People Counting and Tracking with mmWave Radar

Kwok Chun Fung Amos, Ng Chiu Fai, Mak Ka Ho
3035688368, 3035686956, 3035691195
{amoskwok, u3568695, u3569119}@connect.hku.hk
University of Hong Kong
Hong Kong, China

ABSTRACT

People counting and tracking jobs has long been a famous topic in the field of machine learning. However, the privacy concerned in addition to the cost associated with data collecting and model training are severely hindering its availability. As a result, we present a pipeline that leverages low-cost mmWave radar sensor and perform CFAR detection and clustering on filtered range-angle heatmap to output the number of people as well as their position relative to the sensor in our project. In the experiment setting of 0-3 people standing in front of the sensor, our solution achieves mean absolute error of 0.425 for people counting and a RMSE of 0.069 m for position approximation.

1 INTRODUCTION

MmWave is a band of spectrum with wavelengths ranging from 10 millimeters (30 GHz) and 1 millimeter (300 GHz). This characteristic of having such a high frequency yields several advantages in general. For example, mmWave is less prone to interference, has large bandwidth and thus high data transfer rate, has narrow beams and thus great resolution, and finally, it is not prone to extreme weather conditions. All these conditions point to the fact that mmWave is a brilliant tool for people counting and tracking jobs. This is further proven by the research done by Texas Instruments [1]. In their experiment, they made use of velocity information to filter out static objects and configured two squares on the floor to test if the sensor is accurate enough to report correctly if each person is inside or outside the two squares. Another similar research on people counting using FMCW radar [2] introduced a sample flow of signal filtering and the idea of density-based clustering which will also be adopted by our pipeline.

In our project, we performed people counting and tracking by processing and analyzing mmWave signals. We first employed Blackman Harris window to filter raw signal, which is then followed by a Fast Fourier Transform (FFT) and a moving average filter on the frequency spectrum. We implemented Capon Beamforming for angle finding, and Constant False Alarm Rate (CFAR) for peak extraction. Finally, we performed density-based clustering for people counting and position estimation.

2 DESIGN

2.1 Hardware

This experiment is based on DCA1000 with xWR mmWave Sensor for data collection.

Description	
Start frequency (GHz)	77
Frequency Slope (MHz/μs)	29.982

Sampling rate (ksps)	10000
Sample per chirp	256
Number of chips	128
Number of Rx	4
Number of Tx	2
Range Resolution (m)	0.19

Table 1. Parameters setting of DCA1000 with xWR mmWave Sensor

2.2 Filtering

Filtering is a common technique in digital signal processing to filter out noise. In this experiment, we have specifically applied Blackman-Harris window and simple moving average filter.

Minimum 4-term Blackman-Harris window function is a generalization of Hamming family which can achieve a side lobe level of -92 db [3,4]. Its feature to minimize side lobe levels in signal data is effective in reducing amplitude of side lobes because of reflections and interferences of signal responses from target objects.

Simple moving average filter smooths noise fluctuation in signal by taking averages of data in the sliding window. Given that the sensor may capture unnecessary signal response from reflection of objects like tables, chairs, walls, etc., the moving average filter is able to let us focus on human targets that have more significant responses due to small gestures and breathing movements.

2.3 Fourier Transform

Fourier Transform is a common technique to analysis the frequency spectrum of signal data. In the FMCW radar, the intermediate frequency (IF) signal generated by mixing signals transmitted from Tx and captured by Rx. The frequency of IF signal can reveal the range information of detected objects by equation (1):

$$(1): d = \frac{f * c}{2 * S}$$

where d is the distance of object from sensor, f is the frequency of IF signal, c is the speed of light and S is the frequency slope of the chirp signal.

2.4 Capon Beamforming

Beamforming in determining angle of arrival can be understood as comparing the resultant beam that formed by addition of signals in different antennas after adapting to different delays.

In capon beamforming algorithm, the sensor array signal given by:

$$(2): X(t) = A(\theta)s(t) + n(t)$$

where $A(\theta) = [a(\theta_1), \dots, a(\theta_M)]$ is the steering matrix, $a(\theta_n)$ is the steering vector pointing to θ_n and M is the number of angle bins. The angle is finally calculated as:

$$(3): \theta_{capon} = argmin\{trace(A(\theta) * R_x^{-1} * A(\theta)^H)\}$$

where R_x is the spatial covariance matrix of x in a range bin, $A(\theta)^H$ is the conjugate representation.

2.5 Constant False Alarm Rate (CFAR)

CFAR is an adaptive peak finding algorithm. Since the magnitude of frequency response for different targets can be different and the scale of magnitude for different data samples are different, a fixed noise threshold can hardly adapt to every case. CFAR therefore determine the noise level of each cell by its training cells and ignore guard cells to avoid spectral leakage.

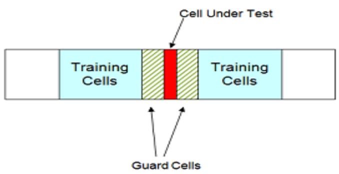

Figure 1. Training and Guard cells example

In this experiment, we applied OS-CFAR that choose the kth element of the sorted levels of training cells as noise level.

2.6 Density-Based Clustering (DBSCAN)

Clustering is introduced to cluster close data points as a human target. DBSCAN form cluster based on density, which is determined by the number of neighbors within the range of ε. It also has the advantage that number of clusters need not to be defined, allowing us to apply to unknown situations.

3 IMPLEMENTATION

3.1 Experiment Setting

In our experiment, we gathered data for four different cases, ranging from 0 to 3 people. For each of these cases, 10 data samples where we stood at slightly varying distances and angles in front of the sensor are collected.

3.2 Pipeline

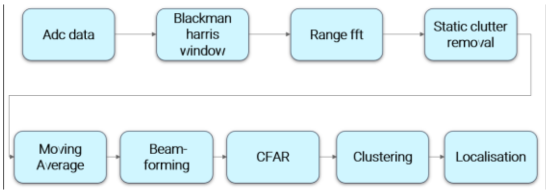

Figure 2. Overview of our solution pipeline

4 EVALUATION

The accuracy of our pipeline was evaluated by comparing the ground truth with the output from density-based clustering.

We began with analyzing the results with signal with no filtering. As shown in figure 3, the result was not satisfactory. The accuracy fluctuated between 0.1 and 0.4 for individual cases, and the total accuracy was only 0.25. Figure 4 also shows the augmented accuracy, where we consider an error window of ±1 person as suggested by the paper [1]. Nevertheless, the augmented accuracy was still only floating between 0.4 and 0.7, with a considerably high mean absolute error (MAE) of 1.425.

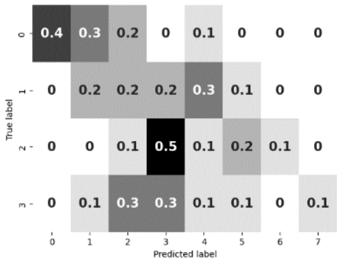

Figure 3. Confusion Matrix of Unfiltered Signals

	Augmented accuracy (consider +-1 difference)	MAE
0 person	0.7	1.1
1 person	0.4	1.9
2 people	0.6	1.7
3 people	0.7	1.2

Average MAE = 1.475

Figure 4. Augmented Accuracy and MAE of Unfiltered Signals

Having recognized the unsatisfactory outcomes from raw signals, we proceeded to interpret the results generated from filtered signals. For our proposed implementation, the raw accuracy has greatly improved to 0.6. In addition, the augmented accuracy almost reached 1 for individual cases, with a comparatively much lower MAE of 0.425. Thus, we could conclude that filtering the raw signals could tremendously elevate the reliability of the signals.

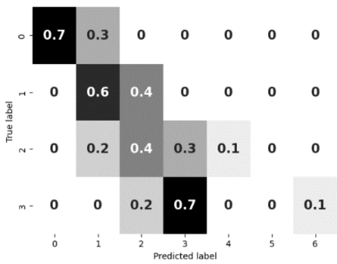

Figure 5. Confusion Matrix of filtered Signals

	Augmented accuracy (consider +-1 difference)	MAE
0 person	1.0	0.3
1 person	1.0	0.4
2 people	0.9	0.7
3 people	0.9	0.5

Average MAE = 0.475

Figure 6. Augmented Accuracy and MAE of filtered Signals

Finally, we performed position estimation on people that are tracked by our pipeline. For each person, by calculating the cluster centroid, we obtained the coordinate of that person with respect to the sensor. An example is illustrated below. We had a sample with 1 person standing 1.4 meters in front of the sensor. After processing the signal of the respective samples and applied clustering, the predicted coordinate of the person is plotted in pink. We could see that the pink dot is close to the ground truth (0, 1.4), concluding that our pipeline is reliable on position estimation of people.

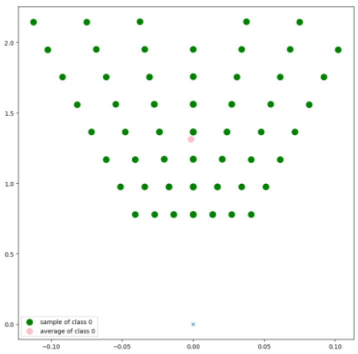

Figure 6. Position Approximation of 1 person example

5 CONCLUSION AND LIMITATIONS

This project successfully found a way to distinguish the number of people within a distance range. After filtering and FFT, beam forming is then performed to obtain the angle of arrival while CFAR is used to perform the peak finding. After CFAR, the result obtained will be clustered to approximate the target location.

During the process, there are a series of limitations. First, given that our experiment data is limited to 3 people, we cannot guarantee the performance in situation with more people. Therefore, further experiment needs to be conducted on more people to see if the algorithm in this project can be effective in all situations. If there are more people to be tested, the problem of body blockage also needs to be considered. For the clustering, there are no big problems when the number of people is small. However, as the number of people grows, there may be problem such as falsely combining relatively close clusters to 1. Second, the environment setting may cause bias to the experiment. For example, there are other users in the testing environment and the reflection of different materials also affect the experiment accuracy. The target object distance is also an issue. On one hand, people standing farther away may seem smaller to the sensors. For these objects, the reflected signals received by the antennas become very weak both in intensity and in energy. Also, the reflected signals of these distant objects may travel in the direction which cannot be captured by the view of the sensors.

In the future, to further improve and upgrade the whole system, the project will be extended to the detection of a crowd of moving people by further applying Doppler FFT. Extended Kalman Filter can also be introduced to realize live tracking of people.

CONTRIBUTION STATEMENT

- Amos: research current solution & available library, raw adc data processing, range-angle heatmap, cfar & dbscan implementation, filter & cfar parameters tuning
- Mak Ka Ho: research current solution & available library, data processing, filtering, and clustering
- Ng Chiu Fai: Research on current solution and available library, data processing, filtering and result interpretation
- Leverage the library functions and referenced capon-beamforming and CFAR implementation from PreSence Team [5]

REFERENCES

[1] Texas Instruments. “People Tracking and Counting Reference Design Using mmWave Radar Sensor”, March 2018

[2] J. Weiß, R. Pérez and E. Biebl, "Improved People Counting Algorithm for Indoor Environments using 60 GHz FMCW Radar," 2020 IEEE Radar Conference (RadarConf20), Florence, Italy, 2020, pp. 1-6, doi: 10.1109/RadarConf2043947.2020.9266607

[3] Smith, J.O. " Blackman Harris Window Family", in Spectral Audio Signal Processing, http://ccrma.stanford.edu/~jos/sasp/Hamming_Window.html, online book, 2011 edition

[4] Harris, F. J. "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform." Proceedings of the IEEE. Vol. 66 (January 1978). pp. 51-84.

[5] Pan E, Tang J, Kosaka D, Yao R, Gupta A. OpenRadar. GitHub repository. Published online 2019. https://github.com/presenseradar/openradar