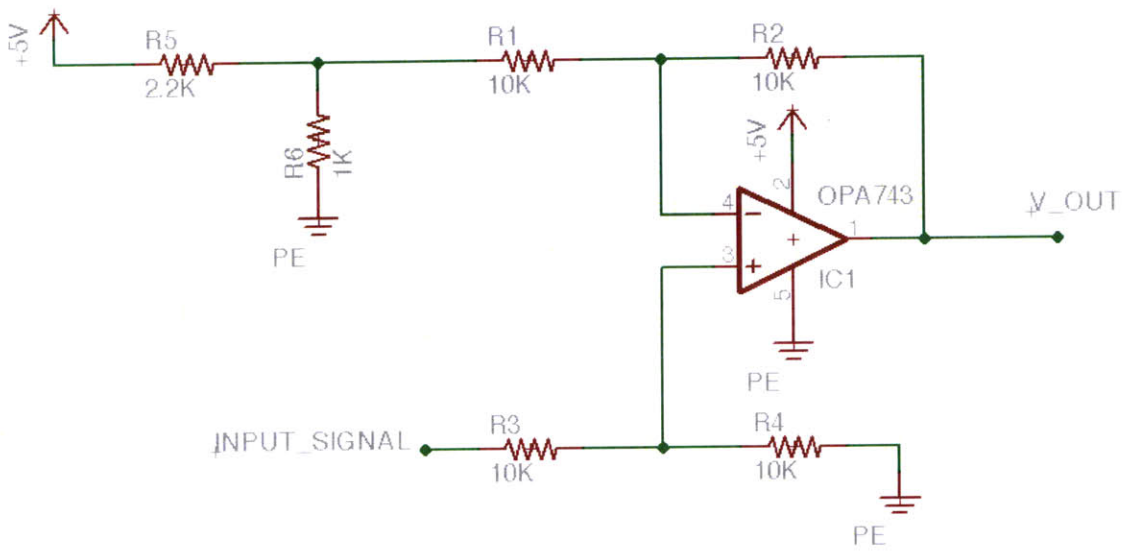

Figure 4-16: Circuit for repeated subtraction step for 7th and 8th bits.

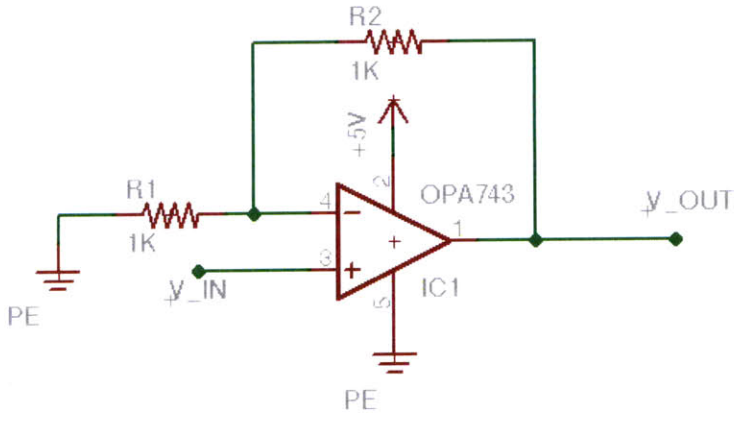

Figure 4-17: Final amplification circuit for 7th and 8th bits.

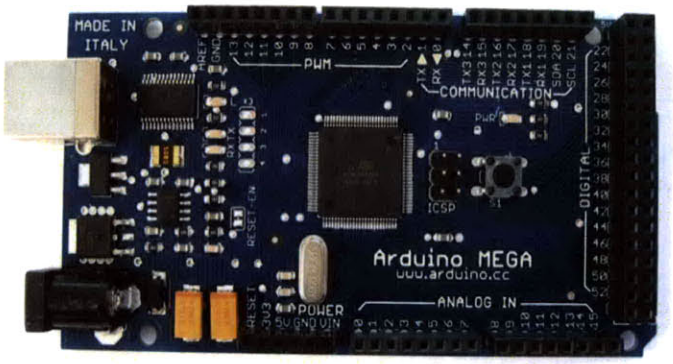

Figure 4-18: The Arduino Mega microcontroller board.

The signals are already in differential form, so a DC value of 1.56 V is subtracted from their signal; the DC value is obtained from a voltage divider applied to the 5 V power supply. This output is then doubled in amplification to zoom in on the oscillating signal. An op-amp amplifying circuit with a gain of two performs the magnification. The signals finally contain the data needed for processing, and is ready for analysis in software.

4.3 Data Collection

4.3.1 Arduino Microcontroller

An Arduino Mega microcontroller, displayed in Figure 4-18, is used to power the circuitry for this system, collect the data, and interface to the computer. It is powered either by a USB connection or a power jack, and was chosen specifically for its 16 analog input pins and 10-bit analog to digital converter (ADC).

4.3.2 Data Conversion

The final eight signals, after being processed in analog circuitry, are sent to the 0-7 analog inputs on the Arduino. The Arduino ADC automatically takes the voltage values of the signals and stores them as a 10-bit number. At this point the Arduino is programmed to either view the data real-time or collect it for post-processing.

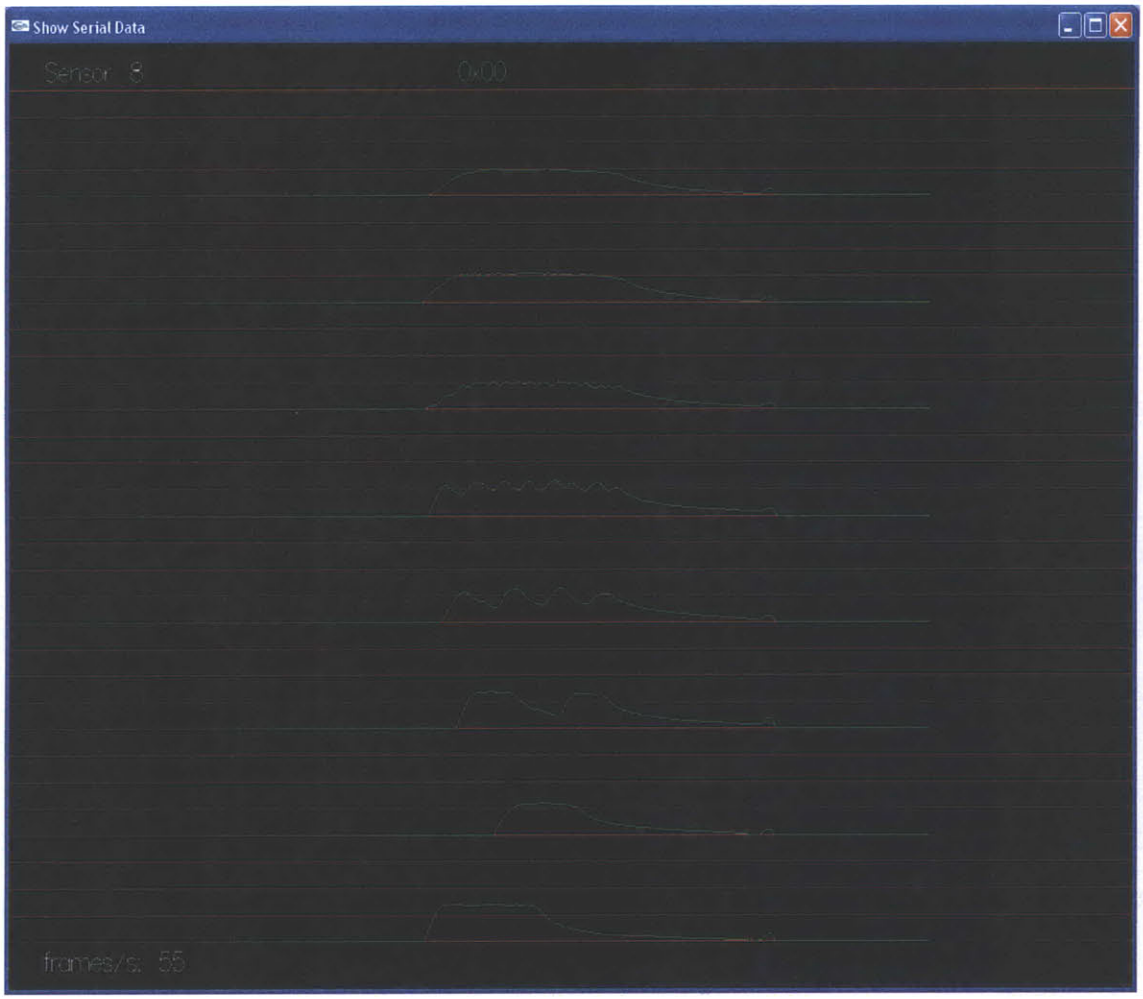

Figure 4-19: The full visualization screen displaying a sample set of 8-bit data.

Real-Time Visualization

A real-time visualization display was written by Dr. Hideaki Nii for the Prakash project [9]. Dr. Nii's visualization system has been modified to be compatible with the visible system. The Arduino writes the ADC values in hex and sends them through the serial-to-usb connection. The visualization system collects the values on the other end in a computer program, and displays the signals on a screen. Either each signal can be viewed independently, or all all signals can be viewed at once. An example of what the visualization screen looks like is depicted in Figure 4-19.

Post-Processing Data Collection

The other option is to collect and store the ADC data for post-processing. The Arduino is capable of displaying values real-time on the computer in a serial monitor, but cannot store them. In order to store the data, a python script [8.1.1] is called from the command terminal. This script extracts the 10-bit data from the Arduino serial monitor and records it as ASCII values ranging from 0-1023. The python script stores these values in a CSV file, which are then imported to MATLAB for software processing.

Chapter 5

Software Client

5.1 Graphical Software Interface

A Model-View-Controller (MVC) design pattern is used for the graphical software client in this system. The user interacts with a controller screen on the PC. The back-end, or model, of the system processes the data behind the scenes and the results are viewed on the controller screen.

5.1.1 User Interface

The software client for this motion tracking system is implemented in MATLAB. The front-end GUI (Graphical User Interface) allows the user to interact with the back-end data processing capabilities of the system. Figure 5-1 depicts a sample window of what a user would see. When first opening the tool, all of the input and output fields are blank. The user must insert the required text strings necessary, and then choose one of two operation modes: "Display 1 Person Data" or "Display 2 Person Data."

One-Person Tracking

If one-person data tracking is desired, the file names for the X-direction data and Y-direction data need to be inserted into their respective input text boxes. Clicking

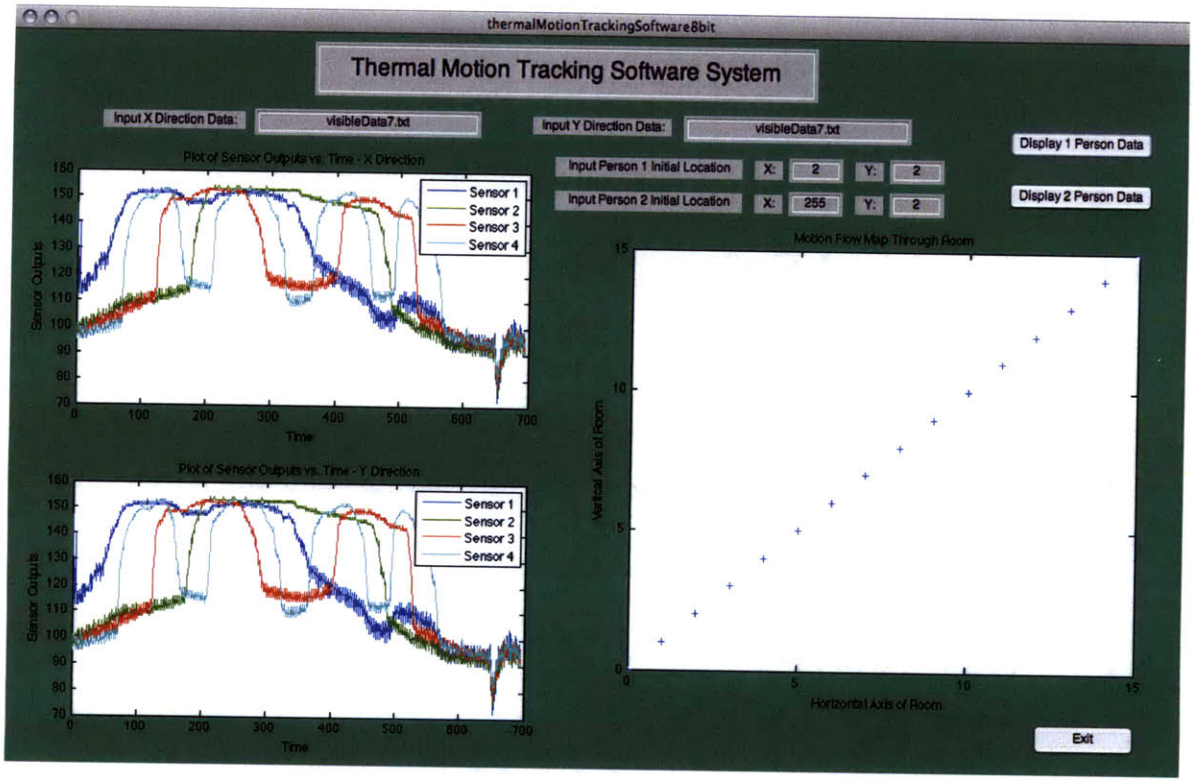

Figure 5-1: Example window of graphical user interface, after data has been displayed on screen. The two graphs on the left display the sensor outputs for the X and Y-direction data. The graph on the right is the motion flow map of the user through the room.

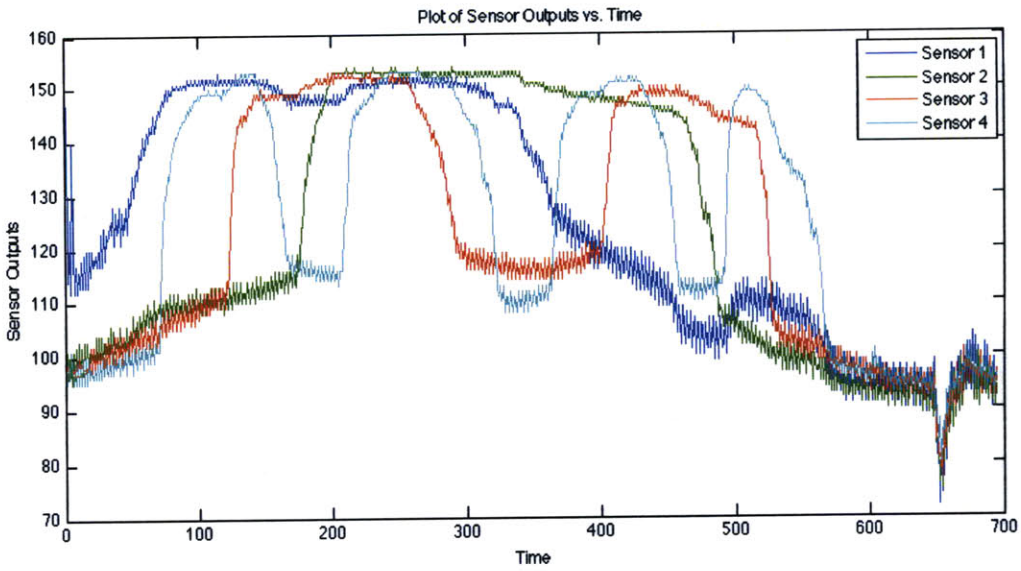

Figure 5-2: Graph of sensor outputs for a sample data collection with the 4-bit system.

on the "Display 1 Person Data" push button will then illustrate the data for the X and Y-directions and the motion flow map of the user through the room. The data is in the form of sensor outputs. The example in Figure 5-1 illustrates a rendering from a 4-bit system, hence the signal outputs of four detectors are shown in each X and Y-direction graph. Figure 5-2 zooms in and shows what one of these graphs would look like. Displaying data from an 8-bit system would consequently reveal the signal outputs of eight individual detectors. There are also two options for the type of motion flow map that is revealed. The user can choose to see a simple flow map that only shows the locations that were visited in the room, there is no notion of direction or walking speed. Figure 5-3 gives an example of this motion tracking map. If more information is wanted by the user, then an alternative motion flow map is available. This selection includes the direction of motion and the user's walking speed in its visualization, as demonstrated in Figure 5-4.

Two-Person Tracking

The second option is to display a motion flow map of two people walking through the room. The two-person tracking algorithm for this system is still in its preliminary

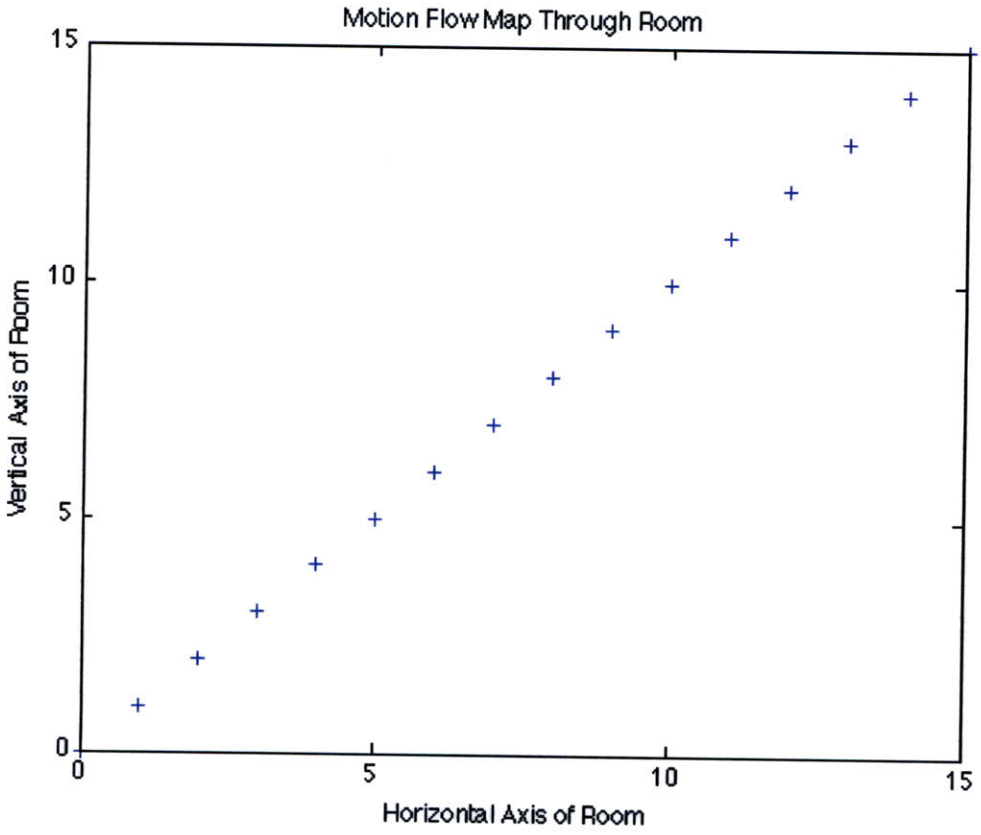

Figure 5-3: Sample data set of user walking diagonally across a room. Motion flow map does not indicate direction or walking speed.

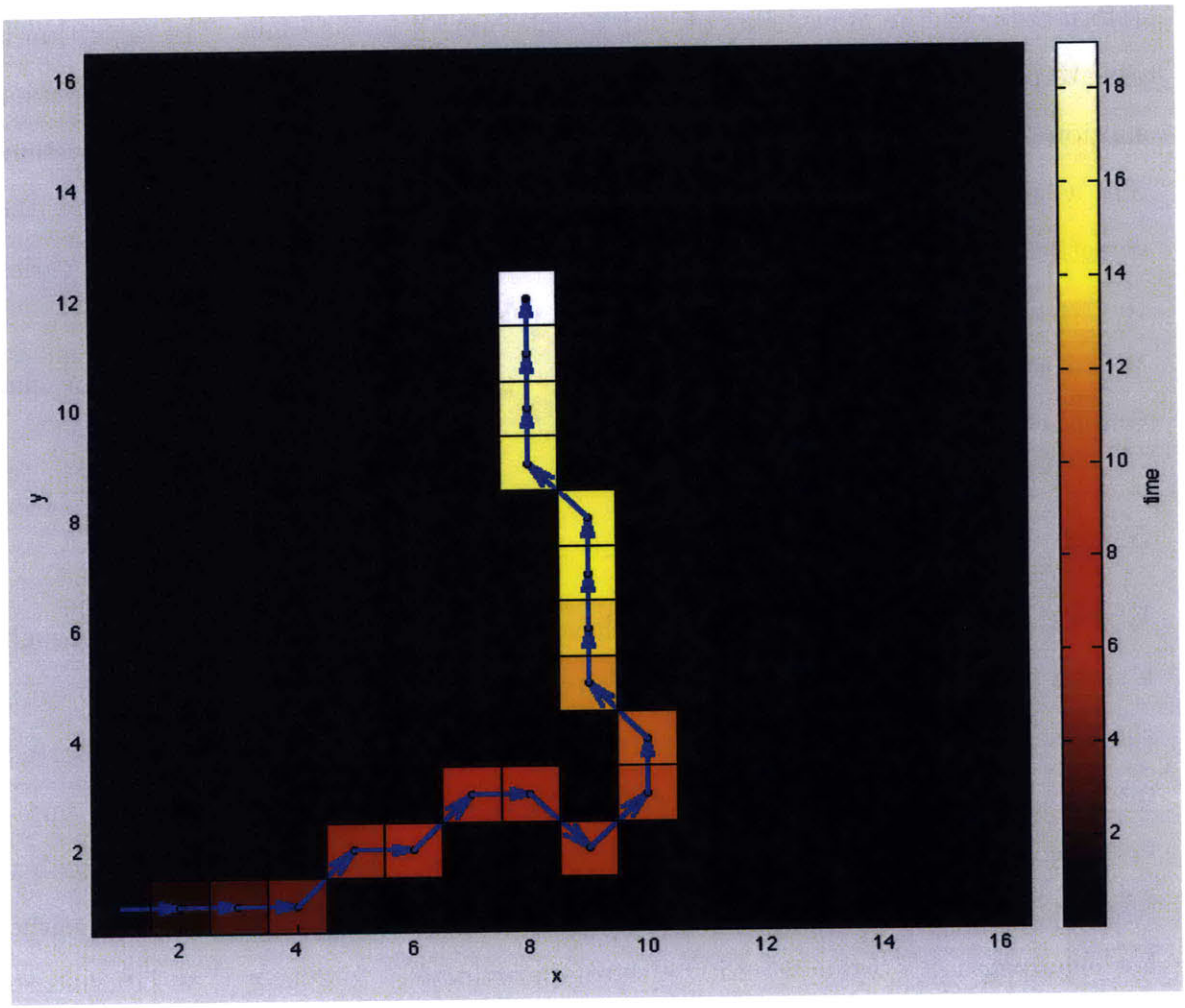

Figure 5-4: Example motion flow map of user through a room. Blue arrows indicate direction of motion and colors of blocks denote time elapsed while user is walking.

stages, and further details can be found in Section 5.2. One of the assumptions needed to track two people simultaneously with this system is the starting location of each person. Therefore, before pushing the button for "Display 2 Person Data," the user needs to input the X and Y coordinate location for each person as well as the file names for X and Y-direction data. In this case, a single data set holds the information for the movements of both people present in the room. The algorithm is able to process this data and sketch the motion flow map of both users. The 2-person motion flow map has a much higher resolution than the one-person tracking system. The visualization for the 2-person tracking has lines with arrows to the show the direction of movement for each person. The colors of the lines change as time varies to give a notion of each user's speed as they walk through the room.

When the user is done interacting with the graphical software client, he/she can simply click on the exit button and the screen will close.

5.1.2 System Back-End

When the software client receives the data selection from the user, it goes through a series of steps to process the data before the results can be displayed. To begin, the data is previously collected and stored in a CSV format file. This file is imported into MATLAB and converted into a matrix. The data is in the form of ASCII numbers, ranging from 0-1023, which is the range of a 10-bit converter. Each number represents a voltage value between 0-5 V, and the representation progresses linearly. For example, a 1 V value would translate approximately to 205, 2 V at 410, and so on. The output voltage of each detector goes up when a person is in front of its zone of detection through the binary mask. This comes across in each signal output as a series of high and low voltages, with the number of oscillations increasing logarithmically as the resolution of each bit increases. In order to decide whether a person is in a sensor's zone of detection or not, a threshold value is set in software. When the voltage level goes above the threshold, the system assumes a person is present in that zone, when it drops below, a person is no longer there. A smoothing average over five data points is used to ensure that there is no noise interference resulting in false

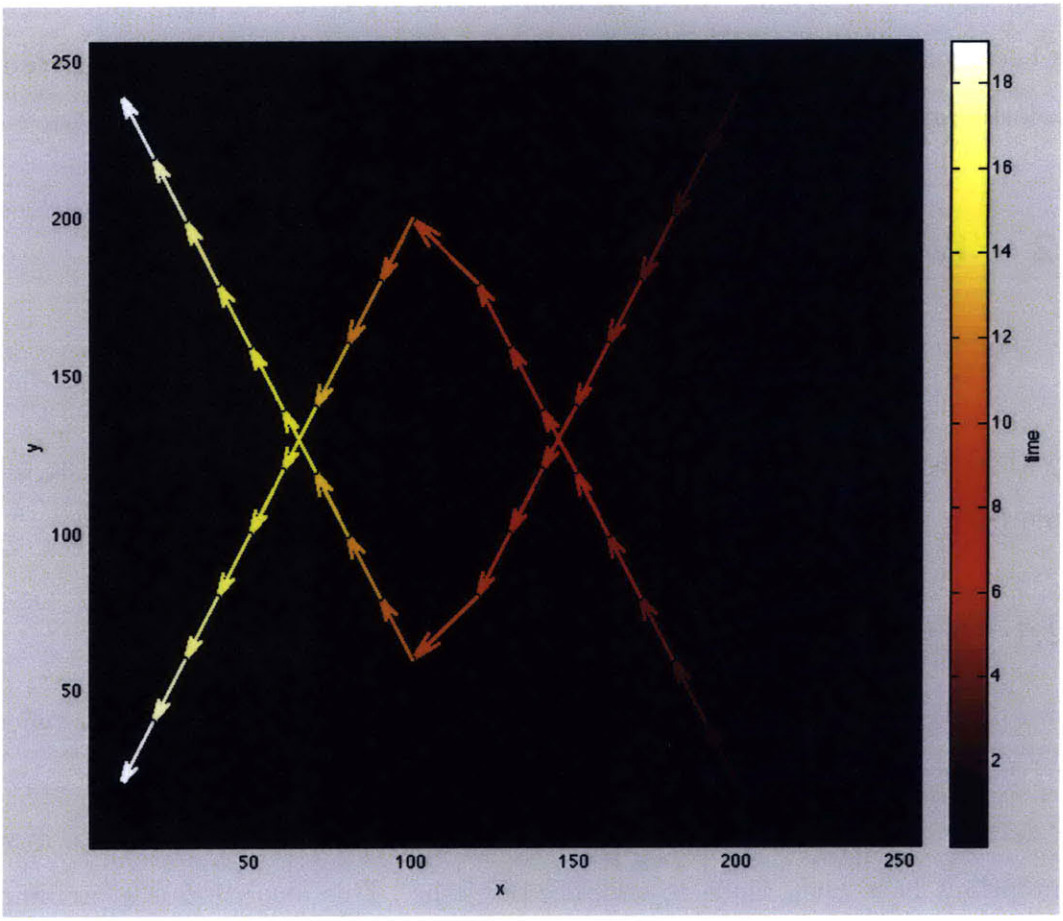

Figure 5-5: Example motion flow map of two users walking through a room. Arrows on the lines indicate direction of motion and colors of the lines represent time elapsed while user is walking.

detections. The process of thresholding the data converts it to binary form. This is done by the Binary Conversion function [8.1.2]. The next step is to convert the binary data into a specific location in the room. The Location Conversion function [8.1.4] is called to achieve this action. A lookup table is stored in the conversion function that holds the 8-bit Gray code in binary. Each row of sensor data is compared to the lookup table to determine which location the person is in. At this point, the data has been converted to a series of location values between 1 and 256 and is ready to be plotted and given to the front-end for display.

5.2 Two-Person Tracking Algorithm

The motion tracking device put forth by this thesis has many benefits over most existing technologies today, but one of its greater disadvantages is its limitation of only tracking one person at a time. If this problem could be overcome, the device would open up to many more possible applications and markets.

5.2.1 Theory

If two people were present in a tracked room simultaneously, then the data collected from the system would contain information on both of their locations. Since decoding this data from scratch constitutes a complex algorithmic investigation of its own, some assumptions have been made to simplify the issue. This algorithm is a preliminary exploration into the possible ways to extract the data for each individual from the collective set. Given the design of the system, the output for every step of data is the logical-OR of each person's location. Therefore, in order to decipher the individual data sets, the algorithm has to work backwards to find out which possible bitwise-OR binary combinations are equal to the known collective data set. The assumption is made that the starting location of each person is known. This can be translated to a real-life setting. A room in a house has a limited number of openings through which a person could enter. The system could be calibrated in advance to know the locations of these doors or hallways, and all activity would have to begin from one

of those positions. If the first point of each person is known, then the algorithm is able to continue deciphering the future locations of each person. All of the data locations are compared in binary until a match is made between the collected data and the bitwise-OR of the two persons being tracked. A brute-force method would test all possible combinations between the two persons and compare it to the data set. However, it was discovered that the logical-OR operation is a heavy constraint in the accuracy of the algorithm. Instead of simply using the brute-force technique, a more proactive solution has to be implemented.

5.2.2 Limitations

Unfortunately, the bitwise-OR operation limits the precision of the 2-person algorithm. For each logical combination of binary values, multiple locations for both persons are possible. The explanation for this stems from the most basic example of a 2-bit sequence.

A	B	$A \bigcup B$
0	0	0
0	1	1
1	0	1
1	1	1

Performing the logical-OR operation on a 2-bit system with sequences A and B produces either a 0 or a 1. A value of 0 never has any uncertainty in this operation because both of the original sequences must have been 0 as well. However, a value of 1 results in a three-way ambiguity. It follows that the number of distinct combinations for A and B that can be derived from A$\bigcup$B depends on how many 1's are in the final output of A$\bigcup$B. When expanding to a 3-bit system, there are 27 ambiguities for each 1. As is evident, this number increases exponentially as the number of bits increases. In order to compute the total number of overlapping possibilities for a combination of sequences, two values are necessary: m, the number of 1's in the A$\bigcup$B sequence, and n, the number of bits in the sequence. The total number of ambiguities can be

found by the equation:

$$\sum_{m=1}^{n} 3^m \binom{n}{m} \tag{5.1}$$

This derivation is exponential in n, the number of bits in the sequence. It proves that the uncertainty created by using the bitwise-OR operation is unavoidable and detrimental to the performance of the algorithm. Specifically because this device is using Gray codes, which have a Hamming distance of one, two persons in closer positions to each other would have increasing possibilities of overlap. A more rigorous approach must be taken to solve this issue.

5.2.3 Proposed Solutions

Narrow Search Frame

In order to avoid the problem set forth by the logical-OR, the algorithm can be further developed. The first enhancement is to narrow the search window when comparing the output combinations. Instead of searching the entire 256x256 frame, only the locations directly neighboring the person's initial position are checked. This is building upon the assumption that the user will always move at a smooth rate [19]. There is no discontinuous or erratic motion: each position for time $t + \delta t$ will always be either in the same position or neighboring the position at time t. Despite this enhancement, there are still overlaps in each frame.

Find Trajectory

An even more rigorous approach is necessary to overcome the ambiguity produced by the bitwise-OR. Instead of assuming only the first position of each user, the first few positions are now known. From this knowledge, the trajectory of each person is determined. It has been shown that when people walk through a room, they most often walk in a straight line toward their destination [19]. With a known trajectory, the algorithm can then find all possible location combinations for the two people, and then see which ones best fit the initial trajectories of each user. Finally, the system is

able to produce motion flow maps of both users as they walk through the same room.

Chapter 6

Conclusion

6.1 Summary and Contributions

This thesis introduces a device that tracks a person's movements through a room and generates alerts in cases of potential emergency. It allows elderly clients to live in the comfort of their homes, without the need for their close family and friends to worry about their health. It also gives banks, museums, or other institutions an option for non-invasive security monitoring. The technology is not only cost-efficient, but also scalable on an order of $log_2(n)$ as opposed to n^2. It is completely passive and requires no markers or tags to be worn by the user, and it functions independent of environmental fluctuations. Finally, the system maintains full anonymity and does not capture any images in its data collection; it is 100% privacy preserving.

Sensor Research

Much research has been done in search of ideal thermal sensors for this device. Such sensors would require two specific characteristics: a far detection range and a quick refractory period. Unfortunately, current thermal sensing technology has not produced a cost-efficient version of this sensor. Typically, sensors can be found with one feature or the other, but rarely both. In order to generate a fully functional proof of concept, the visible spectrum was used instead. Near-IR emitting diodes are used as the detectors, because of their ideal balance between detecting radiated light but not

being overly sensitive to ambient lighting.

Thermal and Visible Spectrum Systems

The key to this system lies in coupling the outputs of the detectors with binary Gray-coded masks. This is what allows the system to divide a room into distinct locations, tracking the user's movement as they walk through the room. For both the thermal and visible spectrum systems, very specific mechanical devices have to be created. These devices must take into account the fields of view of the detectors, the lenses that are used, and have specific optical systems to focus and condition the incoming signals. Once the sensors detect the thermal or visible radiation, the signals need to go through a series of analog processing before the data collection takes place. The processing in analog ensures that the system is acquiring as much information as possible before converting the analog signals to a digital form. Microcontrollers are used to collect the data from the device and store it on a PC. Finally, a software client is available for the user to interact with. The graphical user interface displays the sensor data and motion flow map for the specific data files input by the user.

Two-Person Tracking Algorithm

An exploratory two-person tracking algorithm is also demonstrated in this thesis. The algorithm has to work backwards to decode the possible locations for both persons in the room from the collected data. Due to the limitations of the bitwise-or operation, more rigorous approaches to the algorithm are necessary. In order to increase precision, a narrower search frame was used. When this enhancement didn't suffice, the trajectory of each person was determined as well. The assumptions made by these techniques are: knowing the starting position of each tracked person, expecting each person to move at a smooth rate, and estimating each person's motion trajectory. The combination of these techniques results in a successful two-person tracking algorithm for basic cases.

6.2 Future Work

6.2.1 Software Developments

Plenty of opportunities exist to improve the proposed system. Specific to the applications of monitoring elderly people or security screening, much work can still be done in theory and software. One possible extension is two-person tracking algorithm has been partially implemented as a part of this thesis. This algorithm could be further developed to be more robust and applicable in all probable situations. It could even be expanded to track more than two people in a room at once. Another strong possibility is to build an alert-generating system to signal family members or caregivers when an anomaly is detected by the system. One method of approaching this is to store motion flow maps of a user over time and generate a set of normal movement patterns when he or she is autonomous and in good shape. Then, the system could compare current motion flow maps to the previous set to potentially detect a change in the autonomy of the person, or a break in their general pattern, which would indicate an emergency may have occurred. At this point, it would be simple to produce a system that could send alerts in the form of an email or text message to the right people. Finally, a vital software enhancement would be to convert the data collection and analysis section into a real-time system. It is currently set up as a post-processing tool, but ideally, a real-time motion flow map display with continuous pattern analysis is preferred.

6.2.2 Hardware Improvements

Some work can be done in hardware to enhance this system as well. The project has been completed and successfully tested in the visible spectrum, but the ultimate goal is still to have a passive thermal tracking system. As thermal sensing technology improves in the upcoming years, future detectors can simply be inserted into the mechanical device to have a complete, working system. In the meantime, additional tests with an assortment of diverse optics and sensors could take place to find a

reasonable compromise between the two key factors needed for a successful system. Motion detectors with various focal lengths are a possibility, or adding dynamic masks to achieve a higher resolution with a smaller-scale device.

6.2.3 Prospective Projects

One idea in particular that shows real promise is to combine the theory behind this project with the relatively new technique of compressive sensing. This project would involve a lot of complex mathematical explorations, but holds the potential to increase the applicability of the project by an order of magnitude. For more sparse signals, fewer than than eight detectors are needed to encode the same amount of information. A signal is sparse when there are only a few bright spots in the arena. When this occurs, the dot product of situation A can be stored on one sensor, and the dot product of situation B on another sensor. Since both situations give different dot products, one can differentiate between the two. With compressed sensing, enough information could be encoded in each single-pixel detector to possibly track over ten people in a room at a time.

Chapter 7

Bibliography

[1] S. Roth and M. Black, On the spatial statistics of optical flow, (Proceedings of ICCV 2005), pp 42-49

[2] Yasutaka Furukawa and Jean Ponce, Dense 3D Motion Capture from Synchronized Video Streams, Computer Vision and Pattern Recognition, June 2008

[3] M. Andriluka, S. Roth, and B. Schiele, People-tracking-by-detection and people-detection-by-tracking., (Proceedings of CVPR 2008), pp 1-8

[4] Du Tran and Alexander Sorokin, Human Activity Recognition with Metric Learning, (Proceedings of ECCV 2008)

[5] G. Welch and E. Foxlin, Motion Tracking: No Silver Bullet, but a Respectable Arsenal, (Proceedings of IEEE 2002) Computer Graphics Applications pp 6, 22, 2438

[6] J. Abascal, B. Bonial, A. Macro, R. Casas, and J. L. Sevillano, Ambiennet: an intelligent environment to support people with disabilities and elderly people. (Proceedings of the 10th international ACM SIGACCESS conference on computers and accessibility, October 2008), pp 293294

[7] P. Bahl and V. Padmanabhan, RADAR: An in-building RF-based user location and tracking system, (Proceedings of IEEE INFOCOM, Volume 2, Tel-Aviv, Israel, March 2000), pp 775-784

[8] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'nardo Colucci, HiBall Tracker: High-Performance Wide-Area Tracking for Virtual and Augmented Environments, (Proceedings of ACM Symposium on Virtual Reality Software and Technology, 1999)

[9] Ramesh Raskar, Hideaki Nii, Bert Dedecker, and Yuki Hashimoto, Prakash: Lighting Aware Motion Capture using Photosensing Markers and Multiplexed Illuminators, (Proceedings of ACM Transactions on Graphics, Vol. 26, No. 3, Article 36, July 2007)

[10] R. Want, A. Hopper, V. Falcao, and J. Gibbons, The Active Badge Location System, (Proceedings of ACM Transactions on Information Systems, Vol. 40, No. 1, January 1999) pp 91-102

[11] CODAMOTION, 2007, Charnwood Dynamics Ltd., http://www.charndyn.com/

[12] MOTION ANALYSIS CORPORATION, 2006, Hawk-I Digital System

[13] OPTOTRAK, 2007, NDI Optotrak Certus Spatial Measurement, http://www.ndigital.com/certus.php

[14] PHASE SPACE INC, 2007, Impulse Camera, http://www.phasespace.com

[15] PTI INC, 2006, VisualEyez VZ 4000

[16] VICONPEAK, 2006, Camera MX 40, http://www.vicon.com/products/mx40.html

[17] Gilles Virone, Norbert Noury, and Jacques Demongeot, A System for Automatic Measurement of Circadian Activity Deviations in Telemedicine, (Proceedings of IEEE Transactions on Biomedical Engineering, vol. 49, NO. 12, 1463, December 2002)

[18] Isaac G. Martinez, Autoamtic Gain Control (ACG) Circuits, Theory and Design, University of Toronto, Analog Circuit Design I Term Paper

[19] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, You'll Never Walk Alone: Modeling Social Behavior for Multi-Target Tracking, (Proceedings of ICCV 2009)

[20] Michel Giordani, Mapa: Home Teleactivity Measure, Orange Labs - Business Services

[21] Douglas Lanman and Gabriel Taubin, Build Your Own 3D Scanner: 3D Photography for Beginners, SIGGRAPH '09: ACM SIGGRAPH 2009 courses (Proceedings of SIGGRAPH, 2009), pp 1-87

Chapter 8

Appendix

8.1 Code

8.1.1 Arduino Interface Python Script

```
#
#       arduino library
#
import serial
import time
import string

def start(dev='/dev/tty.usbserial-A6008gxv') :
        SERIAL = serial.Serial(dev,9600, timeout=1)
        return SERIAL

s = start()

str = ""
t0 = time.time()

while 1 :
        str = s.readline()
```

```
        if str != "" :
                print time.time() - t0,",", str[:-1]
#               print str
```

8.1.2 Binary Conversion Function

```
%This code takes a column matrix of voltage values, and converts
%them to a zero or one binary output. The first variable, z, is the
%column matrix. The second variable, T, is the threshold value
%that denotes whether a signal is high or not.

function y = binaryConversion(z,T)
n = length(z);
y = [zeros(n-4,1)];
for i = 1:(n-4)
    w = z(i:i+4);
    if (sum(w>T) == 5)
        y(i+5) = 1;
    end
end
```

8.1.3 4-bit Location Conversion Function

```
function y = fourBitLocationConversion(a,b,c,d)
n = length(a);
%l = length(b);
%m = length(c);
%n = length(d);
y = zeros(n,1);
for i = 1:n
    if      a(i)==1 && b(i)==0 && c(i)==0 && d(i)==0
        y(i) = 15;
    elseif  a(i)==1 && b(i)==0 && c(i)==0 && d(i)==1
        y(i) = 14;
    elseif  a(i)==1 && b(i)==0 && c(i)==1 && d(i)==1
```

```
        y(i) = 13;
    elseif  a(i)==1 && b(i)==0 && c(i)==1 && d(i)==0
        y(i) = 12;
    elseif  a(i)==1 && b(i)==1 && c(i)==1 && d(i)==0
        y(i) = 11;
    elseif  a(i)==1 && b(i)==1 && c(i)==1 && d(i)==1
        y(i) = 10;
    elseif  a(i)==1 && b(i)==1 && c(i)==0 && d(i)==1
        y(i) = 9;
    elseif  a(i)==1 && b(i)==1 && c(i)==0 && d(i)==0
        y(i) = 8;
    elseif  a(i)==0 && b(i)==1 && c(i)==0 && d(i)==0
        y(i) = 7;
    elseif  a(i)==0 && b(i)==1 && c(i)==0 && d(i)==1
        y(i) = 6;
    elseif  a(i)==0 && b(i)==1 && c(i)==1 && d(i)==1
        y(i) = 5;
    elseif  a(i)==0 && b(i)==1 && c(i)==1 && d(i)==0
        y(i) = 4;
    elseif  a(i)==0 && b(i)==0 && c(i)==1 && d(i)==0
        y(i) = 3;
    elseif  a(i)==0 && b(i)==0 && c(i)==1 && d(i)==1
        y(i) = 2;
    elseif  a(i)==0 && b(i)==0 && c(i)==0 && d(i)==1
        y(i) = 1;
    elseif  a(i)==0 && b(i)==0 && c(i)==0 && d(i)==0
        y(i) = 0;
    else y(i) = 0;
    end
end
```

8.1.4 8-bit Location Conversion Function

```
function q = eightBitLocationConversion(Z)
```

```
a = Z(:,1);
b = Z(:,2);
c = Z(:,3);
d = Z(:,4);
e = Z(:,5);
f = Z(:,6);
g = Z(:,7);
h = Z(:,8);

n = min([length(a) length(b) length(c) length(d) length(e)
length(f) length(g) length(h)]);
a = a(1:n);
b = b(1:n);
c = c(1:n);
d = d(1:n);
e = e(1:n);
f = f(1:n);
g = g(1:n);
h = h(1:n);

q = zeros(n,1);

x = dlmread('GrayCoding.csv');
GC = x(:,1:8);

for i = 1:n

    row = [a(i) b(i) c(i) d(i) e(i) f(i) g(i) h(i)];
    temp = repmat(row,size(GC,1),1);
    idx = find(all(temp==GC,2)==1);
    q(i) = idx;

end
```

8.1.5 8-bit Two Person Visualization

```
function z = plot_motion(x,y,x2,y2)
% Reset Matlab environment.
%clear all; clc;

% Define motion history.
N = 256;
x = x+1;
y = y+1;
x2 = x2+1;
y2 = y2+1;


% Create motion history matrix.
H = zeros(N,N);
for t = 1:length(x);
    H(y(t),x(t)) = t;
end

% Create basic motion history plot.
figure(1); clf;
set(gcf,'Name','Motion History');
imagesc(H); set(gca,'YDir','normal');
colormap([0 0 0; hot]); h = colorbar;
set(get(h,'YLabel'),'String','time');
axis equal tight;
xlabel('x');
ylabel('y');

C = hot(length(x));

% Plot motion history using arrows.
hold on;
    for t = 1:length(x)
        if t > 1
```

```
            quiver(x(t-1),y(t-1),x(t)-x(t-1),y(t)-y(t-1),0,...
                'Color',C(t,:),'LineWidth',2,'MaxHeadSize',2.5);
        end
        plot(x(t),y(t),'k.');
        plot(x(t),y(t),'k.');
    end
    for t = 1:length(x)
        if t > 1
            quiver(x2(t-1),y2(t-1),x2(t)-x2(t-1),y2(t)-y2(t-1),0,...
                'Color',C(t,:),'LineWidth',2,'MaxHeadSize',2.5);
        end
        plot(x2(t),y2(t),'k.');
        plot(x2(t),y2(t),'k.');
    end
hold off;

% % Draw lines separating each cell.
% for i = 1:N-1
%     line(i*[1 1]+0.5,[0 N]+0.5,'Color','k');
%     line([0 N]+0.5,i*[1 1]+0.5,'Color','k');
% end
z = [x-1,y-1,x2-1,y2-1];
```

8.1.6 Simulation Data Testing

```
%This block of code takes .csv files imported from excel, that
%contain location data for the x and y directions of person 1 and 2.
%It converts the location numbers into rows of binary gray-coded
%values. The output of this function is then sent to the two-person
%tracking system to test my algorithm. This approach was chosen
%because 1) the best way to test the system for accurate outputs
%is to decide the input 2) this algorithm is meant to be a
%preliminary tracking algorithm, not a robust, final product, and
%3) it would be very difficult to collect accurate data from the
%devices with two people in the room at the same time.
```

```
function [x y] = simDataConverstion(convertData);

GCtemp = dlmread('GrayCoding.csv');
GC = GCtemp(:,1:8);
%GCY = GCX;

simDataTemp = convertData;
%simDataTemp = dlmread('8BitSimData1.csv');
x1 = simDataTemp(:,1);
y1 = simDataTemp(:,2);
x2 = simDataTemp(:,3);
y2 = simDataTemp(:,4);

n = min([length(x1) length(y1) length(x2) length(y2)]);
x1 = x1(1:n);
y1 = y1(1:n);
x2 = x2(1:n);
y2 = y2(1:n);

x1 = repmat(x1,[1 8]);
y1 = repmat(y1,[1 8]);
x2 = repmat(x2,[1 8]);
y2 = repmat(y2,[1 8]);

q = zeros(n,1);

for i = 1:n
    x1(i,:) = GC(x1(i),:);
    y1(i,:) = GC(y1(i),:);
    x2(i,:) = GC(x2(i),:);
    y2(i,:) = GC(y2(i),:);
end

x = x1|x2;
y = y1|y2;
```