Wiffract: A New Foundation for RF Imaging via Edge Tracing

Anurag Pallaprolu
apallaprolu@ucsb.edu
UC Santa Barbara
Santa Barbara, USA

Belal Korany
belalkorany@ece.ucsb.edu
UC Santa Barbara
Santa Barbara, USA

Yasamin Mostofi
ymostofi@ece.ucsb.edu
UC Santa Barbara
Santa Barbara, USA

ABSTRACT

In this paper, we are interested in high-quality imaging of still objects with only received power measurements of off-the-shelf WiFi transceivers. We show that the scattered WiFi signals off of objects carry much richer information about the edges of the objects than the surface points. Based on this observation, we then propose a completely different way of thinking about this imaging problem. More specifically, we propose **Wiffract**, a new foundation for imaging objects via edge tracing. Our approach uses the Geometrical Theory of Diffraction (GTD) and the corresponding Keller cones to image edges of the object. We extensively validate our approach with 37 experiments in three different areas, including through-wall scenarios. We take developing a **WiFi Reader** as one example application to showcase the capabilities of our proposed pipeline. More specifically, we show how our approach can successfully image several alphabet-shaped objects. We further show that our approach enables WiFi to read, i.e., correctly classify the letters, with an accuracy of 86.7%. Finally, we show how our approach enables WiFi to image and read through walls, by imaging the details and further reading the letters of the word "BELIEVE" through walls. Overall, our proposed approach can open up new directions for RF imaging.

CCS CONCEPTS

• **Computer systems organization** → **Robotics**; • **Hardware** → *Sensor devices and platforms*.

KEYWORDS

WiFi Imaging, RF imaging, Edge Tracing, Keller Cones, WiFi Reader

1 INTRODUCTION

The number of wirelessly-connected devices has been growing

rapidly in recent years, making wireless signals, such as WiFi, ubiquitous. This has resulted in a considerable interest in using radio signals beyond communication, and for sensing and learning about the environment. For example, WiFi signals (one of the most widespread forms of wireless connectivity) have been utilized for person identification [19, 35], occupancy estimation [14, 20], and health monitoring [25, 39], among other applications. Overall, WiFi signals have shown promises in the applications where there is motion (e.g., body motion), since extracting information from movements is an easier task. However, imaging details of **still objects** with everyday RF signals, such as WiFi power measurements, has remained a considerably challenging problem due to the lack of motion.

In general, imaging objects is important for many applications, from smart home, to structural health monitoring, to search and rescue, surveillance, and excavation, just to name a few. While cameras can be used for imaging, they fail to do so through occlusions/walls and/or in low-light conditions. As such, if we can image details of objects with cheap ubiquitous WiFi devices, it can open up new possibilities for many applications, and can be complementary to the existing sensors for imaging.

Prior attempts at imaging using WiFi signals (e.g. [13, 33, 38, 42]) mainly rely on the traditional method of back-propagation of the received WiFi measurements of a receiver array to different locations in space (e.g., beamforming), in order to establish the presence/absence of an object at these locations [32]. While this method can provide somewhat useful information at mmWave frequencies (imaging with mmWave is still challenging), it falls short of giving proper images at WiFi frequencies. Other work [24, 27] has collected WiFi training data for the purpose of training a Deep Neural Network for imaging objects. However, their applicability and/or generalizability is bounded due to the limited size of the RF training dataset. Section 2 provides a detailed survey of the state-of-the-art in imaging with RF signals.

In this paper, we are interested in high-quality imaging of still objects with only received power measurements of off-the-shelf WiFi transceivers. One major implicit assumption in traditional imaging is that the object points reflect the wireless signals to the entire receiver array. This may not be valid at lower frequencies (e.g. WiFi) since most surface points can appear quasi-specular and not diffuse, contributing to poor imaging quality. **We then propose a completely different way of thinking about this imaging problem**, which shall enable new possibilities in this area. More specifically, we propose **Wiffract**: a new method to image the edges of the object by utilizing the so-called *Keller cones*. When a wave is incident on an edge point, a cone of outgoing rays emerge according to the Keller's Geometrical Theory of Diffraction (GTD) [18]. We show how we can utilize these cones and their intersection with

the receiver array (i.e., the corresponding conic sections) to trace the edges.

Remark 1: A point on the surface appears as an edge to the incoming wave if the radius of its curvature is small, compared to the wavelength [30]. As such, edges (from a wave's perspective) include not only the visibly sharp points but also other points on the surface of the object that have a small curvature. In this paper, the word "edge" refers to what the wave sees as an edge.

We next discuss our main contributions in more detail:

Statement of Contributions:

- We show that the scattered WiFi signals off of objects carry much richer information about the edges of the objects than the surface points, since edges diffract the incident waves into many directions (dictated by the diffraction cone), while surface points can mainly appear quasi-specular (near mirror-like) at WiFi frequencies.

- We then propose Wiffract, a completely new foundation for imaging still objects via edge tracing. Our approach uses the Keller cones and the corresponding conic sections to infer edge angles, via building proper projection kernels. Once it identifies high-confidence edge points, it then propagates their inferred angles to the rest of the object using Bayesian information propagation.

- We extensively validate our approach with 37 experiments in three different areas, including through-wall scenarios. **We take developing a *WiFi Reader* as one example application to showcase the capabilities of our proposed pipeline since it is a considerably challenging task that was not possible before, to the best of our knowledge.** More specifically, we show how our approach can successfully image several alphabet-shaped objects (30 total experiments drawing from a pool of 18 letters). We further show that our approach enables WiFi to read, i.e., correctly classify the letters, with an accuracy of 86.7%. For classifying our imaged letters, we propose a novel shallow neural network architecture, using the Hough transform, and train it on existing image-based alphabet datasets. Moreover, we show how our approach enables WiFi to image and read through walls, by imaging the details and further reading the letters of the word "BELIEVE" through walls. To the best of our knowledge, this is the first time that commodity WiFi devices can image objects with this level of detail and further read through walls.

2 RELATED WORK

In recent years, a number of papers have tackled the problem of imaging using wireless signals to enable "seeing" objects in space. Some have utilized ultra-wideband mmWave radars for this purpose, e.g., [2, 8, 9, 28, 36]. Such a setup, however, requires the usage of specialized and expensive equipment. Even then, high-quality imaging of objects still remains challenging, even with mmWave signals.

In this paper, we are interested in imaging using low-cost commodity WiFi devices. When imaging with WiFi, most papers rely on the standard back-propagation imaging method, which utilizes the measurements on a receiver grid to form an image of an object by tracing back (beamforming) the received signals to different locations in space [13, 24, 33, 38, 42]. This technique results in low-resolution images that can be informative in terms of detecting

and localizing the object, but do not sufficiently capture any further detail, such as the object's shape or outline. Other techniques [4, 11, 17, 40] use holography/tomography concepts to generate images for cross sections of the objects, by leveraging the penetration capability of the WiFi signals. For these to work, the transmitter and the receiver need to be placed on the opposite sides of the area of interest, which may not always be feasible. Finally, some recent work has collected WiFi training data, pertaining to a number of objects, to train a deep neural network for imaging the same objects [24, 27]. However, these techniques are very specific to the exact configurations/objects they were trained with, and do not generalize well due to the limited size of the training datasets. It is worth noting that while all the aforementioned papers rely on wireless signals in the WiFi band, several use specialized equipment (e.g. USRPs [11, 13, 27] and oscilloscopes [40]) to transmit/receive the wireless signals, thereby logging clean complex baseband WiFi data, which may not be available on commodity WiFi transceivers.

Despite these recent great efforts, high-resolution imaging at WiFi frequencies remains a considerably challenging and unsolved problem. This is the main motivation for the work of this paper, which proposes a completely different way of considering this imaging problem via tracing the edges of the objects.

3 TRADITIONAL IMAGING

Consider the scenario shown in Fig. 1, where a fixed wireless transmitter (located at $\mathbf{p}_t \in \mathbb{R}^3$) emits radio signals which interact with a set of objects (located at $\mathbf{p}_o \in \Theta \subset \mathbb{R}^3$), where Θ is the set of all object locations. The signals scattered from these objects are then captured by a uniform two-dimensional RX grid. Let (x_0, y_0, z_0) denote the location of the first antenna element in the 2-D RX grid, and let Δx and Δz denote the inter-antenna spacing in the x and z directions, respectively. Each receiver element is then located at $\mathbf{p}_r \in RX = \{(x_0 + r_x\Delta x, y_0, z_0 + r_z\Delta z)|r_x \in \{0, 1, 2, \ldots, N_x - 1\}$ and $r_z \in \{0, 1, 2, \ldots, N_z - 1\}\}$, where N_x and N_z are the number of antenna elements in the x and z dimensions, respectively. By using the Born approximation [3], the complex baseband received signal at receiver point $\mathbf{p}_r$ will be:

$$R(\mathbf{p}_r) = \alpha(\mathbf{p}_t, \mathbf{p}_r)g(\mathbf{p}_t, \mathbf{p}_r)+$$
$$\sum_{\mathbf{p}_o \in \Theta} \alpha(\mathbf{p}_t, \mathbf{p}_o)g(\mathbf{p}_t, \mathbf{p}_o)\alpha(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r), \quad (1)$$

where $\alpha(\mathbf{p}_i, \mathbf{p}_j)$ is the amplitude attenuation of the wireless path from point $\mathbf{p}_i$ to point $\mathbf{p}_j$, and $g(\mathbf{p}_i, \mathbf{p}_j)$ is the corresponding Green's function given by $g(\mathbf{p}_i, \mathbf{p}_j) = e^{-j\frac{2\pi}{\lambda}\|\mathbf{p}_i-\mathbf{p}_j\|}$, λ is the wavelength, and $\|.\|$ is the norm of the argument. The first term in Eq. 1 represents the direct path from the TX to the RX, and does not carry any object information. As such, it is typically estimated and subtracted from the received signal as part of the background subtraction process, resulting in

$$\bar{R}(\mathbf{p}_r) = \sum_{\mathbf{p}_o \in \Theta} \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r), \quad (2)$$

where $\tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r) = \alpha(\mathbf{p}_t, \mathbf{p}_o)g(\mathbf{p}_t, \mathbf{p}_o)\alpha(\mathbf{p}_o, \mathbf{p}_r)$.

In traditional near-field beamforming, an image of the object is formed by backward ray tracing of Eq. 2. More specifically, to establish the presence of an object at a test location $\mathbf{p}_m \in \Psi$, where Ψ is the space to be imaged, the complex received signal is projected
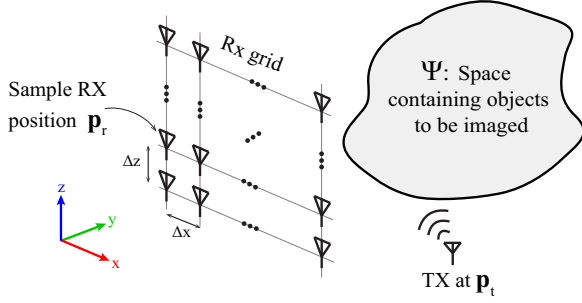
Figure 1: Our imaging scenario: a transmitter emits wireless signals, while a uniform two-dimensional receiver grid makes received power measurements in order to image objects in Ψ.

Figure 2: A surface interacts with (a) a high-frequency incident ray as a rough diffuse surface, and (b) a lower-frequency incident wave as a smooth specular surface.

on the imaging kernel $\kappa(\mathbf{p}_m, \mathbf{p}_r) = g^*(\mathbf{p}_m, \mathbf{p}_r)$, where $(.)^*$ is the conjugation operator [32]. That is,

$$\mathcal{I}(\mathbf{p}_m) = \left| \sum_{\mathbf{p}_r \in RX} \bar{R}(\mathbf{p}_r)\kappa(\mathbf{p}_m, \mathbf{p}_r) \right| = \left| \sum_{\mathbf{p}_r \in RX} \bar{R}(\mathbf{p}_r)g^*(\mathbf{p}_m, \mathbf{p}_r) \right|$$

$$= \left| \sum_{\mathbf{p}_r \in RX} \sum_{\mathbf{p}_o \in \Theta} \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r)g^*(\mathbf{p}_m, \mathbf{p}_r) \right|. \quad (3)$$

It can be seen from Eq. 3 that, if the test location $\mathbf{p}_m$ coincides with one of the object locations $\mathbf{p}_o$, i.e. $\mathbf{p}_m \in \Theta \subset \Psi$, the imaging value $\mathcal{I}(\mathbf{p}_m)$ is maximized due to the alignment of the Green's function $g(\mathbf{p}_o, \mathbf{p}_r)$ with the imaging kernel $\kappa(\mathbf{p}_m, \mathbf{p}_r)$. In other words, the image $\mathcal{I}(\mathbf{p}_m)$ can be written as follows for this case: $\mathcal{I}(\mathbf{p}_m) = |\mathcal{I}_1(\mathbf{p}_m) + \mathcal{I}_2(\mathbf{p}_m)|$, where $\mathcal{I}_1(\mathbf{p}_m) = \alpha(\mathbf{p}_t, \mathbf{p}_m) \sum_{\mathbf{p}_r \in RX} \alpha(\mathbf{p}_m, \mathbf{p}_r)$ is the imaging term whose summation terms add constructively, indicating the presence of an object at location $\mathbf{p}_m$, and

$$\mathcal{I}_2(\mathbf{p}_m) = g^*(\mathbf{p}_t, \mathbf{p}_m) \sum_{\mathbf{p}_r \in RX} \sum_{\substack{\mathbf{p}_o \in \Theta \\ \mathbf{p}_o \neq \mathbf{p}_m}} \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r)g^*(\mathbf{p}_m, \mathbf{p}_r)$$
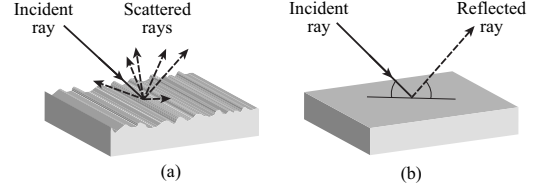
is an interference term arising from the presence of objects at other locations, whose terms are added incoherently. On the other hand, if there is no object at the test location $\mathbf{p}_m$, the $\mathcal{I}_1$ term will be zero, leaving only the $\mathcal{I}_2$ term. Thus, the fundamental principle of this imaging method is that by utilizing a large antenna array, the $\mathcal{I}_1$ term will be large when there is an object at the test location $\mathbf{p}_m$ while the $\mathcal{I}_2$ term will be small due to the addition of a large number of out-of-phase terms.

Remark 2: Given that we are interested in imaging within a few meters from the RX array and with commodity WiFi (e.g. at 5 GHz), our problem of interest is considered near-field.[1]

Shortcomings of traditional imaging: As we discussed in Sec. 2, imaging still objects with off-the-shelf transceivers has proven to be considerably challenging. We next reveal one main reason for this, which then motivates us to come up with a completely different way of thinking about this problem, resulting in the proposed approach of the next section.

It can be easily shown that the imaging approach of Eq. 3 can work well if the RX has a large size. In such a case, the $\mathcal{I}_1$ term becomes high, when there is an object at $\mathbf{p}_m$, while the $\mathcal{I}_2$ term will become negligible. However, an underlying assumption made
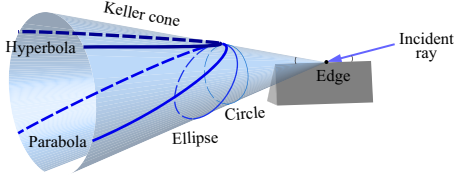
during the analysis of Eq. 3 is that an object scatters the incoming signal in all the directions, thus reaching all the receiver antennas of the array. In other words, it is assumed that $\alpha(\mathbf{p}_o, \mathbf{p}_r)$ is non-zero for all the antenna elements, when there is an object at $\mathbf{p}_o$. If, however, this assumption is not true, then the term $\mathcal{I}_1$ may not be as large as needed to result in a good imaging quality. We next examine the validity of this implicit assumption. In order for a point $\mathbf{p}_o$ on the object's surface to reflect the incident signal to several of the antenna array elements, the surface at $\mathbf{p}_o$ needs to appear diffuse (i.e., rough) to the incoming wave. While this is true at very high frequencies (such as mmWave), at lower frequencies (such as WiFi) the perturbations on the surface of an object are typically small when compared to the wavelength. As such, the same surface can appear much smoother and specular to the incoming wave, only reflecting the signal to one or a very small number of directions. Fig. 2 shows an illustration of this where the same surface acts as a mirror at lower frequencies while it scatters the signal to many directions at higher frequencies. The fact that several everyday surfaces can appear near-specular (i.e. almost mirror-like) to WiFi signals has been established in the literature [37]. As such, one cannot rely on the traditional imaging technique of Eq. 3 when imaging with commodity WiFi transceivers. This then motivates the need for a new approach, as we shall propose in the next section.

Remark 3: Commodity WiFi devices cannot provide reliable phase measurements, especially across a large antenna array, due to multiple sources of synchronization errors [41]. Hence, in this paper, we only rely on the received power measurements at the receiver array for imaging. As such, we not only have to address the aforementioned shortcoming but also have to devise an imaging system that can work with only received power measurements, as we shall see in the next section.

4 PROPOSED IDEA: IMAGING VIA EDGE TRACING

In Sec. 3, we discussed why the traditional imaging results in a poor imaging quality, when deployed with commodity WiFi transceivers, as the surfaces can appear near-specular at lower frequencies. However, edges of an object will interact with the incoming wave in a different manner, as we discuss next.

Keller Cones: When a wave is incident on an *edge* point, i.e. a point at which there is a discontinuity of the object's surface normal direction, a *cone* of outgoing rays emerge according to Keller's Geometrical Theory of Diffraction (GTD) [18]. The angle of the cone is equal to the angle between the incident ray and the edge (which is also the axis of the cone). These diffraction cones are also known in the electromagnetic theory as Keller cones and

---

[1]The near-field region is defined as a sphere of radius D^2/λ around the RX, where D is RX grid's aperture size, and λ is the wavelength [7].

Figure 3: A sample edge interaction and the resulting Keller cone. The intersection between the cone and a plane takes the shape of a circle, an ellipse, a parabola, or a hyperbola.

they have been witnessed in the visible light spectrum [29]. Fig. 3 shows a sample edge interaction and the resulting Keller cone.

Formal Definition of an Edge: As discussed in Remark 1, a surface can appear as an edge to the incoming wave if the radius of its curvature is small relative to the wavelength. More formally, according to [30], an area that does not visually look like a sharp edge can still approximately appear as one to the incoming wave if the radius of its curvature is less than half of the wavelength.

We next show how an outgoing Keller cone leaving an edge point impacts the RX array. More specifically, the RX array elements at the intersection of the RX plane and the corresponding cone are the ones that receive the signal power and thus "see" the impact of that edge point. Depending on the edge orientation, the angle of the incident wave, and the orientation of the RX array plane, this intersection will result in different 2-D shapes, e.g., hyperbola, parabola, ellipse, or circle, formally referred to as **conic sections**, as shown in Fig. 3. Figs. 4 (b)-(e) then show a few example cases of an incident wave interacting with different edges, the resulting Keller cones, as well as the resulting conic sections. As such, while the non-edge points (henceforth referred to as mirror-like points) can appear near-specular at WiFi frequencies, only reflecting the signal to one or a very small number of RX points, the edge points, on the other hand, can provide vital information for imaging since they are visible to a much larger number of RX points at the corresponding conic section.

Let $\mathcal{M}$ and $\mathcal{E}$ denote the set of all the mirror-like points and all the edge points in space respectively. The object space Θ will then be the union of these two sets. An object at location $\mathbf{p}_o$ then reflects/diffracts the incident waves to a *subset* of the receiver array elements, examples of which are shown in Fig. 4. We refer to this subset as the "RX group" of $\mathbf{p}_o$ and denote it by $RX_{\mathbf{p}_o}$, where

$$RX_{\mathbf{p}_o} = \{\mathbf{p}_r \in RX | \alpha(\mathbf{p}_o, \mathbf{p}_r) \neq 0\} \subset RX$$

Conversely, we refer to all the objects that reflect/diffract to a specific receiver point $\mathbf{p}_r$ as the "Object group" of the receiver at $\mathbf{p}_r$, and denote it by $OB_{\mathbf{p}_r}$. That is,

$$OB_{\mathbf{p}_r} = \{\mathbf{p}_o \in \Theta | \alpha(\mathbf{p}_o, \mathbf{p}_r) \neq 0\} \subset \Theta.$$

For a mirror-like point $\mu \in \mathcal{M}$, the size of the set RX_μ is very small, resulting in the received signals carrying a small amount of information about such points (see Fig. 4 (a)). However, the size of the RX group RX_ϵ of an edge point $\epsilon \in \mathcal{E}$ is typically much larger due to the resulting conic section, as shown in Fig. 4 (b)-(e). Next, we shall utilize our observations to develop Wiffract, a new imaging technique that extracts the rich information about the edges in the RX signals, and traces those edges to generate a high-quality image.

4.1 A Foundation for Edge Imaging

In this part, we lay out the foundation of our proposed approach. As motivated by the previous discussion, the edges can provide rich information about the object while the surfaces (i.e. the mirror-like points) are not that informative, especially at WiFi frequencies. **We thus propose a completely different way of thinking about this imaging problem: to image the edges of the object by utilizing the corresponding Keller cones and conic sections.**

Since commodity WiFi devices do not provide reliable phase measurements, we are interested in imaging with only received power measurements. We then start this part by developing an expression for the received signal power.

LEMMA 4.1. *The power (squared magnitude) measurements, after background subtraction[2], can be approximated by,*

$$\bar{P}(\mathbf{p}_r) \approx 2\mathcal{R}\left\{\sum_{\mathbf{p}_o \in OB_{\mathbf{p}_r}} \Lambda(\mathbf{p}_o, \mathbf{p}_t, \mathbf{p}_r) g^*(\mathbf{p}_t, \mathbf{p}_r) g(\mathbf{p}_o, \mathbf{p}_r)\right\}, \quad (4)$$

where $\Lambda(\mathbf{p}_o, \mathbf{p}_t, \mathbf{p}_r) = \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)\alpha^*(\mathbf{p}_t, \mathbf{p}_r)$ and $\mathcal{R}$ is the real part of the argument. See Appendix A for the derivation.

Let us consider the task of imaging an edge location $\mathbf{p}_m$. In order to lay the foundation for our proposed methodology, let us first assume that the RX group, $RX_{\mathbf{p}_m}$, is known for this point, i.e., the corresponding Keller cone and conic section are known. This assumption is merely to facilitate the discussion and will be relaxed shortly. We then propose to use the following imaging kernel:

$$\hat{\kappa}(\mathbf{p}_t, \mathbf{p}_m, \mathbf{p}_r) = g(\mathbf{p}_t, \mathbf{p}_r)g^*(\mathbf{p}_m, \mathbf{p}_r)\mathbb{1}_{\mathbf{p}_r \in RX_{\mathbf{p}_m}}, \quad (5)$$

where $\mathbb{1}_{\mathbf{p}_r \in RX_{\mathbf{p}_m}}$ is an indicator function that is one only if $\mathbf{p}_r \in RX_{\mathbf{p}_m}$ and is zero otherwise.

Consequently, an image at $\mathbf{p}_m$ can be reconstructed by projecting the RX power measurements on to this imaging kernel as follows,

$$\mathcal{I}(\mathbf{p}_m) = \left|\sum_{\mathbf{p}_r \in RX_{\mathbf{p}_m}} \bar{P}(\mathbf{p}_r)\hat{\kappa}(\mathbf{p}_t, \mathbf{p}_m, \mathbf{p}_r)\right|. \quad (6)$$

After some derivations, we can show that $\mathcal{I}(\mathbf{p}_m) = |\mathcal{I}_1(\mathbf{p}_m) + \mathcal{I}_2(\mathbf{p}_m)|$, where $\mathcal{I}_1(\mathbf{p}_m) = \sum_{\mathbf{p}_r \in RX_{\mathbf{p}_m}} \Lambda(\mathbf{p}_m, \mathbf{p}_t, \mathbf{p}_r)$, is the signal term, and

$$\mathcal{I}_2(\mathbf{p}_m) = \sum_{\substack{\mathbf{p}_r \in RX_{\mathbf{p}_m}}} \sum_{\substack{\mathbf{p}_{o'} \in OB_{\mathbf{p}_r} \\ \mathbf{p}_{o'} \neq \mathbf{p}_m}} \Lambda(\mathbf{p}_{o'}, \mathbf{p}_t, \mathbf{p}_r) g(\mathbf{p}_{o'}, \mathbf{p}_r) g^*(\mathbf{p}_m, \mathbf{p}_r)$$

is the interference term. In deriving this, we assume that the distance from the TX to the object space is larger than the system's resolution. The first important thing to note is that instead of using all the RX grid points, we have only utilized the RX grid points that carry information about the edge to be imaged, i.e., the RX grid points that belong to the Keller cone of the edge. This is achieved through the indicator function $\mathbb{1}_{\mathbf{p}_r \in RX_{\mathbf{p}_m}}$. It can also be seen that we have utilized $g(\mathbf{p}_t, \mathbf{p}_r)g^*(\mathbf{p}_m, \mathbf{p}_r)$ as part of our kernel $\hat{\kappa}$ in Eq. 5. The motivation for this design can be seen by examining Eq. 4, as

---

[2]Background signals are typically estimated and subtracted. This can be done through theoretical calculations based on the locations of the TX and RX [13], or more commonly through prior data collection in the absence of objects to be imaged [6, 40]. We note that since we only have received power measurements, background subtraction is based on prior power-only measurements. See Sec. 6 for more discussions on background subtraction.
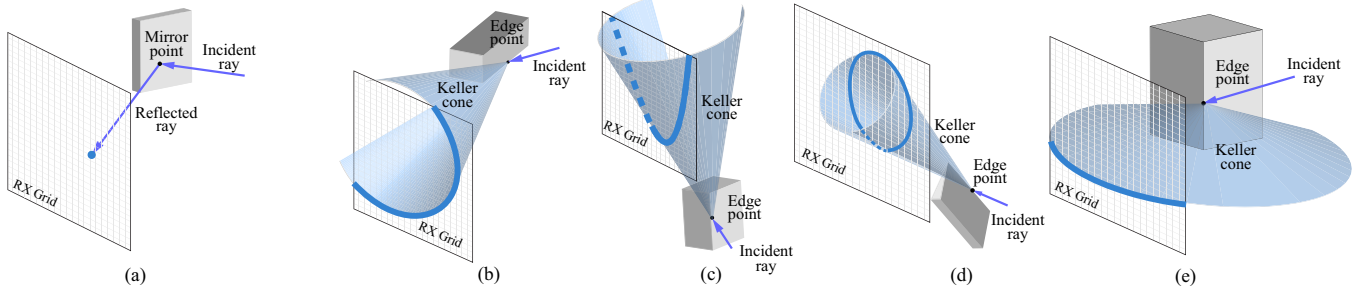
Figure 4: (a) Specularly-reflected waves off of a mirror point, and (b)-(e) diffracted waves off of edges with different orientations. The diffracted waves intersect the RX grid in a conic section.

it results in a co-phased term I_1 when there is an object at $\mathbf{p}_m$. It is worth noting that, compared to using all the RX points, using the indicator function in the kernel will increase the signal to interference ratio. More importantly, if one uses the correct conic section corresponding to the actual edge orientation at $\mathbf{p}_m$, it will result in a much stronger signal value $|I_1|$, which will give us a method to infer the edge orientation, as we shall see next.

So far, we have assumed that the RX group for the edge located at $\mathbf{p}_m$ is known in order to develop our imaging kernel. We next show how to relax this assumption. As discussed, the shape and location of the RX group depend on the location of the TX (which is known), and the orientation of the edge in space, which is unknown. Hence, for each point that we want to image, we propose to test a small set of edge orientations for imaging. Based on the value of I at the corresponding elements of the test set, we then decide if there is an edge at the corresponding point and if so, determine its orientation. Formally, let Φ denote a discrete set of uniformly-spaced angles in $[0, \pi)$, chosen based on a target angular resolution. We then construct a series of hypotheses

$$\mathcal{H}_{\phi_i} = \text{Edge at } \mathbf{p}_m \text{ makes angle } \phi_i \text{ with the +ve x-axis}, \quad (7)$$

where $\phi_i \in \Phi$.

For each edge hypothesis at the imaging location $\mathbf{p}_m$, we need to locate the RX group of an edge with the orientation ϕ_i at $\mathbf{p}_m$, denoted by $RX_{\mathbf{p}_m}(\phi_i)$. Geometrically, it is easy to show that a RX point $\mathbf{p}_r$ belongs to $RX_{\mathbf{p}_m}(\phi_i)$ if it satisfies

$$\left\langle \frac{\mathbf{p}_r - \mathbf{p}_m}{\|\mathbf{p}_r - \mathbf{p}_m\|} + \frac{\mathbf{p}_t - \mathbf{p}_m}{\|\mathbf{p}_t - \mathbf{p}_m\|}, \hat{\mathbf{e}} \right\rangle = 0, \quad (8)$$

where $\hat{\mathbf{e}}$ is a unit vector along the edge axis, and $\langle ., . \rangle$ is the dot product of the arguments. Once we have characterized the set $RX_{\mathbf{p}_m}(\phi_i)$, we have the following edge image under hypothesis $\mathcal{H}_{\phi_i}$:

$$I(\mathbf{p}_m, \mathcal{H}_{\phi_i}) = \left| \sum_{\mathbf{p}_r \in RX_{\mathbf{p}_m}(\phi_i)} \bar{P}(p_r) g(\mathbf{p}_t, \mathbf{p}_r) g^*(\mathbf{p}_m, \mathbf{p}_r) \right|. \quad (9)$$

Once all the hypotheses $\mathcal{H}_{\phi_i \in \Phi}$ are tested for the location $\mathbf{p}_m$, the most likely orientation for the edge at $\mathbf{p}_m$ (if it exists) is declared as $\phi^\star$, where

$$\phi^\star(\mathbf{p}_m) = \arg \max_{\phi_i \in \Phi} I(\mathbf{p}_m, \mathcal{H}_{\phi_i}). \quad (10)$$

We next determine if there is indeed an edge at $\mathbf{p}_m$ (whose angle

is then dictated by $\phi^\star(\mathbf{p}_m)$), by considering a scaled version of I (scaled to have a maximum of 1) as follows:

$$\bar{I}(\mathbf{p}_m) = \frac{I(\mathbf{p}_m, \mathcal{H}_{\phi^\star(\mathbf{p}_m)})}{\max_{\mathbf{p} \in \Psi} I(\mathbf{p}, \mathcal{H}_{\phi^\star(\mathbf{p})})}. \quad (11)$$

If no edge existed at location $\mathbf{p}_m$, the value of the normalized image $\bar{I}(\mathbf{p}_m)$ would be low. Hence, we declare that there is an edge at $\mathbf{p}_m$ if $\bar{I}(\mathbf{p}_m)$ exceeds a threshold I_{th}. The following set is then the set of the points in the image space for which we declare an edge, with high confidence:

$$\mathcal{S} = \left\{ \mathbf{p}; \bar{I}(\mathbf{p}) > I_{\text{th}}, \mathbf{p} \in \Psi \right\}. \quad (12)$$

Remark 4: In this paper, we focus on imaging a 2-D plane that is parallel to the receiver array at $y = y_I$ (see Fig. 6 for an illustration).

Remark 5: Note that the set Φ can be chosen small, for computational efficiency. For instance, in our experiments of the next sections, we choose only 4 angles to find the set of high-confidence locations $\mathcal{S}$.

Until now, we have shown Wiffract as a system with a single TX (fixed at $\mathbf{p}_t$) that illuminates the space of objects Θ. We can easily extend our proposed edge imaging approach to the case of multiple transmitters. In general, having more than one TX can help by illuminating the area from different locations/perspectives. For instance, part of the object area may receive a very weak signal from one TX. Similarly, an object, illuminated by a TX, may be in a *blind region* of the RX array, by which we mean that the scattering from this object may not reach the RX array, for the given TX location. Having multiple transmitters, thus, reduces the chance of such occurrences. It is worth noting that the aforementioned scenarios are not specific to our edge imaging problem and can happen in any general imaging setup.

Next, we show how our proposed edge imaging approach can be easily extended to the case of multiple transmitters. Consider the case where T TXs are located at $\mathbf{p}_{t_k}, k = 1, 2, ..., T$. We can construct an image for each TX using our proposed approach, i.e. we compute $\bar{I}_k(\mathbf{p}_m)$ of Eq. 11 for $k = 1, 2, ..., T$, and subsequently generate a corresponding set $\mathcal{S}_k$ for the k-th transmitter. We then aggregate the high-confidence sets (Eq. 12) of all the TXs into a superset, which we denote by $\mathcal{S}_U = \bigcup_k \mathcal{S}_k$. For the overlapping $\mathcal{S}_k$ points, the highest $\bar{I}_k$ and its corresponding imaged angle will be the final value.

4.2 Bayesian Information Propagation

So far, we have shown how we can find a set of pixels/locations
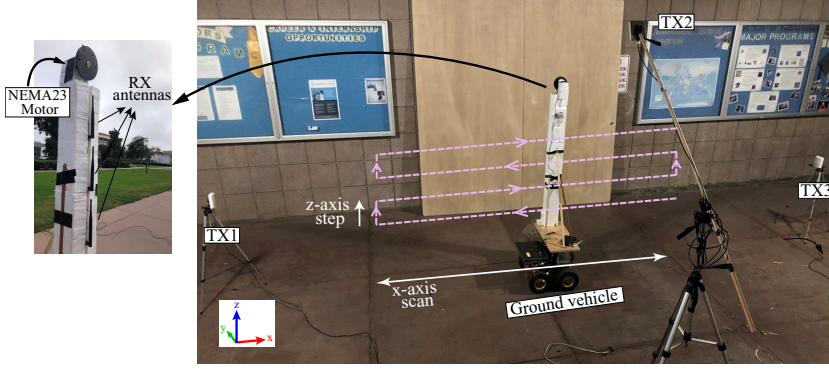
Figure 5: Sample experimental setup: 6 antennas of two laptops serve as receivers while a WiFi card of one laptop is used for transmission. A vertical structure carrying the RX antennas is mounted on a ground robot to synthesize an RX grid in the x-z plane, on which we measure WiFi CSI power measurements from three TX antennas (of one WiFi card) simultaneously.
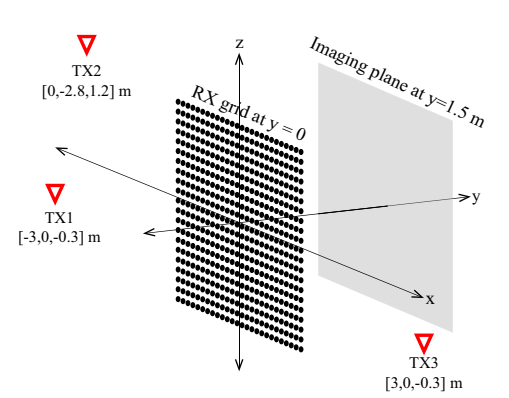
Figure 6: Relative locations of the 3 transmitters, the RX grid, and the imaging plane.

Algorithm 1: Our Proposed Edge Imaging Pipeline

Input: CSI power measurements $\bar{P}(\mathbf{p}_r)$ from T different TXs.

Output: Edge image.

1: Initialize an empty edge image.
2: **for** $k \in \{1, 2, \ldots, T\}$ **do**
3:     **for** $\phi_i \in \Phi$ **do**
4:         Generate an image $\mathcal{I}_k(\mathbf{p}_m, \mathcal{H}_{\phi_i})$, $\forall \mathbf{p}_m \in \Psi$ corresponding to the k-th TX under hypothesis $\mathcal{H}_{\phi_i}$, using Eq. 9.
5:     **end for**
6:     Generate a final normalized image corresponding to the k-th TX, $\bar{\mathcal{I}}_k(\mathbf{p}_m)$, using Eq. 11.
7:     Generate the set of high-confidence points $\mathcal{S}_k$ corresponding to the k-th TX using Eq. 12.
8: **end for**
9: Generate an aggregate set of high-confidence points $\mathcal{S}_U = \bigcup_k \mathcal{S}_k$.
10: Run the Bayesian information propagation algorithm on $\mathcal{S}_U$ and its corresponding imaged angles.
11: **for** $\mathbf{p}_m \in \Psi$ **do**
12:     **if** probability of edge at pixel $\mathbf{p}_m > p_{\min}$ **then**
13:         Add a segment to the edge image at $\mathbf{p}_m$ whose angle corresponds to the most probable state.
14:     **end if**
15: **end for**

---

in space with a relatively high confidence of having an edge and further find their edge angles. Other pixels/locations in the image space Ψ may still have edges that were undetected, for instance due to either being in a blind region, or being overpowered by other edges that are closer to the TXs. In order to deduce information about the presence of edges at the rest of the pixels in Ψ, we observe that there are local dependencies in edges of real-life objects. For instance, we examined a library of 100 random everyday objects (e.g., chairs, tables, cabinets, etc) and found that the probability that an edge to the east of a horizontal edge is horizontal, is 0.905,

while the probability that an edge to the north of a horizontal edge is horizontal, is 0.04. We can then exploit these dependencies by modeling the imaging plane as a Bayesian graph that *propagates* the edge information of the high-confidence set $\mathcal{S}_U$ to the rest of the pixels.

Consider a 3×3 pixel neighborhood where a center pixel c is surrounded by 8 neighboring pixels $n_{c_1}, n_{c_2}, \ldots, n_{c_8}$, with n_{c_1} representing the top left neighbor and the rest representing the other immediate neighbors in a clock-wise direction. The center pixel can either have no edge (i.e., it can be a mirror point or empty), or have an edge making an angle $\phi_i \in \Phi$ with the x-axis, amounting to a state space Γ of $|\Phi| + 1$ possible states. We then have the following *conditional priors*:

$$\Omega\left(\gamma_b^{[n_{c_j}]} | \gamma_a^{[c]}\right) = \begin{array}{l} \text{Prob. that neighbor } n_{c_j} \text{ has state } \gamma_b \\ \text{given center } c \text{ has state } \gamma_a, \end{array} \quad (13)$$

where "state" refers to either having no edge, or having an edge with a specific angle. Then, given the states of the high-confidence points (which are provided by our proposed edge imaging approach of Sec. 4.1), and a graph describing the direction of information flow in the image, the conditional prior Ω serves as the driver to propagate the information from the high-confidence imaged locations to the rest of the pixels. To construct the state probability vector of a high-confidence location $\mathbf{p}_s \in \mathcal{S}_U$, we define a Probability Mass Function (PMF), $\mathbb{P}(\mathbf{p}_s \to \gamma)$ of length $|\Phi| + 1$, given by

$$\mathbb{P}(\mathbf{p}_s) = \left(1 - \bar{\mathcal{I}}(\mathbf{p}_s), \underbrace{0, 0, \ldots, \bar{\mathcal{I}}(\mathbf{p}_s), \ldots, 0}_{\text{Non-zero only at } \phi^\star(\mathbf{p}_s)}\right), \quad (14)$$

where the first element denotes the probability that $\mathbf{p}_s$ has no edge, and the rest of the elements denote the probabilities that $\mathbf{p}_s$ has an edge at different angles with the x-axis. Since $\mathbf{p}_s$ is already associated with precisely one imaged angle $\phi^\star(\mathbf{p}_s)$ (Eq. 10), we set the probabilities for all the other angles to 0. Finally, to describe the direction of information flow in the image, we construct a tree-structured Bayesian graph [21, 26] with the high-confidence imaged edges as root nodes with the PMFs as described in Eq. 14. More
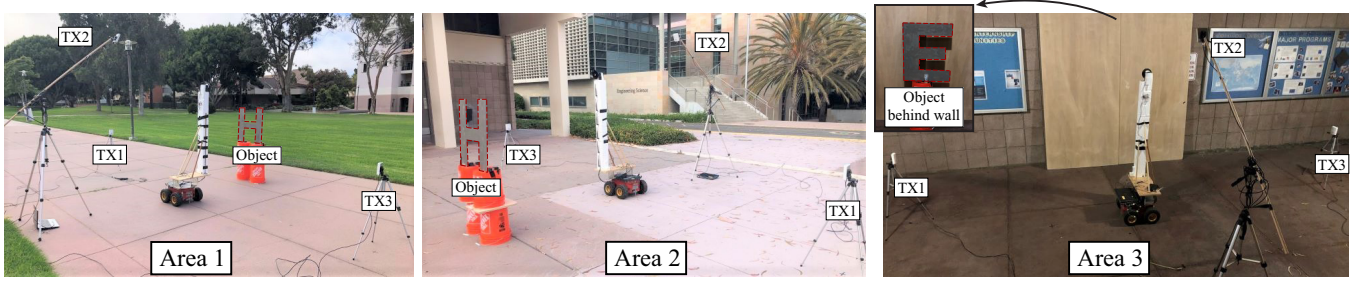
Figure 7: Experiment areas and sample objects. Outlines of objects are highlighted with dashed lines for better display.
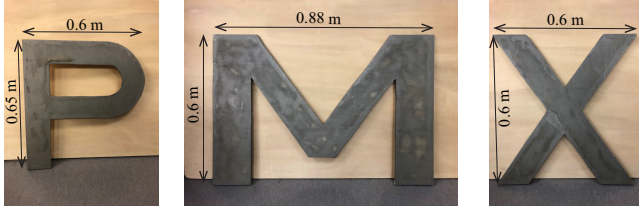

Figure 8: Sample alphabet-shaped objects used for imaging. The objects are made of painted wood, with a surface paint that has a 25 dB less shielding attenuation than aluminum.

specifically, each such node acts as a parent node and claims its 8 neighbors as children. If a neighboring pixel has already been claimed as a child of another pixel, the parent node skips it and claims the other neighbors, thus ensuring that each pixel has exactly one parent. The process then continues recursively, by having the new generation of pixels claim their own unclaimed neighbors as their children. Using the state probabilities of the roots $\mathbb{P}(\mathbf{p}_s \rightarrow \gamma)$, as well as the conditional prior Ω, the information is then propagated via message passing from the roots to the leaf nodes. In order to characterize the prior Ω, we use a library of 100 random everyday objects (e.g., chairs, tables, cabinets, and others). The final output of the Bayesian Network is a PMF for each pixel in space, describing the probabilities of this pixel over the $|\Phi| + 1$ states. If the probability of having an edge at this pixel is greater than a certain threshold $p_{\min}$, an edge is detected and its angle is declared based on the most probable angle state. In general, Bayesian information propagation has been utilized in many applications (e.g., [15, 22]). Due to space limitations, we refer the readers to [21, 26] for more details on the implementation of Bayesian graphs.

Algorithm 1 summarizes the steps of our proposed imaging pipeline.

5 EXPERIMENTAL VALIDATION

In this section, we present experimental results for Wiffract by imaging several objects in three different areas, including through-wall scenarios. We take developing a **WiFi Reader** as one example application to showcase the capabilities of our proposed pipeline since it is a considerably challenging task that was not possible before, to the best of our knowledge. More specifically, we show how our approach can successfully image several alphabet-shaped objects and further enable WiFi to read, i.e., correctly classify the letters. Finally, we use Wiffract to image the details and further read the letters of the word "BELIEVE" through walls. We next start by explaining our experimental setup.

5.1 Experimental Setup

One laptop serves as a transmitter while two laptops serve as the receivers. The laptops have Intel 5300 WiFi cards and we use the CSI-Tool [10] to measure the received power on our receiver antennas at the 5 GHz WiFi band. Each laptop has three transmission/reception ports. We then connect the ports of the TX laptop to three antennas to serve as our transmitters. The six total antennas of the two RX laptops then serve as our receiver antennas and are mounted on a ground vehicle to synthesize a RX grid. Note that since we rely only on power measurements, no synchronization is required among the RX laptops/cards. We next discuss our setup in more detail.

5.1.1 Receiver and Transmitter Details. We mount the 6 omnidirectional antennas of the two RX laptops on an unmanned ground vehicle (see Fig. 5). We place a vertical 1.5 m tall Styrofoam structure on the unmanned vehicle to hold the antennas. As the robot moves in the x direction, it synthesizes a RX grid in the x direction (see Fig. 5). In order to achieve sufficient number of rows along the z-axis, with only 6 receiver antennas, we mount a motor on top of the structure that can move the antennas up and down. More specifically, the motion in the z direction is achieved by a NEMA23 stepper motor that we remotely interface with using a Raspberry Pi. As the robot moves back or forth in the x direction, the motor takes a step in the z direction. In total, the robot moves back and forth 3.5 times in x direction and scans a two dimensional grid of points in the $x - z$ plane, as shown in Fig. 5. It synthesizes a grid of size 2 m in the x direction (spanned by 140 columns), and 1.2 m in the z direction (spanned by 42 rows, i.e. 7 rows per RX antenna) on the plane $y = 0$. On the transmitter side, the three transmission ports of the TX laptop are connected to 3 off-the-shelf half-space panel antennas acting as our transmitters. Fig. 6 shows the details of the TX and RX locations and the imaging plane.

5.1.2 Testing Areas. We carry out extensive experiments in three different areas, shown in Fig. 7. Area 1 is an open area from all four sides, while Area 2 is open from two sides, with the other sides having pillars, walls and other objects. Area 3 is a cluttered and roofed entrance of a building, which we shall use extensively for through-wall experiments.

5.1.3 Objects. We test our proposed pipeline extensively with a total of 37 experiments across all the three areas. Our objects include 18 English alphabet-shaped objects, 3 samples of which are shown in Fig. 8. The objects are made of painted wood, with a surface paint that has a 25 dB less shielding attenuation than aluminum.
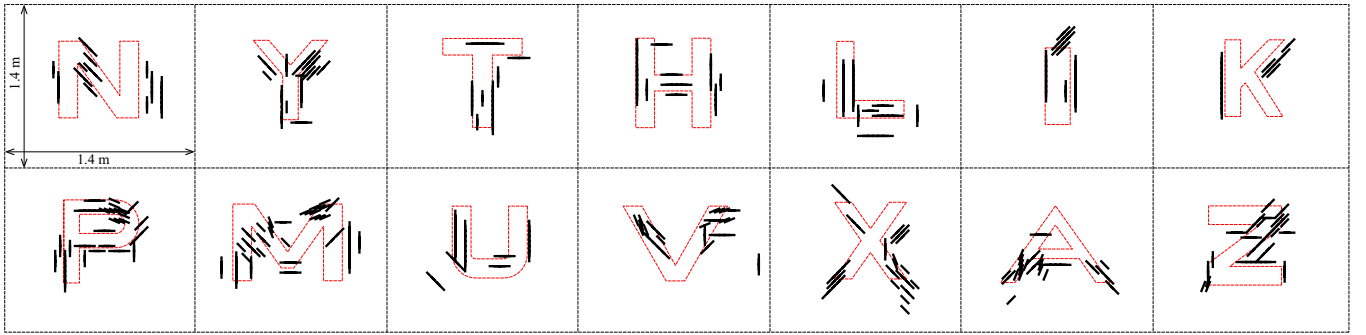
Figure 9: Edge imaging results (output of Algorithm 1) for 14 sample alphabet-shaped objects. Each imaging plane is 1.4 **m** $\times$ 1.4 **m, as marked for the first letter. Dashed-red lines represent the ground-truth location of the objects, while solid black lines represent the edge images. It can be seen that the details of the objects are imaged well.**

5.1.4 Wiffract Algorithm Parameters. The imaging plane Ψ is set at $y_I = 1.5$ m and expands from -0.7 m to 0.7 m in both x and z directions (see Fig. 6). The pixel resolution is further set to 4 cm along both x and z dimensions. In order to estimate the high-confidence locations (i.e., set $\mathcal{S}_U$), we first evaluate Eq. 9 with only 4 hypotheses $\mathcal{H}_{\phi_i}, \phi_i = 0°, 45°, 90°, 135°$ for computational efficiency. Once we find the high-confidence points, we then generate a finer-grained PMF, $\mathbb{P}(\mathbf{p}_s) \in \mathbb{R}^9$, only for $\mathbf{p}_s \in \mathcal{S}_U$, by using 8 equispaced angle hypotheses, and then propagate the corresponding information to the rest of the pixels via our Bayesian information propagation part (Sec. 4.2). When aggregating the edge images of the three transmitters, we give more weight to the TXs with stronger imaging values (I), by lowering their thresholds. In other words, we set the threshold for the k-th TX as

$$I_{\text{th}}(k) = \max\left(p_{\min}, 1 - \frac{\max_{\mathbf{p} \in \Psi} I_k(\mathbf{p}, \mathcal{H}_{\phi^\star(\mathbf{p})})}{\sum_k \max_{\mathbf{p} \in \Psi} I_k(\mathbf{p}, \mathcal{H}_{\phi^\star(\mathbf{p})})}\right),$$

where $p_{\min} = 0.6$ is the minimum threshold for declaring the presence of an edge.

5.2 Sample Imaging Results

As mentioned earlier, we take the challenging task of developing a WiFi Reader as one example application to showcase the capabilities of our proposed pipeline. More specifically, we run 30 experiments to image uppercase English letters in Areas 1 and 2. Fig. 9 shows the final edge images for 14 sample experiments. The ground truth letters are also plotted for comparison. It can be seen that several of the edges are correctly imaged, giving very good representations of their respective letters. For instance, as we shall see in the next part, passing these imaged letters to a neural network that classifies letters results in 86.7% correct classification, which indicates that the imaging process has correctly captured the details of the letters.

5.3 WiFi Can Read!

The results of Sec. 5.2 show that Wiffract can use WiFi signals to image details of alphabet-shaped objects via edge tracing. This motivates us to go a step further and identify the letters themselves based on their detected edges, **enabling WiFi signals to *read*, for the first time**.

To this end, we design a novel shallow neural network that can classify the imaged letters represented by the edge images. The goal

of this step is two-fold. First, instead of solely relying on human visual confirmation to establish the accuracy of our imaging pipeline, we can see how a simple shallow neural network can recognize and classify Wiffract's output. In other words, the neural network can provide a quantitative metric for validating the performance of our imaging pipeline. If a simple shallow network can correctly classify the output of Wiffract, it implies that Wiffract's output image contains sufficient information for classification. Furthermore, by using the last layer's activation data of our predicted class, we show how we can enhance our imaging quality by methodically suggesting edges to improve the original edge image. **We emphasize that the neural network will not have any specific knowledge of the shape, size, or style of the alphabets used in our imaging when being trained. Rather, it can be trained using any of the several existing image-based alphabet datasets (i.e., no need to collect any training data), as we shall see.**

While there exist Optical Character Recognition (OCR) systems that use a CNN-based architecture, such as LeNet [23], they are trained on images that are exclusively made of closed contours, which is not the case for Wiffract's output (e.g., see Fig. 9). As such, we show how to design a novel shallow neural network for alphabet classification. More specifically, we show that a decades-old image processing technique, called the Hough Transform [12], can establish a common ground to compare our edge images and the edges extracted from any existing contour-based alphabet dataset. We then show how to utilize the Hough Transform and train a simple Fully Connected Neural Network (FCNN) for the purpose of reading the letters.

5.3.1 Hough Domain Classifier. We use the Hough Transform [5, 12] to convert our edges in the $x - z$ plane to points in the $\rho - \eta$ domain (or Hough domain), where ρ is the perpendicular distance of the edge line from the origin (taken to be the bottom-left corner of the image), and η is the angle of the edge. A key observation here is that segments belonging to the same extended line are represented by the same point in the $\rho - \eta$ domain, implying that even a sub-segment of the actual edge is as good as the whole edge. The Hough domain representation is then at the core of our proposed classifier, which we shall describe next.

We train a 3-layer FCNN, with approximately 40,000 parameters. For our training set, we use the STEFANN font dataset [31], which has ~900 uppercase font families for each of the 26 letters of the
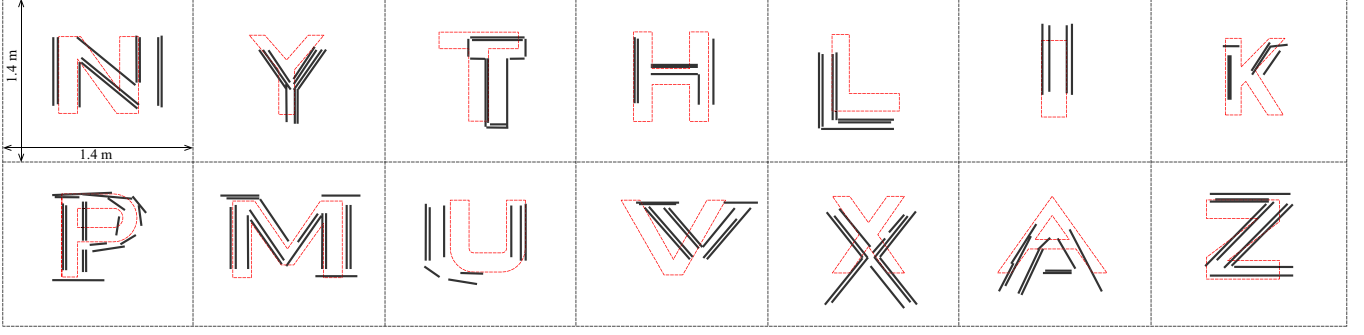
Figure 10: Our final results for imaging 14 sample letter-shaped objects with WiFi signals. The dashed lines represent the ground-truth while the solid lines represent our image.
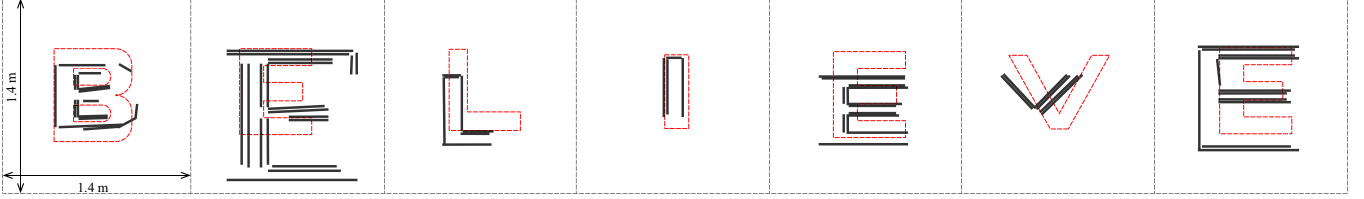
Figure 11: Through-wall reading: Wiffract enabling WiFi to image and read the letters of the word "BELIEVE" behind the wall of Area 3. The dashed lines represent the ground-truth while the solid lines represent our image.

alphabet (all contour-based). We then represent this dataset in the Hough domain, by using the Line Segment Detector (LSD) algorithm [34] to efficiently extract line segments from the fonts in the dataset, and further generate a 2-D histogram of the corresponding $\rho - \eta$ domain representation for each training data point. We scale the range of ρ linearly ($[\rho_{min}, \rho_{max}] \rightarrow [0, 1]$) to make the network invariant to shifting of the origin and we use 64 buckets for the corresponding axis of the histogram. We further set the 8 equispaced angle hypotheses as the buckets for the η axis. As such, the training dataset consists of a total of $23,478$ 2-D histograms, each of dimensions 64×8. We then translate each histogram to a vector in $\mathbb{R}^{512}$ and feed the vector, with its categorical label, to a shallow 3-layer FCNN classifier for training. We repeat the training process across different random seeds to improve the generalizability of our network. Once trained, we generate and vectorize the Hough domain histogram for Wiffract's output edge image and pass it through the network for classification.

We next show how we can use the predicted label of the classifier to further improve our imaging quality. More specifically, we first search the training dataset of the predicted class for a font that is perceived the closest to our input edge image (since the training set has many different versions of the same letter). In order to properly characterize "closeness", we utilize the recent results of [1], which showed that the activation vector of the last layer of a classifier can reliably establish similarity between two objects, i.e., the higher the correlation coefficient between two activation vectors, the higher the chance they are similar. More specifically, let $\bar{L}_{test}$ denote the activation vector of the last layer during classification. For each font f_i in the training set that belongs to the predicted class, we then compute its corresponding last layer's activation vector, $\bar{L}_{train}(f_i)$, by passing its Hough domain histogram to the trained FCNN. We then compute the cosine similarity score $\langle \bar{L}_{test}, \bar{L}_{train}(f_i) \rangle$, for each f_i, and select the one with the highest score as our target font $f^\star$ for

edge completion. In the $\rho - \eta$ domain, whenever there exists a point in the same location for both $f^\star$ and our initial edge image, we then overlay the corresponding edge of $f^\star$ on the original image.
Remark 6: Note that once the classifier has identified the letter, for instance an "H", we cannot simply use any H for our output image since there are many H characters with different fonts, shapes and sizes. Rather, we have proposed a way to improve our own edge image, by using the output of the classifier.

5.3.2 Classification Results. We next discuss our classification results in Area 1 and 2. As mentioned earlier, we have run 30 imaging experiments over these two areas for imaging alphabet-shaped objects (pooled from a total of 18 letters). Our proposed Hough Transform-based classifier was able to correctly predict the letters in 26 out of 30 experiments, resulting in an accuracy of 86.7%. It is worth noting that for a classification problem with 26 categories (the English alphabet), a random guess would have yielded a classification accuracy of only 3.8%. This confirms that Wiffract's edge images carry meaningful information for our classifier to read the alphabet-shaped objects.

We next show the final imaging results of our pipeline after improving the edges based on the classifier's predicted label and the aforementioned proposed method. More specifically, Fig. 10 shows the ground-truth letters as well as our final imaging results, for the sample letters whose edge images were shown earlier in Fig. 9. It can be seen that our proposed pipeline imaged the letter-shaped objects very well, and with an accuracy that was not seen in the literature, when imaging with WiFi, to the best of our knowledge. It is worth noting that the reason Fig. 10 demonstrates such high-quality images is that the proposed edge imaging approach (i.e., Fig. 9) generated a good representation of objects, resulting in a correct classification with a high probability. Once correctly classified, we then methodically improved the original edge image by expanding some of the edges based on the edges of the closest
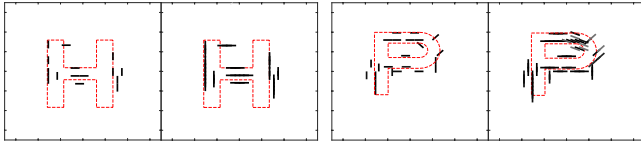
Figure 12: Impact of Bayesian information propagation – edge images (left) before and (right) after this step, for two sample letters. The right images have more connected edges.

object in the detected class, as proposed earlier. In other words, Fig. 9 has already captured crucial information on the object of interest.

Through-wall reading: We next show how our proposed approach can enable WiFi to image and further read through walls. More specifically, consider Area 3 of Fig. 7. We placed the letters of the word 'BELIEVE' behind the wall (one by one) for WiFi to read. Fig. 11 shows our final imaging results for this word. As can be seen, the word is imaged very well. Not only it is easy to identify the letters, but also the details of the letters are imaged well. Furthermore, the classifier correctly identified all the letters. Overall, Wiffract has enabled WiFi, for the first time, to read through walls.

6 DISCUSSIONS & FUTURE DIRECTIONS

In this section, we discuss several aspects related to our proposed imaging approach.

Execution time: On an Intel Core i7-3770 processor, our pipeline takes a total of 25 sec to process the CSI power measurements and generate the high-confidence locations and their corresponding imaged angles, and 0.5 sec to apply the Bayesian network algorithm for information propagation.

Impact of the Bayesian Information Propagation step: In order to isolate and better show only the impact of the Geometrical Theory of Diffraction and the proposed core imaging idea based on exploiting the resulting Keller cones, we next show the imaging results without applying the Bayesian information propagation step (step 10 of Algorithm 1). Fig. 12 shows a comparison of two edge images for two sample letters: "H" and "P". In the left image for each letter, step 10 of Algorithm 1 is not applied, i.e., high-confidence points and their corresponding angles are imaged based on the proposed method of Sec. 4.1. The right image for each letter, on the other hand, applies the additional step of Bayesian information propagation. It can be seen that if two imaged points belong to the same edge, the Bayesian network helps in connecting them, thus improving the imaging quality. However, it can be seen that the imaging results are also decent without applying this step. Our overall classification accuracy drops from 86.7% to 76.7%, without this step. The results of this part confirm that the initial step of Keller cone-based imaging i.e., exploiting the Geometrical Theory of Diffraction, is a key component of Wiffract. But given that the Bayesian network step can further improve the performance and that implementing it only adds 0.5 seconds to the computation time (which was 25 seconds prior to this step as discussed earlier), it justifies using it as part of the proposed pipeline.

Comparison to the state-of-the-art: Traditional near-field imaging utilizes backward propagation for image formation (Eq. 3) [13, 33, 38, 42]. However, as discussed in Sections 3 and 4, this technique results in a poor imaging quality at WiFi frequencies. Fig. 13 shows the results of applying traditional near-field beamforming (using
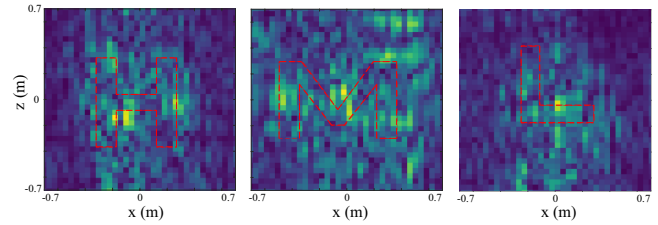
Figure 13: Sample results of the traditional imaging approach for the objects shaped as letters "H", "M", and "L". Comparing these images to Wiffract's corresponding images (Fig. 10) shows that our approach can image the objects with a much higher quality. See the color PDF for best viewing.
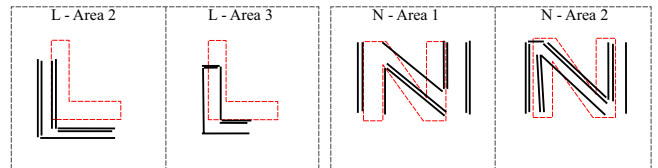

Figure 14: Final edge imaging results for the "L" and "N"-shaped objects in different environments, showing robustness to environmental changes. Red-dashed lines show ground-truth, while black solid lines show our imaging results.

power measurements) to image the letters "H", "M", and "L". It can be seen that the images are very noisy and non-informative. On the other hand, Wiffract has imaged these letters very well, as can be seen in Fig. 9 (or Fig. 10). Finally, by comparing the imaged "H" of Fig. 13 to that of Fig. 12, we can see that the quality of Wiffract images, even before applying the Bayesian information propagation step, is considerably better than that of the traditional imaging approach. This confirms that the core proposed idea of edge-based imaging using the Geometrical Theory of Diffraction can provide a new avenue for imaging still objects.

Impact of different environments: As mentioned earlier, we have tested our approach by imaging 18 uppercase letters across three different areas, where several letters have been imaged in at least two areas. While we have already extensively shown sample results in all three areas, we next compare imaging results across different areas side by side. Fig. 14 shows the results of using Wiffract to image an "L"-shaped and an "N"-shaped object in different environments. As can be seen, our approach performs robustly and consistently across different areas.

Impact of the RX grid size: In this part, we discuss the impact of the RX grid size on the imaging quality of Wiffract. Fig. 15 shows sample outputs of Wiffract when decreasing the RX grid size from 2 m x 1.4 m (original) to 1.5 m x 1 m. While we expect the imaging quality to degrade as we reduce the RX grid size, we can see that Wiffract can still generate images of decent quality. The figure also shows a more extreme case where the RX grid size is significantly reduced to 0.5 m x 0.5 m. It can be seen that the imaging result visibly degrades in this case, as expected.

In general, the size of the RX grid should impact the overall resolution of imaging. A well-known concept in the traditional near-field beamforming is that larger apertures result in better imaging resolution and this principle should also be extendable to
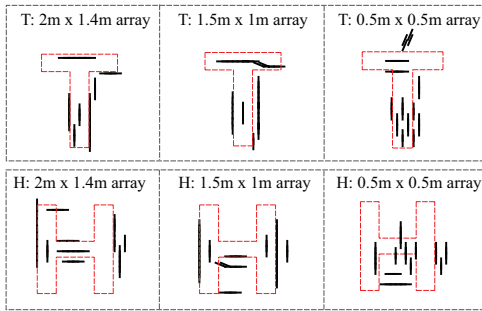
Figure 15: Impact of reducing RX grid size: 1.5m x 1m still results in decent imaging, while 0.5m x 0.5m degrades imaging as expected.


Figure 16: As part of future work, Wiffract can be used for other applications – Two sample results are shown here (output of Algorithm 1). Objects are not drawn to scale for better display.

Wiffract. However, while a closed-form expression for the resolution of traditional imaging is possible (since all the RX points are used to generate the image at all pixels), analysis along similar lines is more challenging for our proposed imaging pipeline since the RX points used for beamforming for different pixels in the imaging space are different. We leave the mathematical characterization of the resolution of Wiffract to future work. Finally, a larger aperture increases the chance of capturing the Keller cones from different edges of the object, which can help reduce the blind regions in the imaging plane, which are the regions whose Keller cones do not reach the RX grid. It is noteworthy that the depth (distance from the RX grid to the object) also affects the blind regions in a similar manner, as it is more likely to capture the Keller cones off of an object's edge if it is closer to the RX grid. Overall, a key parameter in characterizing the impact of an edge on the RX grid is its corresponding conic section as it implicitly takes into account the system parameters such as the size of the RX grid, the distance to the object, and the orientation of the edge.

Extension to imaging other objects: So far, we have showcased Wiffract's performance by imaging several alphabet-shaped objects that were detailed and complex. We had specifically chosen this application since passing the imaged results through an alphabet classifier gave us a clear quantitative metric of our imaging performance. Nevertheless, Wiffract can also be used in other scenarios, to image any other object. To motivate future work, Fig. 16 shows Wiffract's performance when imaging sample daily-life objects, such as a garden fence and a microwave oven (in Area 2 of Fig. 7). It can be seen that Wiffract generated good representations of the objects by tracing their dominant edges. For instance, it could capture the fence's parallel lines even though they are close to each other. Interestingly, for the microwave oven, it could image the inner rectangle, something that would not have been possible with other techniques. As part of future work, one can use Wiffract to image other objects and further pass the imaged results to a classifier for object classification, similar to what we have shown for the letters. Wiffract can, in particular, be useful for detecting cracks and other structural damages, among other applications.

Validation of the underlying Geometrical Theory of Diffraction: In order to directly see the interaction of an incoming wave with an edge and confirm the resulting Keller cones, we carry out a standalone experiment on a single aluminum vertical edge of length 30 cm in Area 2. Fig. 17 (left) shows the schematic of the experiment, where a single TX (labeled as TX1) illuminates the
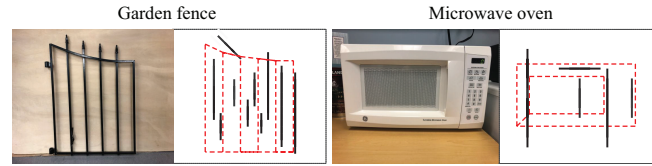
area. The figure also geometrically demonstrates the parts of the receiver grid that are expected to have a non-negligible received power predicted by the Geometrical Theory of Diffraction (i.e., based on the intersection of the corresponding Keller cones and the receiver plane). Fig. 17 (right) then shows the true power of the measured CSI on the RX grid. It can be easily seen that the parts of the RX grid that are affected by the edge are concentrated in the area predicted by the left figure and that we witness a sharp drop in the received power beyond the extents of this part. See the color PDF for a better visualization of this figure. We note that the incident rays are arriving at the edge at an almost perpendicular angle in this case.

Impact of environmental interference: As discussed in Sec. 3, other static objects in the environment do not affect the imaging quality of Wiffract, as their impact is canceled out in the background subtraction step. However, if a person (or object) moves near the transceiver while data collection is underway, there will be a non-trivial impact on our imaging system or, for that matter, on any other static imaging system. To highlight this effect, we carry out an imaging experiment of a vertical aluminum edge, while a person walks near the setup during the RX scanning process. Fig. 18 shows sample outputs for different levels of interference, quantified by the percentage of the time the person moves in the area during the experiment. For instance, 100% means that the person was constantly walking near the transceiver setup throughout the whole experiment. Despite the interference, it can be seen that Wiffract is still able to reasonably image the object. As the level of interference increases, the imaging quality degrades, as expected.

Background subtraction: A generic imaging system should be capable of imaging everything in the environment, including the background. However, imaging static objects with WiFi is considerably challenging. Thus, the radiation from the background is typically subtracted to focus on imaging a simpler scene (or a specific object). In other words, without subtracting the background, one would typically need to image a more complex scene. Moreover, there are additional factors one may need to consider when imaging in a near-field setting (which is the case in this paper), as the TX can be closer to the RX aperture. More specifically, the direct LOS path from TX to RX as well as the ground reflection path, which form the dominant terms in the background radiation, can be space-varying over the RX array and, as such, a simple DC removal would not suffice to remove their impact. One possible solution to compensate for these paths would be to use a model-based approach to estimate their impact e.g., use Friis formula for the direct LOS path and a two-ray model for the ground reflection. In principle, the parameters of such a model can be learned by making a small set of measurements between the TX and RX, placed in the same environment (or in an environment with a similar ground),
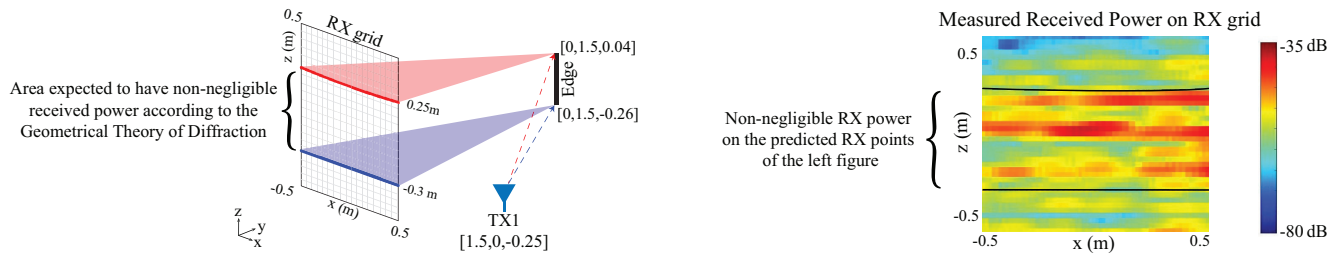
Figure 17: Validation of the underlying Geometrical Theory of Diffraction and the resulting Keller cones. (Left) A single edge is illuminated by TX1. The figure also shows the parts of the receiver grid that is expected to have a non-negligible received power predicted by the Geometrical Theory of Diffraction. (Right) The true power of the measured CSI showing that the diffracted waves primarily affect the area predicted by the left figure. See the color PDF for a better viewing of this figure.

but far from objects. A further investigation of this, in conjunction with near-field imaging, is a potential future work direction.

Impact of the neural network: In Sec. 5.3, we fed Wiffract's output to a shallow neural network classifier, in order to provide an unbiased quantitative metric for Wiffract's performance. As a byproduct, we showed how to use the last layer's activation data of our predicted class to further enhance our imaging quality by methodically suggesting edges to improve the original edge image. In Fig. 16, we showed how Wiffract can be extended to image general everyday objects as part of future work. When imaging general objects, one can generalize the neural network part accordingly. For instance, one can use a general object classifier at the end of the pipeline by taking any of the existing object classifiers and retraining it using the Hough domain representation. If the information of the application is available (for instance imaging household items), one can better tailor the neural network to that subset of objects.

Impact of surface curvature – How sharp should the edge be? As discussed in Sec. 4, an edge is any surface curvature whose radius is less than half of the wavelength [30], and is thus not limited to the visually-sharp edges. For instance, the microwave oven of Fig. 16 has round sides. This then allows our approach to be applied to a wide set of objects. As part of future work, one can further explore the relationship between the curvature of a surface and its impact on our proposed edge-based imaging method.

Applicability to other frequencies/modalities: Even when a surface texture appears diffuse to an incoming signal, one can in principle utilize our edge-based approach, in addition to traditional back propagation-based methods, in order to improve the imaging quality. More detailed investigation of the applicability of the proposed approach to other frequencies or with other modalities (e.g. acoustic signals) is part of future work. Finally, while imaging with WiFi cannot capture the depth information, the proposed approach can be used with a larger bandwidth to get the depth information and image in 3D, as part of future work.

7 CONCLUSIONS

In this paper, we showed that high-quality imaging of still objects, with only the received power measurements of commodity WiFi transceivers, is possible. More specifically, we proposed *Wiffract*, a new foundation for imaging objects via edge tracing. Our approach uses the Geometrical Theory of Diffraction and the corresponding Keller cones to image edges of the object. We extensively validated our approach with 37 experiments in three different areas, including through-wall scenarios, by imaging several alphabet-shaped objects.
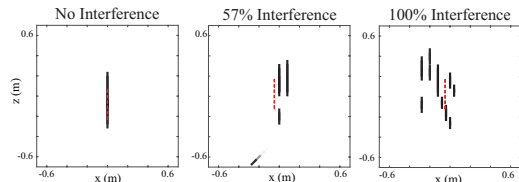
Figure 18: Impact of interference from a moving entity – imaging an edge in the presence of interference caused by a moving person. The interference level is quantified by the percentage of time the person moves near the TX/RX during the data collection process.

We further showed how our approach can enable WiFi to *read* for the first time, i.e., correctly classify the letters, with an accuracy of 86.7%. Finally, we showed how our approach enables WiFi to image and read through walls, by imaging and further reading the letters of the word "BELIEVE" through walls. Overall, our proposed approach can open up new directions for RF imaging.

ACKNOWLEDGMENTS

A DERIVATION OF LEMMA 4.1

The complex baseband received signal at RX location $\mathbf{p}_r \in RX$ is a linear superposition of the direct TX path (from $\mathbf{p}_t$), background reflections (such as those from the ground or any background objects at locations $\mathbf{p}_b \in \mathcal{B}$) and object based reflections (from $\mathbf{p}_o \in \Theta$), and can be written as $R(\mathbf{p}_r) = \alpha(\mathbf{p}_t, \mathbf{p}_r)g(\mathbf{p}_t, \mathbf{p}_r) + \sum_{\mathbf{p}_b \in \mathcal{B}} \tilde{\alpha}(\mathbf{p}_b, \mathbf{p}_r)g(\mathbf{p}_b, \mathbf{p}_r) + \sum_{\mathbf{p}_o \in OB_{\mathbf{p}_r}} \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r)$, where $\tilde{\alpha}(\mathbf{p}_b, \mathbf{p}_r) = \alpha(\mathbf{p}_t, \mathbf{p}_b)g(\mathbf{p}_t, \mathbf{p}_b)\alpha(\mathbf{p}_b, \mathbf{p}_r)$. Assuming the direct path is much stronger than all scattered paths [16], the power of the received signal at $\mathbf{p}_r$ can be written as: $P(\mathbf{p}_r) = C_o + C_{\mathbf{p}_b, \mathcal{B}} + C_{\mathbf{p}_o, OB_{\mathbf{p}_r}}$, where $C_o = |\alpha(\mathbf{p}_t, \mathbf{p}_r)|^2 + \left| \sum_{\mathbf{p}_b \in \mathcal{B}} \tilde{\alpha}(\mathbf{p}_b, \mathbf{p}_r)g(\mathbf{p}_b, \mathbf{p}_r) \right|^2 +$

$\left| \sum_{\mathbf{p}_o \in OB_{\mathbf{p}_r}} \tilde{\alpha}(\mathbf{p}_o, \mathbf{p}_r)g(\mathbf{p}_o, \mathbf{p}_r) \right|^2$, $C_{\mathbf{p}, \mathcal{P}}$ is a cross term given by $C_{\mathbf{p}, \mathcal{P}} = 2\mathcal{R}\left\{ \sum_{\mathbf{p} \in \mathcal{P}} \tilde{\alpha}(\mathbf{p}, \mathbf{p}_r)\alpha^*(\mathbf{p}_t, \mathbf{p}_r)g(\mathbf{p}, \mathbf{p}_r)g^*(\mathbf{p}_t, \mathbf{p}_r) \right\}$ and $\mathcal{R}\{.\}$ is the real part of the argument.

If power measurements are collected in the absence of the object, i.e., in the presence of only the background and the TX, we can perform a similar analysis to get: $P_{bg}(\mathbf{p}_r) = |\alpha(\mathbf{p}_t, \mathbf{p}_r)|^2 + \left| \sum_{\mathbf{p}_b \in \mathcal{B}} \tilde{\alpha}(\mathbf{p}_b, \mathbf{p}_r)g(\mathbf{p}_b, \mathbf{p}_r) \right|^2 + C_{\mathbf{p}_b, \mathcal{B}}$. By subtracting the background measurements from $P(\mathbf{p}_r)$, we then get Lemma 4.1.   □

REFERENCES

[1] H. Cai and Y. Mostofi. Exploiting object similarity for robotic visual recognition. *IEEE Transactions on Robotics*, 37(1):16–33, 2020.

[2] G. Charvat, A. Temme, M. Feigin, and R. Raskar. Time-of-flight microwave camera. *Scientific reports*, 5(1):1–6, 2015.

[3] W. C. Chew. *Waves and fields in inhomogenous media*, volume 16. Wiley-IEEE Press, 1995.

[4] S. Depatla, L. Buckland, and Y. Mostofi. X-ray vision with only WiFi power measurements using Rytov wave models. *IEEE Transactions on Vehicular Technology*, 64(4):1376–1387, 2015.

[5] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.

[6] A. Edelstein and M. Rabbat. Background subtraction for online calibration of baseline RSS in RF sensing networks. *IEEE Transactions on Mobile Computing*, 12(12):2386–2398, 2012.

[7] J. W. Goodman. Introduction to Fourier Optics. *Roberts & Co Publishers, Englewood, Colorado*, 2005.

[8] J. Guan, S. Madani, S. Jog, S. Gupta, and H. Hassanieh. Through fog high-resolution imaging using millimeter wave radar. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11464–11473, 2020.

[9] J. Guan, A. Paidimarri, A. Valdes-Garcia, and B. Sadhu. 3D imaging using mmWave 5G signals. In *2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pages 147–150. IEEE, 2020.

[10] D. Halperin, W. Hu, A. Sheth, and D. Wetherall. Tool release: Gathering 802.11n traces with channel state information. *ACM SIGCOMM CCR*, 41(1):53, Jan. 2011.

[11] P. M. Holl and F. Reinhard. Holography of WiFi radiation. *Physical review letters*, 118(18):183901, 2017.

[12] P. V. Hough. Method and means for recognizing complex patterns, Dec. 18 1962. US Patent 3,069,654.

[13] D. Huang, R. Nandakumar, and S. Gollakota. Feasibility and limits of WiFi imaging. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 266–279, 2014.

[14] O. T. Ibrahim, W. Gomaa, and M. Youssef. CrossCount: A Deep learning system for device-free human counting using WiFi. *IEEE Sensors Journal*, 19(21):9921–9928, 2019.

[15] D. Janakiram, A. Kumar, et al. Outlier detection in wireless sensor networks using Bayesian belief networks. In *2006 1st International Conference on Communication Systems Software & Middleware*, pages 1–6. IEEE, 2006.

[16] C. R. Karanam, B. Korany, and Y. Mostofi. Magnitude-based angle-of-arrival estimation, localization, and target tracking. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 254–265. IEEE, 2018.

[17] C. R. Karanam and Y. Mostofi. 3D through-wall imaging with unmanned aerial vehicles using WiFi. In *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 131–142. IEEE, 2017.

[18] J. B. Keller. Geometrical theory of diffraction. *Journal of the Optical Society of America*, 52(2):116–130, 1962.

[19] B. Korany, C. R. Karanam, H. Cai, and Y. Mostofi. XModal-ID: Using WiFi for through-wall person identification from candidate video footage. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2019.

[20] B. Korany and Y. Mostofi. Counting a stationary crowd using off-the-shelf WiFi. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 202–214, 2021.

[21] M. L. Krieg. A tutorial on Bayesian belief networks. *Defence Science and Technology Organization (Salisbury, Australia)*, 2001.

[22] F. R. Kschischang and B. J. Frey. Iterative decoding of compound codes by probability propagation in graphical models. *IEEE Journal on Selected Areas in Communications*, 16(2):219–230, 1998.

[23] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[24] C. Li, Z. Liu, Y. Yao, Z. Cao, M. Zhang, and Y. Liu. WiFi see it all: generative adversarial network-augmented versatile WiFi imaging. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 436–448, 2020.

[25] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo. Contactless respiration monitoring via off-the-shelf WiFi devices. *IEEE Transactions on Mobile Computing*, 15(10):2466–2479, 2015.

[26] J. Pearl. Fusion, propagation, and structuring in belief networks. *Artificial intelligence*, 29(3):241–288, 1986.

[27] P. C. Proffitt and H. Wang. Static object WiFi imaging and classifier. In *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, pages 1–7. IEEE, 2018.

[28] K. Qian, Z. He, and X. Zhang. 3D point cloud generation with millimeter-wave radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–23, 2020.

[29] Y. Rahmat-Samfi. Keller's cone encountered at a hotel. *IEEE Antennas and Propagation Magazine*, 49(6):88–89, 2007.

[30] R. Ross and M. Hamid. Scattering by a wedge with rounded edge. *IEEE Transactions on Antennas and Propagation*, 19(4):507–516, 1971.

[31] P. Roy, S. Bhattacharya, S. Ghosh, and U. Pal. STEFANN: scene text editor using font adaptive neural network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13228–13237, 2020.

[32] J. G. Ryan and R. A. Goubran. Near-field beamforming for microphone arrays. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 363–366. IEEE, 1997.

[33] S. Vakalis, L. Gong, and J. A. Nanzer. Imaging with WiFi. *IEEE Access*, 7:28616–28624, 2019.

[34] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. LSD: a line segment detector. *Image Processing On Line*, 2:35–55, 2012.

[35] W. Wang, A. X. Liu, and M. Shahzad. Gait recognition using WiFi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 363–373, 2016.

[36] M. E. Yanik and M. Torlak. Near-field 2-D SAR imaging by millimeter-wave radar for concealed item detection. In *2019 IEEE radio and Wireless Symposium (RWS)*, pages 1–4. IEEE, 2019.

[37] F. Zhang, C. Wu, B. Wang, and K. R. Liu. mmEye: Super-Resolution Millimeter Wave Imaging. *IEEE Internet of Things Journal*, 8(8):6995–7008, 2020.

[38] S. Zhang, G. Liu, Y. Keping, and S. Wan. WiImage: Fine-grained Human Imaging from Commodity WiFi. 2021.

[39] M. Zhao, S. Yue, D. Katabi, T. S. Jaakkola, and M. T. Bianchi. Learning sleep stages from radio signals: A conditional adversarial architecture. In *International Conference on Machine Learning*, pages 4100–4109. PMLR, 2017.

[40] W. Zhong, K. He, and L. Li. Through-the-wall imaging exploiting 2.4 GHz commodity WiFi. *arXiv preprint arXiv:1903.03895*, 2019.

[41] Y. Zhuo, H. Zhu, and H. Xue. Identifying a new non-linear CSI phase measurement error with commodity WiFi devices. In *2016 IEEE 22nd International Conference on Parallel and Distributed Systems (ICPADS)*, pages 72–79. IEEE, 2016.

[42] Y. Zi, W. Xi, L. Zhu, F. Yu, K. Zhao, and Z. Wang. WiFi imaging based segmentation and recognition of continuous activity. In *International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pages 623–641. Springer, 2019.