Visible Light Positioning with Unmodified LEDs

James Patrick Noonan III

Department of Computer Science

Boston College

Supervisor

Siddhartan Govindasamy, Sergio Alvarez

In partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science

December 2023

Acknowledgements

I'd like to thank both my Thesis advisors, Siddhartan Govindasamy and Sergio Alvarez. I appreciate all of the time they've given me and the faith they put in me to pursue this project. I'd like to give a special thanks to Siddhartan, however, who graciously invited me into his lab, birthed our project's concept and guided me throughout the process. Without his support, I would not have had such an opportunity nor been able to see it to fruition. I'd also like to thank Brijesh Soni, who was another indispensable source of guidance. It is in no small part because of these three that this work exists today. Finally, I'd like to thank both of my parents who have always supported me throughout all my endeavors. They are the reason for the curiosity and creativity that inspired my pursuit and facilitated the completion of this thesis. Without them, I wouldn't be the person I am today.

Abstract

As industry and academia become increasingly automated, the need for cost-effective, scalable indoor positioning algorithms increases. Current approaches, in the new research field of Visible Light Positioning (VLP), utilize modulated Light-Emitting Diodes (LEDs) for localization. In these works, luminous intensity levels of LEDs are varied rapidly, and the received light levels are used to find the position of a receiver. In this paper, we propose an alternative positioning technique that utilizes visible light, but with unmodified LEDs. Our proposal includes a rudimentary approach, as well as a proof of concept for a cost-effective, easily replicable and robust technique for indoor localization. We leverage the inherent manufacturing variations in LED bulbs, that result in uniquely identifiable light emission, to localize two-dimensional position directly through a simple multiple output regression neural network trained on light samples taken in the target environment, eliminating the need for component decomposition, hardware modification, or triangulation.

Contents

Chapter 1

Introduction

As industry becomes increasingly automated, the need for free-motion indoor navigation becomes more apparent. Take, for instance, a commercial warehouse setting. To automate the organization of product within the warehouse, physical tracks with sensors are commonplace [1]. Such approaches can prove to be costly and rigid. The benefit of Visible Light Positioning (VLP) is that it is both less expensive and allows for free-motion navigation. That is, the robot need not be attached to a physical track to interpret its location. As will be addressed in this chapter, procedures for VLP have been previously proposed, however, they require prior knowledge regarding the system. For that reason, custom LED's and rigs were used. The novelty of our technique is that no customization is necessary; stock LED bulbs, sockets, wiring and power can be used for positioning. Thus, our system proves more cost effective, scalable and less invasive.

1.1 Previous Works

Three competing approaches will be analyzed for their efficacy in localization.

In [2], a procedure was designed to track the one-dimensional distance from LED to sensor within an accuracy of 3 centimeters. The distance was calculated using the prior known LED transmission signal $s(t)$ and the channel attenuation $\alpha(t)$, which was gathered via an initial calibration. For clarity, $\alpha(t)$ represents the effect of the medium, through which the light travels between the bulb and the sensor, on the light itself. A stock LED was used for light emission. To have the LED emit a unique signal $s(t)$, however, a custom driver was designed to output said signal through the LED. More specifically,

their driver utilizes pulse width modulation (PWM) to generate this signal. To implement such a procedure for two or three dimensional triangulation, a minimum of three to four custom drivers would need to be incorporated and each of them must be in the cameras field of vision. An algorithm would also need to be designed to decompose the various signals before triangulation. Extrapolated to a warehouse environment, this approach is unfeasible. A less invasive, and more cost effective approach would be necessary.

A similar procedure was designed in [3] for active positioning of a robot. An LED-ID approach was implemented, wherein the on-off state of each bulb was manipulated in such a way to to create an identifiable ID. A CMOS (complementary metal oxide semiconductor) image sensor was used for photo detection. The procedure exploits the rolling shutter effect of the sensor to identify the unique on-off frequencies of the LEDs. The position of the robot was then triangulated using the understood absolute locations of the LEDs and their phase offsets. This architecture has accuracy under 3 centimeters as well. Once again, the light emitted by the LEDs was modified to allow for unique identification and triangulation. For the reasons mentioned before, this procedure is expensive and invasive when scaled to a larger environment.

A different approach was taken in [4]. In this work, the characteristic frequency (CF) of each LED is identified and used to triangulate the location of the sensor. Each LED first undergoes a process to identify its unique CF. Then, these character frequencies are used to isolate the individual LED contributions from the composite light intensity reading and triangulate the two-dimensional location of the sensor. This approach resulted in an accuracy between 10-30 centimeters. The advantage to such an architecture is that no system modifications are necessary. Instead, each bulb in use must be analyzed prior to use for triangulation. In practice, this may prove inefficient. Consider a warehouse setting once again; sufficiently many pre-analyzed LEDs would need to be detectable from every location for accurate results. Thus the environment must be engineered to account for this. Moreover, the larger the warehouse and the more inventory present, the more pre-analyzed bulbs would be necessary overall. A practical system would require initial analysis for each bulb of first use and for every subsequent replacement bulb in continuous use. Though this approach may not be directly invasive, the warehouse must be designed with foresight and the LEDs must undergo extensive prior analysis, which is expensive, both in time and money.

1.2 Proposal

Our proposal is two-fold. We propose a proof of concept for the uniquely identifiable nature of individual LED light emission, as well as a VLP positioning architecture using unmodified LEDs without the need for triangulation or component analysis. In this work, we first analyze a simple multi-bulb system to provide the basis for our proof, highlighting the distinct way in which two LEDs emit light. Next, our designed environment and sampling procedure are explained. Two machine learning models were then employed for localization: a classification model trained on discrete locations and a multiple output regression model trained on continuous x and y coordinates. The results of our classification model conclude the proof for uniquely identifiable LED light emission, while the results of our regression model enable our proposal for a new approach to VLP. We received a consistent classification accuracy above **90%** and a worst-case **sub-16cm** regression loss.

The novelty of our procedure is that any LED can be deployed directly from its packaging into any multi-bulb system. No prior analysis or modifications to anything in the system is necessary for positioning, making our approach both cost and time efficient. Furthermore, it does not require component decomposition from the observed light intensity. Rather, the system is calibrated to recognize the unique light levels at various locations, which it uses to derive an understanding of the environment as a whole. More specifically, we exploit the excitation of the LED by the rectified power supply, which is distinct at each location. Note: to the best of our knowledge, this work is the first time that this physical effect has been used in VLP. We use the regression model to bypass the need for triangulation, making it robust to positionally-dependent occlusion. Thus, our approach can adapt to any LED-lit environment without the need for customization or triangulation, making it increasingly cost effective, time efficient and robust to occlusion.

Chapter 2

Methods

We hypothesize that two, unmodified, LEDs which were manufactured from the same litter, emit light with sufficient variation to be uniquely identifiable, thus allowing for localization.

2.1 Mathematical Basis

Consider two LED light bulbs: bulb A and bulb B. Let both bulbs be manufactured in the same process and neither be externally modified. We hypothesize that the inherent manufacturing variability, due to the presence of microscopic impurities that affect crystal growth, variation in environmental conditions during manufacturing, etc. [5], would lead to sufficient variation in the way they emit light levels at high frequencies which are imperceptible to the human eye but can be detected using an electronic photodetector. Allow $x_a(t)$ and $x_b(t)$ to represent the light intensities just outside of each, respectively. We have the full-wave rectified power from the outlet, the signal $x(t)$, which is defined as follows:

$$x(t) = V \left|\sin(2\pi f_p t)\right|. \tag{2.1}$$

In North America, $f_p = 60$ Hz and $V = 120$ V. Since $x(t)$ is periodic with fundamental frequency $f_0 = 2f_p$, it can be expressed in terms of a Fourier Series expansion as:

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi j k f_0 t} \tag{2.2}$$

Note that $f_0 = 1/T_0$, where T_0 is the fundamental period. The Fourier Series coefficients in this expansion are known (e.g. see Equation 3.34 of [6]) to be

$$c_k = \frac{2}{\pi(1-4k^2)} \tag{2.3}$$

Let bulb A be represented by the filter $H_a(f)$ and bulb B be represented by the filter $H_b(f)$. We model the electronics following the rectifier and the LEDs in the light-bulbs as linear-time-invariant (LTI) systems. We do this under the assumption that such a system is essentially constant, but note the possibility of a transient period before the bulb is sufficiently warm as well as long-term aging of the bulb, which would affect the transfer functions in practice. The LTI systems are described by the following block diagrams:

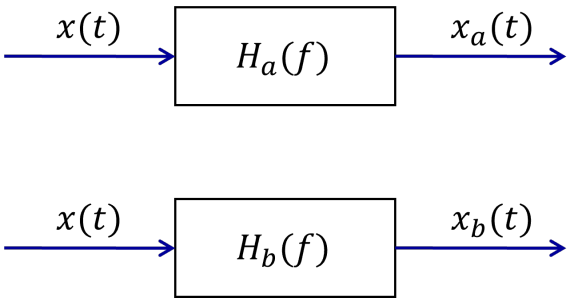

Figure 2.1 Modeling Light Bulbs as Filters

Since we can express $x(t)$ as a sum of exponentials we have

$$x_a(t) = \sum_{k=-\infty}^{\infty} H_a(kf_0)c_k e^{2\pi jkf_0 t} \tag{2.4}$$

$$x_b(t) = \sum_{k=-\infty}^{\infty} H_b(kf_0)c_k e^{2\pi jkf_0 t} \tag{2.5}$$

Furthermore, the corresponding Fourier transforms of these signals are

$$X_a(f) = \sum_{k=-\infty}^{\infty} H_a(kf_0)c_k \delta(f-kf_0) \tag{2.6}$$

$$X_b(f) = \sum_{k=-\infty}^{\infty} H_b(kf_0)c_k \delta(f-kf_0) \tag{2.7}$$

where $\delta(f-f')$ is a unit impulse at $f = f'$. Hence, $X_a(f)$ and $X_b(f)$ are a series of impulses separated by f_0 with different "heights", where the "height" of the k-th impulse of lightbub j is $H_j(kf_0)c_k$, for $j \in \{a, b\}$.

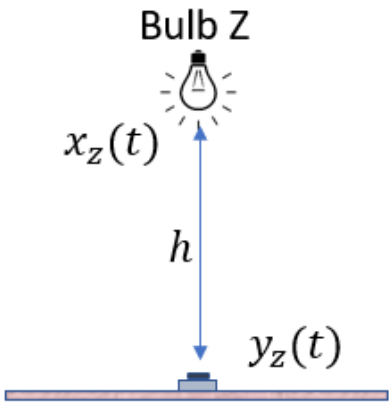

Figure 2.2 Light Collected from Bulb Z

Consider the case of a single bulb system wherein arbitrary bulb Z is placed directly above a photodetector. Let $y_z(t)$ represent the light collected from bulb Z ($x_z(t)$) and let h denote the distance between the photodetector and said bulb. This scenario is depicted in Fig. 2.2.

We can approximate $|Y_z(f)|^2$ using the Welch's Periodogram approach (`pwelch` in MATLAB).

We propose, then, that due to the manufacturing variability of the bulbs, the power spectral density estimates $E[|Y_a(f)|^2]$ and $E[|Y_b(f)|^2]$ sampled from bulb A and bulb B, respectively, would exhibit uniquely identifiable patterns and, similarly, $x_a(t)$ and $x_b(t)$ would be distinct. Here, $E[\cdot]$ denotes the expectation operator. We suspect this variation can be isolated through the series of unit impulses within the respective power spectral densities E$[|Y(f)|^2]$'s.

We can extrapolate this instance to a multi-bulb environment. Consider the 2-dimensional system in Fig. 2.3. In this instance, the photodetector receives input light intensity from both $x_a(t)$ and $x_b(t)$. The received light intensity at said photodetector is the sum of the two light intensity functions after a gain dependent on distance and angle as follows:

$$y(t) = G(r_a, \alpha)x_a(t) + G(r_b, \beta)x_b(t) + n(t) \tag{2.8}$$

where $G(r, \theta)$ is the gain between a lightbulb and a photodetector at distance r and angle θ from the vertical. The noise at the photodetector is denoted by $n(t)$.

Since $x_a(t)$ and $x_b(t)$ are distinct, each $y(t)$ at any 2-dimensional distance d from bulb

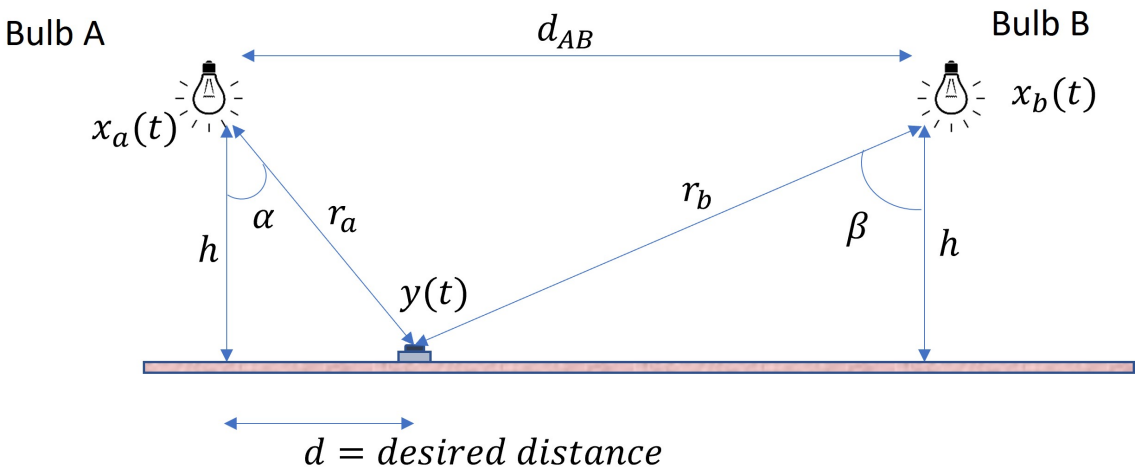

Figure 2.3 2-D VLP Setup for 1D positioning

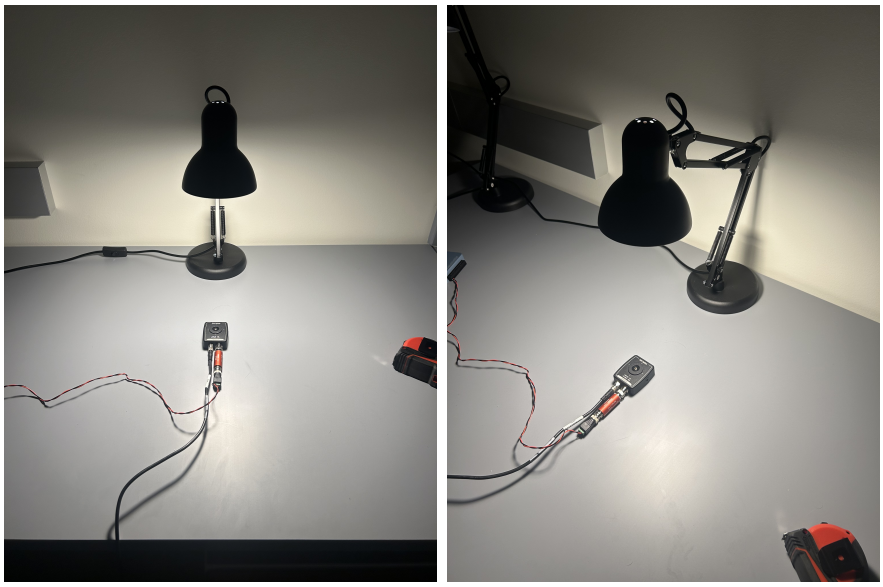

Figure 2.4 Sampling Under Single LED

A will also be uniquely identifiable. That is, each $y_d(t)$ will be distinct for all d from bulb A due to the unique combination of $G(r_a, \alpha)$, $x_a(t)$, $G(r_b, \beta)$, $x_b(t)$ at that location. Therefore, position can be obtained through sampled $y(t)$ and known $y_d(t)$ for locations d.

2.2 Experimental Basis

An experiment was designed to test our hypothesis. Two Sylvania 75W LED bulbs were selected from the same packaging. Denote them Bulb-A and Bulb-B.

First, light was sampled from directly underneath each bulb individually. No other

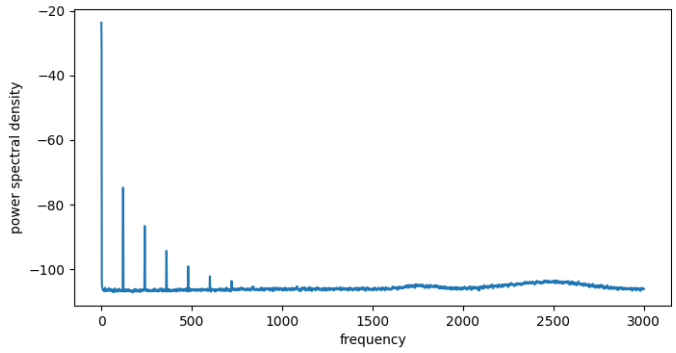

Figure 2.5 Plot of $E[|Y_1(f)|^2]$ in units of dBm (decibles relative to a mW).

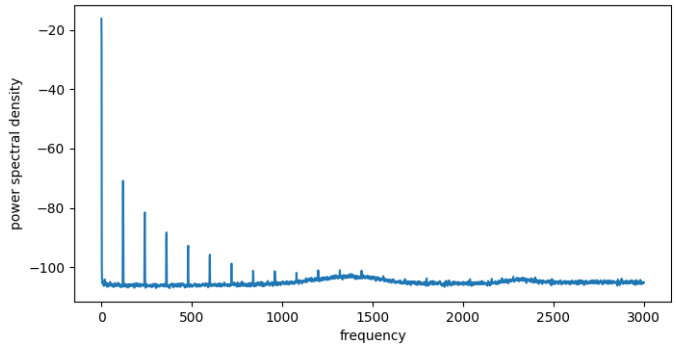

Figure 2.6 Plot of $E[|Y_2(f)|^2]$ in units of dBm (decibles relative to a mW)

light in the room was present at the time of sampling. Thorlab's PDA8A2-Si Fixed Gain Detector and Measurement Computing's USB-231 Data Acquisition Module (Analog-to-Digital Converter) were used for the data acquisition. Unmodified voltage from a wall power supply, as described in section 2.1, was used to power the bulb. The bulb was positioned 35cm above the desk. A photo of the system can be seen in Fig. 2.4. Note: each bulb was sufficiently heated before sampling took place, by leaving the bulb on for 10 minutes before collecting data.

Once again, first consider the single bulb system consisting of a single bulb placed directly above the photodetector with no additional environmental light. Let $y_1(t)$ denote the light sampled from under Bulb-A and $y_2(t)$ denote the light sampled from under Bulb-B. Matlab's `pwelch` function was performed on both $y_1(t)$ and $y_2(t)$ to yield the power spectral densities $E[|Y_1(f)|^2]$ and $E[|Y_2(f)|^2]$. Plots depicting these densities are displayed in Fig. 2.5 and Fig. 2.6.

Now, consider the two dimensional system depicted in Fig. 2.3. A physical system was designed to implement such. The two bulbs were placed 80cm apart, $d_{AB} = 80$cm,

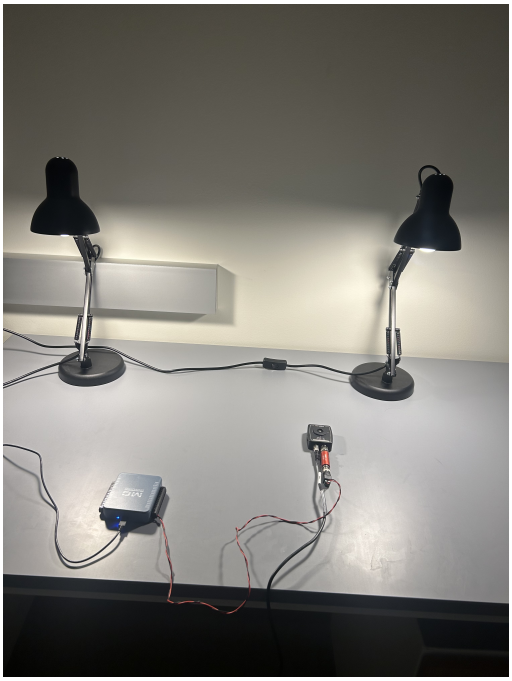

Figure 2.7 Sampling with Sensor 65cm from Bulb-A

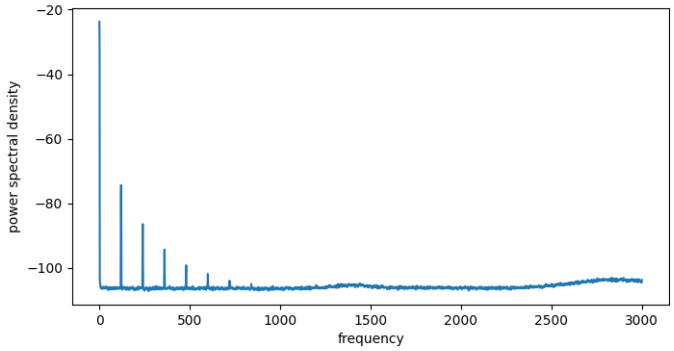

Figure 2.8 Plot of $E[|Y_1'(f)|^2]$ in units of dBm (decibles relative to a mW)

and set to a height of 35cm above the desk, $h = 35cm$. With respect to the orientation of Fig. 2.3, Bulb-A was placed in the rig on the left and Bulb-B was placed in the rig on the right. Light was sampled from two positions: 15cm horizontally from Bulb-A, and 65cm horizontally from Bulb-A (always between the two lights): i.e., d would be 15cm and 65cm, respectively. A photo of the system described in the latter can be seen in Fig. 2.7.

Let $y_1'(t)$ denote the light sampled 15cm from Bulb-A and $y_2'(t)$ denote the light sampled 15cm from Bulb-A. Again, Matlab's `pwelch` function was performed to yield $E[|Y_1'(f)|^2]$ and $E[|Y_2'(f)|^2]$. Plots depicting these densities are displayed in Fig. 2.8 and Fig. 2.9.

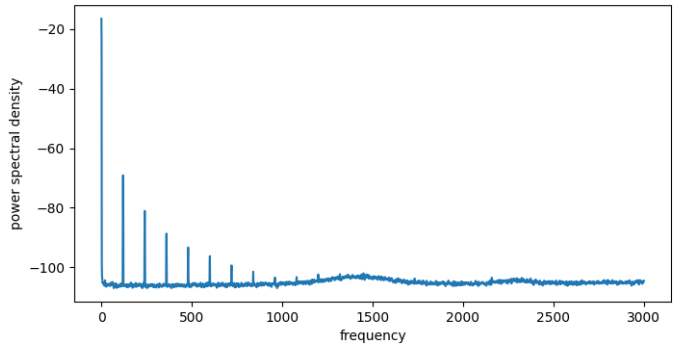

Figure 2.9 Plot of $E[|Y_2'(f)|^2]$ in units of dBm (decibles relative to a mW)

For reference, the first major peak (located at the first index) is the average power output of the sample. The repeating peaks following the initial peak are periodic with fundamental period of 120hz. That is, they are located at 120hz, 240hz, 360hz etc. Note: this is due to the full-wave rectification of the 60hz AC power input, which leads to a 120hz voltage fed to the LEDs.

Per Fig. 2.5, we can categorize $E[|Y_1(f)|^2]$ by its limited prominent peaks and small, right-oriented bumps centered around 1750hz and 2500hz. Similarly, per 2.6 we can categorize $E[|Y_1(f)|^2]$ by the continuity of prominent peaks and larger, middle-oriented bumps centered around 1300hz and 2300hz.

Let P_{y_α} denote the power spectral density output, $E[|Y_\alpha(f)|^2]$, of arbitrary y_α. We can visually deduce through Figs. 2.5, 2.8 and 2.6, 2.9, P_{y_1} and $P_{y_1'}$ display distinct similarities, whereas P_{y_2} and $P_{y_2'}$ do. Qualitatively, the harmonic peaks in P_{y_1} and $P_{y_1'}$ lose their prominence after the 6^{th} peak, whereas those in P_{y_2} and $P_{y_2'}$ continue to have visible prominence through subsequent repetitions.

For a quantitative analysis, some notation must be defined. Consider the estimated power spectral density output of the frequencies harmonic about 120hz, in addition to the output at the 0^{th} element: i.e., the peaks at 0hz, 120hz, 240hz, 360hz... etc. Denote the first x harmonic peaks of arbitrary P_α to be P_α^x. Eg., $P_{y_1}^{15}$ would be the set $E[|Y_1(f)|^2]$ for $f \in (0hz, 120hz, 240hz, ..., 1800hz)$. Denote the element-wise mean absolute difference of the two sets P_α^x and P_β^x to be $D_{\alpha,\beta}^x$. From the data sampled, we have $D_{y_1,y_1'}^{15} = 0.425$, $D_{y_2,y_2'}^{15} = 0.465$ and $D_{y_1,y_2'}^{15} = 3.77$, $D_{y_1',y_2}^{15} = 3.654$, and $D_{y_1,y_2}^{15} = 3.75$. We can then quantitatively deduce the similarities between P_{y_1} and $P_{y_1'}$ as well as P_{y_2} and $P_{y_2'}$. Moreover, these calculations provide evidence for the overall dissimilarity between P_{y_1} and P_{y_2}.

This experiment provides us with a basis for the claim that two similarly manufactured LED's emit uniquely identifiable light and could be used for multi-dimensional positioning.

Chapter 3

Experiment

To simulate a warehouse environment, a PVC pipe frame was constructed in lieu of a warehouse ceiling. The frame contains nine LEDs in a 3x3 grid. Twenty-one locations beneath the frame were selected for sampling. Five instances of 10 minute samples were taken at each location and divided and preprocessed for feature extraction. Features were fed into both a classification algorithm, targeting the discrete locations, and a multiple output regression model, targeting the location's corresponding x and y coordinates. This chapter will discuss the creation of the frame, location selection and sampling process, preprocessing of the data, feature selection and model architectures.

3.1 LED Frame Construction

A frame was designed and constructed to emulate the ceiling of a warehouse. The bulbs were established to have 1 meter separation from the floor (Note: the shallow height was selected to account for use of a low-sensitivity photodetector. A higher-sensitivity photodetector would be necessary in an actual warehouse setting).

The materials for construction are as follows: 30 meters of 1-1/4in. PVC Pipes; 13x 1-1/4in Furniture Grade PVC 5-Way Cross; 12x 1-1/4in PVC Schedule 40 S x S x S Tee; 8x 1-1/4in Furniture Grade PVC 3-Way Elbow; 1x 1-1/4in Ratcheting PVC Cutter; 9x Sylvania 75W A19 LED Bulbs; 9x A19 non-polarized sockets; 9x non-polarized extension cords; Electrical Tape.

Before construction, the 1-1/4in PVC Pipes were cut into 40x 45.5 centimeter pieces and 8x 112 centimeter pieces. The 45.5cm cuts were assembled into a 5x5 grid and the

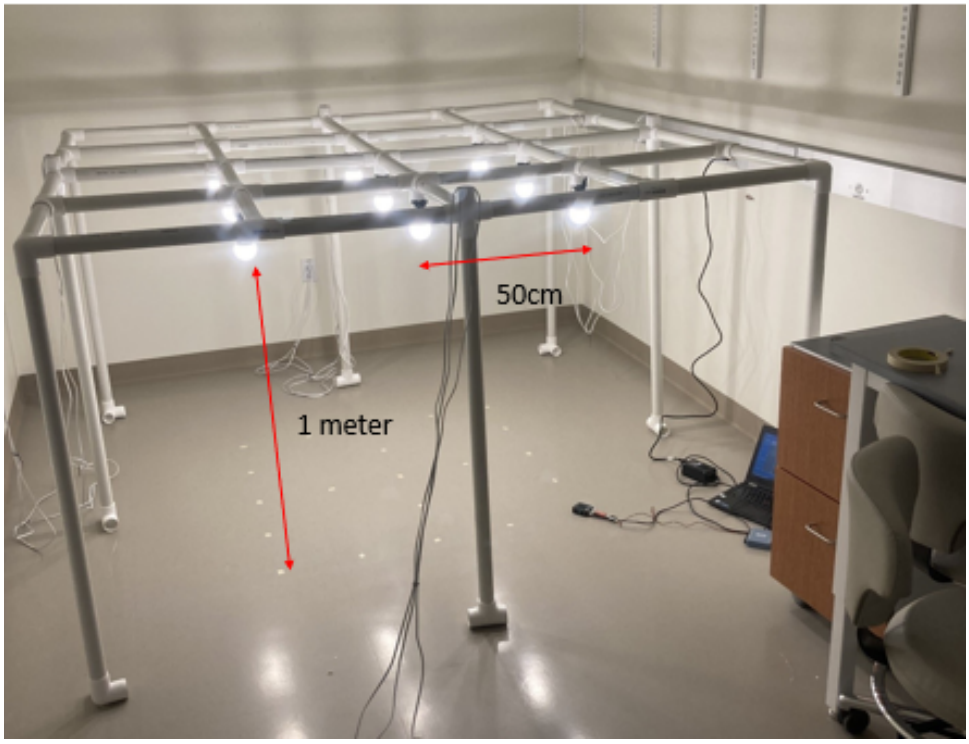

Figure 3.1 Light Frame

112cm cuts were used as the legs. 9x 1-1/4in Furniture Grade PVC 5-Way Cross were placed in the 9 central points of the grid for the LEDs to be suspended from. The other 4x 1/4in Furniture Grade PVC 5-Way Cross were placed on the perimeter, with one opening unfilled to allow for internal wiring. The bulbs and sockets were fastened with electrical tape. A distance of 1 meter was maintained from the ground to the tip of each bulb and 50cm was maintained between the tips of adjacent bulbs. An image of the frame can be seen in Fig. 3.1.

3.2 Location Selection and Sampling

Twenty-one locations were selected for sampling: the locations directly below each light as well as the intermediary points between them. 25 centimeters separate adjacent locations. A photo of the tape marking these points can be seen in Fig. 3.1. The points were subsequently denoted location1, location2, ..., location21 and assigned corresponding x and y coordinates as described in Fig. 3.2. For reference, Fig. 3.2 is oriented with respect to the perspective in Fig. 3.1.

Sampling was done on four occasions. Attempts were made to assure consistency over different sessions. The same room was used for each and the frame position and sample

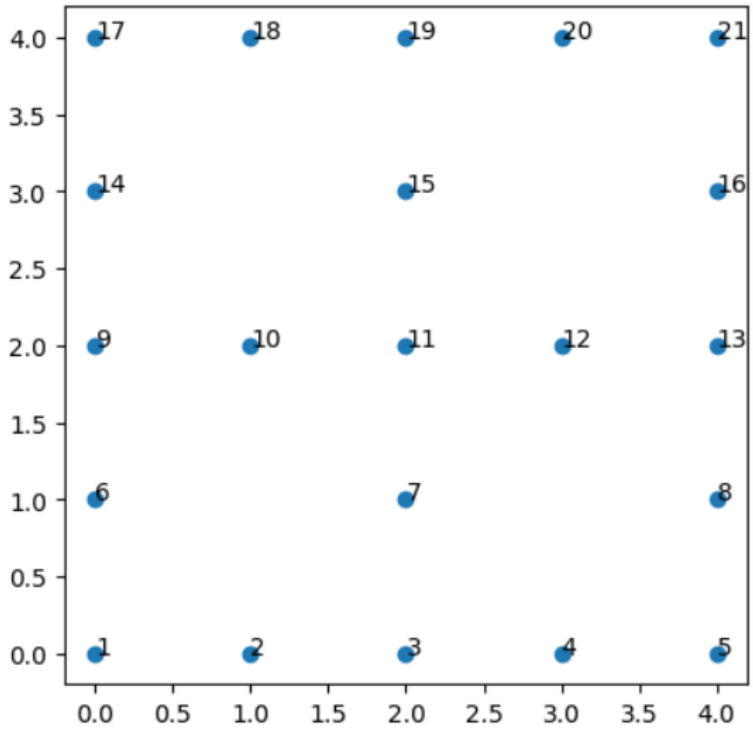

Figure 3.2 Location Coordinates and Position

locations were marked for identical orientation. The room was isolated from external light and held to a temperature within a degree of 67°F. On three consecutive days, 10 minute samples were taken at each location. A combined 630 minutes of sampling was gathered from these days. Two months later, two installments of 10 minute samples were taken at each location on a single day. A combined 420 minutes of sampling was gathered on this day.

Data was collected using Thorlab's PDA8A2-Si Fixed Gain Detector and Measurement Computing's USB-231 Data Acquisition Module (Analog-to-Digital Converter). The sampling process was as follows: the photodetector was placed on the floor, with the sensor positioned at the intended location. The sample rate was set to 6000samples/sec and 3.7 million samples were taken before the data was recorded and the photodetector was moved to the next location.

Once collected, each 10 minute sample was divided into 1 minute samples and exported for preprocessing.

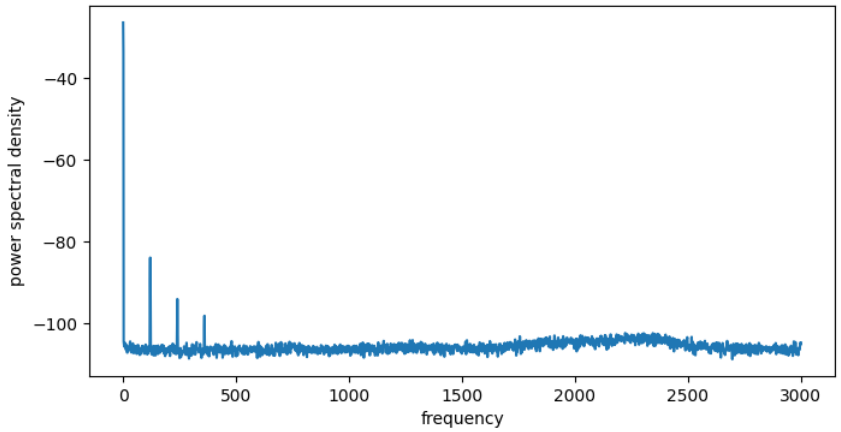

Figure 3.3 $|Y(f)|^2$ Location 1, Round 1

3.3 Preprocessing and Feature Extraction

The photodetector samples light intensity over time. Before being analyzed, the data was converted to the frequency domain. This was done using Matlab's `pwelch` algorithm, which uses Welch's Periodogram approach.

Welch's Periodogram approach performs as outlined in this paragraph. First, The time-domain data is divided into sets of size N data points with M overlapping points. E.g., say a sufficiently large dataset, D, was divided into sets of size 100 with 50 overlapping points. The sets would then be D[0,100], D[50, 150], D[100, 200]... etc.. A Hamming window and Fourier Transform are then applied to each set. Finally, The outputs are squared and averaged to yield the final result.

The welch algorithm was performed on each 1 minute sample with a window size of 5000 data points and overlap of 2500 data points. An example power spectral density estimate can be found in Fig. 3.3. Note: this output is noisier than those in Chapter 2.2 due to the smaller sample size.

Next, features were extracted. The extraction procedure was simple; from each power spectral density estimate P_α of arbitrary sample $y_\alpha(t)$, the set of the outputs from the first 10 frequencies harmonic about 120hz in addition to the output at the 0^{th} index, denoted P_α^{10} (see Chapter 2.2 for notation) were extracted. These comprise the entirety of the features selected. The basis for this selection is outlined in Chapter 2.

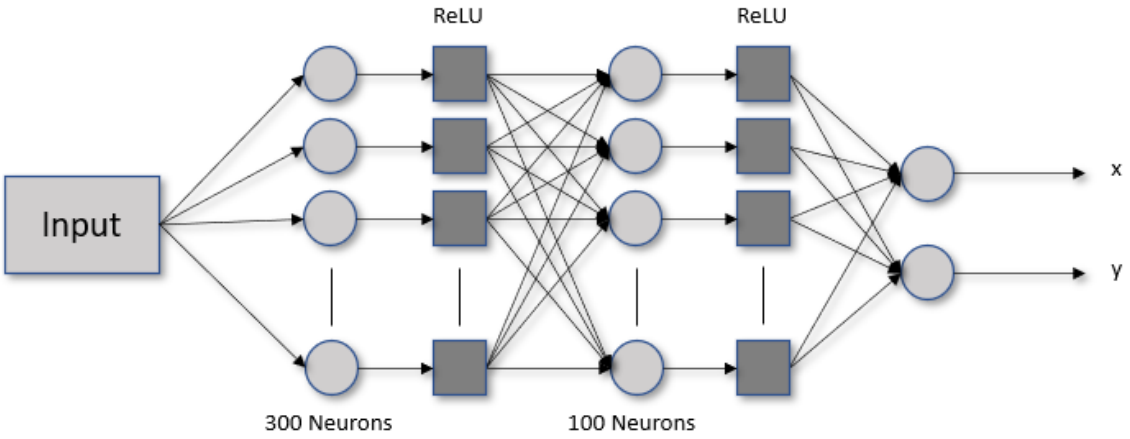

Figure 3.4 Multiple Output Regression Model

3.4 Models

To provide context, the purpose of employing a machine learning approach rather than a traditional signal processing approach is two-fold: (1.) to bypass the need for component decomposition and triangulation and (2.) to increase robustness against variable environmental factors such as positional occlusion from one or more bulbs.

Before advancing to the intended regression model for precise x and y positioning, a classification model was implemented to analyze the efficacy of the features extracted.

The classification model selected was Sklearn's DecisionTreeClassifier [7]. Two protocols were performed separately to train the model: train-test split and k-fold cross validation. For the train-test split protocol, an 80-20 split was selected. The model was trained on the train data and then tested on the test data. Accuracy was recorded and a confusion matrix was created for analysis. K-fold cross-validation was implemented with 10 folds. The resultant validation scores were recorded individually and averaged for a composite accuracy. A standard deviation was also taken for the individual scores.

A multiple output regression model was implemented in Pytorch [8] to allow for intercorrelated training on the continuous x and y output. The architecture is of a simple feed-forward neural network, which consists of 2 fully-connected hidden layers and Rectified Linear Unit (ReLU) activation functions. A diagram of the network can be found in 3.4. Note: the depth and breadth of the model were selected in such a way to be arbitrarily small, which was done to emphasize that a complex model need not be implemented to achieve accurate results.

For training, a k-fold stratified cross-validation protocol was designed and imple-

mented. Due to the limited number of data points, a custom randomization algorithm was designed to create folds that were "well-mixed." This was done to avoid the possibility of all data points for one location being isolated to a single fold, preventing that location from being trained on in the case of that fold being selected for validation (or otherwise validated upon). The algorithm works as follows:

1. All the data for each of the locations is isolated, denoted $data_1, data_2, ..., data_{21}$
2. Each of $data_1, data_2, ..., data_{21}$ is shuffled
3. $\frac{total\_data\_points}{num\_folds \cdot num\_locations}$ data points are selected randomly from each of $data_1, data_2, ..., data_{21}$ to create a single fold
4. the data points selected in (3.) are removed from $data_1, data_2, ..., data_{21}$
5. steps (3.) and (4.) are repeated until $data_1, data_2, ..., data_{21}$ are all empty
6. each of the folds are concatenated and output

The data was randomized in this manner, then the cross validation was performed with 10 folds for 150 epochs. The loss function selected was Mean Absolute Error (MAE) and the optimizer selected was Adam with a learning rate of 0.001. Validation Scores were recorded for analysis.

Chapter 4

Results

Two models were used for positioning: a classification model was imported for discrete location positioning and a multiple output regression model was designed for continuous location positioning. Three collections of data were fed into each model for testing: one containing the 630 minutes of samples collected initially, one containing the 420 minutes of samples collected afterwards and one containing all 1050 minutes of samples. The results are analyzed in this chapter.

For clarity, allow "Collection 1" to denote the original data taken (630 minutes), "Collection 2" to denote the data taken at a later date (420 minutes) and "Composite Data" to denote the full collection containing both Collection 1 and Collection 2 (1050 minutes).

4.1 Classification

As stated in the previous chapter, Sklearn's DecisionTreeClassifier was the classification model in use. The base model was used. Two protocols were implemented to train and test the model: a train/test split and k-fold cross-validation.

4.1.1 Train/Test Split

An 80-20 train-test split was performed on each of the three collections of data using Sklearn's train_test_split algorithm. Note: This algorithm assures that data points from all classes are well mixed in the train and test sets.

Model Accuracies	Collection 1	Collection 2	Composite Data
Classification (train/test split)	0.921	0.917	0.9

Table 4.1: DecisionTreeClassifier Accuracies using Train/Test Split

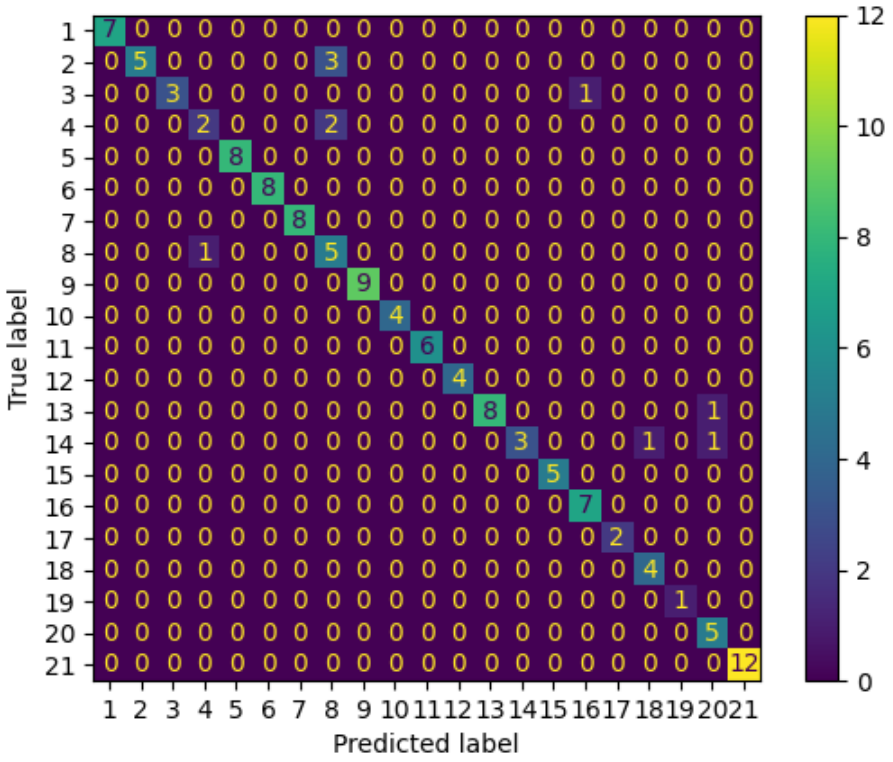

Figure 4.1 Confusion Matrix for Predictions from Collection 1

The model was trained using the training data for each collection, with discrete locations $location_1, location_2, ..., location_{21}$ as labels and tested on the same collection's test data. After being trained, the model predictions were compared to the true labels for accuracy. The proportion of correct predictions can be found in Table 4.1. Note: each location is equally represented in the data, so this overall accuracy estimate is justifiably used.

Confusion matrices were created using the predictions of each model from their respective test sets. The confusion matrix for Collection 1 can be found in Figure 4.1.

Before sampling took place, one concern was that geometrically similar locations in the grid would be indistinguishable. Should that be the case, there would be little basis for the claim that each LED emits uniquely identifiable light.

To explain in more detail, take locations 5 and 17 for example; the location in the

bottom right corner and top left corner, respectively (reference Figure 3.2). These two locations have the same orientation under the grid of LEDs, simply with rotation. In general, each bulbs is positioned directly above locations 1, 3, 5, 9, 11, 13, 17, 19, and 21. Denote these bulbs, Bulb 1, Bulb 3,..., Bulb21. Bulb 17, 19 and 9 look the same to location 17 as Bulb 5, 3, and 13 do to location 5. If each LED emitted identical light, these two locations would receive identical light intensity, making their positions indistinguishable. The same phenomenon would occur for locations 3, 9, 13 and 19, locations 7, 10, 12 and 15 and so on.

As can be seen in the Confusion Matrix, geometrically similar locations are the primary source of confusion. Consider Figure 4.1 for reference. The model predicted location 8 for true value locations 2 and 4. These three are geometrically alike locations. Furthermore, the model also predicted true value location 14 to be locations 18 and 20, which are also geometrically alike locations. Though they are not all included, a similar pattern continues throughout the three matrices.

Despite this fact, however, the models performed with accuracies at or above 90%. More specifically, the model trained on Collection 1 predicted 116/121 correct, the model trained on Collection 2 predicted 77/84 correct and the model trained on Composite Data predicted 189/210 correct. Thus, throughout the three models, over 90% of the samples were not confused with geometrically identical locations. For that reason, we can confidently claim that the light intensity received at each geometrically alike location was not indistinguishable, and therefore, were uniquely identifiable.

4.1.2 Cross-Validation

To re-enforce the validity of the accuracies presented in 4.1.1, k-fold cross-validation was performed on each of the three collections of data.

Once again, Sklearn's DecisionTreeClassifier was used as the model. Sklearn's cross_val_score was used as the algorithm and 10 folds was selected. The average fold accuracies, as well as their standard deviations are depicted in Table 4.2

As can be seen in Table 4.2, the average accuracies remain around 90%. Furthermore, the corresponding standard deviations are relatively low, indicating consistency within the data.

Model Accuracies and Standard Deviation	Old Data	New Data	All Data
Average of Cross-Validation Scores	0.911	0.919	0.877
Standard Deviation of Cross-Validation Scores	0.0345	0.0340	0.0410

Table 4.2: Average Accuracies and Standard Deviations of 10-fold Cross-Validation Using DecisionTreeClassifier

Cross Validation Scores	Collection 1	Collection 2	Composite Data
Av. Total Loss	0.461	0.486	0.562
Av. x Loss	0.552	0.548	0.622
Av. y Loss	0.371	0.423	0.501
STD Total Loss	0.042	0.034	0.050
STD x Loss	0.060	0.068	0.035
STD y Loss	0.040	0.058	0.076

Table 4.3: Results from Multiple Output Regression Model Cross-Validation

4.2 Regression

As described in Section 3.4, the designed regression model is a simple feed-forward neural network consisting of two hidden layers and ReLU activation functions. The loss function used was pytorch's L1Loss function for MAE loss and the optimizer was pytorch's Adam optimizer with a 0.001 learning rate.

A 10-fold cross-validation was implemented, trained for 150 epochs per fold. This procedure was run on each collection of data, with x and y coordinates as targets (reference Figure 3.2 for location coordinates). The results can be found in Table 4.3.

4.2.1 Results Contextualized

For each location, the corresponding x and y coordinates have values within the integer range [0,4]: $x, y \in [0, 4]$. For example, the x, y coordinates of $location_1$ are (4,0), $location_2$ are (3,0) and $location_6$ are (4, 1) etc.

The MAE output from the models have these units, however, they are an average of the loss between the predicted x and true x, and the predicted y and true y. Take a prediction for arbitrary $location_\alpha$ for instance. Say the model predicted (x_p, y_p) and the

Cross Validation Scores (cm)	Collection 1	Collection 2	Composite Data
Av. Total Loss	11.529	12.138	14.040
Av. x Loss	13.791	13.696	15.548
Av. y Loss	9.268	10.580	12.531
STD Total Loss	1.046	0.853	1.261
STD x Loss	1.503	1.706	0.872
STD y Loss	0.997	1.457	1.906

Table 4.4: Results from Multiple Output Regression Model Cross-Validation in Centimeters

true location is (x_0, y_0). The MAE output would be calculated as follows:

$$MAE = \frac{|x_p - x_0| + |y_p - y_0|}{2} \tag{4.1}$$

From this output alone, we can only deduce a rough estimate of the true x and y loss; i.e., we can conclude that, on average, the loss of x or y individually does not exceed $|x_p - x_0| + |y_p - y_0|$. For this reason, individual x and y losses were recorded in addition to the trained upon MAE loss. Hence, the inclusion of these component losses in Table 4.3.

To contextualize the loss in metric units, each adjacent location is separated by a distance of 25cm. Therefore, an average loss of 0.5 would equate to an average radial loss of 12.5cm. The data in metric units can be found in Table 4.4.

For a visual reference, see Figure 4.2. The central point represents the true location, the circle represents the average MAE loss, and the square represents the composite zone of average x and y loss.

4.2.2 Discussion

For clarity, the standard deviations were included in Tables 4.3 and 4.4 to emphasize consistency among the cross-validation scores. Because each is low relative to their mean, we can conclude that the average losses are no fluke.

The goal was sub-decimeter loss. Let the average x loss, the average y loss and the average total loss be denoted L_x, L_y and L_{tot} respectively. Aside from Collection 1's L_y, none of the losses breached this threshold. A more complex regression model would likely need to be designed and trained to achieve such a goal. As can be seen in Figure

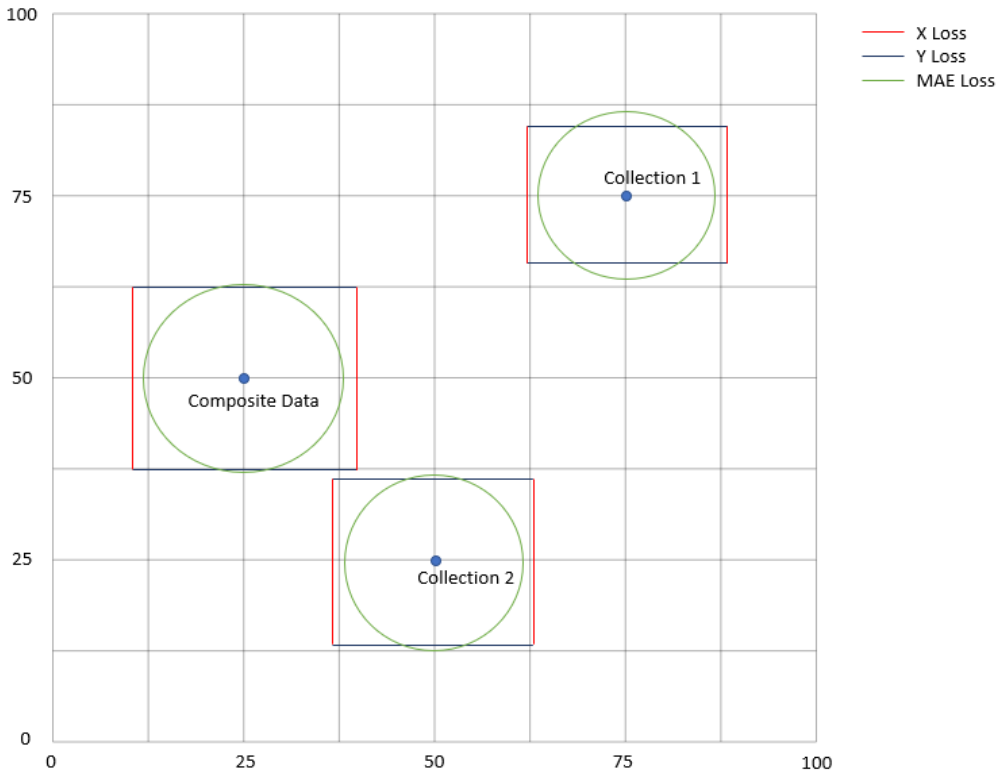

Figure 4.2 Visual of Average Loss in Centimeters

4.2, each of the losses is constricted moderately tightly about the true location, despite the aforementioned weakness. Moreover, upon further inspection, it can be noticed that the zones of loss tend to the true location's unique space. We define a locations "unique space" to be the space containing all points closest to that point. For example, the unique space of the point $p_0 = (75, 75)$ would be the box $x, y \in (62.5, 87.5)$ (reference Figure 4.2). Just minor bleeding from the true location's unique space is exhibited in each of the prediction zones. This leads us to believe that a more diverse set of locations must be sampled and trained upon, in addition to a more complex model, to achieve greater accuracy. Regardless, the current approach must be improved for adequate localization in practice.

Notice, also, that L_y is lower than the L_x in each case. Three potential causes are the model training architecture, greater light emission similarity between bulbs along similar y, and asymmetry in the physically constructed system. Further research would need to be conducted to reach a definitive conclusion.

Furthermore, note that in the case of both the classification model and the regression model, the accuracy was greater for Collection 1 and Collection 2 trained and tested on

individually rather than together. In the case of the cross-validation classification model, the accuracies for Collection 1 and Collection 2 were 0.911 and 0.919 compared to an accuracy of 0.877 for Composite Data. Similarly, the L_{tot} losses for Collection 1, Collection 2 and Composite Data in the regression model were **11.529cm, 12.138cm and 14.040 cm**, respectively. An identical environment was used in sampling for Collection 1 and Collection 2. The only factor of separation was time. Because the model trained on the composite data performed worse than the ones trained on the two components individually, we can conclude that there is variation between the collections. We suspect then, that there is some variation in the way these LEDs emit light as they age. Further research would have to be conducted to reach any definitive conclusions, however.

Chapter 5

Conclusion

This chapter will discuss both the advantages and disadvantages of using the proposed system architecture.

There are two primary disadvantages to this approach to positioning; (1.) extensive calibration must be undergone before practical localization can begin and (2.) such calibrations would be ineffective in a dynamic environment. The regression model must be trained on samples taken within the environment to interpret position. At a fundamental level, a more diverse data set of training sample locations would allow for more accurate positioning. Moreover, however, should the environment be inconsistently laid out, greater diversity of training sample locations would be necessary. For example, in a practical environment, there are certain locations that are especially occluded, such as those nearest storage palettes, containers etc.. The calibration process, therefore, must include sufficiently enough samples in and around such locations to account for the discontinuous nature of the surrounding light levels. As a logical extrapolation, should the environment be dynamic — that is, the environment is constantly changing — static calibrations would be ineffective. Note: the aging or exchanging of bulbs would likely cause a necessity for re-calibration as well. A system would need to be designed to either allow for constant re-calibration or could interpret the effects of a dynamic environment. Further research would need to be conducted to achieve this.

Five advantages to using the the proposed architecture are as follows: (1.) there is no need for component modulation (2.) or environmental design/infrastructure (3.) or component decomposition and triangulation, (4.) the system is robust to general occlusion from one or more LEDs and (5.) the model performs relatively accurately

despite simplicity. Our approach allows the system to interpret its environment through positional light intensity. As evidenced in this work, inherent variation in the way alike bulbs emit light allows for robustness to geometrically identical locations. In theory then, the system could be used with any stock LED bulb, in any orientation, with any light fixtures, cords, drivers, power supply etc.. Thus, there is no need for component modulation or environmental engineering, making this a cost-effective, minimally invasive solution. Furthermore, due to the machine learning approach taken, light levels from each bulb need not be seen by the sensor for accurate positioning. The model trains on the light levels at each location as a whole, therefore occlusion from one or more bulbs should not effect the accuracy. Finally, a minimalistic neural network design was chosen to emphasize the computational efficiency of calibration. This was done with foresight to allow for potential continuous calibration in the future.

Chapter 6

Future Work

6.1 Improved Sampling Methodology

The primary restriction of this work was the sampling method. Each of the 1050 minutes of sampling was done by hand. To reiterate, 10 minute samples were taking at each of the 21 locations 5 times: 3 installments were done in three consecutive days and 2 were done in a single day. This equates to 17.5 hours of live sampling time alone. A minimum of 2 minutes was needed in between each 10 minute sample for movement of the photodetector and recording of the data. Moreover, in an effort to maintain absolute darkness within the environment around the testing LEDs, the sampler had to remain within the testing room. Therefore, a minimum of 21 total hours of active sampling was necessary for the collection of this limited data. Note: an attempt was made to automate the collection using a TurtleBotWafflePi, however, due to an incompatibility issue between the required software and included hardware, it was unsuccessful. As a result of the time-expense, this procedure limited both the total samples taken as well as the number of sample locations selected.

To advance this work further, an automated sampling procedure must be implemented. We suggest a LiDAR based approach. More specifically, we suggest a robot with LiDAR based self-positioning capabilities outfitting with a photodetector and an absolute clock. The LiDAR based positioning approach performs well in various environments [9] and an absolute clock would allow us to connect the positional data with the visible light data. Should an automated sampling method be materialized, we would be able to increase the total number of samples as well as the locations sampled from. The

result would increase general robustness of the model. We believe this to be a critical step before our positioning approach can be used in practice. Not only would model accuracy improve, but for it to be robust to spatial inconsistencies, such as partial occlusion, a greater diversity of sample locations must be incorporated in the train data.

6.2 Study on Factors for Unique Light Emission

This work gave us the answer to the question, "do same-manufacturing LED's emit uniquely identifiable light?," but it did not tell us why. We believe a study intended to understand why these LED's emit identifiable light would be necessary to create a practical, real-world system for positioning in this manner. The study would provide us with a basis for what specific variation is being exploited. Thus, allowing for positioning architecture to be designed with intent. For obvious reasons, without such, an optimal system would be unfeasible.

Furthermore, once the unique factor is isolated, a further study could be conducted on the change in light emission over time. As stated in Section 4.2.2, our models trained on data collected in close temporal proximity performed better than the one trained indiscriminately of temporal relation, despite all other factors being consistent. This leads us to conclude that the identifiable way in which each LED produces light alters as they age. This phenomenon could be accounted for, should the unique factor be understood.

References

(1) Roach, J., *The quest to deploy autonomous robots within Amazon fulfillment centers*; Amazon Science: 2022.

(2) Raes, W.; De Strycker, L.; Stevens, N. In *2018 15th Workshop on Positioning, Navigation and Communications (WPNC)*, 2018, pp 1–5.

(3) Guan, W.; Chen, S.; Wen, S.; Tan, Z.; Song, H.; Hou, W., *High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning*; 2; IEEE: 2020; Vol. 12, pp 1–16.

(4) Bastiaens, S.; Deprez, K.; Martens, L.; Joseph, W.; Plets, D., *A comprehensive study on light signals of opportunity for subdecimetre unmodulated visible light positioning*; 19; MDPI: 2020; Vol. 20, p 5596.

(5) Bornoff, R.; Mérelle, T.; Sari, J.; Di Bucchianico, A.; Farkas, G. In *2018 24rd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, 2018, pp 1–6.

(6) McClellan, J. H.; Shafer, R. W.; Yoder, M. A., *DSP First*, 2nd; Pearson: 2016.

(7) Pedregosa, F. et al., *Scikit-learn: Machine Learning in Python*, 2011; Vol. 12, pp 2825–2830.

(8) Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A., *Automatic differentiation in PyTorch*, 2017.

(9) Holmberg, M.; Karlsson, O.; Tulldahl, M., *Lidar Positioning for Indoor Precision Navigation*, 2022, pp 359–368.

(10) Ghassemlooy, Z.; Alves, L. N.; Zvanovec, S.; Khalighi, M.-A., *Visible light communications: theory and applications*; CRC press: 2017.

(11) Bowles, R., *Organized Chaos: Behind The Scenes of Amazon's Inventory Management System*, 2021; Vol. 3.

(12) Welch, P., *The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms*; 2; IEEE: 1967; Vol. 15, pp 70–73.