EE292Q - Final Project Report

mmWave Through-Wall Imaging

Akshata Patil, Jesica Gonzalez, Lilian Chan, Yiting Wu

June 12, 2023

1 Theory of Sensor Operation

1.1 Overview

The Frequency Modulated Continuous Wave (FMCW) radar is widely utilized for target ranging and velocity estimation. It operates by transmitting a continuous signal that undergoes linear frequency variation over time. This type of signal is called a chirp.

When the transmitted signal hits a target, it reflects back and is received at the sensor end with varying frequencies. The working principle of FMCW radar relies on two factors: the Doppler effect and the delay time of the reflected signal. The Doppler effect causes a frequency shift as a result of relative motion between the radar sensor and the target. By analyzing this frequency shift, the radar system can accurately determine the velocity of the target.

1.2 Key Components

The FMCW Radar consists of the following components:

1. **Synthesizer**: Generates the chirp signal.
2. **Transmitter (TX) Antenna**: Transmits chirp signals.
3. **Receiver (RX) Antenna**: Captures the reflected chirp signal from the target.
4. **Frequency Mixer Component**: Combines TX and RX signals to generate an intermediate frequency (IF) signal.

1.3 Doppler Data Processing

For a single object at a distance d, the IF signal will be $x = Asin(2\pi f_0 t + \Phi_0)$

- $f_0 = \frac{S2d}{c}$
- $\Phi_0 = \frac{4\pi d}{\lambda}$

(A: amplitude, t: time, S: slope, d: distance, λ: wavelength, c: speed of light)

When there are multiple objects, the IF signal undergoes Fourier transform to separate different tones. A frequency spectrum is produced with separate peaks for different tones. Each peak corresponts to an object at a specific distance.

1.4 Radar Equations and Target RCS

The radar range equation is the fundamental model for estimating the received power in a given scenario. A higher target RCS, the equivalent area seen by the radar, would result in a higher received power.

$$P_r = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4}$$

(P_r: received power, P_t: transmit power, G_t: transmitter gain, G_r: receiver gain, λ: carrier wavelength, σ: mean RCS of the target, R: range from radar to target)

2 Advantages and Disadvantages of Sensor Modality for Specific Application

2.1 Advantage: Penetration ability for through-wall imaging

The mmWave radar sensor operates in millimeter wave frequency ranges, allowing it to detect objects and provide distance, speed, and angle information. Its small wavelengths enable sub-millimeter range accuracy and penetration through materials like walls, plastic, and clothing, even in adverse conditions. The sensor's non-contact nature makes it ideal for indirect measurements.

2.2 Advantage/Disadvantage: Dynamic sensor parameter

There are many parameters that can be modified in the mmWave sensor. For instance, the chirp slope, the number of chirps, and the sampling frequency. These dynamic parameters allow us to set the hardware environment most accurately for a specific application. However, these parameters are also limited by the hardware. For example, the chirp slope has a relatively tight boundary, so does the center frequency. Hence, we can only improve the resolutions to a certain extent bounded by the hardware limitations.

2.3 Disadvantage: Cluttering

In our application, there is much cluttering that adds difficulty to single out and detect target objects. Since the cluttering results from waves bouncing off surfaces such as the wall and the ceiling, it is not the true system noise. Hence, it cannot be removed simply by changing the SNR ratio. How much cluttering depends on the field of view, which relies on the number of receivers and their positions. It is hard to manipulate these components in the sensor, hence difficult to perfectly filter out cluttering on the sensor hardware end.

3 Experimental Results and Comparison to Theoretical Expectations

3.1 Experimentation

We used a wall made out of plastic-like material of around 5 mm clamped to the floor. The wall covered the entire height of the target, which for the experiment was one of the team members. We conducted several experiments varying the number of people walking behind the wall, as well as their walking speeds and paths. In one case, each person walked back and forth from the radar, in the other case they walked tangentially to it, and in the last case, they walked randomly. Our sensor always faces the wall perpendicularly. We collected data across 200 and 300 frames, and converted raw binary data to radarCube and later to Range-Doppler arrays with MatLab.

3.2 Pre-Processing: Filter and Convolution

We experimented with different pre-processing of the raw data as well as clustering algorithms. To remove the wall from our data, we first remove zero Doppler by taking the first frame as the calibration frame and subtracting each frame with it. This method would remove all static elements in the scene.

Next, because there is much cluttering around the target object, we set a filter to remove weaker received waves and only focus on the concentrated power, which is the signal directly reflected by our target. To apply the filter, we look at the maximum and median power across the entire frame and set a threshold accordingly. We also experimented with convolution methods, which reduced the resolution by blurring the received signals. Convolution worked best with K-Means clustering algorithm, but not with our final solution (mentioned below), hence it was not adopted in the end.

Lastly, we explored several clustering algorithms. These include FindPeak, K-Means Clustering with Elbow Method and Find Connected Components. Out of the three algorithms, FindPeak with Euclidean distances performed the best. The clustering algorithms and results are explained in the following sections.

3.3 Clustering: FindPeaks and Euclidean Distances

After filtering and convolving the data, one of the clustering methods was a combination of finding the maximum peaks, then only keeping one peak when a group of them are within a certain distance of each other.

For the FindPeaks code, all the peaks in the data are found that are at least of the FindPeaks threshold multiplied by the maximum peak value. This usually meant that we were finding all the peaks that were around 90-95% (different thresholds are optimal for detecting different amounts of people) of the maximum power threshold.

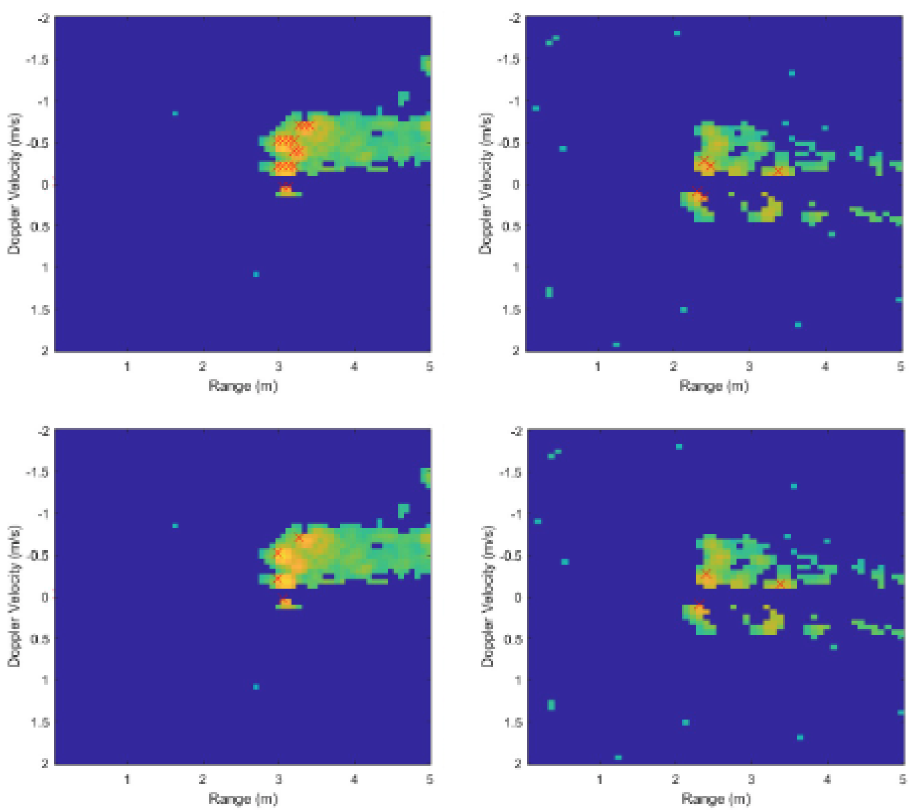

Figure 1: The above figure shows the effects of the FindPeaks and Euclidean distances code. The top 2 figures show the data after all the peaks have been found. To the bottom are the results of the narrowing down the peaks through the Euclidean distances code.

After finding all the maximum peaks, the Euclidean Distances method is applied. In this method, for every peak found, it checked the Euclidean distance between that peak and every other peak. For every peak that is within a certain distance of another (0.3m to 0.4m, depending on how many people we are optimizing detection for), it removed all but one. See Figure 1 for a visual on how these two algorithms worked visually.

Euclidean distance threshold	Find Peaks Threshold	Single Target Detection Accuracy	Two Targets Detection Accuracy	Three targets Detection Accuracy
0.3	0.95	0	1	0
0.4	0.92	1	1	0.22
0.3	0.9	0	0.5	0

Table 1: Accuracy results for the FindPeaks + Euclidean distances algorithms.

Additionally, after running tests, we discovered optimal Euclidean Distance and FindPeaks thresholds for 1 and 2 people. With the Euclidean threshold at 0.4 and FindPeaks threshold at 0.92, we can detect 1 and 2 people with near 100% accuracy (see Table 1). However, we had trouble finding optimal values for detecting 3 people (22% accuracy with the parameters above).

3.4 Clustering: K-Means Clustering with Elbow Method

Another clustering technique is the K-means clustering with the elbow method. K-means clustering is an interactive algorithm that attempts to split the dataset into a k number of clusters. It does this by randomly initializing K amount of clusters, then adjusting these clusters until each centroid is in equilibrium. The elbow method is used to determine the most optimal amount of centroids (K) for K-means. To do this, for every possible K-value, we calculate the sum of the squared distance between the centroids and each point. These are then plotted to determine the K-value where the graph begins to straighten out.

We experimented with different cutoff values for K-means clustering. For this experiment, the data was pre processed through 2 3x3 convolutions, then a filter to filter out values below (median power + max power)/2. As shown in Table 2, 0.88 was the optimal cutoff value for detecting 2 and 3 targets at a decent rate. This method was overall great for detecting 2 or 3 people, but not 1.

Cut Off Value	Single Target Detection Accuracy	Two Targets Detection Accuracy	Three targets Detection Accuracy
0.8	0.5	0.2	0
0.88	0	0.8	0.89
0.9	0	0.6	0.88
0.95	0	0	1

Table 2: Experimenting with different cutoff values for K-means clustering

We also experimented on the optimal preprocessing method. In these experiments, the K-means cutoff was set to 0.88. As seen in Table 3, the 3x3 convolution filter had the best results, with a 0.8 two target detection accuracy and 0.1 three target detection accuracy.

Preprocessing		Single Target Detection Accuracy	Two Targets Detection Accuracy
Conv 3x3	Filter	0	0.8
Conv 2x2	Filter	0	0.4
Filter	Conv 3x3	0	0.8
Filter	Conv 2x2	VERY BAD	

Table 3: Experimenting with different preprocessing methods with a k-means cutoff of 0.88

3.5 Clustering: Find Connected Components

Similar to K-Means Clustering, Find Connected Components looks for how many fully connected pixel (array element) clusters there are on the Range-Doppler plot whose power value exceeds a certain threshold. For this algorithm to work, the filtering threshold must be set tight enough as to eliminate any cluttering that connects two different target clusters.

Preprocessing		Single Target Detection Accuracy	Two Targets Detection Accuracy	Three targets Detection Accuracy
Filter with (median+max)/2	Conv 2x2	0.5	0.6	0.56
Filter with (median*7+max*5)/12	Conv 2x2	1	1	0.44
First filter + Conv 2x2	Filter 3 range + Conv 2x2	0.5	0.8	0.22

Table 4: Experimenting with different preprocessing methods

3.6 Comparison with Theory

After post-processing of removing zero Doppler, the wall became hardly visible. Theoretically, if a person moves perpendicular to the radar, it can be more challenging to detect them in comparison to those cases in which they are moving back and forth from the radar module. This was observed on the Doppler-range plots, where the signal didn't seem to move as much due to the relatively low velocities at which the people behind the wall walked. In addition, from the experiments, when one person passed right next to another person their radar signals appeared to be almost identical, producing detection ambiguity.

Additionally, in theory, we know that targets at a further distance have less reflected power. This can be observed in our Range-Doppler array. However, the difference is not obvious since the distance is limited to around 5m in the classroom. As a result, we found that filtering with the entire frame and filtering with respect to different ranges yield similar results.

4 Sensor Parameters Influence on Results

The main radar equations for range and velocity are:

- $d_{res} = \frac{c}{2B}$
- $d_{max} = \frac{c}{2S}F_s$
- $v_{res} = \frac{1}{2N_cT_c} * \frac{c}{f_{RF}}$
- $|v_{max}| = \frac{1}{4T_c} * \frac{c}{f_{RF}}$

(B: chirp bandwidth, F_s: IQ sampling frequency, N_c: number of chirps per frame, T_c: chirp periodicity, f_{RF}: center frequency of the chirp)

To minimize range resolution as well as decrease the maximum range as per our application, we experimented with different slopes, the number of chirps, and the sampling frequency. Initially, we increased the slope to decrease the range resolution (d_{res}) and the maximum unambiguous range (d_{max}). We saw an improvement in d_{res} as the bandwidth increased with slope. However, this adjustment had no effect on d_{max}. Our second attempt with lowering the IQ sampling rate (F_s) succeeded in decreasing d_{max}. However, as the IQ sampling rate decreases, more aliasing is introduced. Hence, there is a limitation to the smallest d_{max} we can reach without having too much aliasing, which is around 22m.

As for velocity, we attempted to increase the number of chirps to lower the velocity resolution (v_{res}), but we encountered a limitation of maximal 255 chirps imposed by the sensor itself. Nevertheless, increasing the chirp number and decreasing the sampling frequency both helped with lowering the Doppler velocity resolution.

Additionally, another consideration is angular resolution (SA), which is the minimum distance between two targets to be distinguished, and is determined by:

$$S_A = 2R * sin\frac{\theta}{2}$$

where θ is the antenna bandwidth (determined by sensor design) and R is the slant range from the target to the sensor. As targets move further away from the sensor, the bigger the angular resolution is, and the harder it is to distinguish between the two objects. In our experiment, since maximum target distance is limited, we do not have a clear observation of this phenomenon.

5 Limitations on System Performance

Based on the experiments conducted, it was observed that a cluttered environment, such as a crowded lab, can impede the detection capabilities of the system. This is because signal reflections from various parts of the scene can introduce environmental noise, hindering accurate detection. Furthermore, if two individuals are closer to each other than the resolution of the system, a range-Doppler ambiguity may arise, affecting the system's ability to distinguish between them.

If instead of having a wall made out of plastic, we have a material with a high dielectric constant such as metal, then the radar signal will be reflected on the surface of the wall and it wouldn't be possible to detect objects behind the wall. Additionally, the system faces limitations in detecting objects or people that move at speeds lower than the Doppler velocity resolution. Therefore, it is necessary to adjust the system parameters to suit the specific application requirements.

Apart from these, the radar also had constraints on the available bandwidth and the number of chirps which were taken advantage of. Furthermore, reducing the sampling rate leads to a trade-off between improved resolution and the undesired occurrence of aliasing. Another limitation lies in the accuracy of the post-processing algorithms employed, as their effectiveness varies depending on the number of people in the scene. This discrepancy arises due to the tuning of parameters specific to each clustering method utilized.

6 Future Work

One potential avenue for future development is the real-time operation of the system. In this scenario, the system's bandwidth and sensitivity would set theoretical limitations, as they determine its ability to swiftly capture and process data. Additionally, the algorithms presented in this study would require optimization to enhance detection accuracy and minimize computational requirements for data processing. Moreover, it is crucial to conduct further testing to calibrate the system for diverse environmental conditions outside the laboratory. Lastly, a real-time visualization interface should be developed to allow users to conveniently observe detected targets, their properties, and potentially classify them.

Furthermore, it is worth noting that due to the limitations of the lab setting, only a plastic wall was available for experimentation. As a result, we were unable to explore the effects of different wall materials. Speculatively, since the sensor's signal power is influenced by the material it traverses, various materials would exhibit distinct power returns. However, static signals would be eliminated with zero Doppler, minimizing the differences as long as the power is sufficient to penetrate the wall twice. Conversely, if the wall material is metallic, it would cause substantial attenuation and greatly diminish or even entirely eliminate the system's object detection capability. Additionally, greater material thickness would result in increased attenuation and reduced penetration depth. Therefore, wall thickness and material type would significantly impact the accuracy of detection and the required power for detecting objects behind the wall.

Another potential future direction is to use machine learning to detect objects in the video as a whole instead of individual frames. When the data is looked at as a whole video, the movement of the objects makes it much easier to discern how many people are detected. When each individual frame is analyzed, noise and aliasing make it much harder to detect the exact amount of people. To do this, we can use a variety of convolutional neural networks for video object detection, such as Fast R-CNN or YOLO. Some potential limitations for this method would include the higher computational power required and slower speed to process data.