OCCUPANCY ESTIMATION THROUGH VISIBLE LIGHT SENSING (VLS)

By
MOHAMMAD AL MESTIRAIHI
Bachelor of Science in Computer Engineering
AL YARMOUK UNIVERSITY
IRBID, JORDAN
2008

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER of SCIENCE
May, 2019

OCCUPANCY ESTIMATION THROUGH VISIBLE LIGHT SENSING (VLS)

Thesis Approved:

Dr. Sabit Ekin

Thesis Advisor

Dr. Keith Teague

Dr. Carl latino

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my outstanding advisor Dr. Sabit Ekin for his invaluable instructions and endless guidance, patience and words of wisdom throughout the course of my study. His perspicacity, support, assistance, and understanding have been a numerous source of motivation for this work. His gigantic experience in his field of study has greatly benefited me in producing new approaches and fulfilling them. He has taught me to be a significant, sincere, kind, and polite researcher. He has given me something to strive for technically and personally. Isolated from professional and academic advice, I also came to know him as a wonderful person. Thank you Sabit for your charitable assistance and supervision.

Next, I would like to thank my committee members Dr. Keith Teague and Dr. Carl Latino for all the valuable advice and insights they provided. I would like to express my gratitude to both of them for accepting to be on my committee. I also would like to give my great appreciation to the great friends I met here at Oklahoma State University. Next, I would like to express my deepest gratitude to my family back home in Jordan, my wife for her love, encouragement, support, blessing, and prayers.

Furthermore, I would like to thank Oklahoma State University for presenting me with this once in a lifetime occasion. State of the art labs and excellent faculty and staff along with welcoming people made me feel right at home.

Name: MOHAMMAD AL MESTIRAIHI

Date of Degree: May, 2019

Title of Study: OCCUPANCY ESTIMATION THROUGH VISIBLE LIGHT SENSING (VLS)

Major Field: ELECTRICAL AND COMPUTER ENGINEERING

Abstract: Visible light is everywhere around in our daily life. The part of the electromagnetic spectrum that is visible to the human eye is called visible light; it is in the range between 350-750 nm, which is roughly between 400-750 THz in terms of frequency. Visible light sensing and communication (VLS and VLC) are considered two of the most emerging fields in sensing and wireless communication areas, where light emitting diodes (LEDs) are used as the transmission unit (T_x). LEDs have a number of advantages, some of which are extended life expectancy, illumination, lower energy dissipation, and eco-friendly. As a result, visible light can be used in several sensing applications in our life such as occupancy estimation.

In this thesis, a new occupancy estimation method based on VLS is presented. A visible light source (e.g., LED) is utilized as the transmitter and a photo-detector (PD) used as a receiver, forming a visible light sensing system. Depending on the number of people in the room crossing the line-of-sight LOS (between the light source and PD), the received power at the receiver change. Consequently, probability density function (PDF) and cumulative distribution function (CDF) of the received power at the receiver change. First, a theoretical analysis of the received power is developed to incorporate the impact of room occupancy on the PDF and CDF of the received power. Second, these received power PDF and CDF expressions are compared with simulation results. Both results are in perfect agreement that verifies the theoretical analysis. In addition, Kullback-Leibler divergence (KL-divergence) method to analyze measurement data to detect the number of occupants in an environment. In this method, the stored PDF of the received power in the database is compared with the measured received power PDF, which reveals the estimated room occupancy. It was shown how a slight variation in room occupancy can dramatically alter the statistics the received power. Theoretical analysis and simulations are performed. In addition, we have conducted experiments in the Wireless Communication and Sensing Research Lab (WCSRL) located in Engineering South (ES) 408 at Oklahoma State University. As a future work, we are planning to study the impact of scattering on estimation accuracy. Multiple LEDs and PDs (i.e., multiple transmitters and receivers) can also be considered in future tests. Developing a complete system that can control and regulate buildings HVAC (heating, ventilation, and air conditioning) and lighting to improve sustainability and energy efficiency will be one of our promising research direction in the future.

Keywords: Occupancy Estimation, Visible Light Sensing, Statistical Analysis, KL-Divergence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Introduction to Visible Light Sensing

People have been using visible light as a method of communication for years and years, which is still a reservoir of numerous benefit in the field of wireless communication. Before Thomas Edisons invention of the light bulb in the 19th century, people used fire smoking as a communication medium, and after Edisons invention, new communication ways by using light were deployed [2]. Alexander Graham Bell was the first one who implemented the concept of using the communication medium by his invention of photo-phone in 1880, which is a machine used to transmit voice signals using visible light rays. Bell directed sunlight with a reflector and then spoke into a device that echoed the mirror [6]. The vibrating rays were picked up by the detector at the receiving end and decoded back into the voice signal, the identical scheme as the telephone performed with electrical signals. The two main problems that Bell could not solve were the inability to generate a valuable carrier frequency, and environmental dilemmas for instance fog and rain. With the invention of LED, the use of light as a communication medium has been popular, and it has started to gain more research attention. Visible light is in the frequency range between 350-750 THz in the electromagnetic spectrum as shown in Figure 1.1, which is 10,000 larger than the entire RF (radio frequency) spectrum [7]. The availability of visible light everywhere around us makes the use of it as a communication and sensing medium a promising research issue; therefore, extra installment is not required which has a great impact on reducing the cost [8].

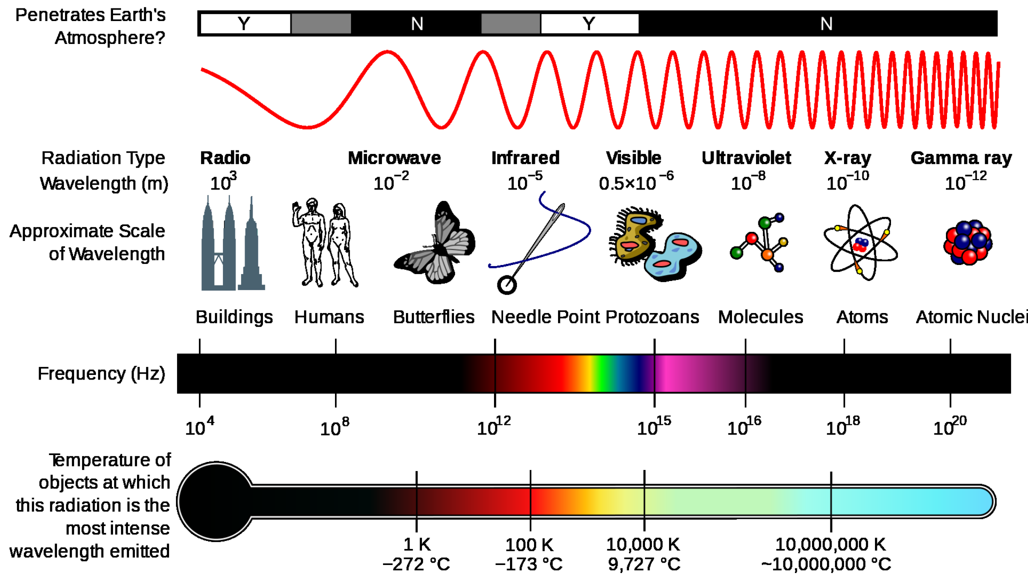


Figure 1.1: Electromagnetic Spectrum [1]

.

Occupancy estimation techniques have plenty of numerous applications in diverse fields. These methods essentially concentrate on counting the number of people inside a particular region, also known as crowd counting. One of the most significant application is smart-building management, where a control system can optimize the power consumption by regulating the heating and cooling depending on each room occupancy [9], [10]. Over and above, occupancy estimation can be used to have an idea about the places that attract more customers in large retail stores or shopping malls. This information can be useful for marketing purposes, advertising campaigns, and emergency exit plans.

The current state of the art for occupancy estimation can be classified mainly into two categories, namely: device-based and device-free techniques. Device-based techniques use a device on the human target to detect the number of humans in a certain space as done in [5] and [11]. In [5], the authors proposed to use Bluetooth Low Energy (BLE) tags that can be distributed to the people, then a smart-phone will scan the Bluetooth tags and update the people count through an application. The authors performed tremendous experiments during the extremely crowded outdoor environment where the results showed that a 90% detectability rate could be achieved. However, this technique has a lot of limitations especially the issue of distribution of the Bluetooth tags, and the range and power consumption of the tags. Therefore, in this thesis, we will focus on the device-free techniques where no carried devices are required for the occupancy estimation algorithm to work. In [11], the authors performed a new hybrid mechanism for occupancy estimation based on utilizing two kinds of information, which are static received signal strength indicator (RSSI) obtained from radio frequency identification (RFID) tags and dynamic RSSI taken by scattered deployed sensors.

Device-free occupancy estimation techniques utilize two main technologies, RF (radio frequency)-based and imaging-based (camera) technologies. In case of imaging-based occupancy estimation, researchers in [12], [13], and [14] make use of the camera systems to estimate the occupancy of a certain area by performing image processing algorithms. In [13], the authors proposed a new technique for crowd counting using feature extraction algorithms derived from a video based cluster. In [14], the authors estimated occupancy inside a region using a camera by dividing the area into sectors. After that, image processing techniques are used to extract features from the images and estimate the total occupancy. They were able to achieve up to 92% accuracy rate. One of the limitations of these methods is that cameras have a certain field of vision, in order to cover a wide space, extra cameras will be required which will add a large amount of processing and cost to the system. Moreover, the quality of the received image depends mainly on the environment condition changes like smoke, fog, and lighting. Thus, better results will be obtained in case of higher quality images acquired by the camera (highly environment dependent method). In addition, the privacy issue is a concern when collecting pictures or videos of people without their permission.

A lot of studies have discussed the possibility of using RF-based technologies for occupancy estimation as in [2] (Wi-Fi-based), [3; 15], and [4] (Long Term Evolution (LTE-based). In [2], authors have estimated the total number of people based on the RSSI between only one transmitter (T_x) and one receiver (R_x) connected to the same Wi-Fi network. In addition, they formulated a mathematical model to describe the human motion and they used this model to generate a theoretical formula for the RSSI probability density function (PDF) used in the occupancy estimation algorithm. They were able to estimate up to 9 people in indoor and outdoor environments with an accuracy up to ($88\% - 95\%$). In [3], knowing that channel state information (CSI) is highly sensitive to variations in environment and people movement, the authors performed theoretical and experimental analysis of their device-free occupancy estimation system, their results showed that a direct and monotonic relation exists between the number of people and CSI variations. Lastly, using LTE in occupancy estimation is not as popular as Wi-Fi-based occupancy estimation because Wi-Fi is a more developed and mature technology than LTE. However, the authors in [4] have proposed a very attractive technique for occupancy estimation based on the variations in the LTE signals specifically the reference signal received power. Their results showed how the number of people in an indoor environment can be related to changes in the LTE signal by varying the position of the LTE receiver.

In this thesis, we introduce the idea of using visible light sensing (VLS) in crowd counting and estimation. The EM spectrum varies from long waves with very low frequency to gamma rays with high frequency and consequently includes wavelengths extending from numerous kilometers for long waves down to the size less than an atom for gamma rays [16]. A different way to inspect visible light is the colors it provides. The colors that can be produced by visible light within various and small frequency band are called pure spectral colors. Table 1.1 in page 5 shows the approximate spectral colors of visible light [1]. However, VL is still not utilized by applications other than illumination. Our system utilizes VL for occupancy estimation purposes and does not interfere with other applications unlike RF-based techniques. Since light sources (LEDs) are available everywhere and power efficient [8], our method does not need to have a dedicated transmitter (available light source will be used as the transmitter). The VLS is a cost-effective novel approach and is not yet fully explored in research. Nonetheless, in [17], the authors suggested the using VLS in several applications such as, Human-Computer Interaction (gesture recognition), and indoor localization.

Color	**Wavelength**	**Frequency**
Violet	380-450 nm	668-789 THz
Blue	450-495 nm	606-668 THz
Green	495-570 nm	526-606 THz
Yellow	570-590 nm	508-526 THz
Orange	590-620 nm	484-508 THz
Red	620-750 nm	400-484 THz

Table 1.1: Approximate spectral colors of visible light.

As shown from Figure 1.1 the problem of low bandwidth in the RF spectrum is resolved by the use of VL due to the ability to use large bandwidth. In VLC, the receiver can receive only light that befalls in the same room and thus, it has the immunity security issues that occur in the RF communication system. Consequently, VLC can be used as a source of communication and illumination; it saves extra power compared to RF. Keeping in view the above benefits, VLC is one of the encouraging candidates because of its characteristics of non-licensed channels, high bandwidth, and low power dissipation. From the above points, the idea of VLC is to complement RF wireless systems, which lack satisfactory bandwidth, hence a low data rate.

1.2 Applications of Visible Light

The following are short list of some of the applications for visible light sensing and communication in our daily life.

1. Power saving and distribution inside buildings: the idea of using VL in regulating heating, ventilation, air conditioning (HVAC) and lightnings is one of the most outstanding applications for VL because it defiantly benefits in diminishing power while it keeps the same level of comfort to people.

2. Medical applications: the approach of using VL in medical purposes is one of the most significant research fields that we in WCSRL focus in. As we know, using Wi-Fi or RF is not a great scheme in hospitals because it might produce some dilemmas to patients, for instance, the rays generated from an RF machine will cause so many problems for a patient experiencing high blood pressure moreover, it might cause some interference with the medical machines and equipment. It has been shown in [18] that RF might influence the work of some machines in hospitals such as pacemakers, apnea monitors, electrically powered wheelchairs, etc.

3. Acoustic communication: underwater communication is still considered one of the obstacles in RF communication that could be resolved using the concept of VL. The tremendous cost for underwater communication using RF, the tough deployment, and high power consumption cause data rate degradation problem. Consequently, VL can be deemed as a desirable replacement approach [19].

4. Public places: one of the most beneficial applications for VL in the outdoors conditions is the vehicle to vehicles communication (V2V), which is very helpful in the pre-crash system. Using LED can accommodate potential to have communication going from traffic light to the car when there is a collision ahead of time. This could avoid likely later accidents.

Some other uses for VL could be reducing catastrophic issues like the exposure of obstacles in airplanes motors and parts so that it can be mended on time. Fatigue crash can be seen using VL and this will help in reducing any accidents with less human effort, on the other hand, human eye could not see fatigue crash. Meanwhile, heavy thunderstorm or fog, moving the aircraft for grounding is not a straight-forward operation, but it could be manageable with the aides of this technology because information regarding climate fluctuations could be transferred within VL [20], [21].

1.3 Visible Light Challenges

Since VL still in the beginning steps, it has many critical dilemmas and constraints that need to be approached. First of all, visible light can go through most objects in our daily life and this could be a security advantage and could be a coverage disadvantage. On the other hand, if you look under the table you can still see even though there is no visible light source, and this might be considered a limitation because it can prevent our signal from broadcasting to other places. Furthermore, obtaining an ideal LED that can be used for communication and illumination as well is very costly. In addition, multipath distortion is one of the prominent problems that need to be addressed for visible light since RXs can receive various images of the signal concurrently and this might cause interference. Also, the interference from sunlight is considered an issue for visible light. Lastly, lights need to be on for communication and this is the case for vast majority of industrial, retail, and commercial environments, during daylight we do tend to switch off lights so that we need to find another source for power if we desire to use VL during the day.

1.4 Research Areas in Visible Light

The following are some of the active research areas of visible light:

1. Energy saving in VL: the concept of using energy efficiently and intelligently is

known as power saving. Energy conservation diminishes the demand for energy services and can result in an eco-friendly environment, economic advancements, more prestigious and longer life quality [22].

2. Designing hybrid systems, VL systems and RF systems: in some cases, using a hybrid system will result in more efficient and powerful impacts since it will benefit from both the advantages of both VL and any other methods that could be performed in the system [17].

3. Channel modeling, understanding, and transmission algorithms in VLS: knowing different information and data about the channel state information (CSI) and how to model these kinds of data in the most useful way and utilise it in VLS is still one of the most active research areas.

4. V2V communication: it is an automobile technology invented to enable automobiles to "talk" to each other. This idea can be implemented using VL [23].

5. Localization addressing problems using VLS: visible light positioning systems aims to estimate the position of the objects by utilizing LED [24].

6. Occupancy estimation and localization using VLS.

We will be working in occupancy estimation, localization, and sensing using VL, we have taken some measurements and did some processing for the data to get useful information that can lead to determining the total number of occupants inside an environment. We also derived a mathematical model to describe our set up environment.

1.5 Organization of Thesis

This thesis will be divided into four main chapters: in Chapter I, an introduction to visible light sensing and communication, applications for visible light, challenges, and research ideas will be presented. State of the art occupancy estimation techniques will be presented in Chapter II. In Chapter III, examining the proposed model for occupancy estimation based on visible light sensing will be presented. Conclusions and future research directions will be addressed in Chapter IV.

CHAPTER II

State of the Art Occupancy Estimation Techniques

In this chapter, we are going to take a detailed look at occupancy estimation techniques and the definition of occupancy estimation. Determining the number of people inside a particular region is known as occupancy estimation or crowd counting. During the last century, researchers have proposed so many techniques to do that, all of them can be categorized as:

- Occupancy estimation using video based techniques [12], [13], and [14].
- Occupancy estimation using non-video based techniques such as RF, Wi-Fi, long term evolution (LTE), and Bluetooth [2], [3; 15], [4].

In video-based schemes or sometimes it is called device free schemes, people do not carry any devices or sensors with them, it is easy to install, and repair, also it is affordable. The main dilemma for those techniques are that cameras have to work in the line-of-sight (LOS) in order to find the accurate occupants, otherwise, it will not work. Besides, the quality of the obtained image depends essentially on the environmental variations like smoke, fog, and rain, so that if the obtained image has high quality then we will get ideal outcomes. Otherwise, it will not give us high-quality outcomes. Furthermore, some people consider taking pictures or videos of them without permission a privacy issue [12].

The authors in [12] introduced an innovative approach to address the difficulty of predicting the number of people in surveillance views. The proposed method combines a MID (Mosaic Image Difference) based foreground segmentation algorithm and a

HOG (Histograms of Oriented Gradients) based head-shoulder detection algorithm to implement an exact evaluation of occupancy estimation in the detected region. In the same paper, the MID-based foreground segmentation module gives effective blocks for the head-shoulder exposure module to identify heads and calculate the number of people. Diverse analyses are handled and impressive outcomes express the effectiveness of their approach. In [14], image segmentation and feature extraction using genetic algorithms and machine learning techniques have been used to determine the total number of occupants inside a closed area. They calculate the performance metrics like precision rate, recall rate, and accuracy and compare their results with previous results.

On the other hand, in [2] the authors have estimated the total number of people based on the received signal strength indicator (RSSI) between only one transmitter (T_x) and one receiver (R_x) located under the same Wi-Fi as shown in Figure 2.1. In addition, the authors have discovered a mathematical equation to express the human movement and they used this mathematical equation to verify that their approach works fine. They were able to count up to 9 people in indoor and outdoor scenes.

A different approach for crowd estimation is based on the variations in the Channel State Information (CSI), since CSI is extremely sensitive to changes in environmental fluctuations, this idea has been used in estimating the number of people inside a closed-door area. The authors in [3] introduced a device free crowd counting (FCC) procedure based on variations in CSI. They did a theoretical and experimental investigation for the suggested scheme, their results show that a direct and monotonic relationship exists between the number of people and CSI variations. They have shown that their approach works well compared to other approaches in terms of accuracy, scalability, and reliability. They compared their work to [25] and showed that their approach outperforms the one in [25].

Figure 2.2 shows the crowd estimation results for both Sequential Counting Par-

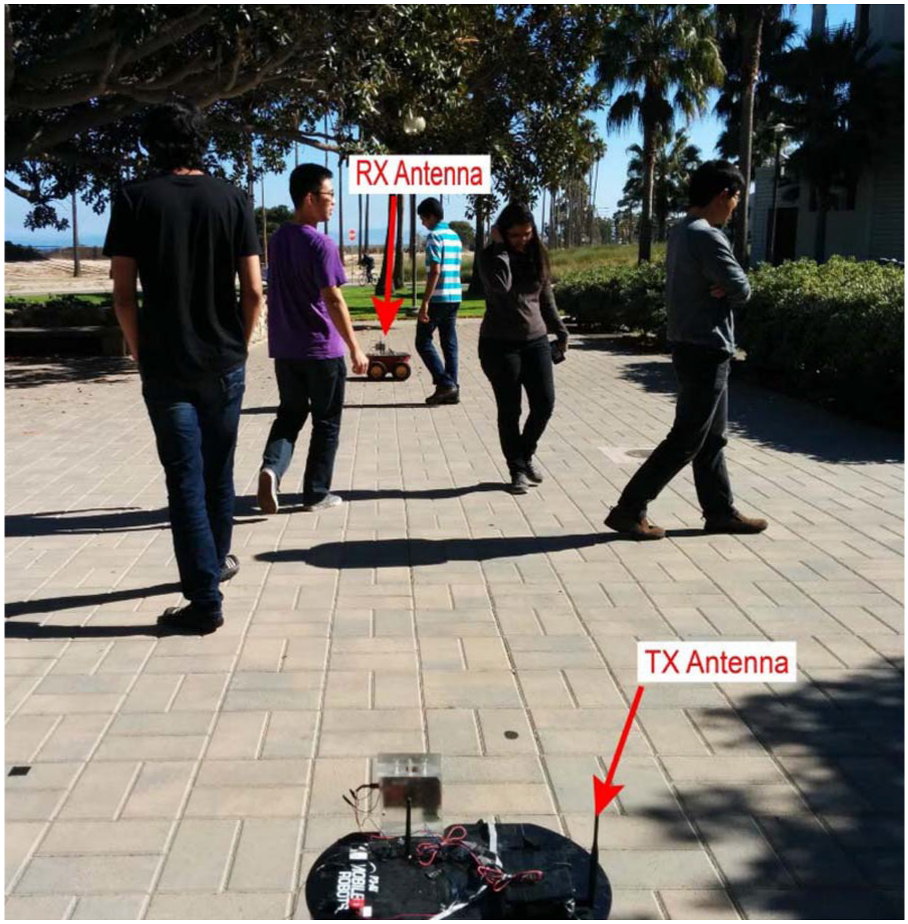
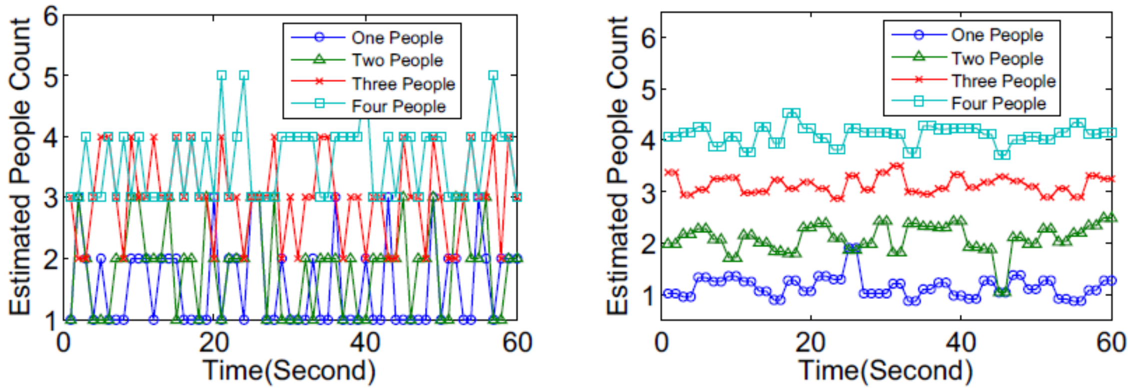


Figure 2.1: Transmitter and receiver to determine the total number of people [2].

allel Localization (SCPL) and FCC respectively, it can be easily noticed that the estimated number of people usually fluctuates and never gives accurate results in SCPL. The authors shown that 45% results are accurate. On the other hand, FCC shows more stability due to the intensive devices needed in SCPL links.

On the other hand, the use of long term evolution (LTE) signals for physical analysis ideas has gained limited attention than the use of Wi-Fi probably for the pragmatic reason that Wi-Fi technologies are around for a long time. Nevertheless, this does not mean that LTE is not used in crowd estimation. In [4], the authors have offered a highly attractive procedure that is used in occupancy estimation in indoor environments. In fact, they have done a number of experiments to judge the relation between the number of people and the variations in the LTE signals. Their



(a) SCPL (b) FCC

Figure 2.2: Comparison results from [3].

conclusions show that they were capable to distinguish the correlation between the number of people and the changes in the LTE signal. They did not give an accurate amount for the number of people but they were able to do some classification. Figure 2.3 shows their results, All of the beforehand discussed algorithms can be viewed as device free based approaches, one of the most attractive and familiar device-based strategies is the use of Bluetooth to count the number of people inside an environment as in [5]. The authors suggested using Bluetooth low energy (BLE) tags that can be scattered to the people, after that using smartphone equipped with Bluetooth reading application, the authors calculated the number of people. They did an extensive set of experiments lasted for the most gathering and crowded 5 days in the year (The Hajj), there will be around 5-6 million people in an area as small as Stillwater, Oklahoma). The results show that 90% detectability rate could be achieved. See Figure 2.4 for the system model used in [5].

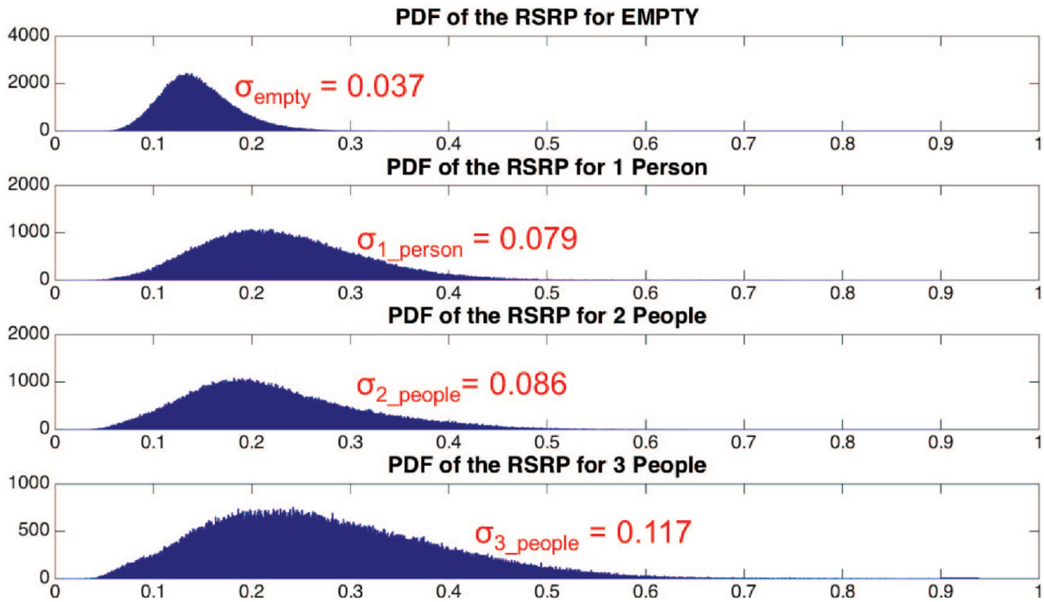


Figure 2.3: PDF for received signal for different number of people [4].

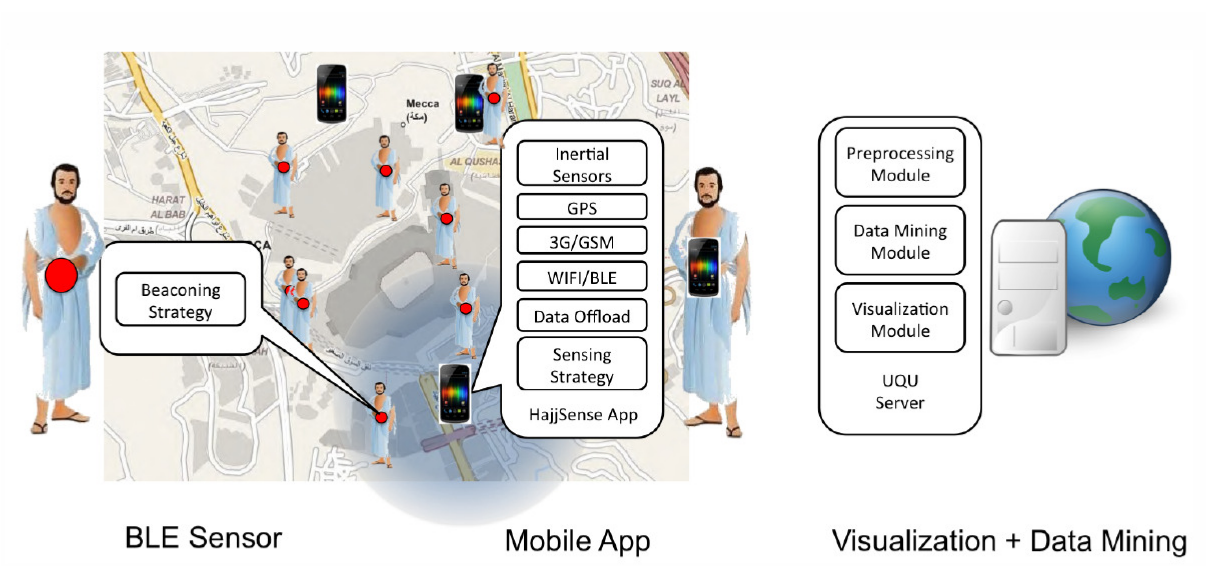


Figure 2.4: System architecture [5].

CHAPTER III

Occupancy Estimation through Visible Light Sensing

3.1 Introduction

During the last few years, the average energy consumption in the world is getting very huge and growing rapidly. In fact, Oklahoma is one of the most productive states for gas in the US and so many studies show that the average energy consumed per person (per capita) is greater than 75% of the other states [26].

As all of us know, the world population is increasing rapidly and the demands for comfort and building services is growing dramatically, which also leads to the exhaustion of energy resources and adverse environmental impacts. Because of that, it is the job of leading organizations to raise awareness about the increasing rate of energy consumption and contribute toward mitigating it by being more efficient in the use of energy. Consequently, an essential need for effective, adaptive as well as feasible solutions that decrease energy consumption in buildings while maintaining the current levels of comfort is prominent. The use of visible light as a source of communication can help in the reduction of consumed energy by trying to control heating, ventilation, and air conditioning in large buildings (HVAC). The way we control this is by trying to predict the total number of people inside a building. On the other hand, occupancy estimation has so many advantages in the business and marking. For instance, in the big retail stores like malls we can know which palaces are more crowded than others, so we can make more announcements over the crowded places to help attract the customers attention. Another important place in need of occupancy estimation using VLS is airports. Knowing the crowded places in airports can help in managing and directing more staff toward these places. We can do more announcements and advertisements to attract visitor attention. The main advantage for visible light is that there are no extra installment jobs needed, it is everywhere around so, we can benefit from it and increase its utilization as much as we can [8; 6].

3.2 System Model

Our system model is presented in Figure 3.1, where we showed a number of people randomly walking in the environment, some of them crossing the line-of-sight (LOS) and the others are not. In this system model, we described a simple case where some people block the light beams. In this case, the light detector calculates the received light intensity from the reflected light beam when someone passes by and interrupts the beam. The received light intensity will decrease by a noticeable amount. Once the PDF and CDF of the received power are found, using simple signal processing algorithms and machine learning tools, we will be able to estimate the number of occupants in the room. Although finding a mathematical model for human walk is a challenging problem and not in the scope of this thesis, we assumed that people in our environment are walking in a simple probabilistic motion model as introduced by [2] (for simplicity). Based on this model, we derived a mathematical model for the received signal power, which will be impacted by the number of people crossing the LOS[1]. After building a database for the possible PDFs and CDFs of the received power under different number of occupant scenarios, the PDF obtained is compared with the database to select the best estimate for the occupancy of the room.

[1]Although there are some components from the non line-of-sight (NLOS) at the receiver, we assumed that the LOS is the main contributor to the received signal (for simplicity of the theoretical analysis).

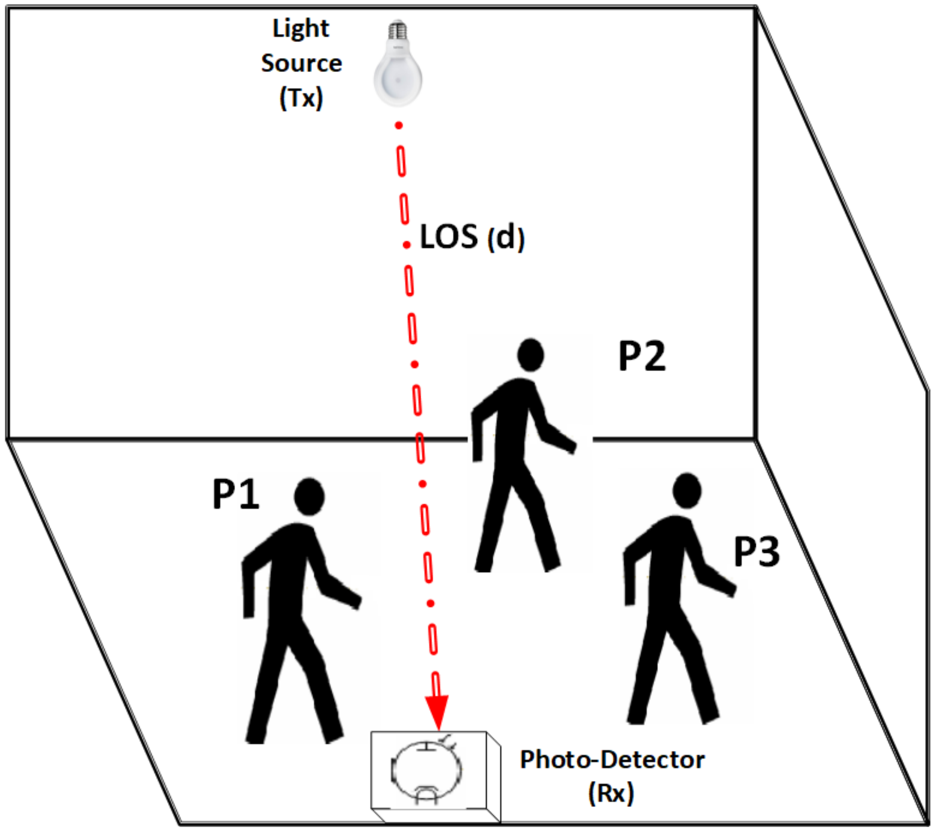


Figure 3.1: Occupancy estimation system model.

3.3 Theoretical Analysis

In this section, we will provide theoretical analysis for the received light power. Following that, we provide the simulation results. For a tractable mathematical analysis, we assume there is no scattering and multipath fading. We started our derivation with the received power as follows

$$y = Cd^{-\gamma} + \omega, \tag{3.1}$$

where y is the received signal power, C is a constant, which incorporates the losses, transmit power, etc, d is the distance between the transmitter and the receiver, γ is the path loss exponent, which has direct relation to the environmental conditions, and ω is the additive white Gaussian noise (AWGN) with mean (μ) and variance (σ^2).

Figure 3.2: Taking Measurements.

PDF expression of ω is given by

$$f_W(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\omega-\mu)^2}{2\sigma^2}}. \tag{3.2}$$

Since there are two possible values for the received power, the path loss exponent follows the Bernoulli random distribution with the values of no-crossing as γ_0 and crossing as γ_c. On the other hand, the probability mass function (PMF) (P_r) of k number of people crossing among N people can be modeled by Binomial distribution, i.e., the probability of k successes in N independent experiments. Therefore, the PMF is given as

$$\begin{aligned} P_r(k; N, P_c) &= P_r(X = k) \\ &= \binom{N}{k} P_c^k (1 - P_c)^{N-k}, \end{aligned} \tag{3.3}$$

where $k = 0, 1, 2 \ldots N$ where P_c is the probability of crossing. In [2], the authors modeled the casual walking and found that the asymptotic probability of a cross (P_c) of a single person can be characterized as $P_c = \frac{2d_{step}}{\pi L}$, assuming d_{step} is the total distance traveled by the occupants and L is the length of the testing area (room). The probability of no-crossing is given by

$$\begin{aligned} P_r(\gamma_0) = P_r(X = 0) &= \binom{N}{0} P_c^0 (1 - P_c)^{N-0} \\ &= (1 - P_c)^N. \end{aligned} \tag{3.4}$$

Similarly, the probability of at least one crossing can be obtained as

$$\begin{aligned} P_r(\gamma_c) &= P_r(X \geq= 0) \\ &= 1 - P_r(\gamma_0) = 1 - (1 - P_c)^N. \end{aligned} \tag{3.5}$$

Therefore, the PMF of path loss exponent is

$$\gamma = \begin{cases} \gamma_o, & (1 - P_c)^N \\ \gamma_c, & 1 - (1 - P_c)^N \end{cases} \tag{3.6}$$

From the theory of mixed random variables [27], it is known that for a discrete random variable X with range $R_X = x_1, x_2, x_3, ...$ and PMF $P_x(x_k)$, we can define the (generalized) PDF by using Delta Dirac function as

$$f_X(x) = \sum_{x_k \epsilon R_x} P_x(x_k)\delta(x - x_k). \tag{3.7}$$

Hence, the PDF of γ can be given as

$$f_\gamma(\gamma) = (1 - P_c)^N \delta(\gamma - \gamma_0) + (1 - (1 - P_c)^N)\delta(\gamma - \gamma_c). \tag{3.8}$$

Now, let $x = g(\gamma) = Cd^{-\gamma}$ then, using the transformation of random variables [27], the PDF of x can be obtained as

$$f_X(x) = f_\gamma(g^{-1}(x)) \left| \frac{\partial}{\partial x} g^{-1}(x) \right|, \tag{3.9}$$

given

$$\gamma = g^{-1}(x) = \log_d\left(\frac{C}{x}\right), \tag{3.10}$$

and

$$\frac{\mathrm{d}\gamma}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}g^{-1}(x) = \frac{1}{x\ln(d)}. \tag{3.11}$$

We have

$$f_X(x) = f_\gamma\left(\log_d\left(\frac{C}{x}\right)\right)\left|\tfrac{1}{x\ln(d)}.\right| \tag{3.12}$$

Then, by substituting (3.8) in (3.12),

$$f_X(x) = \left|\frac{1}{x\ln(d)}\right|\left\{(1-P_c)^N\delta\left(\log_d\left(\frac{C}{x}\right)-\gamma_0\right)+\right.$$
$$\left.(1-(1-P_c)^N\delta\left(\log_d\left(\frac{C}{x}\right)-\gamma_c\right)\right\}. \tag{3.13}$$

Finally, given the PDF of x and ω, we can now derive the expression for the PDF of received power:

$$y = Cd^{-\gamma}+\omega = x+\omega, \tag{3.14}$$

which is nothing but convolution of PDFs, hence the PDF of y is given by

$$f_Y(y) = f_X(x) * f_W(\omega) = \int_{-\infty}^{-\infty} f_X(x)f_W(y-x)dx, \tag{3.15}$$

substituting (3.2) and (3.13) into (3.15), we get:

$$f_Y(y) = \int_{-\infty}^{\infty}\left|\frac{1}{x\ln(d)}\right|\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-x-\mu)}{2\sigma^2}}$$
$$\left\{(1-P_c)^N\delta\left(log_d\left(\frac{C}{x}\right)-\gamma_0\right)\right.$$
$$\left.+(1-(1-P_c)^N))\delta\left(log_d\left(\frac{C}{x}\right)-\gamma_c\right)\right\}dx, \tag{3.16}$$

by using the U-substitution and the shifting property of the Dirac Delta function [28], the PDF of the received signal power can be obtained as

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\Bigg\{(1-P_c)^N e^{-\frac{(y-\mu-Cd^{-\gamma_0})^2}{2\sigma^2}} + \\ (1-(1-P_c)^N)e^{-\frac{(y-\mu-Cd^{-\gamma_c})^2}{2\sigma^2}}\Bigg\}, \quad (3.17)$$

Finally, by integrating (3.17), we can readily obtain the CDF expression as

$$F_Y(y) = \frac{1}{2}\Bigg\{(1-P_c)^N\left(1+\phi\left[\frac{(y-\mu-Cd^{-\gamma_0})}{\sqrt{2}\sigma}\right]\right) + \\ (1-(1-P_c)^N)\left(1+\phi\left[\frac{(y-\mu-Cd^{-\gamma_c})}{\sqrt{2}\sigma}\right]\right)\Bigg\}, \quad (3.18)$$

where $\phi(x)$ is the error function of x defined by [29]:

$$\phi(x) = \frac{2}{\pi}\int_0^x e^{-t^2}dt. \quad (3.19)$$

In statistics, such distributions are called as finite mixture, also known as mixture distributions [30].

3.4 Occupancy Estimation Algorithm

First of all, we need to generate PDFs and CDFs database[2] for each occupancy scenario. After finding the received power PDF and CDF during the simulation. We have to compare this PDF with the PDF in the database using two methods, KL-divergence and Euclidean distance. The nearest PDF in the database which minimizes the distance with the obtained PDF during the simulation will give us the estimated occupancy.

[2]Measurements are collected for a known number of people to generate the PDFs.

The KL-divergence is the logarithmic difference between two PDFs as shown in (3.20), which also can be called the entropy estimation. Fundamentally, KL-divergence measure how one PDF deviates from another. For discrete probability distributions f_1 and f_2, the KL-divergence from Q to P is defined as shown in (3.20), [31], which can be denoted by $D_{KL}(f_1||f_2)$.

$$D_{KL}(f_1||f_2) = -\sum_i f_1(i) \log \frac{f_2(i)}{f_1(i)} = \sum_i f_1(i) \log \frac{f_1(i)}{f_2(i)}. \tag{3.20}$$

On the other hand, the Euclidean distance is the straight line connecting two points p and q. In Cartesian coordinates, if $p = (p_x, p_y)$ and $q = (q_x, q_y)$ are two points in Euclidean space, then the distance (d) from p to q is given by [31]

$$D(p, q) = \sqrt{(q_x - p_x)^2 + (q_y - p_y)^2}. \tag{3.21}$$

3.5 Simulation and Analysis Results

The parameters which we used in our simulations[3] are as follow:

- Probability of crossing the LOS, $P_c = 0.2$.
- Noise variance (power), $\sigma_{dB} = -10$ dB.
- Noise mean, $\mu_{dB} = 4$ dB.
- Channel constant, $C_{dB} = 10$ dB.
- Path loss exponent when there are no people crossing the LOS, $\gamma_0 = 1$.
- Path loss exponent when there is at least one person crossing the LOS, $\gamma_c = 4$.

[3] These parameters have been selected for illustration purposes. The results/insights will not change with these numbers. It is expected to get more realistic parameters when have some real measurements data.

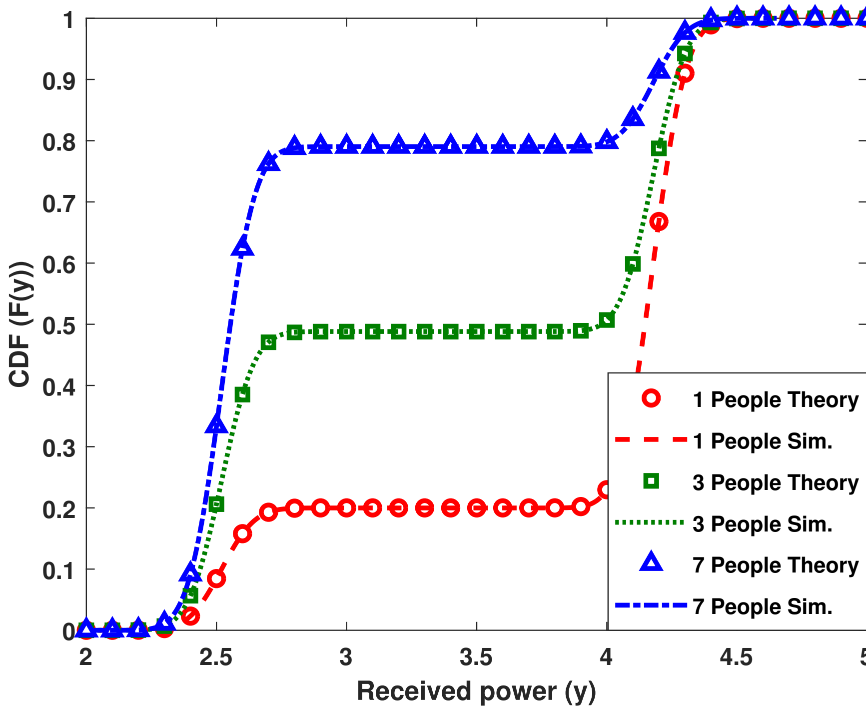


Figure 3.3: Comparison results between the theoretical and simulated CDF.

In Figure 3.3, a comparison results between the theoretical and simulation CDF for different number of people is shown. It is observed that both the simulated (dashed lines) and the theoretical (circles) CDFs are on top of each other that proves the validity of our approach.

In Figure 3.4, a comparison results between the theoretical and simulation PDFs for different number of people is presented. Similarly, the theoretical and simulation results are matching with each other in case of CDFs. Furthermore, as the room occupancy varies, the PDF and CDF of the received power changes. As expected, higher number of people increases probability of crossing the LOS, which decreases the received power. Therefore, the first peak (modal) in the PDF is higher and the second will be lower. Similarly, if there are fewer people crossing the LOS, then the probability of the received power will be higher, which indicates the first peak will be lower and the second peak will be higher. These two scenarios are clearly shown in Figure 3.4 for different number of occupants.

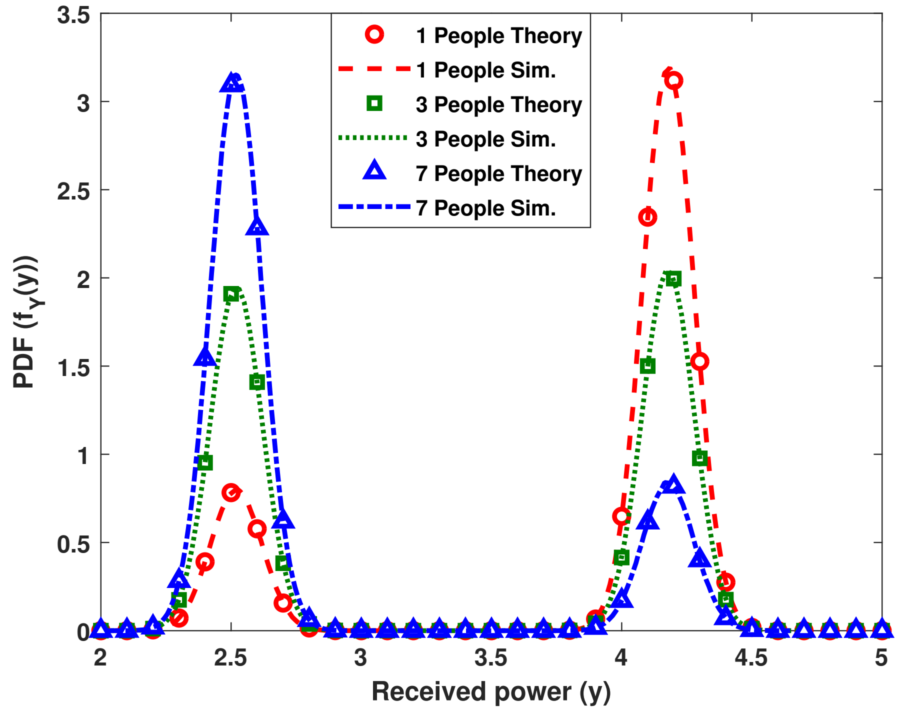


Figure 3.4: Comparison results between the theoretical and simulated PDF.

In Figure 3.5, the KL-divergence for a different number of people is shown. As we mentioned earlier, the argument that minimizes it will be our approximate occupants. It is shown for four different cases how the algorithm will estimate the occupancy correctly depending on the minimum value of KL-divergence between the database PDFs and the measured PDF.

In Figure 3.7, we showed how the mean square error (MSE) decreases as the number of samples used in the PDF calculation increases, this implies that our algorithm requires a large number of samples to be used for the occupancy estimation. Moreover, KL-divergence is more accurate than Euclidean distance when using less than 1,000 samples in the estimation which can be explained by the fact that KL-divergence is used to measure the difference between PDFs while Euclidean distance is a general equation to measure the distance between two points in space.

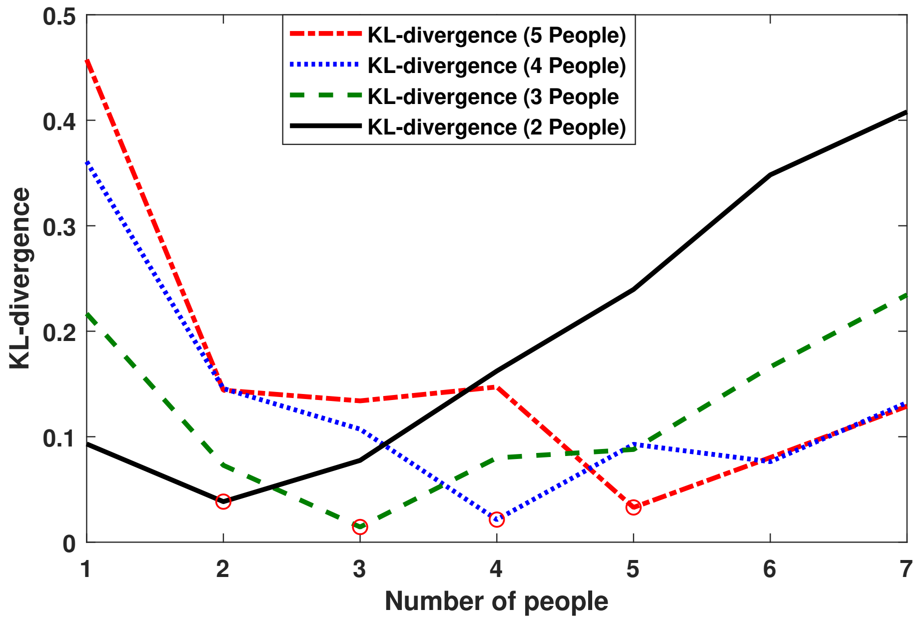


Figure 3.5: KL-divergence for different number of people.

3.6 Experimental Results

In this section, we will show our experimental results for PDFs, CDFs, and some of the statistical parameters that we have investigated. Before we start talking about the results, lets talk in detail about the hardware equipment's that were used in taking the measurements:

1. The Raspberry Pi is an inexpensive, credit-card sized computer that is connected to a computer monitor or TV, with a keyboard and mouse. It allows people of all ages to investigate computing and to discover how to program in languages like Scratch and Python. Its proficient of doing everything you would expect a desktop computer to do, from browsing the internet and playing high-definition video, to making spreadsheets, word-processing, and playing games [32].

2. Photo-Diode (PDA-100A2): The PDA-100A2 Photo-diode Array Detector shown in Figure 3.8 is an optical detector able of estimating the absorbance spectrum

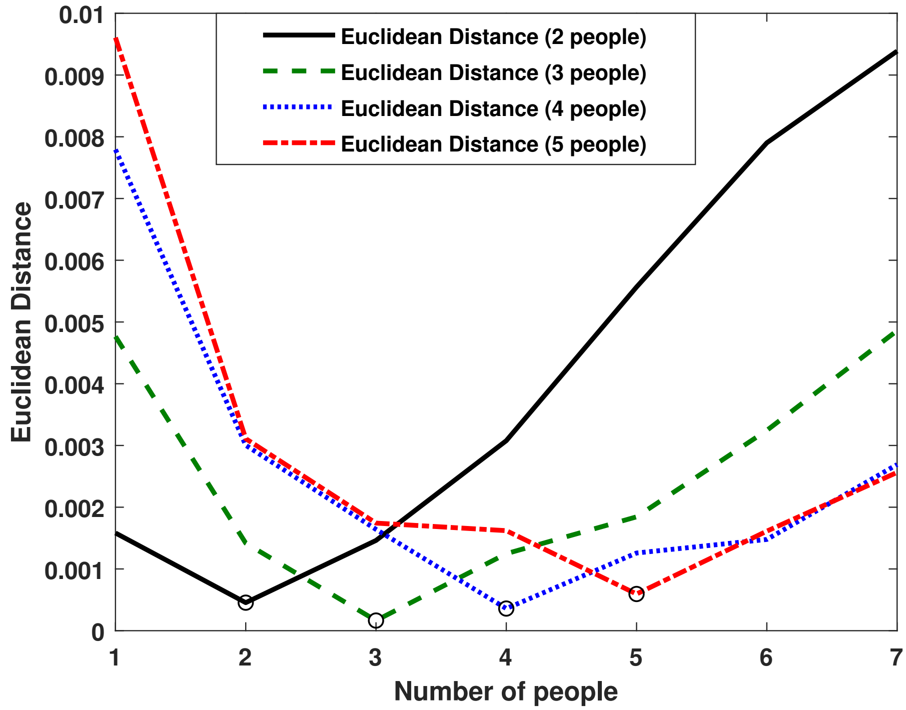


Figure 3.6: Euclidean Distances for different number of people.

from 190 nm to 800 nm. We have used this photo-diode as our (R_x) ,[33].

3. LED Flash Light: hand-held rechargeable zoom-able LED light has been used in data gathering process as a sender source (T_x) [34].

In Figure 3.9, a comparison results between the experimental CDF for a different number of people is shown. It is observed that as the number of people increased the probability of crossing the LOS will increase resulting in a higher CDF as proved in this figure.

Besides, the mathematical expectation or sometimes known as the mean which can be defined as is the central value of a discrete set of numbers: specifically, the sum of the values divided by the number of values of the data is shown in Figure 3.10, it can be easily shown as the number of people crossing the LOS gets bigger, the mean of the data gets smaller.

In probability theory and statistics, the variance is the expectation of the squared

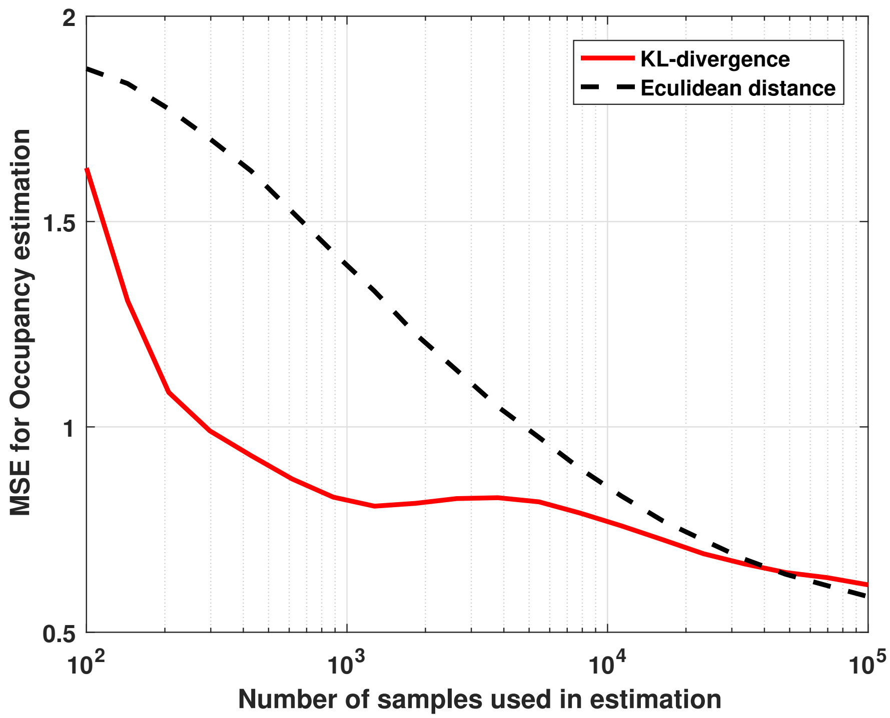


Figure 3.7: The MSE in the occupancy estimation when applying KL-divergence and Euclidean distance techniques using 50,000 iteration for each case.

difference of a random variable from its mean. Informally, it measures how far a set of (random) numbers are spread out from their mean. In Figure 3.11, the variance for different number of people is plotted. As the number of people crossing the LOS increased the variance of the data increased. The same idea can be applied to the standard deviation which can be defined as a measure that is used to quantify the amount of variation or dispersion of a set of data values, in other words it is the square root of its variance and it is used to measure confidence in statistical conclusions. In Figure 3.12, the standard deviation of the data is shown.

Figure 3.8: Photo-detector used in the project.

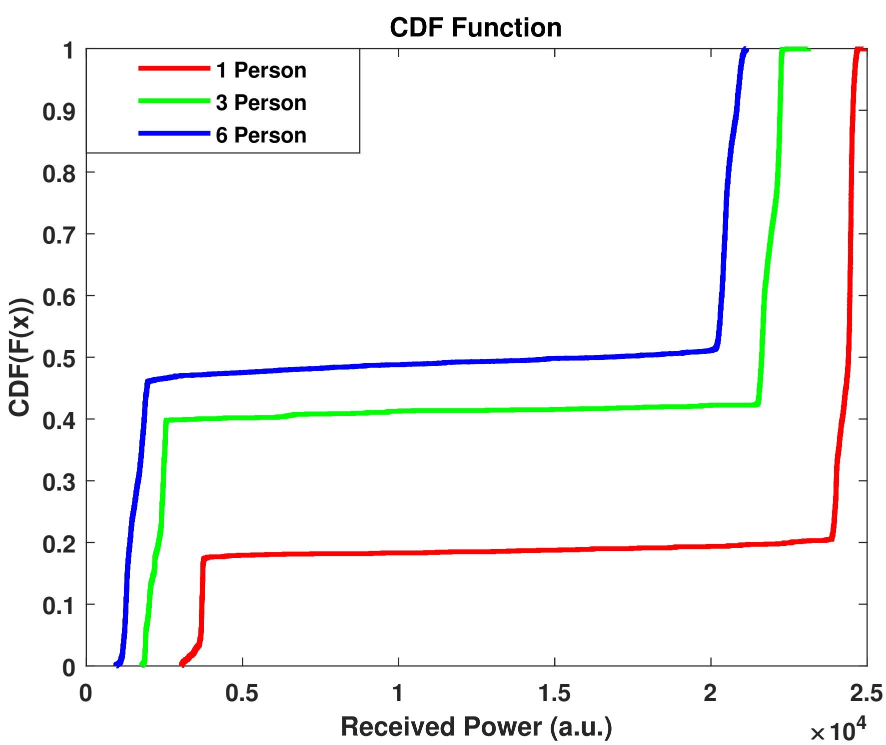


Figure 3.9: CDF of the received power for different number of people.

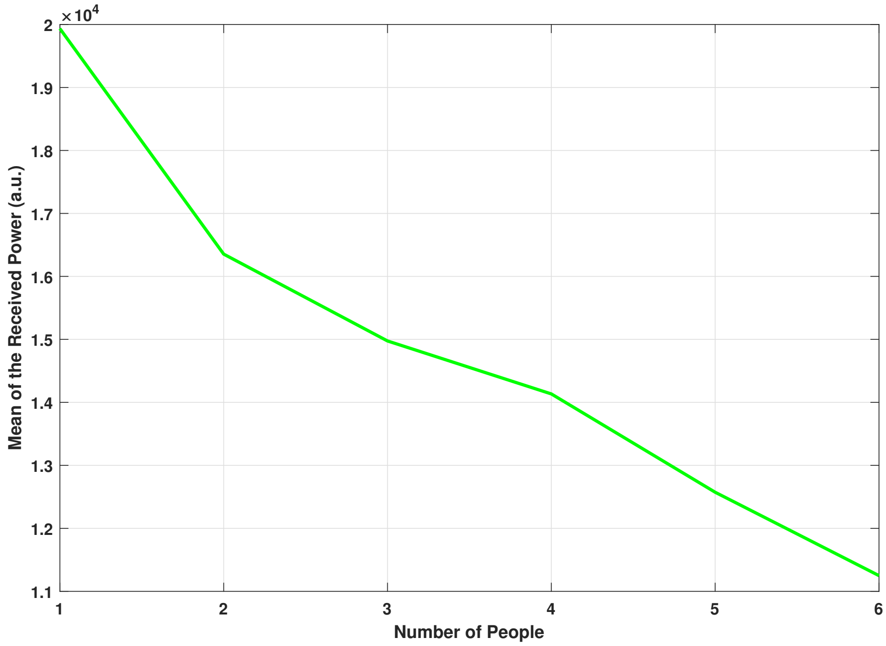


Figure 3.10: Mean of the data.

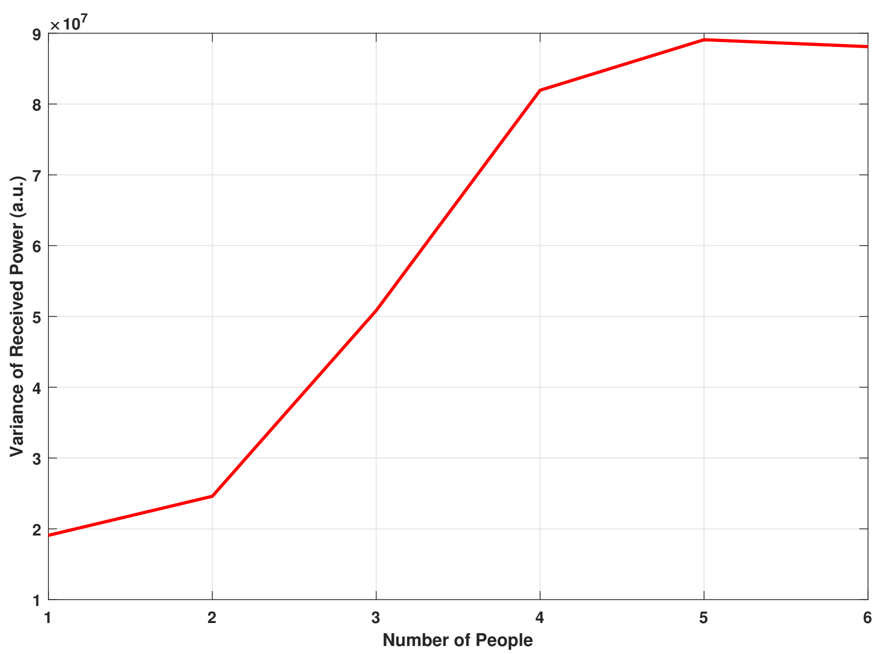


Figure 3.11: Variance of the received power versus number of people.

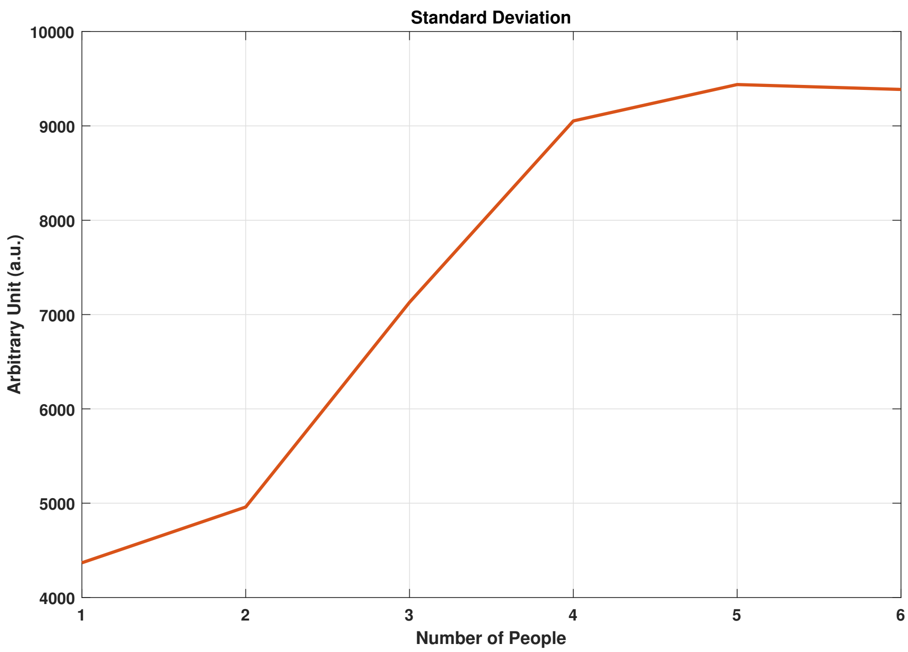


Figure 3.12: Standard deviation of the data.

CHAPTER IV

CONCLUSION AND FUTURE WORK

4.1 Conclusions

In this research, an innovative system for occupancy estimation using only VLS is presented. Chapter 1 gives a brief introduction about VLS, its definition, applications, limitations, and some of the research ideas for visible light sensing. The state of the art occupancy estimation techniques were presented in Chapter 2. In Chapter 3, a novel technique for crowd counting and occupancy estimation using only VLS is proposed. Conclusion and future work were presented in Chapter 4.

LED is utilized as the transmitter and Photo-diode as the receiver. A mathematical model for the PDF and CDF of the received power is derived. Also, we have used the KL-divergence and the Euclidean distance algorithms as an estimation methods for the total number of occupants by stating that the argument that minimize the difference between the database PDFs and measured data PDF is the estimated number of occupants. Our simulation and theoretical results for both the PDFs and the CDFs were matching. Finally, it was shown that KL-divergence performs is more accurate than Euclidean distance when using less than 1,000 samples in the occupancy estimation. Furthermore, it was shown that our experimental results for both the PDFs and the CDFs were comparable to the ones in the simulation and the theoretical consequences.

4.2 Future Work

The following are some of the fundamental concepts that could be examined and studied more in detail and provide more outcomes to the field of visible light sensing and communication:

- The use of various algorithms for occupancy estimation using visible light sensing and performing a comparison to see which one is the most applicable in real life: as discussed in the earlier chapters, different algorithms may be implemented using visible light to find occupancy estimation.

- Develop a comprehensive scheme that can regulate HVAC and lighting: power saving is one of the most important concerns researchers aspiring to achieve in several sections of research. Developing a complete system that can control and manage HVAC using VLS based on occupancy estimation is considered as future work.

- Studying the scattering effect of visible light sensing: this is one of the most important issues in VLS and VLC. In this thesis, we have scattering effects caused by the surrounding objects in the field and these objects reflect light from the T_x. We will study the effects of scattering in the up coming future works.

- Solving the synchronization problem between various receivers: using multiple PDs and study the RSSI and see the impacts for using multiple receivers on the system are other future research directions.

BIBLIOGRAPHY

[1] "Electromagnetic spectrum," Mar 2019. [Online]. Available: https://en.wikipedia.org/wiki/Electromagnetic_spectrum

[2] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1381–1393, 2015.

[3] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using WiFi," in *Infocom, 2014 proceedings ieee*. IEEE, 2014, pp. 361–369.

[4] S. Di Domenico, M. De Sanctis, E. Cianca, P. Colucci, and G. Bianchi, "LTE-based passive device-free crowd density estimation," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[5] A. Basalamah, "Sensing the crowds using Bluetooth low energy tags," *IEEE Access*, vol. 4, pp. 4225–4233, 2016.

[6] T. Komine, "Visible light wireless communications and its fundamental study," Ph.D. dissertation, Ph. D. dissertation, Keio University, 2005.

[7] S. Fuada, A. P. Putra, and T. Adiono, "Analysis of received power characteristics of commercial photodiodes in indoor LoS channel visible light communication," *Int. J. of Advanced Computer Science and Applications*, 2017.

[8] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED Based Indoor Visible Light Communications: State of the Art." *IEEE communications surveys and tutorials*, vol. 17, no. 3, pp. 1649–1678, 2015.

[9] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building*. ACM, 2010, pp. 1–6.

[10] T. A. Nguyen and M. Aiello, "Energy intelligent buildings based on user activity: A survey," *Energy and buildings*, vol. 56, pp. 244–257, 2013.

[11] D. Zhang, Y. Yang, D. Cheng, S. Liu, and L. M. Ni, "COCKTAIL: An RF-based hybrid approach for indoor localization," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[12] M. Li, Z. Zhang, K. Huang, and T. Tan, "Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008, pp. 1–4.

[13] S. Choudhary, N. Ojha, and V. Singh, "Real-time crowd behavior detection using SIFT feature extraction technique in video sequences," in *Intelligent Computing and Control Systems (ICICCS), 2017 International Conference on*. IEEE, 2017, pp. 936–940.

[14] A. K. Chandran, A. Subramaniam, W. C. Wong, J. Yang, and K. A. Chaturvedi, "A PTZ camera based people-occupancy estimation system (PCBPOES)," in *Machine Vision Applications (MVA), 2017 Fifteenth IAPR International Conference on*. IEEE, 2017, pp. 145–148.

[15] E. Cianca, M. De Sanctis, and S. Di Domenico, "Radios as sensors," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 363–373, 2017.

[16] A. Cailean and M. Dimian, "Current Challenges for Visible Light Communica-

tions Usage in Vehicle Applications: A Survey," *IEEE Communications Surveys & Tutorials*, 2017.

[17] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *ieee communications surveys & tutorials*, vol. 17, no. 4, pp. 2047–2077, 2015.

[18] H. Bassen, "Radiofrequency interference with medical devices. a technical information statement." vol. 17, pp. 111–114, 05 1998.

[19] I. F. Akyildiz, D. Pompili, and T. Melodia, "Challenges for efficient communication in underwater acoustic sensor networks," *ACM Sigbed Review*, vol. 1, no. 2, pp. 3–8, 2004.

[20] F. Khan, S. R. Jan, M. Tahir, and S. Khan, "Applications, limitations, and improvements in visible light communication systems," in *Connected Vehicles and Expo (ICCVE), 2015 International Conference on.* IEEE, 2015, pp. 259–262.

[21] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, "Estimating crowd density in an RF-based dynamic environment," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3837–3845, 2013.

[22] O. Zehner, "Unintended consequences of green technologies," *Green technology. Sage, London*, pp. 427–432, 2011.

[23] H. Abuella, S. Ekin, and M. Uysal, "Vildar: a novel speed estimation system using visible light in vehicles," in *2017 IEEE 38th Sarnoff Symposium.* IEEE, 2017, pp. 1–6.

[24] C. Zhang and X. Zhang, "Visible light localization using conventional light fixtures and smartphones," *IEEE Transactions on Mobile Computing*, 2018.

[25] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. An, "SCPL: Indoor device-free multi-subject counting and localization using radio signal strength," in *Information Processing in Sensor Networks (IPSN), 2013 ACM/IEEE International Conference on.* IEEE, 2013, pp. 79–90.

[26] A. E. Outlook *et al.*, "Energy information administration," *Department of Energy*, vol. 92010, no. 9, pp. 1–15, 2010.

[27] H. Pishro-Nik, "Introduction to probability, statistics, and random processes," 2016.

[28] G. B. Arfken and H. J. Weber, "Mathematical methods for physicists," 1999.

[29] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Courier Corporation, 1964, vol. 55.

[30] G. McLachlan and D. Peel, *Finite Mixture Models.* John Wiley & Sons, 2004.

[31] M. M. Deza and E. Deza, "Encyclopedia of Distances," in *Encyclopedia of Distances.* Springer, 2009, pp. 1–583.

[32] T. Magazine, "Raspberry pi 3 is out now! specs, benchmarks & more," 2016.

[33] I. Thorlabs. (2019) Photo Diode Detector kernel description. [Online]. Available: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/28829-31898-03_PDA_Photo_Ara_Dect.pdf

[34] I. ALPHA TEK. (2019) Floodlight TACTICAL FLASHLIGHT kernel description. [Online]. Available: https://www.amazon.com/RECHARGEABLE-ZOOMABLE-Floodlight-Spotlight-FLASHLIGHT/dp/B01A03CJGY

[35] S. Arnon, *Visible light communication.* Cambridge University Press, 2015.

[36] L. U. Khan, "Visible light communication: Applications, architecture, standardization and research challenges," *Digital Communications and Networks*, vol. 3, no. 2, pp. 78–88, 2017.

[37] D. Gujjari, "Visible light communication," 2012.

[38] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: indoor location sensing using active RFID," *Wireless networks*, vol. 10, no. 6, pp. 701–710, 2004.

[39] G. Simon, G. Zachár, and G. Vakulya, "Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 9, pp. 2337–2348, 2017.

[40] N. Kumar, "Visible light communication based traffic information broadcasting systems," *International Journal of Future Computer and Communication*, vol. 3, no. 1, p. 26, 2014.

[41] Y. Wang, "Indoor Localization Based on Visible Light Communication," Ph.D. dissertation, 2017.

[42] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Pervasive computing and communications (PerCom), 2013 IEEE international conference on.* IEEE, 2013, pp. 193–200.

[43] Y. Yang, J. Hao, J. Luo, and S. J. Pan, "Ceilingsee: Device-free occupancy inference through lighting infrastructure based LED sensing," in *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on.* IEEE, 2017, pp. 247–256.

[44] S. Vijay and K. Geetha, "A survey on visible light communication appliances used in inter-vehicular and indoor communication," *International Journal of Applied Engineering Research*, vol. 11, no. 7, pp. 4893–4897, 2016.

[45] S. Kullback, *Information theory and statistics.* Courier Corporation, 1997.

VITA

Mohammad Al Mestiraihi

Candidate for the Degree of

Master of Science

Thesis: OCCUPANCY ESTIMATION THROUGH VISIBLE LIGHT SENSING (VLS)

Major Field: Electrical and Computer Engineering

Biographical:

Personal Data: Born in Janin Alsafa, Irbid, Jordan on January 25, 1985.

Education:

Received the Master of Engineering degree from Jordan University for Science and Technology, Irbid, Jordan, 2011, in Computer Engineering

Completed the requirements for the degree of Master of Science with a major in Electrical and Computer Engineering, Oklahoma State University in May, 2019.

Education:

Received the B.S. degree from Al Yarmouk University, Irbid, Jordan, 2008, in Computer Engineering