Noisy Object Detection using Deep Learning (CS-3099)

Aishwarya Das
Ashoka University
Submitted to Dr. Debayan Gupta

Abstract— **Radio waves are electromagnetic radiation with wavelengths longer than the infrared light. Radio waves have been used to detect objects using RADAR and SONAR, they are typically of very high frequency and involve a lot of signal processing. We explore how Neural networks when combined with signal processing help us detect the presence of an object at lower band frequencies.Using COTS components, we present a method to simply detect the presence of an object when kept in between the line of sight of a transmitter and receiver using radio waves.**

I. Introduction

The inspiration for this project comes from Wi-Vi[1], a SDR based project that uses advances in MIMO techniques to detect the presence of a person behind the wall of a room. It operate in the same band specturm as the WI-Fi which is 2.4 Ghz. The goal of our project is ambitious, we want to use devices that operate in the same band frequency as Wi-Fi and use it to detect the presence of people in a room. The goal is to be able to do that with just COTS components rather than military grade equipment. We also aim to use a combination of signal processing and neural networks to achieve the task. The previous approaches in this field have been done by using heavy signal processing and military grade expensive components. To achieve the goals of our project, we propose a three phase plan:

- Use a transmitter and a receiver to build a Bistatic Radar which detects when objects are kept in the line of sight. Once this is done, use multiple transmitters and receivers to detect objects when they are kept not in the line of sight.
- Detect the shape of the object and the location of the object in the given room. Some additional hardware modifications might be required.
- With the learning's from the first two phase, figure out a way to detect the presence of people, in a given room.

The specifics for the third phase haven't been decided, but most likely we will be treating the human body as an antenna and then using a combination of neural networks and signal processing to detect its presence.

II. Setting Up the Experiment

We interfaced our RTL-SDR with the computer and connected it to a dipole antenna. Initially we used **GQRX** to visualise and check if the SDR was receiving any band frequency. SDR along with our dipole antenna served as a receiver.

We used Raspberry Pi as our transmitter, this is achieved by installing the **rpitx**[2] library and a male to female cable in the pin 18 of RasPi. This enabled us to transmit a maximum frequency of 433 Mhz. The transmission frequency is limited due to Raspberry Pi's hardware. **GQRX**[3] proved to be a useful software for checking if the SDR is receiving any data or not but otherwise its difficult to collect data using the software.

The SDR connected to the dipole antenna and the Raspberry Pi form a Bistatic Radar. A Bistatic radar is a radar system comprising of a transmitter and receiver that are separated by a distance comparable to the expected target distance. In this our target distance was kept in between the RasPi and SDR in the line of sight. Both the devices were kept at a distance of 1.5m.

As **GQRX** wasn't helpful for us to collect data, we tried interfacing our SDR with MATLAB.

MATLAB consists of all libraries required for SDR interfacing and collection of data. When data collection from MATLAB was started, the process was taking too much time. We eventually had to forego MATLAB and settle with the python library for RTL-SDR.

The crux of our experiment is to detect changes in the spectrogram of the data received by our SDR. We guess that when an object is kept in between our devices, there is some change in the spectrogram and when there is no object kept in between our devices, the spectrogram is different as our receiver receives more data packets. We thus aim to build a binary classifier to distinguish between both the cases.

As mentioned before, we collected data using the RTL-SDR library in Python[4]. Data was collected in the form of I/Q values stored in CSV files. The experiment was conducted in a room where there was minimal electrical and human activity. Our sampling rate was set to 2.6 million samples/sec and the center frequency was kept at 433 MHz. Each reading came out to be 5 seconds long and over 10000 samples of raw I/Q data was collected. This dataset however came out to be 74 GB and was only for one case.

Later one of our teammates, generated a small dataset of a few images. Power Spectral Density plots were generated instead of spectrogram as they take less time to generate. When the images were trained for the first time, our accuracy came out to be around 65%. Recently our teammate trained the data on a deep neural network with 6 convolutional layers[5] and 2Dense layers and reported an accuracy of 95%. The object kept in detection for our Bistatic Radar is a 1 litre plastic bottle filled with water.

III. Key Terms

- **IQ Data** : Quadrature signals, also called IQ signals, IQ data or IQ samples, are often used in RF applications. They form the basis of complex RF signal modulation and demodulation, both in hardware and in software, as well as in complex signal analysis.

  A pair of periodic signals are said to be in quadrature when they differ in phase by 90 degrees. The in-phase or reference signal is referred to as ”I” and the signal that is shifted by 90 degrees (the signal in quadrature) is called ”Q”. [6]

  If the phase difference between two sinusoids is 90 degrees (or $\pi/2$ radians), then these two signals are said to be in quadrature. An example of this is the sine wave and the cosine wave. By convention, the cosine wave is the in-phase component and the sine wave is the quadrature component [7]. The capital letter ”I” represents the amplitude of the in-phase signal, and the capital letter ”Q” represents the amplitude of the quadrature signal.

  Software Defined Radio (SDR) systems use these concepts extensively because the baseband I&Q signals are often represented as discrete time sampled data. Therefore, digital signal processing (DSP) can be used to literally define the transmitter and receiver characteristics including filtering, modulation and demodulation, AGC, etc.[7] SDR receivers often feature a baseband bandwidth of a few hundred kHz or more, giving the ability to perform a wide variety of functions including wide bandscope and spectrogram functions, as well as being able to simultaneously monitor and demodulate several signals of different types at once.

- **Software Defined Radio** : Radio components such as modulators and tuners are traditionally implemented in analogue hardware components[8].The advent of modern computing and analogue to digital converters allows most of these traditionally hardware based components to be implemented in software instead. Hence, the term software defined radio. This enables easy signal processing and thus cheap wide band scanner radios to be produced.

As mentioned before, we are using the RTL SDR for our project. RTL-SDR is a very cheap 25$ USB dongle that can be used as a computer based radio scanner for receiving live radio signals in your area (no internet required). Depending on the particular model it could receive frequencies from 500 kHz up to 1.75 GHz.[8] Most software for the RTL-SDR is also community developed, and provided free of charge.

- **Spectrogram** : A spectrogram is a visual way of representing the signal strength, or loudness, of a signal over time at various frequencies present in a particular waveform. Not only can one see whether there is more or less energy at, for example, 2 Hz vs 10 Hz, but one can also see how energy levels vary over time. In other sciences spectrograms are commonly used to display frequencies of sound waves produced by humans, machinery, animals, whales etc. as recorded by microphones.

  In our experiment, the spectrogram has been created using FFT (Fast Fourier Transform). Creating a spectrogram using the FFT is a digital process. Digitally sampled data, in the time domain, is broken up into chunks, which usually overlap, and Fourier transformed to calculate the magnitude of the frequency spectrum for each chunk.

  Each chunk then corresponds to a vertical line in the image; a measurement of magnitude versus frequency for a specific moment in time (the midpoint of the chunk). These spectrum or time plots are then "laid side by side" to form the image or a three-dimensional surface, or slightly overlapped in various ways, i.e. windowing. This process essentially corresponds to computing the squared magnitude of the short-time Fourier transform (STFT) of the signal) [9].

- **Power Spectral Density** : When the energy of the signal is concentrated around a finite time interval, one may compute the energy spectral density[10]. Power Spectral Density (or simply power spectrum), which applies to signals existing over all time, or over a time period large enough (especially in relation to the duration of a measurement) that it could as well have been over an infinite time interval. The power spectral density (PSD) then refers to the spectral energy distribution that would be found per unit time, since the total energy of such a signal over all time would generally be infinite.

- **Fourier Transform & FFT** : The Fourier transform (FT) decomposes a function (often a function of time, or a signal) into its constituent frequencies. The Fourier transform of a function of time is a complex-valued function of frequency, whose magnitude (absolute value) represents the amount of that frequency present in the original function, and whose argument is the phase offset of the basic sinusoidal in that frequency. The Fourier transform of a function f is traditionally denoted $\hat{f}$ and is given by :

  $$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} dx$$

  A Fast Fourier Transform (FFT) is an algorithm that computes the discrete Fourier transform (DFT) of a sequence, or its inverse (IDFT). Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. [11] The DFT is obtained by decomposing a sequence of values into components of different frequencies. An FFT rapidly computes such transformations by factorizing the DFT matrix into a product of sparse (mostly zero) factors. An FFT rapidly computes such transformations by factorizing the DFT matrix into a product of sparse factors. As a result, it manages to reduce the complexity of computing the DFT.

IV. Literature Review

A. See through walls with Wi-Fi!

This seminal paper shows that Wi-Fi can extend our senses, enabling us to see moving objects through walls and behind closed doors. In particular, we can use such signals to identify the number of people in a closed room and their relative locations. We can also identify simple gestures made behind a wall, and combine a sequence of gestures to communicate messages to a wireless receiver without carrying any transmitting device.[1]

They use MIMO interference nulling to eliminate reflections off static objects and focus the receiver on a moving target. The paper also 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 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. To this end, 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 the flash effect.

In any 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 is due to the significant attenuation which electromagnetic signals suffer when penetrating dense obstacles. This is called the flash effect.

Wi-Vi bridges state-of-the-art networking techniques with human-computer interaction. It motivates a new form of user interfaces which rely solely on using the reflections of a transmitted RF signal to identify human gestures.

B. 3D Tracking via Body Radio Reflections

This paper introduces Wi-Track, a system that tracks the 3D motion of a user from the radio signals reflected off their body. It works even if the person is occluded from the Wi-Track device or in a different room[12]. Wi-Track can also provide coarse tracking of a body part. In particular, the user may lift 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.

Wi-Track's motion tracking works as follows. The device transmits a radio signal and uses its reflections to estimate the time it takes the signal to travel from the transmitting antenna to the reflecting object and back to each of the receiving antennas. WiTrack then uses its knowledge of the position of the antennas to create a geometric reference model, which maps the round trip delays observed by the receive antennas to a 3D position of the reflecting body.

Wi-Track localizes the center of a human body to within a median of 10 to 13cm in the x and y dimensions, and 21 cm in the z dimension. It also provides coarse tracking of body parts, identifying the direction of a pointing hand with a median of 11.2°.

Wi-Track bridges a gap between RF-based localization systems which locate a user through walls and occlusions, and human-computer interaction systems like Kinect, which can track a user without instrumenting her body, but require the user to stay within the direct line of sight of the device.

C. FindIT - Human motion detection

This paper presents a system utilizing narrow band software defined radio to detect the moving human through walls, and give some motion details, such as motion orientation which includes relative moving direction. To achieve high accuracy, FindIt applies Short Time Fourier Transform (STFT) and statistical methods to received signals.

FindIT consists of a hardware platform for sending and receiving signals and removing flash effect, and a software that controls the transmission and runs detection algorithms. It's able to identify the moving human, recognize relative moving direction and motion orientation[13]. In addition, it can be used in different places by introducing machine learning methods so as to achieve adaptive detection.

FindIT detects human motion by using STFT. STFT is generally applied into Ultra-wide bandwidth signals, so it can finish the computation rapidly due to narrow bandwidth meanwhile it provides us with both time and frequency information. At last, we get range data, which is compared with the pre-set threshold to detect whether there is a moving human, by statistical methods.

V. Conclusion

From our experiments we conclude that with an accuracy of 95 % on a small test data is not enough. The original plan of generating 10,000 samples needs to be implemented to get a general purpose accuracy. We also need to take the Radar Equation into account and the size of the cross section of the object in detection. These things were not taken into account, when we were performing our experiment.

The antenna in use for this experiment is a bi-pole antenna and the object is kept in line with the antenna and the Raspberry Pi, the gain is not strong enough to give us a good Signal to Noise ratio (SNR). While evaluating our dataset, a thorough SNR should be calculated for all data points and the ones below a certain threshold should be discarded.

Our transmitter had a maximum transmission of 433 MHz. Since the frequency is low, this led to a poor quality of data being collected. One way the frequency could be increased is by using a down-converter. If not for a down-converter, we plan on using a Function generator to generate a Sine-Wave of 1.7 Ghz, amplify the signal using an amplifier and transmit it using a suitable antenna.

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

[3] https://gqrx.dk/

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

[5] Fukushima, K. (2007). Neocognitron.

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

[7] https://in.tek.com/blog/what's-your-iq—about-quadrature-signals...

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

[9] http://zone.ni.com/reference/en-XX/help/371361E-01/lvanls/stfspectrogramcore/details

[10] P Stoica R Moses (2005). ”Spectral Analysis of Signals” (PDF).

[11] Van Loan, Charles (1992). Computational Frameworks for the Fast Fourier Transform. SIAM.

[12] F.Adib,Z.Kabelac,D.Katabi,andR.C.Miller.3DTracking via Body Radio Reflections. In Usenix NSDI, 2014.

[13] FindIt: Real-time Through-Wall Human Motion Detection Using Narrow Band SDR, Zhizhao Zhang