ISM Report: Line of sight object Detection using RF Signals

Pratyush Kumar
Ashoka University
Submitted to: Prof. Debayan Gupta

Abstract— **In this report, I discuss the steps taken by our group to build a system capable of detecting object places in line-of-sight (LOS) of a transmitter and receiver. The project involved interfacing with a Software-defined Radio (SDR) to collect data in the RF signals in the form of In-phase and quadrature, and training a Machine Learning/ Deep Learning model to build a binary classifier capable of predicting if an object is present in the Line-of-Sight between the transmitter and receiver.**

I. Introduction and Related work

Radio Waves form the basis of Radar systems widely in air crafts and other industrial applications to detect objects and avoid obstructions. However, the use of powerful systems is limited to the government. Radar uses signals that have can frequencies in the order of hundreds of Gigahertz. Such high frequency is harmful in short range and around humans. Researchers have been successful in using Radio Waves falling in the L-Band to detect objects, humans, movement, and gestures. For this, they have used setup that either resembles a bi-static radar or a Passive Radar. [1], [2], [3], [4], and [5] among others use RF signals to detect objects in indoor and through-wall setups.

However, all of these approaches depend heavily on sophisticated signal processing techniques which requires for all the processes from data collection to the detection process to be hard-coded using formulas.

In this, first, phase project, we aim to build a system that does away with all the signal processing steps required for detection and replace it with Machine Learning techniques. Our goal is to identify if it is possible to approximate the complex physical processes involved in signal processing with Neural Networks, Support Vector Machines and other common function approximation methods.

Following this, we will move towards detecting multiple objects using RTL-SDR. The third phase of this project is to interface with Wi-fi signals in the environment and test out the feasibility of object, human and activity recognition using these.

II. Procedure/ Research Process

Initial process involved using interfacing with the RTL-SDR[6] to understand how it works. For this, each of the team members used an off-the-shelf SDR software like GNU Radio[7], Cubic SDR[8], GQRX[9].
Following this, Raspberry Pi[10] was used as a transmitter. In order to make a transmitter, our teammate Yash More installed **rpitx**[11] and used a male to female cable in the pin 18 of Raspberry Pi to transmit at a maximum of 433 MHz. The transmission frequency is limited due to hardware limitations of Raspberry Pi which requires it to transmit at a maximum of 433 Mega Hertz.

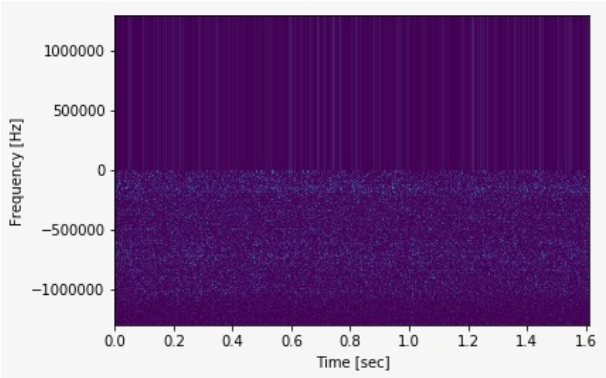
Initial research and brainstorming led to the conclusion that the setup can be either in the form of a passive or an active radar. Finally, we settled on using the active radar setup since we had a transmitter, but my experiments with simply using the receiver to detect presence of object in very close range to the antenna ($\geq$ 50 cm) showed positive results. The environment used for this was fairly noise-free and the object used were bigger in size when compared to a water bottle.

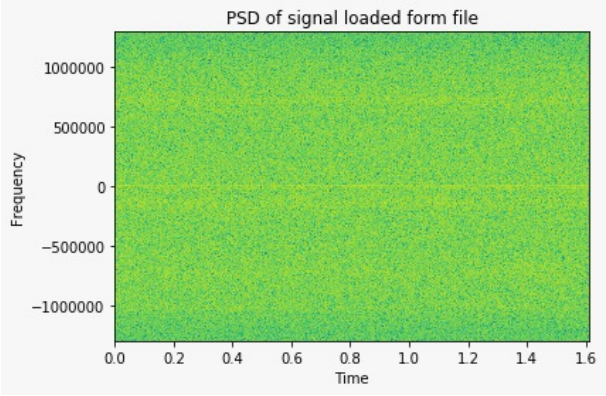
Following this, Aishwarya Praveen proposed MATLAB[12] to be a useful IDE for our whole process as it consisted all of the required libraries for collecting RF Signal data directly from the RTL-SDR, and making Spectrograms out of In-phase and Quadrature (I/Q) data. Attempts to use MATLAB for this process failed since while testing the whole setup along with teammates Aditya Singh and Aishwarya Praveen, it became apparent that MATLAB could not handle large sample sizes and was also unable to take multiple samples continuously.

Finally, PyRTL-SDR [13], a python library capable of successfully interfacing with RTL-SDR to collect I/Q data and generate Power-Density plots and Spectrograms was used. Spectrogams are used as they are capable of displaying any change in energy in the signal over time. This would allow us to identify if there is any change in power of the signal from when there is no object places in between the transmitter and the receiver and when there was a 1 litre plastic water bottle was places in between them.

Code was written to collect data in the form of I/Q data in python and store it in the form of Comma-separated values. A fairly Noise-proof room was selected for data collection. Transmitter and Receiver (RTL-SDR's and antenna) were kept 1.5 meters apart. Sampling rate was set to 2.6 M/S, i.e., 2.6 million samples per second, and the center frequency to 433 MHz. Each reading was kept to be 5 seconds long. A total of 10000 samples of raw I/Q data for without object was collected. This lead to a data set with raw I/Q data of 74 Giga Bytes for only one case.

In the light of little time for processing and the substantially large amount of time required for data collection, our teammate Aditya Singh decided to settle for a much smaller data set which had a couple hundred images of power density plots (as power-density plots are easier to generate compared to Spectrograms) for both with and without objects.
Below is an example of spectrograms that were generated from the I/Q data:





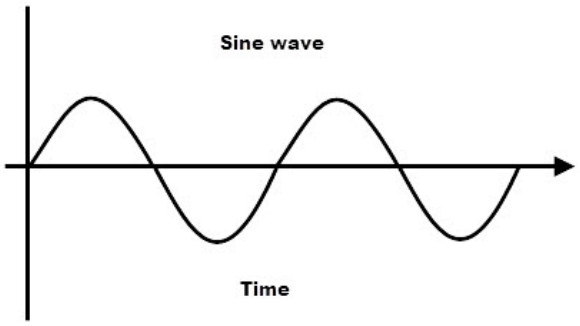
Following this, Aditya Singh and Vedansh Priyadarshi trained a Deep Learning model.
Again, recently, Aditya Singh trained a Deep Network with 6 Convolutional Layers [14] and 2Dense Layers on SFFT and reported an accuracy of .95 on the test set.

III. Key terms and Hardware

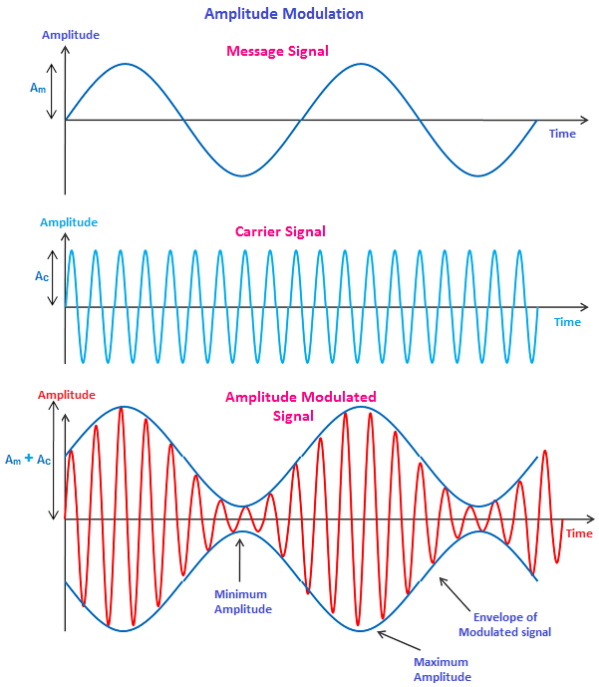
The key components in our current approach was In-phase and Quadrature data (I/Q data), Rasoberry Pi and the RTL-SDR. Below we describe them.

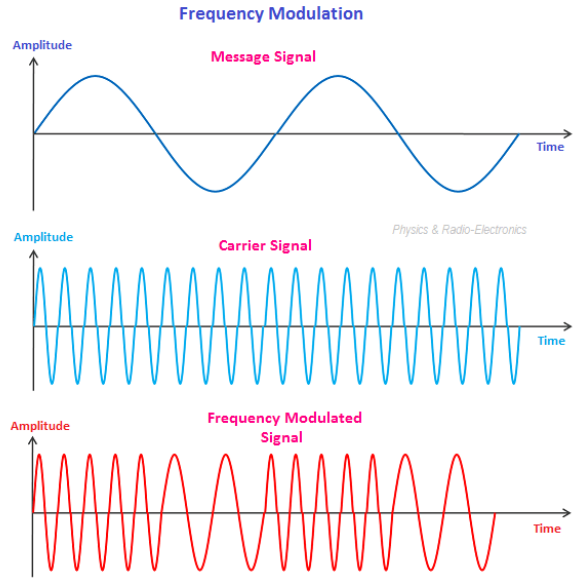
1. In-phase and Quadrature(I/Q) data: An unmodulated RF carrier wave is simply a sine

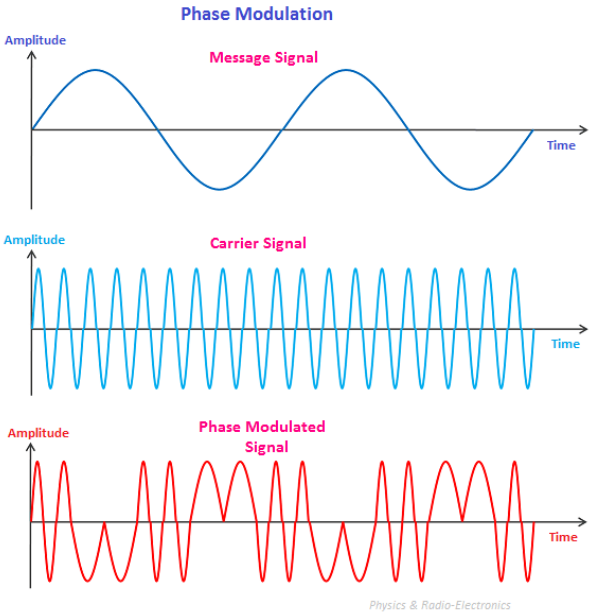
wave[]:



It can be represented as $V(t) = Asin(2\pi ft+\phi)$, where A is the amplitude of the wave, f stands for the center frequency, t for time period and and ϕ for the phase[15]. In order for an RF carrier to transmit information, it must be modulated using one of the many wave modulation techniques available, i.e., amplitude modulation (AM) which involves varying the amplitude of the wave, frequency modulation (FM) which involves varying frequency of the wave, and phase modulation (PM) which involves changing the phase of the wave [16]. Figures below depict AM [17], FM [18], PM [19] when applied to a carrier wave:







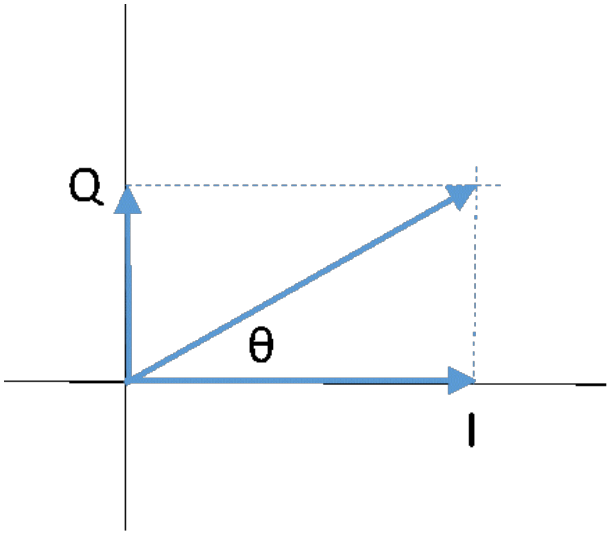
The carrier wave is periodic, and there can be two types of periodic signals: **sine** and cosine, although they are the same thing as :

$$cos(\omega t) = sin(\omega t + 90)$$

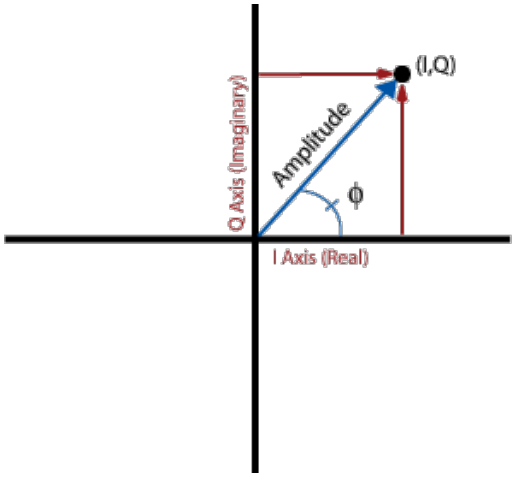
[20]. By convention, I signal is a cosine waveform, and the Q signal is a sine waveform and the 90-degree shift is the source of the names for the I and Q data where I refers to in-phase data (because the wave is in phase) and Q refers to quadrature data (because the wave is offset by 90 degrees) [21]. It can also be

viewed as: sine wave is at the quarter point, (90 out of the 360 degree cycle) or is said to be in quadrature-phase to the cosine wave. So it came to be that we call the cosine the in-phase or I and sine the quadrature-phase, or Q channel [22].

Now, if we consider I and Q to be two different values and plot it in a complex plane as the figure below with I on the x-axis and Q on the y-axis, we can calculate the hypotenuse using the Pythagoras theorem which gives us the amplitude (A). The angle Θ and amplitude can be used to represent I and Q as I.cos(θ) and Q.sin(θ) [21].



The new representation of I and Q allows for them to be used as polar coordinates and have a rectangular representation [21][23].

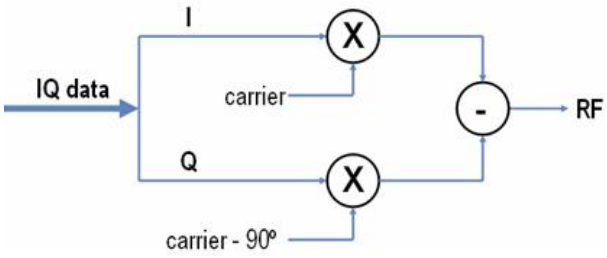


So, I and Q can now be represented as the complex number $I + iQ$ where I is the real part and Q is the complex part.

In RF signal transmission, a base band signal[] is modulated with the information that needs to be transmitted and then added to a carrier wave of higher frequency for transmission. An RF signal with any type of modulation can be created with the appropriate I and Q base band signals. By mixing an RF signal with local oscillator signals in quadrature, I and Q base band signals can be created.[21] .

$$Modulated\ RF = Icos(2ft) + Qsin(2ft)$$

This versatile property of I/Q data allows to control the amplitude, frequency, and phase of a modulating carrier sine wave by simply manipulating the amplitudes of separate I and Q base band signals [][]. This can easily be done by an I/Q modulator which has the following design []:



The rectangular form of I/Q Data is chosen as practical hardware design concerns make I and Q data the better choice as practical hardware design concerns make I and Q data the better choice. Precisely varying the phase of a high-frequency carrier sine wave in a hardware circuit according to an input message signal is difficult. A hardware signal modulator that manipulates the amplitude and phase of a carrier sine wave would therefore be expensive and difficult to design and build, and, as it turns out, not as flexible as a circuit that uses I and Q wave forms[22].

I and Q data contain all the information transmitted via the RF signal.Getting these from the modulated RF signal is fairly easy. Following the above representation of a modulated sine wave, we can write the full equation of the modulated wave in the form of I/Q data as [24]:

$$Acos(2\pi ft+\phi) = Acos(2\pi ft)cos(\phi) - Asin(2\pi ft)sin(\phi)$$

above follows the trigonometric expansion of $cos(a+b)$. Here, $Acos(2\pi ft)$, $Asin(2\pi ft)$ represent I and Q. In order to get the information encoded in I and Q back at the receiver end, we need to do the following simple steps:

$$Acos(2\pi ft+\phi)cos(2\pi ft)$$
$$=2Acos(2\pi ft)cos^2(2\pi ft)$$
$$-2Asin(2\pi ft)sin(2\pi ft)cos(2\pi ft)$$
$$=Acos(2\pi ft)+[Acos(2\pi ft)cos(2ct)$$
$$-Asin(2\pi ft)sin(2ct)]$$

Similarly,

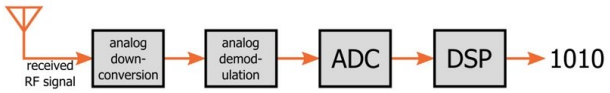
$$-2Acos(2\pi ft+\phi)sin(2\pi ft)$$
$$=-2Acos(2\pi ft)cos(2\pi ft)sin(2\pi ft)$$
$$+2Asin(2\pi ft(t)sin^2(2\pi ft)$$
$$=Asin(2\pi ft)+[-Asin(2\pi ft)cos(2(2\pi ft))$$
$$-Acos(2\pi ft)sin(2(2\pi ft))]$$

Finally I and Q can be obtained back by low-pass filtering[] the above two results:

I = low-pass filter($Acos(2\pi ft+\phi)cos(2\pi ft)$)
Q = low-pass filter($-2Acos(2\pi ft+\phi)sin(2\pi ft)$)
Because of above mentioned properties, I/Q data is widely used.

2. Raspberry Pi: Raspberry Pi[] is a small single board computer. For this project, we used a Raspberry Pi $2+$ Wifi as a transmitter. It uses a 1.2GHz 64-bit quad-core Arm Cortex-$A53$ CPU, has $1GB$ RAM, integrated $802.11n$ wireless LAN, and Bluetooth 4.1.

3. RTL-SDR: Radio components such as modulators, de-modulators and tuners are traditionally implemented in analogue hardware components. Invention of analogue to digital converters and advancement in technology allows most of these traditionally hardware based components to be implemented in software instead.[] A software-defined radio (as in, the device itself) is, thus, an RF communication system that has software system which can replace these hardware components and provide easier access to signal-processing functionality. Below is a block diagram of the receive path for an SDR [16]:



The RTL-SDR[6] stems from mass produced DVB-T TV tuner dongles that were based on the RTL2832U chipset. It has a maximum sampling rate of 3.2 MS/s (mega samples/ second) but a 2.56 MS/s is a stable as it does not drop samples. It has ADC resolution of 7 bits(effective), i.e., the Analog to digital converter can produce a maximum of 2^7 distinct output codes. The receive frequency is approximately 24 MHz $-$ 1766 MHz with a maximum bandwidth of 3.2 MHz although 2.8 MHz is stable.

IV. Mistakes,Conclusions &Future Work

Not considering the recent model accuracy achieved by Aditya Singh, below are a few points which could be some of the factors on why we did not get high quality data.

1. Antenna: The antenna [25] we used for transmission is a male to female jumper cable. The cable can be considered as a uni-pole antenna that transmits in all directions. Since our object is places in front of the transmitter, the gain of the antenna is not high enough such that enough power is reached to the object. To explain gain, I need to explain two more terms. Firstly, an **Isotropic Antenna** is an ideal antenna that transmits signal at same intensity in all directions. Building up on that, a **Directional Antenna** is one that sends out energy in one single direction. Gain can be written as how well we can take the power of the antenna and concentrate it in one direction. The gain of a directional antenna is the amount of energy a directional antenna sends out in one direction more than that sent by an isotropic antenna.
A directional antenna would have allowed for much better data collection as it would have increased the Radar Cross-section[26] of the object we are trying to detect.

2. Frequency of transmitted signal: The power of the transmitted signal depends on the frequency. Our transmitter had a maximum transmission frequency of 433 MHz which has definitely hampered the Radar Cross-Section[26] at the object. This finally led to not much information being reflected back to the SDR. This impacted the quality of model that was trained.

Moving ahead, for this phase, we plan on using a Function generator to generate a Sine wave of frequency 1.7GHz, amplify the signal and transmit using a suitable antenna. This will allow us to do away with the Downconverter obstacle we were facing in the recent past and move ahead with experiments quickly.

V. LITERATURE REVIEW:

A. 3D Tracking via Body Radio Reflections:

This paper introduces **WiTrack**, a system capable of tracks the 3D motion of a user using radio reflections that bounce off a person's body. It works through walls and occlusions, but does not require the user to carry any wireless device. WiTrack can also provide coarse tracking of a body part. In particular, the user may lift his or her hand and point at objects in the environment; the device detects the direction of the hand motion, enabling the user to identify objects of interest. Writers of the paper claim WiTrack to be the first device that can achieve centimeter-scale accuracy in tracking the 3D motion of a human based on radio reflections off her body.Witrack, according to the paper, had three applications:
1. 3D tracking of human motion through a wall.
2. Elderly fall detection
3. Gesture detection and its application in controlling appliances/ interacting with electronics.

WiTrack operates by transmitting an RF signal and capturing its reflections off a human body. It tracks the motion, fall and gesture by processing the signals from its received antennas using the following three steps:

1. Time-of-flight (TOF) estimation: WiTrack first measures the time it takes for its signal to travel from its transmit antenna to the reflecting body, and then back to each of its receive antennas. The time obtained is called TOF. This is obtained through FMCW transmission.Here, CW stands for continuous wave which refers to a sinusoidal carrier wave that can be switched on and off. FMCW (Frequency-Modulated Continuous Wave) is a special type of continuous wave that can change its operating frequency during the measurement: that is, the transmission signal is modulated in frequency (or in phase). Possibilities of Radar measurements through runtime measurements are only technically possible with these changes in the frequency (or phase)[]. In this, a signal is transmitted, which increases or decreases in frequency periodically. When an echo signal is received, that change of frequency gets a delay δ (by runtime shift)

like to as the pulse radar technique. To identify the distance from a reflector, FMWC compares the carrier frequency of the reflected signal to that of the transmitted signal. Since the carrier frequency is changing linearly in time, delays in the reflected signals translate into frequency shifts in comparison to the transmitted wave. Therefore, by comparing the frequency difference between the transmitted signal and the received signal, one can discover the time delay that the signal incurred, which corresponds to the TOF of that signal. Writers of the paper built an FMCW system that sweeps a total bandwidth of 1.69 GHz from 5.56 GHz to 7.25 GHz, and transmits at 0.75 milliwatt.

To distinguish a human's reflection from reflections off other objects in the environment, furnitures and walls, the writers ingeniously use the fact that since these reflectors are static, their distance to the WiTrack device does not change over time, and therefore their induced frequency shift stays constant over time. The power from these static reflectors is eliminated by simply subtracting the output of the FFT in a given sweep from the FFT of the signal in the previous sweep. This process is called background subtraction because it eliminates all the static reflectors in the background. Here, a sweep is one transmission of wave from transmitter to receiver. This makes it easier to see the much weaker reflections from a moving human. Following this, Witrack also needs to eliminate the signals that bounce off of humans, then bounce off of a static object and comes back to the receiver. This is called Dynamic multi-path. The idea for eliminating dynamic multi-path is based on the observation that, at any point in time, the direct signal reflected from the human to WiTrack device has travelled a shorter path than indirect reflections. Because distance is directly related to TOF, and hence to frequency, this means that the direct signal reflected from the human would result in the smallest frequency shift among all strong reflectors after background subtraction. In implementation, the final reading is average over five consecutive sweeps. Averaging allows boosts the power of a reflection from a human while diluting the peaks that are due to noise. This is because the human reflections are consistent and hence add up coherently, whereas the noise is random and hence adds up incoherently.

Finally, in order to deal with Noise, Witrack rejects impractical jumps

in distance estimates that correspond to unnatural human motion over a very short period of time, uses its tracking history to localize a person when he or she stops moving, and uses Kalman filters to smooth distance estimates.

2. 3D Localization: After receiving TOF for each of its antennas, WiTrack uses the geometric position of its antennas to localize the moving body in 3D. The TOF estimate from a receive antenna defines an ellipse whose foci are the transmit antenna and the receive antenna. It is straightforward to generalize the argument to localizing in 3D. Specifically, in a 3D space, the round-trip distance defines an ellipsoid whose two foci are the transmit antenna and one of the receive antennas. The intersection of two ellipsoids would define an arc in the 3D space, and hence is insufficient to pinpoint the 3D location of a person. However, by adding a third directional antenna, a unique solution in 3D that is obtained within the beam of all the directional antennas. While the minimum number of Rx antennas necessary to resolve a 3D location is three, adding more antennas would result in adding extra robustness to noise.

3. Fall Detection and Pointing: WiTrack detects a fall by monitoring fast changes in the elevation of a human and the final elevation after the change.To detect a fall, WiTrack requires two conditions to be met: First, the per- sons elevation along the z axis must change significantly (by more than one third of its value), and the final value for her elevation must be close to the ground level. The second condition is the change in elevation has to occur within a very short period to reflect that people fall quicker than they sit.

WiTrack can also differentiate an arm motion from a whole body motion; it can track the motion of raising ones arm, localize the initial and final

position of the arm, and determine the direction in which the arm is pointing. To track such a pointing gesture, WiTrack needs to distinguish between the movement of the entire body and the motion of an arm. To achieve this goal, they use the fact that the reflection surface of an arm is much smaller than the reflection surface of an entire human body. The size of reflection surface is estimated from the spectrogram of the received signal at each of the antennas. color).
If the reflector is large, its parts have slightly different positions from each other; hence, at any point in time the variance of its reflection along the y-axis is larger than that of an arm movement. WiTrack uses this spatial variance to detect body part motion from a whole body motion.
Following this, Witrack estimates direction using a three step process that involves Segmentation – to determine the start and end of a pointing gesture, Denoising – to obtain a clean estimate of the round trip distance of the arm motion as a function of time, for each receive antenna, and finally to Determine the Pointing Direction, Witrack performs robust regression on the location estimates of the moving hand, for which it uses the start and end points of the regression from all of the antennas to solve for the initial and final position of the hand, then repeats it for the dropping motion of the hand and finally predicts the final direction as the middle direction of the two.

Results: WiTracks median location error for the line-of-sight experiments is 9.9 cm, 8.6 cm, and 17.7 cm along the x, y, and z dimensions respectively. In comparison, the median location error in the through-wall experiments is 13.1 cm, 10.25 cm, and 21.0 cm along the x, y, and z dimensions. WiTrack localizes the center of the human body to within 10 to 13 cm in the x and y dimensions, and 21 cm in the z dimension.

From 11 users and over 133 experiments, WiTrack distinguishes a fall from standing, walking, sitting on a chair and sitting on the floor with an accuracy of 96.9% (the F-measure is 94.34%). It estimates the pointing direction with a median of 11.2 degrees and a 90th percentile of 37.9 degrees.

B. Multi-Person Localization via RF Body Reflections:

This paper presents WiTrack2.0, a multi-person localization system that operates in multi path-rich indoor environments and pinpoints users locations based purely on the reflections of wireless signals off their bodies. environment. It does so by disentangling the reflections of wireless signals that bounce off their bodies. Furthermore, it neither requires prior calibration nor that the users move in order to localize them. Witrack 2.0 builds on Witrack, its predecessor and therefore uses the same concepts for TOF. It extends the idea to enable multiple people tracking. Algorithmically, WiTrack2.0 has two main components:
1. Multi-shift FMCW, a technique that enables it to deal with multi-path effects
2. Successive Silhouette Cancellation (SSC), an algorithm that allows WiTrack2.0 to overcome the near-far problem.

The following sections describe these components:
1.Multi-Shift FMCW: Multi-path is the first challenge in accurate indoor localization. As this person moves, this multi-path reflection also moves with him and survives the background subtraction step. In single-user localization, one may eliminate this type of multi-path by leveraging that these secondary reflections travel along a longer path before they arrive at the receive antenna. Specifically, by electing the smallest TOF after background subtraction, one may identify the round-trip distance to the user. However, the above invariant does not hold in multi- person localization since different users are at different distances with respect to the antennas, and the multi-path of a nearby user may arrive earlier than that of a more distant user, or even interfere with it.
For this, Witrack 2.0 uses the same concept of Time-of-flight (TOF) as WiTrack. As the number of users increases, we need TOF measurements from a larger number of Transmitter(Tx)-Receiver(Rx) pairs to localize them, and extract their reflections from multi-path. Now, all TOFs are mapped to form a TOF profile. This process produces a heatmap where the x and y axes correspond to the plane of motion. The heatmap consists

of half-ellipses. As the profiles of each Tx-Rx pair is overlaid, the result shows how the noise and multi-path from different antennas averages out to result in a dark blue background. This is because different Tx-Rx pairs have different perspectives of the indoor environment; hence, they do not observe the same noise or multi-path reflections. As a result, the more we overlay heatmaps from different Tx-Rx pairs, the dimmer the multi-path effect, and the clearer the candidate locations for the two people in the environment. we are able to localize the two users using TOF measurements from five Tx-Rx pairs. Combining these measurements together allows us to eliminate the multi-path effects and localize the two people passively using their reflections.

2. Successive Silhouette Cancellation: Another obstacle in the muliple people setup is that reflections of a nearby user are much stronger than reflections of a faraway user or one behind an obstruction. To deal with this near-far problem, rather than localizing all users in one shot, WiTrack2.0 performs Successive Silhouette Cancellation (SSC) which consists of 4 steps:

1.SSC Detection. In the first step, SSC finds the location of the highest power reflector in the 2D heatmap.

2. SSC Re-mapping: This maps a persons location to the set of TOFs that would have generated that location at each transmit-receive pair.The signal transmitted from the transmit antenna will reflect off different points on the persons body before arriving at the receive antenna. Thus, the persons reflections will appear between some TOF_min and TOF_max in the TOF profile at the Rx antenna. TOF_min. The closest point on the persons body is the one that corresponds to the shortest round-trip distance to the Tx-Rx pair, where the round-trip distance is the summation of the forward path from Tx to that point and the path from that point back to Rx.Similarly, TOF_max is bounded by the round-trip distance to point on the persons body that is furthest from the Tx-Rx pair.the x and y coordinates of the furthest point are determined by the persons location from the SSC Detection step. The z co-ordinate is taken as the person's feet location.

3. SSC Cancellation: The next step is to use TOF_min and TOF_max to cancel the persons reflections from the TOF profiles of each transmit-receive pair. To do that, we take a conservative approach and remove the power in all TOFs between TOF_min and TOF_max within that profile. This relies on the fact that multi-shift FMCW provides a large number of TOF profiles from many Tx-Rx pairs.Hence, even if we cancel out the power in the TOF of a person with respect to a particular Tx-Rx pair, each person will continue to have a sufficient number of TOF measurements from the rest of the antennas.
We repeat the process of computing TOF_min and TOF_max with respect to each Tx-Rx pair and cancelling the power in that range until we have eliminated any power from the recently localized person.

4. Iteration: We proceed to localize the next person. This is done by regenerating the heat maps from the updated TOF profiles and overlaying them. WiTrack2.0 repeats the same cancellation procedure for all people in the heat map.

Following this, WiTrack 2.0 performs some optimization to improve SSC which include:
– a focusing step for each user to refine his location estimate.

– applies a Kalman filter and performs outlier rejection to reject impractical jumps in location estimates that would otherwise correspond to abnormal human motion over a very short period of time.

–To disentangle multiple people who cross paths, we look at their direction of motion before they crossed paths and project how they would proceed with the same speed and direction as they are crossing paths.This helps us with associating each person with his own trajectory after crossing.

WiTrack2.0 can track gestures even when they are simultaneously performed by multiple users.

Specifically, by exploiting SSC focusing, it zooms onto each user individually to track his gestures.By tracking the trajectory of each moving hand, it can determine its pointing direction.

3. Localization based on Breathing: Writers of the paper use the fact that a static person moves slightly due to breathing. Specifically, during the process of breathing, the human chest moves by a sub-centimeter distance over a period of few seconds. The key challenge is that this change does not translate into a discernible change in the TOF of the person. However, over an interval of time of a few seconds (i.e., as the person inhales and exhales), it would result in discernible changes in the reflected signal. Therefore, by subtracting frames in time that are few seconds apart, it is possible to localize the breathing motion. It is shown in the paper that a static, seated person can accurately be localized using a subtraction window of 2.5 seconds. However, this long subtraction window will introduce errors in localizing a pacing person. To accurately localize both static and moving people, WiTrack2.0 performs background subtraction with different subtraction windows. To localize moving users, it uses a subtraction window of 12.5 ms. To localize static human beings a subtraction window of 3-6 seconds is considered under the assumption that normal adults inhale exhale over a period of this interval. As a result, breathing users reflections pop up at different instances, allowing WiTrack 2.0 to detect and localize them.

Results: WiTrack2.0's median location error in tracking the motion of four users when it is in the same room as the subjects is 8.5 cm in x and 6.4 cm in y for the first user detected, and decreases to 15.9 cm in x and 7.2 cm in y for the last detected user. In through-wall scenarios, WiTrack2.0 can accurately localize up to three users with median location error for these experiments being 8.4 cm and 7.1 cm in xy for the first detected user, and decreasing to 16.1 cm and 10.5 cm in x/y for the last detected user. WiTrack2.0s breathing-based localization accuracy goes from a median of 7.24 and 6.3 cm in x/y for the nearest user to 18.31 and 10.85 cm in x/y for the furthest user, in both line-of-sight and through-wall settings. For 3D Gesture Detection, the median orientation error in goes from 8.2 degrees to 12.4 degrees from the first to the third person, and from 12 degrees to 16 degrees in θ.

C. See Through Walls with WiFi!:

The objective of this paper is to enable a see-through-wall technology that is low-bandwidth, low-power, compact, and accessible to non-military entities. This paper introduces two main innovations:
1. It shows how one can use MIMO interference nulling to eliminate reflections off static objects and focus the receiver on a moving target.

2. It shows how one can track a human by treating the motion of a human body as an antenna array and tracking the resulting RF beam.

The paper introduces Wi-Vi, a see-through-wall device that employs Wi-Fi signals in the 2.4 GHz ISM band. Wi-Vi limits itself to a 20 MHz-wide Wi-Fi channel, and avoids ultra-wideband solutions used today to address flash effect(Reflections off the wall overwhelm the receivers analog to digital converter (ADC), preventing it from registering the minute variations due to reflections from objects behind the wall. This behaviour is called the flash effect). It uses a smaller 3-antenna MIMO radio: device: two of the antennas are used for transmitting and one is used for receiving. It also employs directional antennas to focus the energy toward the wall or room of interest.

Wi-Vi achieves the proposed functionality by the following steps:

1. Eliminating the Flash Effect: In a through-wall system the signal reflected off the wall, i.e., the flash, is much stronger than any signal reflected from objects behind the wall. This problem is exacerbated by two other parameters: first, the actual reflected signal is significantly weaker since it depends both; second, in addition to the direct flash caused by reflections off the wall, through-wall systems have to eliminate the direct signal from the transmit to the receive antenna, which is significantly larger than the

reflections of interest. Wi-Vi uses interference nulling to cancel both the wall reflections and the direct signal from the transmit to the receive antenna, hence increasing its sensitivity to the reflections of interest. In the initial phase, Wi-Vi uses the Channel-State information from both the receivers and calculate a ratio, p. Now, this along with the initially transmitted preamble is transmitted. by the end of this phase Wi-Vi has eliminated the signals reflected off all static objects as well as the direct signal from the transmit antennas to the receive antenna. If no object moves, the channel will continue being nulled. However, since RF reflections combine linearly over the medium, if some object moves, its reflections will start showing up in the channel value.

The signals due to moving objects behind the wall are too weak and it's hard to discern the signal due to moving objects since it will be immersed in the receivers hardware noise. Thus, Wi-Vi next boosts the transmitted signal power thereby increasing the overall power that traverses the wall. After boosting the transmit power, residual reflections which were below the ADC quantization level become measurable. Such reflections from static objects can create significant clutter in the tracking process if not removed. To address this issue, Wi-Vi performs a procedure called iterative nulling. The objective of iterative nulling is to null the signal again after boosting the power to eliminate the residual reflections from static objects. Wi-Vi leverages the fact that errors in the channel estimates are much smaller than the channel estimates themselves, and uses this observation to refine its estimates. For this, Wi-Vi iterates between these two steps to obtain finer estimates for both h1(from receiver) and h2 (from receiver), until the two estimates h converge.

2. Identifying and Tracking Humans:
Here, the writers use an called inverse synthetic aperture radar (ISAR) to emulate an antenna array. In ISAR, there is only one receive antenna; hence, at any point in time, the receiver captures a single measurement. However, as the target moves, he/she samples the received signal at successive locations in space, as if we had a receive antenna at each of these points. Furthermore, because of channel reciprocity, successive time samples received by Wi-Vi correspond to successive spatial locations of the moving target. Hence, Wi-Vi effectively receives in time what an antenna array would receive in space. By treating consecutive time samples as spatial samples, Wi-Vi can emulate an antenna array and use it to track motion behind the wall.

However, with multiple humans, the noise increases significantly. On one hand, each human is not just one object because of different body parts moving in a loosely coupled way. On the other hand, the signal reflected off all of these humans is correlated in time, since they all reflect the transmitted signal. The lack of independence between the reflected signals is important. Wi-Vi uses the smoothed MUSIC algorithm to disentangle correlated super-imposed signals. MUSIC aims at aim at identifying the spatial angle of the signal. by projecting on the null space and taking the inverse norm MUSIC achieves sharper peaks. MUSIC can also be used detect single human beings.

For Gesture encoding, writers of the text modulate as a 0 bit is a step forward followed by a step backward; a 1 bit is a step backward followed by a step forward. Decoding the above gestures is fairly simple and follows standard communication techniques.

Results: Wi-Vi can identify whether there is 0 or 1 person in a room with 100% accuracy; two humans are never confused with 0 or 1. However, Wi-Vi confused 2 humans with 3 humans in 15% of the trials, whereas it accurately identified their number in 85% of the cases.Wi-Vi correctly decoded the performed gestures at all distances less than or equal to 5m. It identified 93.75% of the gestures performed at distances between 6m and 7m. At 8m, the performance starts to degrading, leading to correct identification of only 75% of the gestures. Finally, Wi-Vi could not identify any of the gestures when the person was standing 9m away from the wall.

D. A Wireless Passive Radar System for Real-Time Through-Wall Movement Detection:

This paper reports advancements that allow our software-defined through-wall radar to operate in real time and with increased Doppler sensitivity to obtain additional target information with a real Wi-Fi signal. This system is able to measure both reference and surveillance signals through walls, enabling a standoff and stand-alone system without the requirement to connect to an internal reference source. A technique to develop a Doppler event history that has enabled recognition of different types of body motions is also described.
The whole system is described as follows:

Signal Process & **SDR System:** To obtain target Doppler and range information, the recorded reference and surveillance channel data undergo discrete cross-ambiguity function (CAF) processing, This processing stage is computationally expensive for high-bandwidth signals. To achieve real-time operation without loss in performance, the system employs two strategies:

– The batch processing allows the range-Doppler surface to be generated in a computationally efficient manner.

– The pipeline processing allows key signal and data handling procedures to be carried out simultaneously. This process is similar to the parallel computing process which facilitates fast computation on multiple cores in a computer. To optimize processing in the wireless passive radar system, the authors mention that they divided their multistage algorithm into three subroutines:

(a) reading the recorded sample data from the random access memory

(2) performing cross-ambiguity processing on this data using our batch processing technique

(3) applying the CLEAN DSI algorithm

Because the three subroutines are implemented simultaneously, the total processing time for the system is dependent on which subroutine takes the longest time to complete.

2. Batch Processing: The CAF processing is computationally expensive as it involves large volume of long sequence cross-correlations and Fourier transforms. This includes dividing the recorded synchronized reference and surveillance signals into isometric data segments. Following this, the selected surveillance portion is cross-correlated with the corresponding reference portion. The results of each cross-correlated portion pair is combined in a matrix and finally, a Fourier transform is applied to each column of the matrix. The range-Doppler surface can then be obtained by applying CAF to output from previous step. The DSI is then carried out using the CLEAN algorithm.

3. Direct Signal Interference Suppression Using CLEAN: This section introduces a simplified interference cancellation method, which is modified from the CLEAN algorithm in is introduced. The purpose of the CLEAN algorithm is to mitigate the impact of DSI and stationary clutter.

4. Target Detection: Using above processing methods, a low-interference range-Doppler surface can be obtained to identify target responses. A two-stage target identification routine is developed for picking the target out from this range-Doppler surface in real time. The technique first identifies one or more areas of pixels in the range-Doppler surface display where the targets have a high likelihood of appearing. Then, a threshold is applied to assist in deciding on which data pixel is likely to be the target. The target is more likely to be located in an area where there are higher-strength signal returns on the range-Doppler surface. But a pixel with a higher power value may not necessarily be a target, because there may still be some impact from interference and clutter. An information display method is used to facilitate accurate detection of person. A target that has a large cross section will generate a larger detection index than a smaller target. Once a target is detected, the system makes a record of the detection. The real-time system outputs the Doppler record continuously. By using

this record, time-Doppler characteristics can be obtained and displayed. This Doppler event history gives a clear view of target movement status varying with time and can be used to characterize a range of movements.

Results: During the stoopingstanding gesture cycle, only the upper body is moving and approaching or leaving the antenna. Predictably, a positive and negative Doppler is generated. The Doppler record shows a sharper characteristic than in the other motions.
In the leftright swing experiment, the subject was facing the wall, pointing toward the surveillance antenna on the other side and swinging the torso from left to right. This results in initially positive Doppler followed by a significant negative Doppler trough as the torso swings through the complete return cycle. This is then followed by positive Doppler again as the torso swings upright.
During the squatting standing gesture cycle, the Doppler record of this gesture cycle shows a combination of positive and negative excursions.
In the final, Back-Forward Swing, the result is a sinusoidal positive and negative Doppler.

Some papers that have not been included in this literature review dealt with using Channel State Information to make a Doppler-shift surface and using Machine Learning on this to detect the presence of humans in an indoor setting [27]. This organization [28] appears to be using the same concept. i.e. 5GHz Wi-Fi signals in its devices that promises to detect human motion in a room.

REFERENCES

[1] Adib, Fadel, and Dina Katabi. See through Walls with WiFi! Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM - SIGCOMM '13, 2013, doi:10.1145/2486001.2486039.

[2] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller. 3D Tracking via Body Radio Reflections. In Usenix NSDI, 2014.

[3] Adib, Fadel, Zachary Kabelac, and Dina Katabi. "Multi-Person Localization via RF Body Reflections." 12th USENIX Symposium on Networked Systems Design and Implementation." 4-6 May, 2015, Oakland, California, USENIX, 2015.

[4] B. Tan, K. Woodbridge and K. Chetty, "Awireless passive radar system for real-time through-wall movement detection," in IEEE Transactions on Aerospace and Electronic Systems, vol. 52, no. 5, pp. 2596-2603, October 2016, doi: 10.1109/TAES.2016.140207.

[5] H. C. Yildirim, L. Storrer, M. V. Eechkhaute, C. Desset, J. Louveaux and F. Horlin, "Passive Radar based on 802.11ac Signals for Indoor Object Detection," 2019 16th European Radar Conference (EuRAD), PARIS, France, 2019, pp. 153-156.

[6] https://www.rtl-sdr.com/about-rtl-sdr/

[7] https://www.gnuradio.org/

[8] https://cubicsdr.com/

[9] https://gqrx.dk/

[10] https://www.raspberrypi.org/products/

[11] https://github.com/F5OEO/rpitx

[12] https://www.mathworks.com/products/matlab.html

[13] https://pypi.org/project/pyrtlsdr/

[14] Fukushima, K. (2007). "Neocognitron". doi:10.4249/scholarpedia.1717

[15] https://in.tek.com/blog/

[16] https://www.allaboutcircuits.com/textbook/radio-frequency-analysis-design/radio-frequency-demodulation/understanding-i-q-signals-and-quadrature-modulation/

[17] https://www.physics-and-radio-electronics.com/blog/amplitude-modulation/

[18] https://www.physics-and-radio-electronics.com/blog/frequency-modulation/

[19] https://www.physics-and-radio-electronics.com/blog/phase-modulation/

[20] http://complextoreal.com/tutorials/

[21] http://whiteboard.ping.se/SDR/IQ

[22] http://www.ni.com/tutorial/4805/en/

[23] https://www.tek.com/blog/calculating-rf-power-iq-samples

[24] https://dsp.stackexchange.com/questions/7612/i-and-q-channels/76137613

[25] https://www.youtube.com/watch?v=p2ify-2DQdElist=PLUJAYadtuizA8RC2Qk8LfmiWA56HZsk9yindex=14

[26] https://www.youtube.com/watch?v=a53Cg3KUTt4list=PLUJAYadtuizA8RC2Qk8LfmiWA56HZsk9yindex=9

[27] S. Yousefi, H. Narui, S. Dayal, S. Ermon and S. Valaee, "A Survey on Behavior Recognition Using WiFi Channel State Information," in IEEE Communications Magazine, vol. 55, no. 10, pp. 98-104, Oct. 2017, doi: 10.1109/MCOM.2017.1700082.

[28] https://www.celeno.com/wifi-doppler-imaging