Topological Object Localization in a Visible Light Communications Application

---

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment
of the requirements for the degree

Master of Science

by

Emmanuel Oluwadurotimi Denloye-ito

December 2018

APPROVAL SHEET

This Thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science

Author Signature: _____

This Thesis has been read and approved by the examining committee:

Advisor: Dr. Maïté Brandt-Pearce

Committee Member: Dr. Scott Acton

Committee Member: Dr. Stephen Wilson

Committee Member: _____

Committee Member: _____

Committee Member: _____

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

December 2018

UNIVERSITY OF VIRGINIA

Topological Object Localization in a Visible Light Communications Application

by

Emmanuel Oluwadurotimi Denloye-Ito

A thesis submitted in partial fulfillment for the degree of Masters of Science

in the

Charles E. Brown Department of Electrical Engineering

August 2, 2018

UNIVERSITY OF VIRGINIA

Abstract

Charles E. Brown Department of Electrical Engineering

Masters of Science

by Emmanuel Oluwadurotimi Denloye-Ito

In this thesis, we develop a technique for object localization in an up-link optical wireless system for, most generally, indoor applications. In contrast to other works (e.g. [1]), wherein the authors exploit measures such as the line of sight peak power (LOS), the second power peak of the channel's impulse response, and the delay between these aforementioned measures to localize an actively transmitting object, we instead localize a *passively* sensed object via a hitherto, to the best of our knowledge, novel methodology. We employ a device from a relatively new research sub-field of computational mathematics called topological data analysis (TDA). In particular, we employ a methodological framework called persistent homology(PH), which in short, is useful for identifying topological structure, by identifying "holes" within a data-set at several scales of analysis. In particular, we utilize the "holes" identified by PH to provide a count of the number objects in an optical wireless system. In general, PH can provide an overall assessment of the global structure and connectivity of a data-set. PH is also useful because, in addition to the simple aforementioned qualitative guide, it can be utilized, given certain permitting conditions, to localize and potentially provide a complete geometric description of objects in an indoor application while obviating the need for potentially expensive explicit shape models of various objects within a room. In summary, this thesis introduces a methodology for object localization, enumeration and description in an up-link wireless optical system for indoor applications with the aid of PH. We demonstrate the viability of our approach through several standard simulations and assess the implications of their outcomes for determining the character of a path for future work.

Acknowledgements

I would like to thank the groups and organizations that have supported me so graciously throughout my time as a University of Virginia graduate student. I would most especially like thank my advisor Dr. Maïté Brandt-Pearce for her patience, stunning brilliance and eagerness to help me overcome trying academic and non-academic issues. I would also like to give a special thanks to Dr. Scott Acton who is another professor in the Electrical Engineering department. His digital signal and image processing classes were a treat, and I enjoyed joining his weekly group meetings. I also thank him for all the academic support he offered. Of course, organizations such as the Commonwealth Center for Manufacturing Engineering and The Jefferson Trust Fund deserve a very special mention for the educational and financial support they bestowed to me. Additionally, I am grateful for my masters' committee, who were willing to offer their time and sage minds towards the development and final production of my masters' thesis. Finally, I would like to thanks my parents, Dr. Adeyemi Abiodun Adekoya and Dr. Sulola Ojuolape Adekoya, for their unwavering guidance and patience. Their support was vital for helping me complete my thesis.

Of course, I should also thank the many students and faculty members that I crossed paths with at the University of Virginia. Though there may be too many to list in this section, I have certainly not forgot who you are.

Contents

List of Figures

List of Tables

List of Abbreviations

GPS Global Positioning System

IoT internet-of-things

IR Infrared

LOS Line of Sight

PD Photo-Detector

PH Persistent Homology

TDA Topological Data Analysis

VLC Visible Light Communications

I dedicate this thesis to my loving parents, Dr. Adeyemi Abiodun Adekoya and Dr. Sulola Ojoulape Adekoya. Additionally, I dedicate this document to Mr. Letshaga "Levi" Matlala, a close family friend, who has recently left behind his loving wife, Audrey, and three wonderful children, Tsholo, Neo and Kabelo.

Chapter 1

Introduction

1.0.1 Opening Remarks

In this document, we present an up-link optical wireless *passive* object/obstacle localization system for indoor applications. There is a high demand nowadays for localization services in many applications like robotics, unmanned aerial vehicles, Internet of Things (IoT) applications and self-driving cars. In this thesis, our analysis focuses upon *passive* object localization. When we use the term *passive* as in aforementioned case, we are referring to situations wherein a robot, for example, is attempting to create an "image" or map of an environment, in a similar manner that may be utilized in sonar, from a set of measurements. While a GPS service may have robust localization capabilities in outdoor applications, its oft poor coverage in indoor environments, as well as its poor accuracy indoors, provides a strong impetus towards crafting a robust indoor localization system by innovatively using GPS services or capabilities or utilizing another technology and service. More importantly, at least for our purposes, GPS would not be the best choice for passive object localization in indoor environments. In this thesis, we address this problem by introducing a new optical wireless-based technique that uses particular features of an infrared IR up-link channel impulse response for

object localization. At a certain level of analysis, we are simply "photographing" a room filled with static, immobile objects. A map, produced by the collection of "photographs", could be employed in a variety of scenarios where human intervention would either be costly or impossible. For example, consider manufacturing companies that might want to automate a process that would build a map of their inventory stored in a warehouse, whose objects might be geometrically intricate, and spread across a large and structurally complex room.

1.0.2 Motivation: Visible Light Communications

As already mentioned in the opening paragraph, indoor positioning system, irrespective of the supporting technology, is of great interest in situations whereby passive sensing of the environment is the most achievable means of localizing and characterizing objects in an indoor environment. Though GPS works very well for situations such as map services for mobile phones, navigation for cars/ships and planes, GPS performance can be severely degraded in indoor environments and more importantly GPS requires active objects in order to provide and perform localization service(s).

There have been a number of approaches for active localization ([4], [5], [6], [7], [8], [9]) based on wireless technologies such as, but not limited to: WLAN, RFID, cellular, UWB, Bluetooth. Whilst these methods can deliver positioning accuracies from tens of centimeters to several meters, they still suffer from the interference issues that can arise when signals within the radio-wave portion of the electromagnetic spectrum are utilized for communication-based purposes. For example, radio-wave signals can interfere or disable certain kinds of medical devices and also congest the limited bandwidth available to wireless devices which are constrained by their design to use the portion of the electromagnetic spectrum

that they must share with the other similarly operating (at least in their usage of the currently available bandwidth) devices.

However, in contrast to these techniques for active localization, we also consider the approaches that are similar, at least in their ultimate goal, to ours. (RADAR or LIDAR).

Visible Light Communications (VLC) is considered an attractive solution for indoor positioning systems because this burgeoning technology offers certain *key* advantages over its competition. In summary, VLC is, essentially, a technology for providing communication services over a medium composed of modulated optical wave-forms. In particular, these modulated optical wave-forms are sent via light emitting diodes, fluorescent fixtures or other lighting fixtures capable of propagating the desired modulated optical wave-form. *Firstly*, VLC is less costly to setup since there is a broad set of lighting sources (fluorescent lamps, light emitting diodes, etc.) that can switch fast enough to send coded messages. This cost is even more appealing when one considers that with these available light sources, such VLC systems typically expend only a little more energy than what would be consumed if the light was used for its commonplace, pedestrian purpose. This is important, as it implies that the cost and burden of a typical VLC system deployment and maintenance may be manageable to bear. *Secondly*, VLC systems typically do not introduce any of the interference effects commonly encountered with radio-wave transmissions heretofore mentioned. This makes VLC an appealing option in locations where radio-wave signals are either strictly prohibited or highly undesirable. *Finally*, for reasons most relevant to the topic at hand, a VLC system typically comes with a promise to provide better positioning accuracy (typically measured in centimeters, or millimeters) than its competing technologies due to less signal degradation from multi-path effects and interference from other wireless devices. This effect is rather pronounced when one considers the manifold sources of multi-path fading and interference that can exist in a sufficiently

confined space, such as a generic 5 meter by 5 meter by 3 meter room.

1.0.3 Motivation: Localization via Visible Light Communications

Up until to this point, we have merely mentioned a few reasons for using VLC for positioning. We would like to extend the analysis and make, at the very least, a contribution that will put localization and efforts towards object description towards the fore of the conversation. An initial attempt at determining shape and geometric configuration of objects might use an explicit model for modeling each object. It could argued that using an explicit geometric model might be useful and faster in certain cases. However, with the aid of tools from topological data analysis TDA, we have found that it is possible to conduct an analysis of shape without an explicit model for each shape. In topology, one essential notion that under-grids its analysis is the concept of a homeomorphism. A homeomorphism, in simple terms, entails an equivalence between topological spaces (a sphere, for instance) that is determined by a bi-continuous function between the two. We shall describe this further in a later section. However, what is important to remember at this point, is that it is possible to have an analysis without an explicit model and TDA provides a vehicle for this analysis.

1.0.4 Scope of this Thesis: Objectives and Contribution

In this thesis, we explore how topological data analysis tools can be applied to problems in VLC, particularly in indoor applications. Over the development of this thesis, we briefly touch upon how these methodologies can be utilized in a more generalized context. This generalized context, which will be established with the

aid of concepts and methodologies from topological data analysis, will hopefully give some insight into the following hypotheses:

1. To what extent can tools borne from research pursuits in topological data analysis TDA be used to discover the geometric attributes of "object(s)/obstruction(s)" from infrared data?

2. Following, the above point, qualitatively explore how the algorithm utilized can be adapted/extended to reconstruct the *actual shape* of an "object/obstruction".

3. As mentioned in the opening paragraph of this section, we consider how our current application of topological data analysis may inform future research endeavors; either in VLC and other areas such as image processing/machine learning.

These items broadly outline the overarching themes we hope this document's reader will appreciate whilst consuming this document. Although, at first encounter, a connection between VLC and TDA may seem tenuous, we shall allay this concern in due time.

In this thesis, we have three essential objectives:

1. We demonstrate that topological data analysis can be effectively used to solve an important problem in the burgeoning and active field of research called Visible Light Communications VLC.

2. We discuss and evaluate the utility of our positioning algorithm in a VLC system.

3. We discuss viable extensions of our current algorithm and consider the possible future directions someone could extend the current work.

Chapter 2

Topology and Homology Background

2.1 Opening Remarks

In this section, we review the essential theoretical concepts from topology in addition to some of the practical adaptations of these concepts. To clarify, in the latter part of this section, we are not reviewing existing applications of topology based upon similar methodologies as employed in this document. Instead, we are simply reviewing the necessary practical, or *engineering-based*, solutions that even a recently "green" practitioner would need to know.

2.2 Topological Data Analysis (TDA)

We suggest that the insights provided by topological data analysis are difficult to emulate with other methods. We first begin by remarking upon certain notions most prevalent and important in any discussion about topology. An extremely thorough review of these concepts can be found in these text(s) [10], [11].

2.3 Basic Introduction to Topology

In a generic sense, topology is the study of particular mathematical structures, hereafter referred to as topological spaces, the invariant properties of these spaces and the maps over and between these abstract objects. A topological space is defined by two primal objects, a set $\mathcal{X}$ and a topology $\mathcal{T}$ over that set. A topology $\mathcal{T}$, to be precise, is a collection of subsets of $\mathcal{X}$, oft referred to as open subsets. Any given collection $\mathcal{T}$ will be characterized in the following ways:

(1) $\mathcal{X}$ and $\emptyset$ are open.

(2) The union of any family of open subsets is open.

(3) The intersection of any finite family of open subsets is open.

Though this definition/properties may seem uninspiring at this point, it is important to keep this essential construction in mind when we discuss the more pertinent aspects of topology such as simplicial complexes, homology and most saliently, *persistent homology*. In addition, we have forgone discussions about other topological concepts, however we shall introduce them if and when deemed necessary.

The most venerated invariant in topology is the homeomorphism. A homeomorphism essentially is a bi-continuous map between topological spaces. A simple example of this phenomenon is the popular homeomorphism between a mug and a donut. This example is particularly illustrative because it is easy to visualize how modifications of either object by *particular* deformations can be applied to obtain the other object. To aid in the visualization, we added a (Figure 2.1) few snapshots of the process that would "convert" a mug into a donut. These deformations, by consequence of the bi-continuity of a homeomorphism (technically this would be require some additional assumptions about the underlying homeomorphism, namely that this *infinitely* differentiable map ($\mathcal{C}^\infty$ between topological

manifolds), prevents any sort of tearing or self intersections of either object. An interesting consequence of just considering this particularly simple example is that both objects, in a sense, represent the same underlying "abstract" mathematical structure. Though topology is a vast field in which other important invariants are considered, a homeomorphism is without a doubt, one of the most fundamental types of maps between any two topological spaces.

Given this seemingly austere description of a topological space, it may not seem immediately obvious how these ideas connect to the sort seen in analysis typically conducted in image processing or computer vision. However, such a hesitation to weave "old" and well-worn mathematical ideas without utilizing an evidentiary based approach would be ill-advised.

Even though we described what a topological space is and its essential building blocks, the given description does not really provide compelling motivation for applications. And although homeomorphisms between topological spaces can help to motivate that some plausible topology of some set, in some qualitative sense, codifies its "shape", there are some important considerations left unanswered. First, how do we establish that a homeomorphism exists between non-contrived topological spaces. Now, on the face of it, embarking upon such a venture would be difficult. Attempting to discover a precise homeomorphism would require having perfect knowledge of both topological spaces and also knowledge of how to build a map between them. It is easier to make such an arrangement with contrived

(a) (b) (c) (d) (e)

FIGURE 2.1: A depiction of the process of a "mug" topologically morphing into a "donut". Courtesy of https://en.wikipedia.org/wiki/User:LucasVB/Gallery.

examples, but when dealing with real data-sets, our hopes are likely to be dashed. Fortunately, there are finite approximations to topological spaces (under a specific sets of conditions) which are amenable to computation. More importantly, these finite approximations also are imbued with a related topological invariant, called homology. Homology, as it turns out, is invariant over homeomorphisms and can be calculated by a computer. We shall delve more into homology in the coming section.

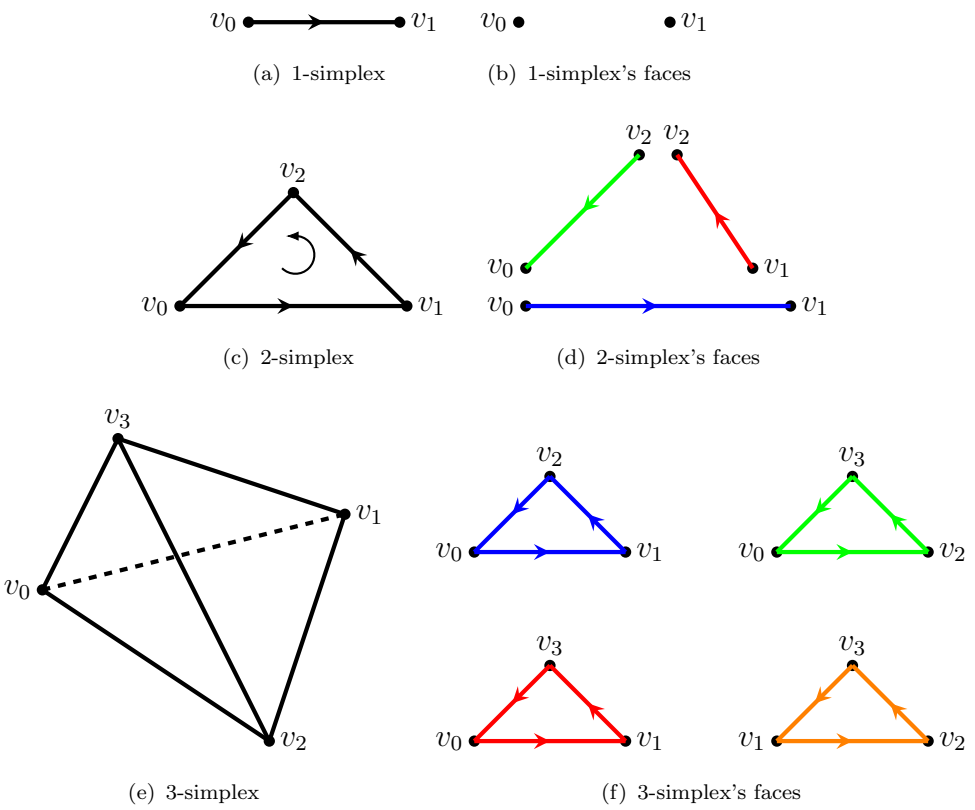
2.4 Simplicial Complexes, Homology, Persistent Homology

In this section, we describe the relevant aspects of algebraic topology that will be important within our work. For a more thorough description of the concepts described and more, please consult [10].

2.4.1 Simplicial Complexes

We shall now describe what a simplicial complex is as well as its essential building blocks. A simplicial complex is a set K, together with a collection S of subsets of K called simplices such that for all $v \in K, \{v\} \in S$ and if $\tau \subseteq \sigma \in S$ then $\tau \in S$. We call the sets $\{v\}$ the vertices of K. When it is clear from context what S is, we refer to set K as a complex. We say $\sigma \in S$ is a k simplex of dimension k if $|\sigma| = k + 1$. If $\tau \subseteq \sigma$, τ is a face of σ, and σ is a coface of τ. An orientation of a k-simplex σ , $\sigma = \{v_0, \ldots, v_k\}$, is an equivalence class of orderings of the vertices of σ, where $(v_0, \ldots, v_k) \sim (v_{\tau(0)}, \ldots, v_{\tau(k)})$ are deemed equivalent if the sign of τ is 1. With this basic machinery, it is not a stretch to see that low-dimensional k-simplicies, when $0 \leq k \leq 3$, are vertices, edges, triangles,

and tetrahedron. Exemplar drawings of the last one to three simplicies (edges, triangles, tetrahedrons) can be found in Figure 2.2.



(a) 1-simplex (b) 1-simplex's faces

(c) 2-simplex (d) 2-simplex's faces

(e) 3-simplex (f) 3-simplex's faces

FIGURE 2.2: Some simple simplicies and their faces.

2.4.2 Chain Complexes

With the definitions of a simplex and simplicial complex given, we now turn our attention to a consequent construction, a *chain group* [11]. A chain group C_k is a set that has elements which are called k-chains. These elements are defined as the finite formal sum of k-dimensional simplicies in the following way:

$$\sum_{\alpha} k_\alpha \sigma_\alpha \tag{2.1}$$

where $k_\alpha \in \mathbb{Z}$ and α is an index over simplicies. This formal sum can be thought of as a collection of k-simplicies with integer multiplicities encoded by the coefficients

k_α. Alternatively, the chain group C_k is also the free abelian group on oriented k-simplicies.

The *boundary operator*, $\delta_k : C_k \to C_{k-1}$ maps k-chains to the sum of $(k-1)$-dimensional faces of its k-cells, which is necessarily a $(k-1)$-chain. The boundary operator is a homomorphism that acts linearly on a chain c. A homomorphism is a structure-preserving map across two identical algebraic structures. It can essentially be likened to other structure-preserving maps like isomorphisms or homeomorphisms that have been alluded to earlier in this document.

For a given simplex $\sigma = [v_0, \ldots, v_k]$, we have:

$$\delta_k \sigma = \sum_i (-1)^i \Big[v_0, v_1, \ldots, \hat{v}_i, \ldots, v_k\Big] \tag{2.2}$$

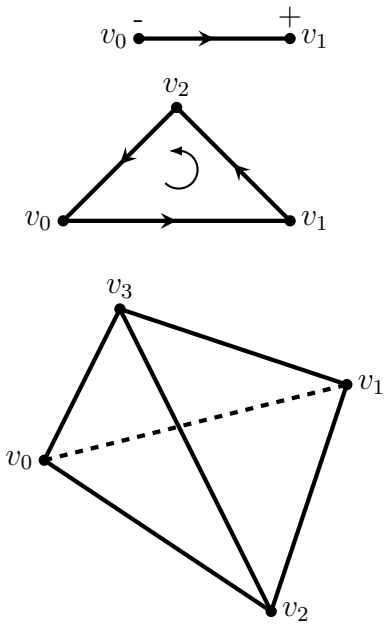
where $\hat{v}_i$ indicates that the term v_i is missing from the sequence defining that simplex. A given term in 2.2, $[v_0, v_1, \ldots, \hat{v}_i, \ldots, v_k]$, is referred to as the boundary of $[v_0, v_1, \ldots, v_k]$. We show some examples of how to construct a boundary of a simplex below in Figure 2.3 [1].

Repeated applications of the boundary operator 2.2 connect chain groups in the following way:

$$\cdots \to C_{k+1} \xrightarrow{\delta_{k+1}} C_k \xrightarrow{\delta_k} \cdots \tag{2.3}$$

The kernel and image of a boundary operator are called the *cycle*, $Z_k = \ker\delta_k$, and the *boundary*, $B_k = \mathrm{im}\delta_{k+1}$, groups. An important property of the boundary operator is a quite intuitive one, namely that the boundary of a boundary is always empty. This can be formally stated in terms of the boundary operator as $\delta_k \circ \delta_{k+1} = 0$. This is readily apparent by also applying 2.2 twice to some

[1] This diagram is very similar to the one found on pg.105 of [11]. Although this diagram was recreated from scratch, we feel it is necessary to give credit to the original author/creator.



$\delta[v_0, v_1] = [v_1] - [v_0]$

$\delta[v_0, v_1, v_2] = [v_1, v_2] - [v_0, v_2] + [v_0, v_1]$

$\delta[v_0, v_1, v_2, v_3] = [v_1, v_2, v_3] - [v_0, v_2, v_3] + [v_0, v_1, v_3] + [v_0, v_1, v_2]$

FIGURE 2.3: Applications of equation 2.2 upon the simplicies from Figure 2.2.

simplex σ. Since its proof, courtesy of [11], is rather short, we include it here for completeness.

Theorem 1. The composition of $\delta_k : C_k \to C_{k-1}$ and $\delta_{k-1} : C_{k-1} \to C_{k-2}$ is zero.

Proof. Given some $\sigma \in C_k$, which can be expressed as $\Big[v_0, \ldots, v_k\Big]$, we have $\delta_k(\sigma) = \sum_i (-1)^i \Big[v_0, v_1, \ldots, \hat{v_i}, \ldots, v_k\Big]$. When we apply δ_{k-1} to $\delta_k(\sigma)$, we have

$$\begin{aligned} \delta_{k-1}\delta_k \quad = \quad & \\ & \sum_{j<i} (-1)^i (-1)^j \Big[v_0, v_1, \ldots, \hat{v_j}, \ldots, \hat{v_i}, \ldots, v_k\Big] \\ + \quad & \sum_{j>i} (-1)^i (-1)^{j-1} \Big[v_0, v_1, \ldots, \hat{v_i}, \ldots, \hat{v_j}, \ldots, v_k\Big] \end{aligned}$$

These summations cancel because after switching i and j in the second sum, the terms in latter become the negative of the terms in the first. □

Consequently, it is possible to note that

$$B_k \subseteq Z_k \subseteq C_k. \tag{2.4}$$

This relationship will be especially important when we discuss homology. Please note, that the boundary operator will necessarily be the null map whenever the dimension of C_k demands that it be the empty set ($k = 0$).

2.4.3 Homology

The k^{th} homology group $H_k = {}^{Z_k}/_{B_k}$ is of interest because of the previous set inclusions determined by (2.4). As has been already stated, for any boundary homomorphism, we have $\delta_k \circ \delta_{k+1} = 0$. Therefore, for any p-chain in B_k, these chains, hereafter referred to as boundary chains, form a subgroup of the cycle group Z_k. Once we take the quotient of Z_k by B_k, we discover that its elements are classes of *homologous* cycles. These classes of cycles are defined as $c + B_k = \{c + b \quad | \quad b \in B_K\}$. One can visualize the relationship between the groups Z_k, B_k and C_k as:

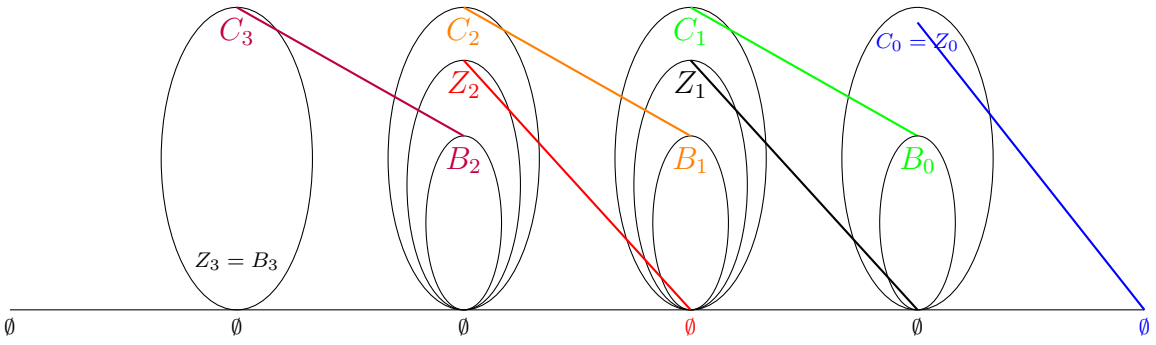


FIGURE 2.4: Chain, cycle, boundary groups and their images under the boundary operators. Courtesy of [2].

In general, the k^{th} Betti number is the rank of H_k, which is.

$$\text{rank } H_k = \text{rank } Z_k - \text{rank } B_k$$

Alternatively, for the group H_k, if we assume that the coefficients for the chains come from some field F, then in general we have

$$\begin{aligned} H_k &= F \oplus F \oplus \ldots \oplus F \\ &= F^{\beta_k} \end{aligned}$$

where $\oplus$ is the direct sum. Then it is true that H_k is a vector space over F and its rank is β_k. The rank of these groups are typically called Betti numbers, and for a H_k, β_k is called the k^{th} Betti number. For any simplicial complex X representing some topological space, we write $H_k(X)$ and $\beta_k(X)$ for the k^{th} homology group and k^{th} Betti number. These values are defined for any integer k, but if the dimension of the underlying topological space X is n, then the only possibly interesting homology groups are defined for values of k such that $0 \leq k \leq n$. Consequently, for values of k outside of this range H_k is the empty set and $\beta_k = 0$.

When a topological space, typically a topological manifold, can be described by some simplicial complex, homology is a useful algebraic tool used to discover particular algebraic properties (its homology groups and their Betti numbers). The Betti numbers of the homology groups counts the number of k-dimensional holes and these holes effectively correspond to actual voids within a topological space. For a more thorough explanation of homology, please consult either [10] or [11].

2.4.4 Persistent Homology

Persistent homology enables an investigator to incorporate certain measurements in the description of topological features. In particular, persistent homology enables her to analyze the evolution of homological features (their "births" and

"deaths") within a data set analyzed over several scales appropriate for a given set of measurements.

Before we embark on a full explanation, we first describe the concept of persistence for single variable functions (We tersely reproduce the explanation given in [12]). Let us consider some smooth function $f : \mathbb{R} \to \mathbb{R}$. As a brief reminder, a smooth function is generally a function that is differentiable up to any order e.g. the exponential function e^x. An x is a *critical point* and $f(x)$ is a *critical value* of f whenever $\frac{d}{dx}[f(x)] = 0$. Whenever $\frac{d^2}{d^2x}[f(x)] \neq 0$ at a critical point, we label that critical point as *non-degenerate*. We further constrain the quality of f by requiring that it only have non-degenerate critical points with distinct critical values. Thus, each critical point is either a local minimum or a local maximum. Now consider the *sublevel* sets

$$\mathbb{R}_t = f^{-1}(-\infty, t].$$

It is readily observable that as we increase t from $-\infty$, the connectivity of $\mathbb{R}_t$, remains the same except when we pass a critical value. At local minimums, a new component is added to the *sublevel* set and at a local maximum two components merge into one.

These critical points are paired in the following manner: When a local minimum appears, then that representative local minimum represents a *new* component. However, when a local maximum appears, we merge two already existing components by pairing the found maximum with the higher (or younger) of the two local minima on either side of the maximum (provided they exist of course). Consequently, the other minimum represents the effective component from the merger.

When x and y are paired in the aforementioned manner, we define the pair's *persistence* as $f(y) - f(x)$. We can use a *persistence diagram* to code persistence by mapping each pair to a point $(f(x), f(y))$ such that the coordinates are the

corresponding critical points of the function f. Within the diagram, points that are above half space specified by $x_1 = x_2$. Persistence, as put forth, can be visibly seen as the vertical distance to the diagonal of the half space $x_1 = x_2$.

Given this definition of persistence for single variable functions, we now present an analogous definition for simplicial complexes. Given a simplicial complex K, we now consider a nested sequence of sub-complexes, whereby a sub-complex is a subset of simplicies is closed under the face relation, referred to as a *filtration* of K. This filtration can be written as:

$$\emptyset \subseteq K_0 \subseteq K_1 \subseteq \cdots K_m = K. \tag{2.5}$$

The sub-complexes are like the sublevel sets encountered before. In particular, for our application, the "sublevel" sets that we form from our data-sets unveil the data at gradually increasing scales. A homology class α is *born* at K_i if it is not found in the image of the inclusion map $K_{i-1} \subseteq K_i$. This can be further extended by observing that if α is born at K_i, it *dies* entering K_j if the image of the inclusion map $K_{i-1} \subseteq K_{j-1}$ does not contain the image of α but the image of the inclusion map $K_{i-1} \subseteq K_j$ does. The persistence of α can be written as $j - i$. Given this scenario, we can consider the number of p-dimensional homology classes that are born at or before K_i and are still alive at K_j. These necessarily includes classes that never die through the filtration. The number just referenced is the rank of the map between homology groups:

$$f_p : H_p(K_i) \to H_p(K_j). \tag{2.6}$$

So, in essence persistent homology is tracing how homology classes "survive" through different portions of the filtration through maps such as f_p. We can go back to regular homology by fixing the entire filtration at a single point, of course.

However, the primary difference is that we are defining maps between compatible homology groups in *different* simplicial complexes.

2.5 Simplicial Complex Formation/Generation

There are a number of different methods for generating simplicial complexes extant in the literature. In this section, we review several of the methodologies used and provide a cost benefit analysis of each method in terms of their applicability and efficiency.

2.5.1 Cech Complex

Given a point set X in some metric space and equipped with a parameter $\epsilon > 0$, the Cech complex [13], K_ϵ is defined as the simplicial complex whose simplicies are formed as follows. For every subset $S \subset X$ of points, categorize S as simplicial, if for every $\frac{\epsilon}{2}$ ball formed about each point in S, there are a non-zero number of intersections. The number of balls that have non-trivial intersection is the dimension of the complex S. This definition obviously satisfies the standard definition of a simplicial complex because any set $S^{'} \subset S$ is clearly also a simplex.

An obvious question that arises from such a construction is if there are conditions that make it "equivalent" to the underlying topological space X. Fortunately, it can be shown by the Nerve Theorem[11] that when the homotopy types, another topological invariant, of X and C_ϵ are the same then, homology is unable to distinguish between these two. Unfortunately this renders a persistent homological analysis from being carried out. We deem it necessary to discuss the complex because this construction type could be deemed by someone unaware as a straightforward way to build a simplicial complex. In addition to an inability to perform a

persistent homological analysis, this particular construction is also computationally expensive. In fact, the computing time for the entire complex is generally exponential! This explosive asymptotic computational resource usage precludes analysis of any worthwhile data-set that may contain even just a thousand points (a small figure for most *worthwhile* data-sets). We shall need to consider other simplicial complex generation methods.

2.5.2 Vietoris Rips Complex

The Vietoris Rips [14] complex is developed in a similar fashion as the Cech complex, except that instead of adding a k simplex only when there are k common intersections between balls of radius ϵ (> 0) as before, we only consider pairwise intersections between such balls. We denote this complex as VR_ϵ.

Even though being fairly straightforward, do our previous concerns about topological relevancy still hold? We observe that for all $\epsilon > 0$, the following holds

$$C_\epsilon \subset VR_\epsilon \subset C_{2\epsilon} \tag{2.7}$$

However, despite the reduction in the number of intersections needed to be evaluated during the simplex's construction, attempting to compute the Vietoris complex when the number of points in a data-set exceeds a certain practical threshold is computationally expensive. In particular, we have, over the course of using the tool, determined that for the number of points (on the order 10^4 to 10^6) typically found in the data-sets of this thesis, the computational time becomes prohibitively expensive.

2.5.3 Witness Complexes

The last and most appealing simplicial complex construction algorithm we discuss is the witness complex. In light of the sheer amount of memory that the other simplicial complex construction algorithms would require, it is quite fortunate that there exists this simplicial complex construction method. The simplest and most straightforward description, the simplicial complex is based upon a subset of landmark points of the data-set. Interestingly enough, no other input parameters are required. The remaining points are "witnesses" to the landmark points chosen as members of simplicies in a prescribed manner. We start by designating a distance matrix D, that is $n \times N$ matrix of non-negative entries, where n is the number of landmark points and N is the number of points. This matrix, as its namesake implies, contains the euclidean distance between i^{th} landmark point $(1 \leq i \leq n)$, where the landmark point set is denoted by L, and the j^{th} point $(1 \leq j \leq N)$ in the entire data-set, denoted by X. The witness complex $W_\infty(D)$, with vertex set $\{1, 2, \ldots, n\}$ is given as: Each i that satisfies these conditions is

1. An edge between points a and b exists in $W_\infty(D)$ if there exists a data point $i \leq i \leq N$ such that $D(a, i)$ and $D(b, i)$ are the smallest entries in the i^{th} column of D.
2. Consider that there exists some p-simplex $\sigma = [a_0 a_1 \ldots a_p]$, whose faces all belong to $W_\infty(D)$. Then σ belongs to the $W_\infty(D)$ iff a data point $1 \leq i \leq N$ exists such that $D(a_0, i), D(a_1, i), \ldots, D(a_p, i)$ are the smallest $p + 1$ entries in the i^{th} column of D.

FIGURE 2.5: Procedure for building a witness complex

called a "witness" to the existence of σ.

Choosing the landmarks is another matter that is necessary to discuss. There are two main methods for choosing landmarks, random or *maxmin*. Since the former is easy to grasp, we shall explain the latter. Maxmin is an inductive procedure that begins by selecting some landmark point $l_1 \in Z$ randomly. By induction, if

the set $l_1, l_2, \ldots, l_{i-1}$ have been chosen, then we choose $l_i \in Z\{l_1, l_2, \ldots, l_{i-1}\}$ as the data point that maximizes the function

$$\rightarrow \min D(z, l_1), D(z, l_2), \ldots, D(z, l_{i_1}) \tag{2.8}$$

where D is the ambient or assumed metric. This process is iterated until the desired, as given by some criteria, number of landmark points have been chosen. This simple procedure is striking because the landmarks it finds usually provide a good "cover" of the data-set. However, it should be noted that whilst this selection process is fast due its inherent greediness, we have seen that in some instances, this process does not capture the best landmarks. We shall demonstrate this by asking the reader to consider the following situation.

Let there be two hollow 2-spheres of equal radius, $r > 0$ units, whose centers are $d > 0$ units apart. For fixed r, the variation of the numerator of the ratio $\frac{d}{r}$ will produce scenarios where either the spheres are totally separate and do not intersect one another or the spheres overlap. It should be noted that within this small example, we have two sampling processes for points upon the hollow sphere. For simplicity, we sample, with uniform distributions, the θ and ϕ angles of each sphere. Instead, we could also linearly sample these two angles. Irrespective of the sampling procedure, we find that it is rather straightforward to get the "expected" answer for the persistent Betti numbers, i.e. the number of "holes" (where "holes" is intuitively interpreted), when the two spheres are well separated, in other words $\frac{d}{r} \gg r$. However, when the two spheres are not well separated, $\frac{d}{r} \approx 1$ or $\frac{d}{r} < 1$, there are two possibilities to consider here. Either the spheres are almost touching or the spheres overlap one another and produce a 2-dimensional void within the overlap. Though it may seem straightforward to deduce the Betti numbers of the simplicial complex (via the witness complex), it turns out that this is not so simple. Depending on the landmark selection, we could get vastly different landmark point sets and consequently, the resulting homology groups may not

reflect the underlying data-set. This is especially true when the landmark point selection process is random. This unexpected result provides the groundwork and inspiration for the application of persistent homology to particular problems we discuss later in this document.

Chapter 3

Methodology & Results

3.1 Opening Remarks

This chapter contains the essential details of our experiments to determine the viability of utilizing persistent homology for passive object localization. Specifically, we first discuss the simulations we have constructed to test the efficacy of using persistent homology to perform localization in a standard VLC application setting. We follow a similar protocol as found in other established works([3],[4] and [1]). We should clarify that when we say we follow a similar protocol, we do so by replicating some essential elements of their methodologies. Before we discuss our results, we now give a short description of the experiments we seek to carry out. In summary, we are simulating a scenario where there are multiple transmitting nodes (that use LEDs) in a standard office room and that there multiple receivers on the ceiling. We trace these rays and sample their trajectories to collect a dense set of points. It is the analysis of these collected points via PH that may allow us to ascertain the location of passive objects (a chair or another piece of furniture e.g.) in the room.

Before we delve any further, here is a listing of our parameters can be found in Table 3.1. An explanation of each parameter is produced here. For each transmitter, we simulate its height in the room, up-link wavelength, and LED transmit power. Each of these parameters have a rather straightforward definition. We also include transmitter grid resolution as a way to specify the number of transmitters in the room. For the receivers on the ceiling, we simulate the receiving surface area of the photo-detector, as it readily apparent that there is finite amount of area that can receive incoming radiation for each receiver. Also, we must consider that the receiver can also receive incoming radiation over a certain range of angles. This is specified by the half angle. We also vary the number and location of receivers in the room. Each receiver has the same parameters that have already been mentioned (height, surface area, field of view). Given the specification of the transmitter and receiver parameters, we must also specify the room's dimensions. We choose a fairly standard size in all of our experiments, as this models a standard office room. Of course, there are rooms of varying and interesting size, but we postulate this room type (the office setting) is more commonly found than an indoor stadium or conference room. Within our experiments, we vary the ray sampling rate, the number of objects and their shape type. In regards to the last parameter, we keep the shape cylindrical in all our of experiments. However, we should say that other shape types are also of interest. We could choose to pursue that line of investigation in future work.

Our tests include different types and numbers of objects that we hypothesize that our method is capable of positioning and localizing to a application-dependent and satisfactory degree. In order to utilize persistent homology, in a similar spirit to other applications of persistent homology evident in the literature ([15],[16],[17], [18]), we first must build a point cloud from the available measurements we have between the photo-detectors on the ceiling and the transmitters on the ground (see Figure 3.4). An example of a point could generated during our experiments is given

here. While we expand upon this setup in a later section, we remark here that we

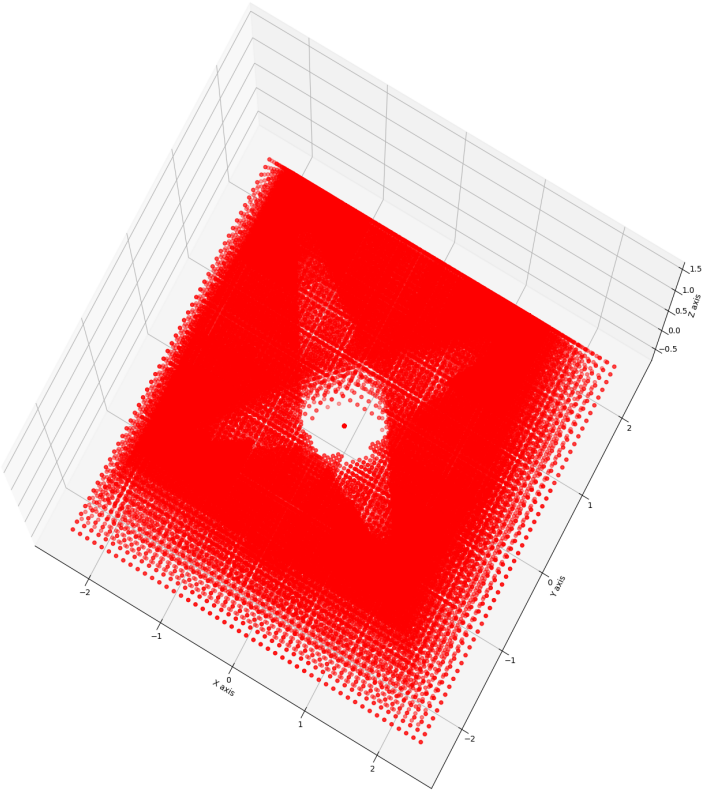


FIGURE 3.1: Infrared Ray Samples based upon Figure 3.2

make a simplifying assumption of minimal multi-path fading and interference that would be consequent of the signal emanations from the infrared transmitter(s) in the experiment. With this simplifying assumption, we trace a ray emanating from a transmitter and determine whether, given the obstructions/objects in the room, it reaches one of the Q photo-detectors on the ceiling. If this ray does reach a photo-detector, in other words, if there exists a LOS path from one of the transmitters to at least of the photo-detectors on the ceiling, then several points

(of course, not including the last) along this ray are sampled and collected in the aforementioned point cloud. It should be noted that the points collected in this process are necessarily *mostly* not points on the object. On the contrary, the points by consequence of the sampling process are necessarily not on the surface or inside the objects (provided that innards should be noteworthy).

A rather straightforward deduction of this process shows that regardless of the object placed in the room, if the transmitters are only placed on the ground, then the point cloud will necessarily exclude points on the top(s) of an object and less importantly, but worthy of consideration, exclude points contained in certain cavities along an object's "torso".

This unfortunate circumstance can be remedied by placing transmitters on several levels above the floor. A staggered layout of the levels is desirable as it ensures that the same point in the room is either never sampled twice along different rays or at the very least minimally re-sampled. However, re-sampling a point more than once may be not totally undesirable. Even though we do not consider this potential complication in this thesis, consider that we also were to consider an object's specularity.

In all of our experiments, instead of collecting data from one transmitter on the ground, we are collecting data from a densely and deterministically laid grid of transmitters on the ground. By having many transmitters, we can collect several pieces of data. We can predict, given we have a reasonable formulation/description of the channel, the received signal strength for each photo-detector (see the examples in Figure 3.2 and Figure 3.3). The plots serve as important important visual aid for a number of reasons. Firstly, the collection of these plots essentially give the viewer a quick overview of the light that each photo-detector is capable of receiving. Consequently, when we later build the point cloud and draw conclusions about the calculated persistent homology and what it may mean, we can

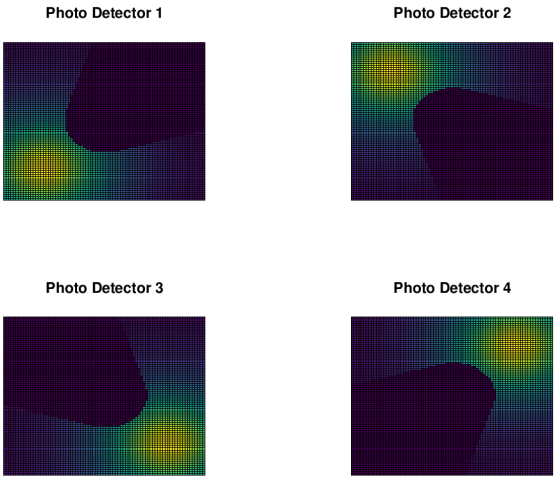


FIGURE 3.2: Received Signal Strength of Line of Sight Rays for Four Photo Detectors and One Object

refer back to these diagrams to understand later results. This can be seen, even at this stage, because the point cloud we later build is effectively an amalgamation of these plots, and the density of points in the point cloud is directly a function of plots like these. For instance, consider this point cloud (Figure 3.4). This point cloud, in particular, is a function of Figure 3.2. It can be readily observed that near the center of this diagram, the vacancy left behind is not quite as cylindrical as would have been expected.

Within this diagram, for $q \in \{1, Q\}$, PD_q is an individual photo-detector, $S(t)$ is an infrared transmitter, and $v^q(\theta, t)$ is the channel impulse responses measured at each PD_q. Currently, at this part of the analysis we do not integrate the information from a hypothesized or estimated channel model into our simulations and consideration. We also do not erroneously make the assumption that such information is unimportant; however, an integration based upon a well-informed evaluation of the channel model could aid in other object-localization related tasks. Instead, we trace the trajectories of rays emanating from transmitting LEDs placed about the room (near the ground) and make enough simplifying assumptions about the propagation of these rays, namely we do not consider

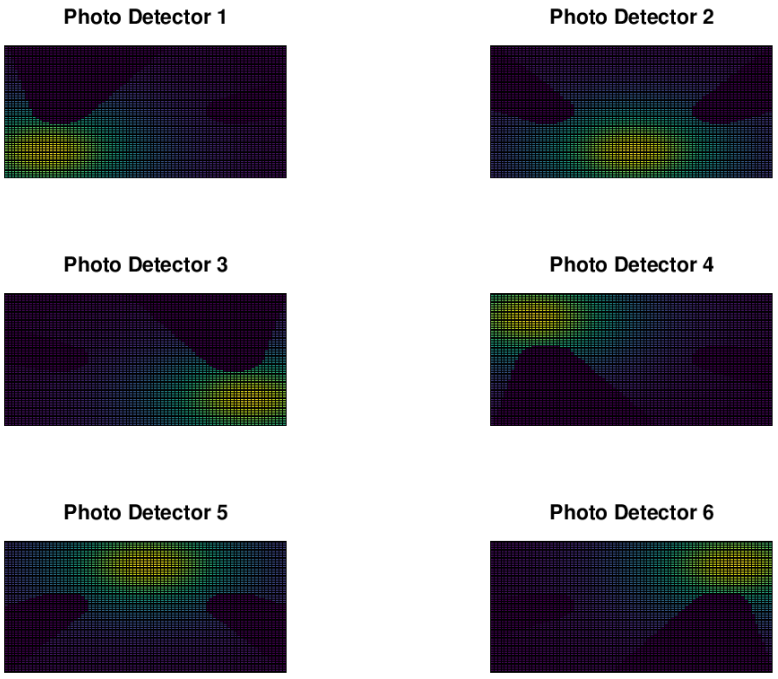


FIGURE 3.3: Received Signal Strength of Line of Sight Rays for Six Photo Detectors and Two Objects

reflections and constructive/destructive interference nor the effect of the room's properties upon the propagation and reflection of each ray. We primarily do this because of the small size of the room and because of the general properties of light rays in such an environment. Given the dimensions, we assume that second order effects (as discussed above) can be ignored given the strong power of the light beam across the dimensions of the room.

We shall enumerate the objects we test and discuss why, at this juncture, these objects are reasonable approximations for the more difficult objects we could possibly simulate, i.e. a complete description of a human body or room furniture. Within our experiments, the objects that we place within the room are cylinders, and cylinders with differing radial profiles. The last category is important to include because we have stated before that topology, when used to describe common shapes in a particular "abstract" form, should produce equivalent qualitative and

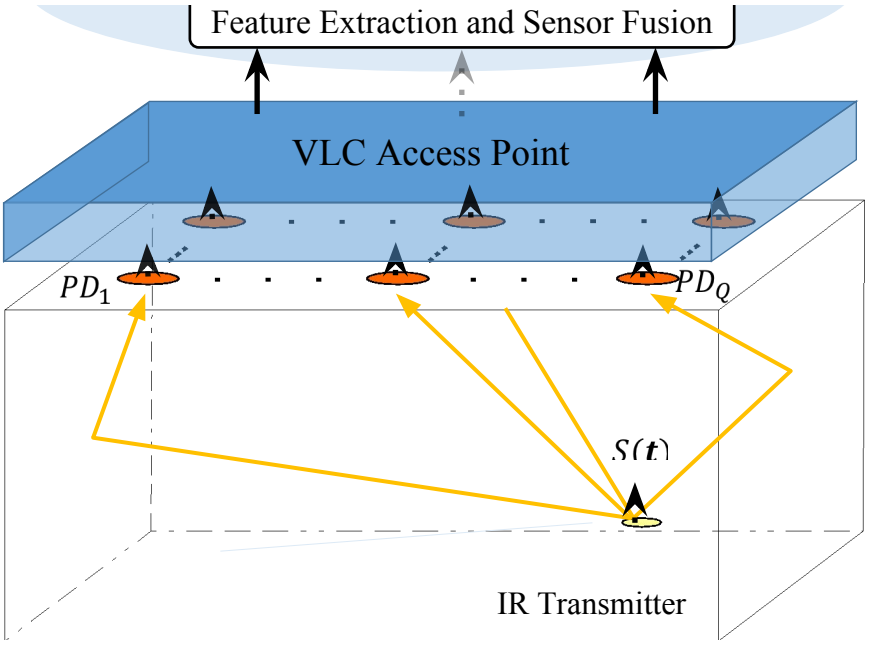


FIGURE 3.4: System configuration for visible light communication up-link system and the impulse responses based on it. Courtesy of [3].

quantitative (with the aid of persistent homology) descriptions. We also include an elliptical cylinder and a "squeezed" cylinder as possible shapes to test:

Elliptical cylinder

$$x = r\cos(\theta)$$
$$y = \pm\frac{b}{a}\sqrt{a^2 - x^2}$$
$$z = z$$
$$r = r_0$$
$$\theta \in \{0, 2\pi\}$$

"Squeezed" cylinder

$$x = x$$
$$y = y$$
$$z = z$$
$$r = r_0$$
$$r' = \sqrt{x^2 + y^2}\left[\frac{(z-\mu)^2 + \psi}{(r_0 + \psi)^2}\right]$$
$$\theta \in \{0, 2\pi\}$$
$$x' = r'\cos(\theta)$$
$$y' = r'\sin(\theta)$$
$\mu \in \mathbb{R}$ - cylinder midpoint
$\psi \in \mathbb{R}$ - a "push" parameter

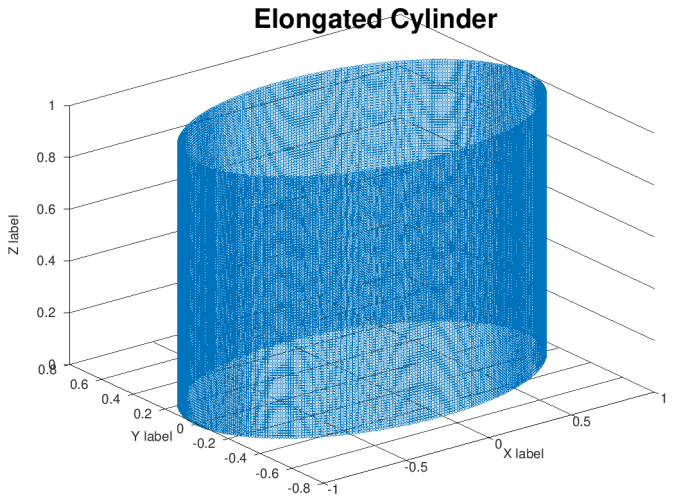


FIGURE 3.5: Elongated/"Elliptical" Cylinder

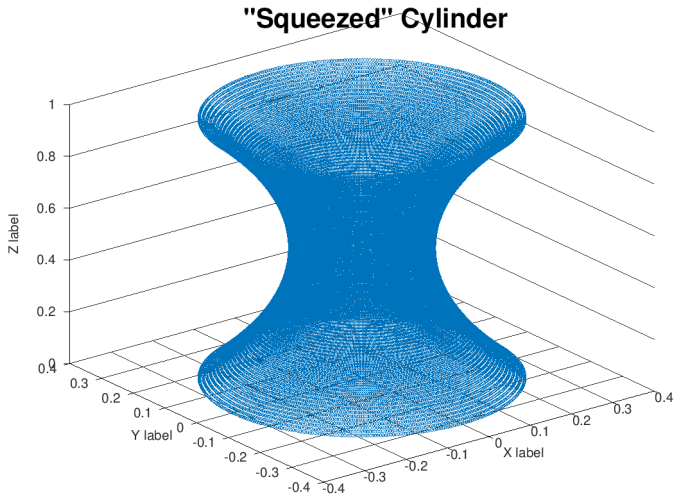


FIGURE 3.6: "Squeezed" Cylinder

We have produced Figures (Figure 3.5 and Figure 3.6) of an elliptical and "squeezed" cylinder. Following this, we detail the parameters we employ in the topological data analysis. Specifically, earlier in the document, we discussed the different types of simplicial complex constructions available for analysis. In our experiments, we utilize the witness complex, introduced in [19], because whilst this simplicial complex generally does a good job of not generating an exponentially large number of simplicies to process (unlike its counterparts discussed before),

the loss of detail inherent in the complex's construction is tolerable because of the smart landmark point selection as discussed in an earlier section.

Accordingly, we look at different types of grid and ray samplings and their effect on the computational complexity of building and deriving persistent homology calculations. We remark upon the trade-offs between the ability to localize objects and the resolutions of the constructed data-sets. We also consider that it is worth mentioning that there have been theoretically sound efforts to actually localize basis chain elements of a homology group in a data-set [20]. We are inspired by their work, but we must remark that our primary goal within this document is to localize with simple set of coordinates (its center point e.g.) whereas they are trying to find a localized basis of the homology group. Though the goals are similar, they are clearly distinct. We foresee that using localized homology bases could be useful for other tasks such as shape description.

3.2 Experiment Design & Setup

Our experimental setup is inspired by the methodology of [3]. However we depart from their contributions, by not utilizing fingerprints based upon on the line of sight (LOS) peak power of the impulse response and the delay between LOS. In spite of our differences, let us first highlight the similarities.

The visible light communication system we consider for analysis is essentially a simulation of a typical indoor office space. There are white LED fixtures fixed on strategic locations across the ceiling that transmit down-link data to their targets. There are also infrared transmitters that transmit up-link signals to the aforementioned photo detector (PDs) on the ceiling. This scenario is depicted in Figure 3.4. The qth PD is fixed on the ceiling at position $\left(x^{(q)}, y^{(q)}, z^{(q)}\right)$, where $q \in \{1, \ldots, Q\}$ and is facing downwards. We also model objects in the rooms

by keeping account of object's boundary coordinates. An object's coordinates are marked as $\left(x^{(\phi)}, y^{(\phi)}, z^{(\phi)}\right)$ where $\phi \in \{1, \ldots, N_\Phi\}$ and N_Φ is the number of objects in the room that we wish. For the sake of simplicity, we are making rather substantial assumptions about the how an infrared signals will interact with an object within the room. In particular, we do not specify whether the object in question has a matte finish, is translucent or has some other distinctive quality which would markedly affect an infrared signal's refractory nature. We assume that an object will affect a PD's ability to detect the trajectory of infrared ray from certain positions within the room. This simple assumption permits us to simply consider which rays actually "reach" the PDs on the ceiling. Given that we simulate each ray path by tracing a straight line between its start and end points, we are ale to sample the rays N_r times along the vector from the point of origin of the infrared signal to the qth PD on the ceiling.

3.2.1 Simulation Parameters

In table 3.1, we list the basic set of simulation parameters to create the synthetic data-sets used within this work. We vary the room's parameters and the object parameters (shape, number and position). We also vary the number of objects within the room and their geometries. In particular, we vary the number of objects in our experiments from one to three, and we also list their locations for each test.

3.2.2 Effectiveness Metric(s)

We evaluate the effectiveness of the algorithm by assessing the size of by comparing the size and "center of mass" of each the bases of the 1st homology group found along the filtration. The first homology group corresponds to closed loops found within the data-set that persist across a significant portion of the filtration.

Transmitter Parameters	**Value**
Height	0.85 m
Up-link wavelength	950 nm
Transmitter Grid Resolution (same along each x, y, z axis)	0.25m
LED transmit power, P_T	10 mW
Receiver Parameters	**Value**
Surface area of the Photo-Detector, A_{PD}	1 $cm\times$ 1 cm
Height	3 m
Photo-Detector Field of view (Half Angle)	70°
Photo-Detector 1 location	(150 cm, 150 cm, 300 cm)
Photo-Detector 2 location	(350 cm, 150 cm, 300 cm)
Photo-Detector 3 location	(150 cm, 350 cm, 300 cm)
Photo-Detector 4 location	(350 cm, 350 cm, 300 cm)
Room Parameters	**Value**
Room size (width $\times$ height $\times$ height)	$5 \times 5 \times 3$ m^3
Miscellaneous Parameters	**Value**
Ray Sampling Rate	{5,10,20...100}
Object Number	{1...4}
Object Shape Types	{cylinder, elliptical cylinder 3.5, "squeezed" cylinder 3.6} (object parameters specified when necessary)

TABLE 3.1: Simulation Parameters of Ray Tracing Channel Model

In our case, our filtration is built by building sets of simplicial complexes, each parameterized by the a parameter that corresponds to the nearness of a point to its neighbors. When this parameter is low, points tend to be grouped by themselves. As this parameter is increased, a "persisting" grouping of the data set begins to emerge, until eventually all points are part of a single group. We seek to find loops within that "persisting" portion of the filtration as that signifies a meaningful portion of the data set.

We also evaluate the accuracy of this loop by evaluating a measure similar in construction and spirit to the Dice similarity coefficient [21]. The Dice coefficient is typically employed in image segmentation algorithms to evaluate the effectiveness

of a segmentation algorithm relative to some ground truth. Its formulation is:

$$\mathsf{Dice}(A, B) = 2\frac{|A \cap B|}{|A \cup B|} \tag{3.1}$$

where A, B are some subsets of an image, and $|\cdot|$ is the cardinality of a set (for an image, this is typically the effective area or number of pixels of a connected region). It should be clear from this definition that whenever $A = B$, then the Dice coefficient will attain its maximal value of 1. In contrast, when $A \cap B = \emptyset$, the measure will clearly attain its minimal value of 0. Depending upon the type of chain we assess, there can be some ambiguity in determining the ground truth. This is more clearly seen to be the case when using 1-chains, as 1-chains are essentially closed single parameter curves (contours). However, we've found that in our experiments, that the variability in the contour's orientation about the z-axis is generally not a major cause for concern.

We wish to show that even with a naive application of PH, we can quantitatively assess about the connectivity of the underlying space of sampled rays in our simulation.

3.3 Results

Now we show the results of several simulations that we have run. As stated, we show the accuracy of the localization method under a variety of circumstances. We measure this with respect to the capabilities of our chosen algorithm.

There are several important parameters we focus upon to support the argument that persistent homology is a capable tool for localizing objects in our simulation. We assess the transmitter resolution, in other words, the number of transmitters we employ in our experiments that lay very near to the ground (as a reminder, their height is '85 cm). Generally, we noticed that after a certain

Experiment #	**# of Objects**	**Object Parameters**	**Average Maximum Minimum Dice Coefficient**
1	1	Cylinder - Radius = 1m Center = (2.5m, 2.5m) Height = 2m	**Average** = 0.5062 **Maximum** = 0.73195 **Minimum** = 0.35393
2	1	Cylinder - Radius = 0.5m Center = (2.5m, 2.5m) Height = 2m	**Average** = 0.43022 **Maximum** = 0.70946 **Minimum** = 0
3	1	Cylinder - Radius = 0.1m Center = (2.5m, 2.5m) Height = 0.75m	**Average** = (Not Applicable) **Maximum** = (Not Applicable) **Minimum** = (Not Applicable)
4	1	Cylinder - Radius = 0.1m Center = (2.5m, 2.5m) Height = 0.5m	**Average** = (Not Applicable) **Maximum** = (Not Applicable) **Minimum** = (Not Applicable)

TABLE 3.2: Single Object Results: Each over **5 trials** and **Transmitter Resolution** of 0.1m. Here we are varying the object's size and assuming the parameters from Table 3.1.

threshold, there are diminishing returns in terms of localization capability and accuracy. Secondly, we vary the number of photo-detectors on the ceiling. For a single object of a sufficiently large size, four well placed photo-detectors is generally adequate. However, once we increase the number of objects in the room, we observe that now the "receiver resolution", by this we mean to say the relative placement of them upon the ceiling, starts to become an important factor. In all of our experiments, we utilize the witness complex because it is an efficient yet accurate simplicial complex type to build and derive its persistent homology. All of the code was compiled/run in Octave 4.4.0 and Java 8. In particular, the persistent homology portions is from a popular Java library called Javaplex [22] which also offers other topology related codes.

Based on these parameters, we showcase what can happen when the object size, object location, number of photo-detectors, and number of objects and photo-detectors vary. We also show which resolutions are sufficient for an accurate analysis in either the single or multi object scenarios.

We also include a diagram of persistent[1] for the first case in Table 3.2 to showcase what PH actually detects. This is shown in the form of what's called a Persistent bar-code interval diagram in the literature[23]. Reading these bar-codes is actually rather straightforward. For each dimension that we measure persistent homology, we include a diagram. The purpose of the x-axis of each diagram is to parameterize the filtration. The y-axis is simply incidental and does not actually measure anything. It is just there simply because there can multiple homology classes at a given dimension and it is easier to separate them along the y-axis. The length of each line, hereafter bar-code, is a visualization of the persistence of a homology class. In each diagram featured in this document, an extra upper dimension is always included as a sanity check.

3.3.1 Parameter Exploration

We run several experiments, as aforementioned, to show the effects of varying the important parameters highlighted above.

3.3.1.1 Object Size & Location

In this section, we analyze how effective our algorithm is at localizing a single object in the room. Most importantly, we show what happens as we vary the size and location of a single cylinder. When we look at the results contained within Table 3.2, we see that the radius of the cylinder has a more or less inferred effect upon the quality of the recovered 1-chains. Although, 0.1m resolution is a rather modest transmitter resolution to use (a tenth of meter is about the width of standard mug), we see that if the object has a large enough radius and height (as in the first two rows), then we are able to extract 1-chains that at the very least surround the cylinder (see Figure 3.8). We also get other seemingly spurious

[1] A persistent diagram does connote something else within the literature.

1-chains. Those 1-chains are a manifestation of the fact that not all 1-chains are necessarily detected at all filtration values. This sort of occurrence is natural, of course. Consider, if the ray sampling resolution were more fine . Although, we do not show an example, if the ray sampling resolution were instead sufficiently coarse, then the distances between a set of values at $z = z_0$ and at $z = z_1$ would be large enough such that simplicies would not form across these planes. As a result, there might be lots of chains which never survive the entire filtration. Worse yet, they might actually survive the entire filtration value but not actually indicate anything substantial. Remedying this issues is fairly straightforward, simply set various resolutions such that the witness complex is "forced" to see clumps forming over adjacent z-levels at even low filtration values.

Despite a resolution of a tenth of a meter, which already seems quite high, is it possible to achieve even better Dice coefficient values? Looking at the persistent diagrams and their accompanying 1-chains surrounding each cylinder (Figure 3.7, Figure 3.8, Figure 3.9, Figure 3.10, Figure 3.11,), there is a clear relationship between the necessary amount of transmitter resolution and the object size. If the goal is to simply localize the cylinder, then modest resolution values are sufficient. However, if localization of an object also entails discovering a model which clearly articulates an object's geometrical complexity, then a finer resolution may be more appropriate. This is outside the scope of this thesis and may be further expounded upon in a subsequent document.

It is encouraging and validating that there is a nice correspondence between Figure 3.7 and Figure 3.8. If we focus on the second diagram ,of the persistent bar-codes, from the left, we notice that there are three line intervals (bar-codes) and correspondingly there are three prominent 1-chains (loops) in the accompanying figure. This shows us that it is possible to localize objects with a straightforward application of persistent homology.

Notice that the last two rows of Table 3.2 are void of any Dice values. This is purposefully left there to indicate that objects of sufficiently small height and size (radius) can be missed by our setup. This can be remedied by employing photo-detectors that are placed at other locations about the room other than the ceiling. For instance, photo-detectors, can be placed along the other side walls of a room. Additionally, we can also consider placing transmitters at other z-levels besides the ground. As an alternative to photo-detector replacement, these objects can also be discovered by increasing the ray sampling resolution. In either case, it is clear that there are limits to the degree of accuracy one can obtain.

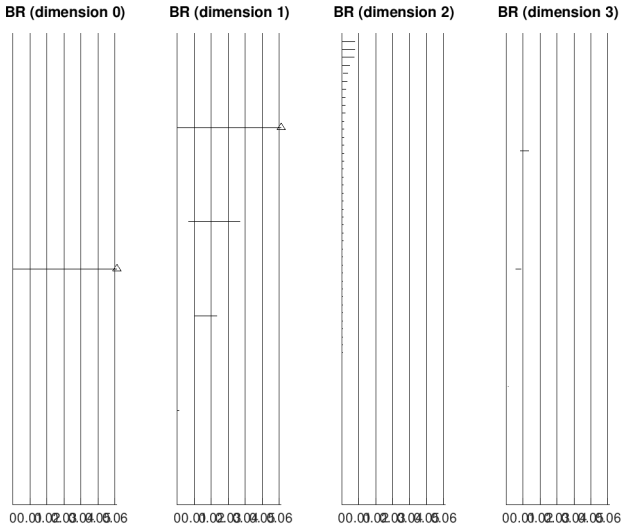


FIGURE 3.7: Persistent bar-codes for of the trials of the first row of Table 3.2

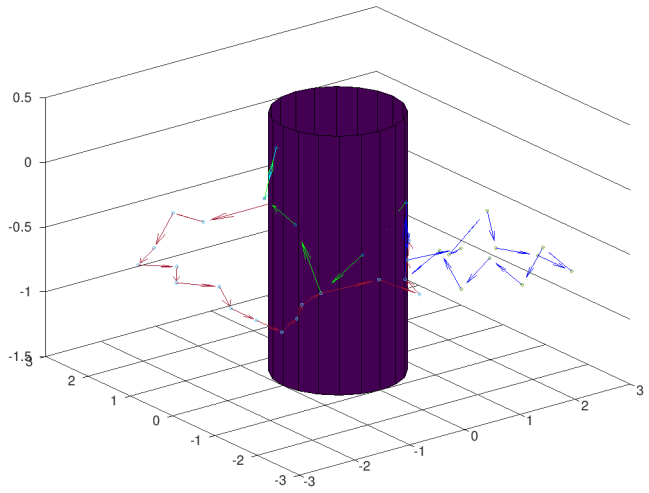


FIGURE 3.8: The 1-chains corresponding to the persistent bar-codes of Figure 3.7

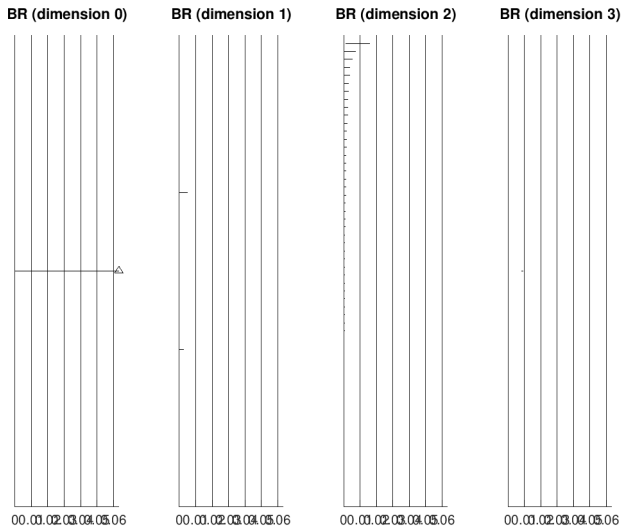


FIGURE 3.9: Persistent bar-codes for of the trials of the second row of Table 3.2

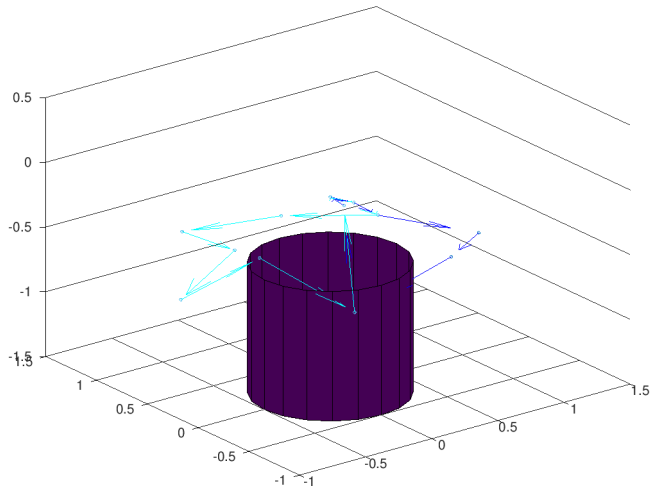


FIGURE 3.10: The actual chains corresponding to the persistent bar-codes of Figure 3.9

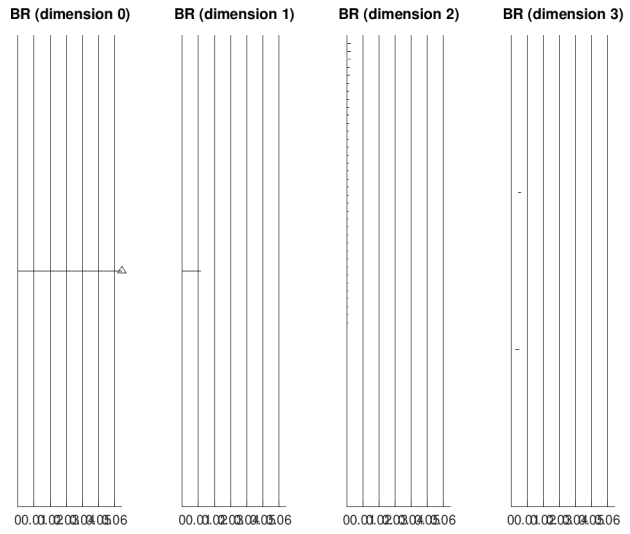


FIGURE 3.11: Persistent bar-codes for the trials of the third row of Table 3.2

3.3.1.2 Multiple Objects

Experiment #	**# of Objects**	**Object Parameters**	**Average Maximum Minimum Dice Coefficient**
4	2	Cylinder - **Radius** = 0.75m **Centers** {(1m, 2.5m), (4m, 2.5m)} **Heights** {2m, 2m}	**Average** {0.14347 0.39995} **Maximum** {0.71734 0.71905} **Minimum** {0 0}
5	3	Cylinder - **Radius** = 0.75m **Centers** {(1m, 2.5m), (2.5m, 2.5m), (4m, 2.5m)} **Heights** {2m, 2m, 2m}	**Average** {0 0.29428 0.12717} **Maximum** {0 1.03895 0.31792} **Minimum** {0 0 0}

TABLE 3.3: Multiple Object Results: Each over **5 trials** and **Transmitter Resolution** of 0.1m. In this table, we vary the number of objects.

In this section, we analyze our localization algorithm's efficacy when there are multiple objects in the room. In particular, we place several cylinders, that is three, four, or five, in the room and observe. For all of these experiments, we keep the number of photo-detectors on the ceiling at a low number of six. We derive a very similar table, Table 3.3, as compared to Table 3.2, however, instead of varying the size of each object, we simply pick a standard size for each object and then vary the number of photo-detectors. We also choose the same resolution as last time. We deemed this to be sufficient because in the previous section, we had varied size and it was rather apparent that decreasing the size of the object led to better localization. In the interest of time, we focused on the easier path of analyzing the efficacy of persistent homology at detecting multiple objects in one of the the simpler cases. Also, there is another difference between this Table 3.3 and Table 3.2. When you look at the last column, the average/maximum/minimum coefficient are sets instead of singletons. Each member of the set corresponds to

the result for each object. Consequently, the size of the corresponding set for the second row of Table 3.3 has two members while the last has three.

Even though we the sample size here is rather row (there are only two rows in Table 3.3), each instance was averaged over 5 trials. The essential difference between both sets of trials is the number of objects and the distance between them. Consequently, we seek to observe the relationship of inter-object separation on distinguishing and localizing each object. We found, fortunately, in each case, all objects were distinguished. This can be readily verified by looking at the surrounding 1-chains in Figure 3.13 and Figure 3.15. However, the appearance of the 1-chains in each of the preceding diagrams is of note. In addition, the persistent bar-codes (Figure 3.12 and Figure 3.14) also suggest something revealing.

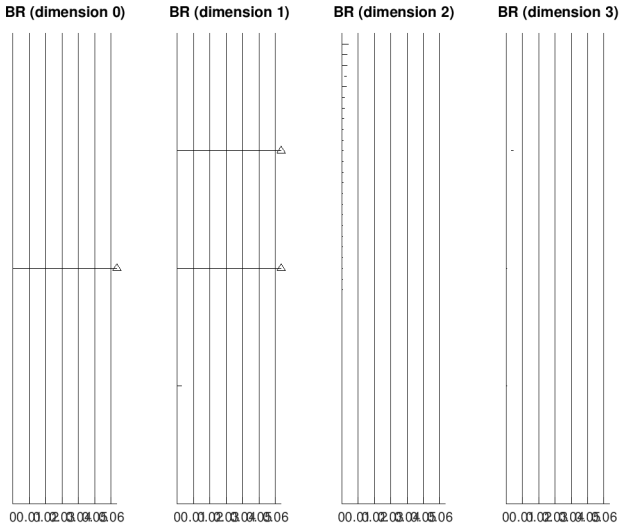


FIGURE 3.12: Persistent bar-codes for the trials of the first row of Table 3.3

The appearance of the 1-chains are generally unremarkable, save for their approximate location. However, in all of the previous cases, these chains were generally situated below the top of the cylinder. In Figure 3.15, we can see some of the chains have strange configurations, which could suggest that the underlying simplicial complex may either have points with undesirable locations or that these particular points' inclusion in the homology calculation is dominates over points

unjustly. If you refer back to the Figure 3.1, we note that there are few points near the top of cylinder (there is a faint cross near the center). Although, the 1-chains are generally correct, this observation about the formation of the point cloud is something that should be kept in mind for future work.

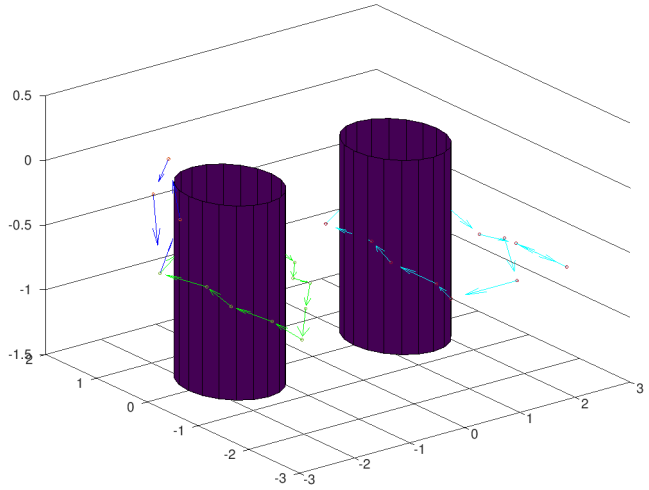


FIGURE 3.13: The 1-chains corresponding to the persistent bar-codes of Figure 3.12

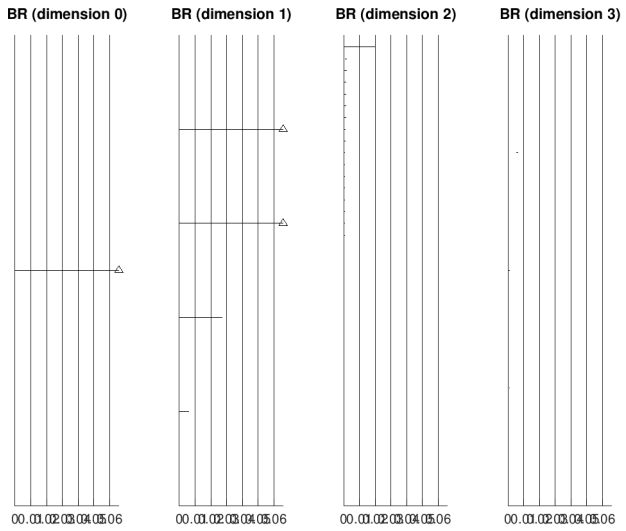


FIGURE 3.14: Persistent bar-codes for the trials of the second row of Table 3.3

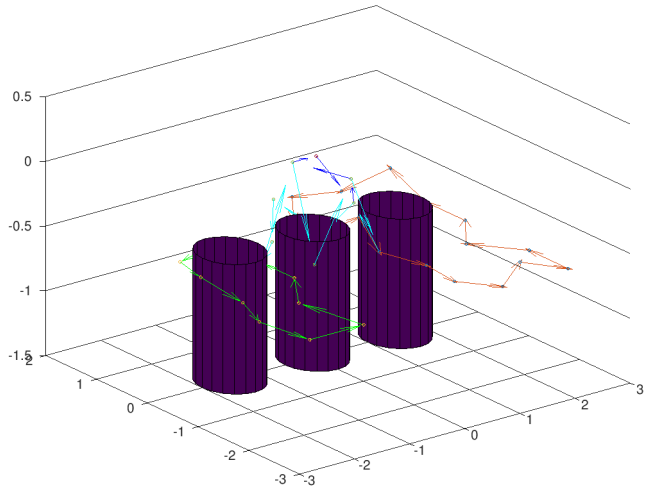


FIGURE 3.15: The 1-chains corresponding to the persistent bar-codes of Figure 3.12

3.3.1.3 Photo-Detector Number

Experiment #	**# of Objects**	**# of Photo-Detectors**	**Object Parameters**	**Average Minimum Maximum Dice Coefficient**
6	3	4	**Height** = 2m **Radius** = 0.5m **Center** = (2.5m,2.5m)	**Average** {0 0.00020482 0} **Minimum** {0 0 0} **Maximum** {0 0.00614445 0}
7	3	6	**Height** = 2m **Radius** = 0.5m **Center** = (2.5m,2.5m)	**Average** {0.045443 0.29956 0.14119} **Minimum** {0 0 0} **Maximum** {0.31937 0.88652 0.32339}
8	3	8	**Height** = 2m **Radius** = 0.5m **Center** = (2.5m,2.5m)	**Average** {0.094684 0.26289 0.11911} **Minimum** { 0 0 0} **Maximum** {0.93186 0.89957 1.0767}

TABLE 3.4: Varying Number of Photo-Detectors: All at a Transmitter Resolution 0.25m, over 3 trials

In this section, we analyze our localization algorithm's efficacy when more than four photo detectors placed upon the room ceiling. We demonstrate this with multiple objects at sufficient resolution, in similar situations as in the previous sections. Just like in Table 3.3, Table 3.4, the last column contains sets instead of single items for the same reason as discussed in the previous section. In this section, we do not display the cylinder with the accompanying 1-chains simply because we wish to only include as many figures as necessary. Since the localization accuracy, as given by the correlated low values of the dice coefficients for each respective object, is sufficiently high, we do not produce the diagrams.

What is important about the results is that there can be diminishing results

with an increasing number of photo-detectors. Whilst this may seem to be a dead-end, we note that in these experiments that we only place photo-detectors on the ceiling. In a realistic scenario (an actual office room), there would be no reason that there would not be led lights placed almost anywhere in the room. Given this observation, we should consider, in future work, placing photo-detectors in varying locations in some optimal matter. The selection of an objective function to place these photo-detectors would be an interesting venture to explore; possibly with tools from machine learning.

In each of the persistent bar-codes produced for these experiments, we do observe that in the third diagram from the left (dimension 2), there are lots of non-persistent components. This is unsurprising, given that increasing the number of photo-detectors necessarily produces more data points. As such, there is a spread of variable-density zones along certain paths in the data-set, and this necessarily produces spurious components in the persistent bar-code diagrams.

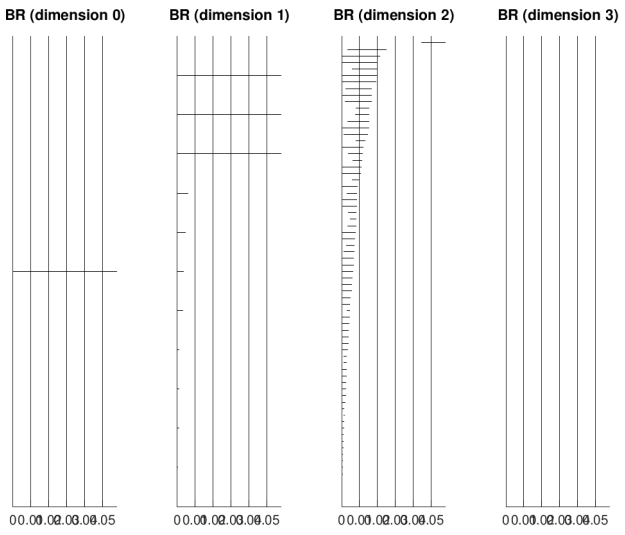


FIGURE 3.16: Persistent bar-codes for the trials of the first row of 3.4

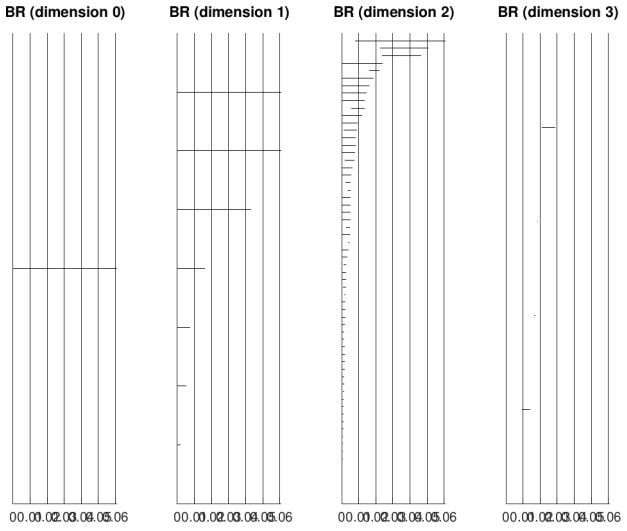


FIGURE 3.17: Persistent bar-codes for the trials of the second row of 3.4

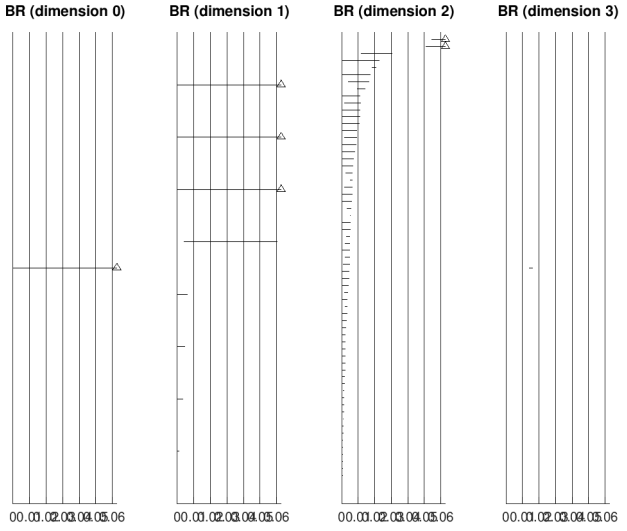


FIGURE 3.18: Persistent bar-codes for the trials of the second row of 3.4

3.3.1.4 Grid Resolution and Chain Identification

Experiment #	**# of Objects**	**Transmitter Resolution**	**Object Parameters**	**Average Minimum Maximum Dice Coefficient**
9	1	0.75m	Height = 2m Radius = 1m Center = (2.5m,2.5m)	**Average** 0.5062 **Minimum** 0.35393 **Maximum** 0.73195
10	1	0.5 m	Height = 2 m Radius = 1 m Center = (2.5m,2.5m)	**Average** 0.43022 **Minimum** 0 **Maximum** 0.70946
11	1	0.25 m	Height = 2 m Radius = 1 Center = (2.5m,2.5m)	**Average** NA **Minimum** NA **Maximum** NA
12	1	0.1 m	Height = 2 m Radius = 1 m Center = (2.5m,2.5m)	**Averge** NA **Minimum** NA **Maximum** NA
13	1	0.05 m	Height = 2 m Radius = 1 m Center = (2.5m,2.5m)	**Average** NA **Minimum** NA **Maximum** NA

TABLE 3.5: Various Transmitter Resolutions

In this section, we consider the effect of changing the resolution from very coarse values to fine values. It is possible, without actually looking at the number of chains drawn out in their geometric form, to determine that the chosen transmitter resolution is appropriate for discovering out anything of value in the room. We have compiled the persistent bar-codes of each scenario in the rows of

Table 3.5. What becomes clear as one goes through the Figures 3.19, 3.20, 3.21, 3.22 and 3.23 is that increasing the resolution has a dramatic effect upon the number of chains discovered and each chain's persistence. In an earlier part of this document, we had commented about how the simplicies "clump" together under different resolutions, and these bar-codes show the evolution of the simplicial complexes over different resolutions. These set of diagrams show compelling evidence that the simplicial complex is sensitive to the "clumpiness" of the data and the practitioner needs to be acutely aware of this phenomenon.

We also include some results that give the positioning error of the 1-chain. We compute this by comparing the centroid of the 1-chain against the centroid of the cylinder. Due to some numerical/computational issues, we report these results for a particular set of resolutions.

Experiment #	**# of Objects**	**Transmitter Resolution**	**Object Parameters**	**Inter-centroid distance**
14	1	0.75m	Height = 2m Radius = 1m Center = (2.5m,2.5m)	0.46470 m
15	1	0.1m	Height = 2m Radius = 1m Center = (2.5m,2.5m)	0.48861 m
16	1	0.15m	Height = 2m Radius = 1m Center = (2.5m,2.5m)	0.15374 m
17	1	0.175m	Height = 2m Radius = 1m Center = (2.5m,2.5m)	0.20327 m

TABLE 3.6: Transmitter Resolutions versus inter-centroid distance.

Fortunately, this small result in Table 3.6, though lacking in some regards, does illustrate concretely that increasing the resolution is associated with better localization of the object in question. As the the resolution drops, there is accordingly also a corresponding drop in inter-centroid distance as would be expected.

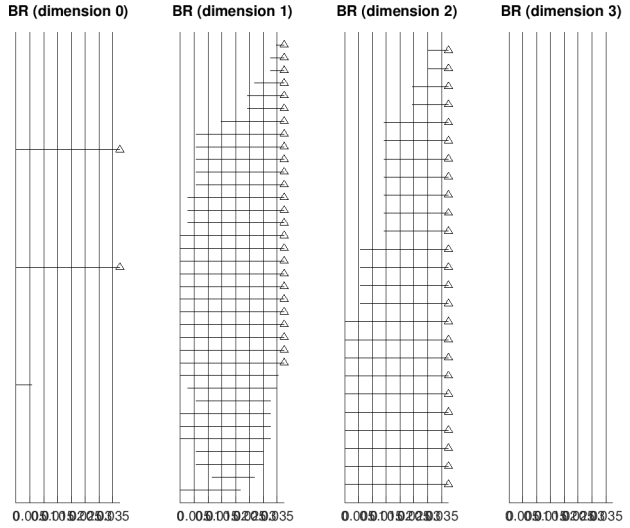


FIGURE 3.19: Persistent bar-codes for the trials of the first row of 3.5

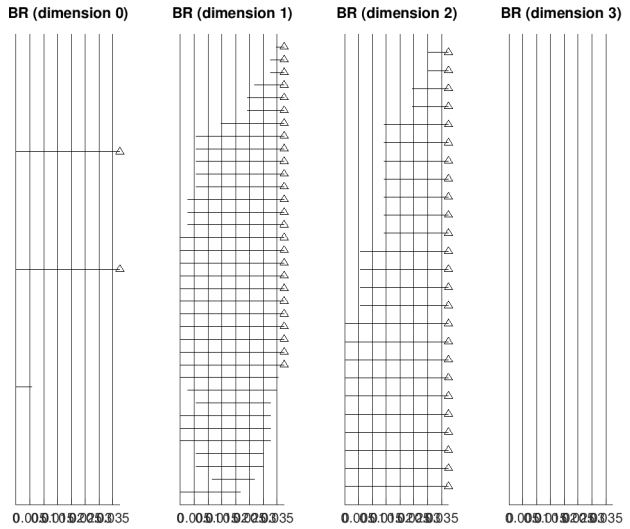


FIGURE 3.20: Persistent bar-codes for the trials of the second row of 3.5

3.3.2 Discussion

What should be clear from these graphs and results is that it is possible to perform localization with the aid of persistent homology. Since it is capable of detecting equivalence classes of cycles, the analysis entailed only requires a sufficiently informative rendering of the phenomenon in question. However, as demonstrated by the Dice coefficient, there are some issues with the cycles detected. Since we

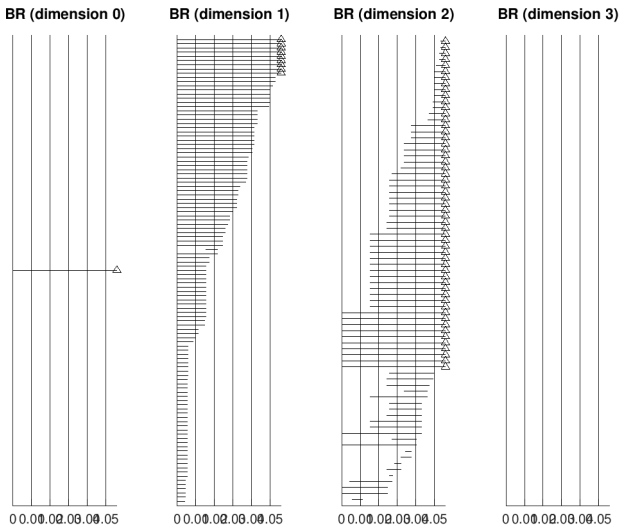


FIGURE 3.21: Persistent bar-codes for the trials of the third row of 3.5

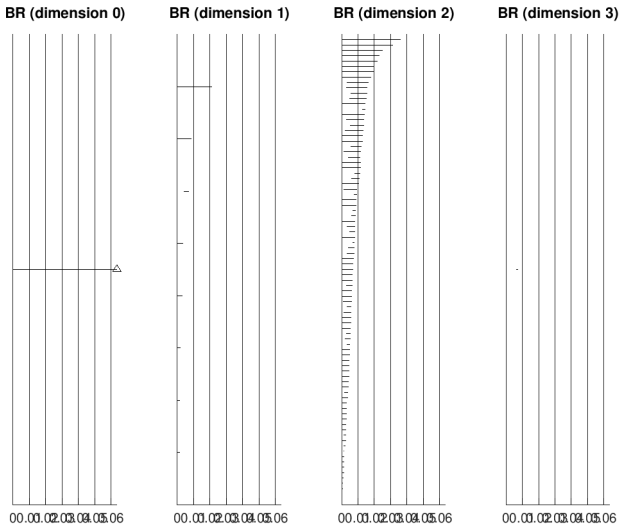


FIGURE 3.22: Persistent bar-codes for the trials of the fourth row of 3.5

are finding equivalence classes of cycles, it is rather difficult to guarantee that we get the tightest cycle possible to cover each "hole" (each "hole" here is a representation of the passively detected object). To recall, each hole in the experiment is supposed to represent the presence of an object. There are two main strategies that immediately come to mind for addressing the tightness of the retrieved contours. First, we can increase the number of samples, either by increasing the

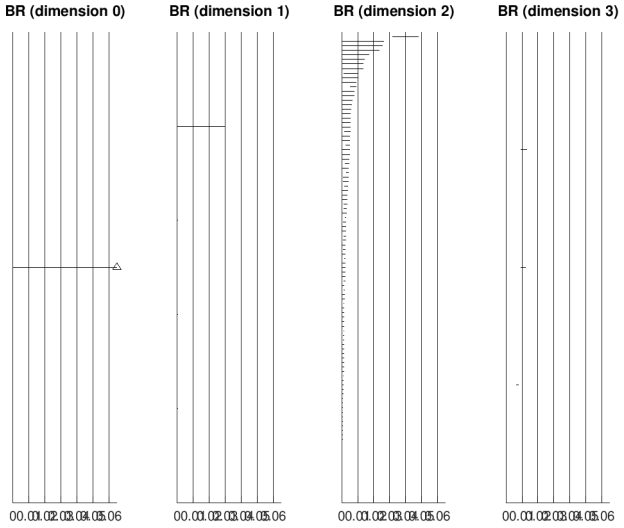


FIGURE 3.23: Persistent bar-codes for the trials of the fourth row of Figure 3.5

transmitter resolution, by increasing the ray sampling resolution or thirdly by increasing the number of photo-detectors. Although, the computational cost of each these options can be high, at the very least, the last of these options is something that can be achieved in an actual VLC scenario.

Chapter 4

Future Work & Conclusion

4.1 Concluding Remarks

In this thesis, we have studied the utility and usefulness of persistent homology for localization problems in visible light communications networks. This problem is important as localization is an important building block for "higher level analysis" that could be considered. We also show that a naive application of persistent homology, though useful, needs to be extended/improved upon to provide the best localization results. We envision that our work forms a sturdy scaffold for future researchers to build upon.

4.2 Future Work

Given our results and the subsequent discussion sections, we envision that it may be possible to refine our methods for more than just object localization/positioning. We could consider how to perform object tracking or how to build a model for determining an object's shape that we detect. These efforts would be useful

for more involved projects based in Visible Light Communications. If it were possible to determine the likely shape of an object bathed in the light rays of a VLC network, then there could be potential security applications.

Bibliography

[1] Hamid Hosseinianfar, Mohammad Noshad, and Maite Brandt-Pearce. Positioning for visible light communication system exploiting multipath reflections. *2017 IEEE International Conference on Communications (ICC)*, May 2017. doi: 10.1109/icc.2017.7996815. URL http://dx.doi.org/10.1109/icc.2017.7996815.

[2] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Proceedings 41st Annual Symposium on Foundations of Computer Science*. doi: 10.1109/sfcs.2000.892133. URL http://dx.doi.org/10.1109/sfcs.2000.892133.

[3] Hamid Hosseinianfar and Maite Brandt-Pearce. Analysis of Indoor Uplink Optical Communication Positioning System Exploiting Multipath Reflections. *arXiv preprint arXiv:1804.09360*, 2018.

[4] Ching-Sheng Wang, Chun-Hong Huang, Yong-Si Chen, and Li-Jie Zheng. An implementation of positioning system in indoor environment based on active RFID. *2009 Joint Conferences on Pervasive Computing (JCPC)*, Dec 2009. doi: 10.1109/jcpc.2009.5420212. URL http://dx.doi.org/10.1109/jcpc.2009.5420212.

[5] Junyang Zhou, K.M. Chu, and J.K. Ng. Providing Location Services within a Radio Cellular Network Using Ellipse Propagation Model. *19th International Conference on Advanced Information Networking and Applications*

(AINA'05) Volume 1 (AINA papers), 2005. doi: 10.1109/aina.2005.286. URL http://dx.doi.org/10.1109/aina.2005.286.

[6] Le Thanh Son and Po Orten. Enhancing Accuracy Performance of Bluetooth positioning. *2007 IEEE Wireless Communications and Networking Conference*, 2007. doi: 10.1109/wcnc.2007.506. URL http://dx.doi.org/10.1109/wcnc.2007.506.

[7] Yuhong Liu and Yaokuan Wang. A novel positioning method for WLAN based on propagation modeling. *2010 IEEE International Conference on Progress in Informatics and Computing*, Dec 2010. doi: 10.1109/pic.2010.5687594. URL http://dx.doi.org/10.1109/pic.2010.5687594.

[8] Koichi Kitamura and Yukitoshi Sanada. Experimental Examination of a UWB Positioning System with High Speed Comparators. *2007 IEEE International Conference on Ultra-Wideband*, Sep 2007. doi: 10.1109/icuwb.2007.4381077. URL http://dx.doi.org/10.1109/icuwb.2007.4381077.

[9] W. Zhang and M. Kavehrad. Comparison of VLC-based indoor positioning techniques. *Broadband Access Communication Technologies VII*, Jan 2013. doi: 10.1117/12.2001569. URL http://dx.doi.org/10.1117/12.2001569.

[10] James R Munkres. *Elements of algebraic topology*, volume 2. Addison-Wesley Menlo Park, 1984.

[11] Allen Hatcher. Algebraic topology. *Cambridge, UK*, 2002.

[12] Herbert Edelsbrunner and John Harer. Persistent homology-a survey,. *Contemporary mathematics*, 453:257–282, 2008.

[13] Robert W Ghrist. *Elementary applied topology*, volume 1. Createspace Seattle, 2014.

[14] L. Vietoris. Über den höheren zusammenhang kompakter räume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen*, 97 (1):454–472, Dec 1927. ISSN 1432-1807. doi: 10.1007/BF01447877. URL https://doi.org/10.1007/BF01447877.

[15] Barbara Di Fabio and Claudia Landi. A Mayer-Vietoris formula for persistent homology with an application to shape recognition in the presence of occlusions. *Foundations of Computational Mathematics*, 11(5):499, 2011.

[16] Herbert Edelsbrunner and Dmitriy Morozov. Persistent Homology: Theory and Practice. *European Congress of Mathematics, Kraków, 2 – 7 July, 2012*, page 31–50, 2012. doi: 10.4171/120-1/3. URL http://dx.doi.org/10.4171/120-1/3.

[17] Paul Bruillard, Kathleen Nowak, and Emilie Purvine. Anomaly Detection Using Persistent Homology. *2016 Cybersecurity Symposium (CYBERSEC)*, Apr 2016. doi: 10.1109/cybersec.2016.009. URL http://dx.doi.org/10.1109/cybersec.2016.009.

[18] Herbert Edelsbrunner. Persistent Homology in Image Processing. *Lecture Notes in Computer Science*, page 182–183, 2013. ISSN 1611-3349. doi: 10.1007/978-3-642-38221-5_19. URL http://dx.doi.org/10.1007/978-3-642-38221-5_19.

[19] Vin De Silva and Gunnar E Carlsson. Topological estimation using witness complexes. *SPBG*, 4:157–166, 2004.

[20] Afra Zomorodian and Gunnar Carlsson. Localized homology. *Computational Geometry*, 41(3):126–148, 2008.

[21] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[22] Andrew Tausz, Mikael Vejdemo-Johansson, and Henry Adams. Javaplex: A research software package for persistent (co) homology. *Software available at http://code. google. com/javaplex*, 2, 2011.

[23] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.