Exploring gesture recognition with low-cost CW radar modules in comparison to FMCW architectures

Alan Bannon, Richard Capraru, Matthew Ritchie
Electronic and Electrical Engineering Department
University College London
London, WC1E 7JE
Email: alan.bannon.15@ucl.ac.uk

***Abstract*—Radar-based hand gesture recognition is an area receiving a significant amount of interest in recent years due to the rapid increase in the availability of low-cost low-footprint RF sensors. The most common configuration is Frequency Modulated Continuous Wave (FMCW), whereas Continuous Wave (CW) radar is not receiving as much attention. In this paper we explore the use of extremely low cost CW radar modules for gesture recognition. In doing so a set of signal processing electronics is developed, implemented, and used to supply the resulting signal to PC audio input for recording. A dataset of gestures was recorded and gesture recognition accuracy was compared to FMCW recordings to show that CW systems can provide a high accuracy for gesture recognition at a very low cost.**

I. Introduction

Radar systems have origins in applications that focus on the ability to detect large targets at long ranges. This typically requires complex systems that, due to cost, are not widely manufactured or accessible for low cost sensing activities. With the advent of modern electronics and signal processing it is now possible to create very compact RF sensor modules that are able to sense subtle movements over short ranges, resulting in exciting new application areas.

The current significant push for autonomous vehicles is producing a need for compact radar sensors that provide part of the vision solution of a self-driving car. Additional new areas of radar sensing include both medical and human machine interaction (HMI) tasks that can be performed with low cost radar modules. Gesture recognition is part of these emerging areas of radar research, an example project, Google's Soli radar, being one particularly high-profile case [1].

Micro-Doppler is the additional signatures imparted onto the reflected signal back to the radar that a target generates. This movement creates a signature which was coined as Micro-Doppler by researcher V. Chen [2]. These signatures are in addition to the bulk velocity and are created by vibration, rotation and other subtle movements. For example a person may walk forwards at 3 m/s but as they move at this speed their arms and legs oscillate back and forth. Gait analysis has in fact been researched extensively as a suitable application of Micro-Doppler [3], [4]. In addition there has been a recent growth of research evaluating the use of Doppler data to monitor human vital signs without the need for continuous contact between the sensor and the subject [5].

Micro-Doppler has recently been considered as a means of recognising hand gestures. [6] evaluated the ability of a Deep Convolutional Neural Network to recognise hand gestures with a 24 GHz Frequency Modulated Continuous Wave (FMCW) radar module with accuracies of 99% shown. Other researchers looked to apply a interferometric radar sensor in order to distinguish hand gestures [7]. The information found from the interferometric configuration were found to be useful as part of the discrimination process although further work is needed to fully quantify their effectiveness.

Continuous Wave (CW) radars are used for some niche applications that require the ability to sense the movement but not the range of the target. They operate by sending only a single tone and measuring any shift in the returned signal that is caused by the Doppler effect of a moving target. For example police speed cameras can apply this technology to simple extract the velocity of a car that is in the main focus of the sensor. Hand gesture recognition may be a task well suited to CW radar sensing, as it is often true that the hand gesture is the dominant target within the antenna beam and the actual range to the hand may not be necessary, depending on the movements selected. As part of this work a compact CW radar module was selected and integrated into a USB or battery powered device that output signals that can be sampled by a audio card in a PC. The module selected was a CDM324 which is a 24 GHz motion sensor device that only has one transmit and one receive channels.

Previous work has explored the classification of hand gestures using a 24 GHz FMCW radar with 2 GHz of bandwidth [8]. Following on from this analysis evaluated the effect of removing and spatial information once the gesture had been identified in range [9]. This showed high success rates for classification even when no range information was utilized as part of the classification process.

This work is part of a wider push to generate a shared repository of micro-Doppler signature that can be utilized by the radar research community to improve classification techniques. A recently launched website Dop-Net.com has been setup to allow the exchange of radar datasets and the data shown within this publication will be made available there [10].

This paper aims to introduce a simple extremely low cost integrated CW radar module that has been applied to the role of recognizing various human hand gestures. These results are then compared to FMCW radar measurements of the same hand gestures to contrast the success rates for the two different type of radar architectures. Section II introduces the hardware used and how it was developed into a standalone module. Section III shows the data captured and the results of the classification process. Finally Section IV concludes from the findings shown and discusses future work.

II. Hardware

A. CW Modules

Recently there has been an explosion in the availability of extremely low cost CW radar sensors. Frequently the end use case is presence or motion sensing; performing the same role as the older Passive Infrared (PIR) sensors often used in intruder prevention systems. Radar hardware, however, produces far more data than comparable IR sensors.

One such CW radar sensor is the CDM324, see Fig 1. Available from many suppliers and frequently cloned, the CMD324 operates at around 24GHz and provides a single in-phase output channel. The sensor operates on a supply voltage of 5V, making it simple to integrate with many other systems. By far the most outstanding feature of the sensor is the price; the CMD324 can frequently be found for as little as $6. The sensor is fundamentally similar to that presented in [11], although given the lack of documentation available it is not possible to draw a link between that work and this specific module.

Frequently one of the most costly components in a radar system is the Analogue to Digital Converter (ADC) used to digitize the signal for further processing. However, as the CDM324 outputs the raw beat frequency, the range of frequencies of interest top out at around 1kHz. At these frequencies we can consider the use of standard consumer hardware used for audio recording, such as the line-in available on almost all PCs. Thus the need for a dedicated ADC for our system can be completely avoided, further simplifying the system and reducing cost. However, for this approach to be viable, the very low voltage level signals the CDM324 outputs contain high frequency and near-DC noise, and so some signal processing electronics has been implemented to perform this function.

B. FMCW Radar

The comparative FMCW radar system that has been used for this work is the Ancortek SDR-KIT 2400AD2. This is a 24 GHz devices that has up to 2 GHz bandwidth (although this was set to 750 MHz for the data used within this publication) and was set to a 1 ms chirp period. It has +13 dBm power and used 14 dBi horn antennas. The sensor has a standalone GUI to control and capture data or can be commanded within a Matlab interface to capture signals. The radar has one transmit antenna and two receive antennas, but only the co-polarised (H-pol) channel was used for the purposes of this dataset. For data gathering repeated hand gestures were made approximately 30-40 cm away from the radar over a long continuous period. The single data file generated was then cut into individual gesture actions for feature extraction and classification processing.

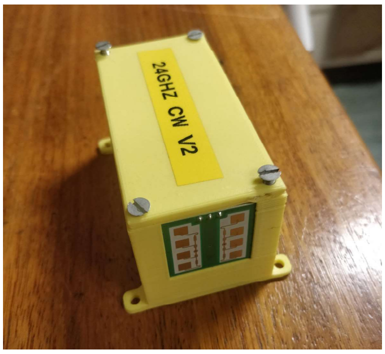

Fig. 1: CDM324 CW Radar Module within 3D printed housing

C. Signal processing electronics

In order to achieve the goal of recording the raw in-phase beat frequency signal with commercial audio recording equipment the signal must be filtered of high frequency and near-DC noise and amplified 30 dB.

The upper cut-off required for the filter can be determined by considering the Doppler equation, see Equation 1. Given the 24 GHz transmission and an angle of incidence relative to motion of 0, we find that the upper frequency limit is around 100 Hz. This limit was experimentally confirmed by observing the unprocessed signal of several gestures and noting that the signal over interest never exceeded around 80 Hz.

$$f_D = 2f_0 \cdot \frac{v}{c_0} \cdot \cos\alpha \tag{1}$$

where:

f_d = resultant Doppler frequency
f_0 = incident wave frequency
v = velocity of observed object
c_0 = speed of light
α = angle between incident wave and observed object's motion

The lower frequency cut off was simply set to be as close to DC as practically possible in order to capture very slow motions, while still eliminating the majority of near-DC noise. In order to both amplify the signal, and provide the necessary filtering, two identical cascaded active band pass filters are used. With upper and lower cutoff frequencies set to 4.7 Hz and 102 Hz respectively, both with a gain of 15 dB. Cascading the filters provides a higher Q-factor, lowers the gain requirement of each individual filter and compensates for the 180 degree phase offset generated in each filter.

Thus the overall effect is a band-pass filter with a passband of 4.7 Hz to 102 Hz with a gain of 30 dB. Finally, this results in a signal that is of the correct magnitude and frequncy to be digitise by commercial audio recording equipment, see Figure 4.

D. Recording units

To aid data gathering the sensor and signal processing electronics was housed in a 3D printed case with power provided either by a USB socket or battery, see Fig 1. The processed signal was provided via a 3.5 mm audio out connection into a PC audio line-in.

E. Extending to full complex capture

Fundamentally, a major drawback of the CDM324 is that it provides only a single in-phase output. Thus while the speed of motion in the view of the sensor may be determined, the direction of motion cannot be determined. To change this we would need both the in-phase and quadrature-phase outputs for a full complex IQ signal capture.

The RSM2650 sensor provides this in an almost identical form factor to the CDM324. Again running at 5V, and costing in the region of $25, the RSM2650 provides two outputs for in-phase and quadrature-phase signals. The signal processing electronics for the RSM2650 are simply a duplication of that used for the CDM324 to cover both output channels, see Fig 2.

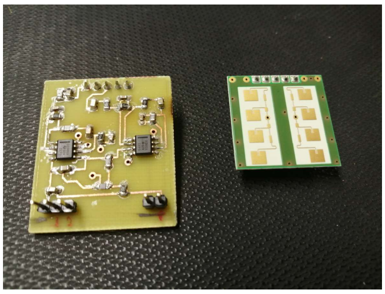

Fig. 2: RSM2650 CW Radar Module within signal processing electronics

III. Data and Results

A. CW recorded hand gestures

In order to evaluate how effective this radar is in recognising gestures, a new database of 5 individuals completing 4 separate gestures with over 400 repetitions was recorded. The 4 gestures were chosen to match an existing database of gestures captured using the Ancortek FMCW radar. A comparison of traditional FMCW type of Radar (which is capable of providing both range and Doppler information of targets) and the proposed newly integrated CW module radar are the main focus of this papers outputs. A key consideration is that the CW module used is much less complex and costly than the FMCW equivalent.

For the CW dataset each gesture was made directly in front of the radar at a distances of approximately 30 cm - 40 cm, which is the same as the FMCW geometry used. The system was then initiated to capture 3 seconds of 1 gesture from

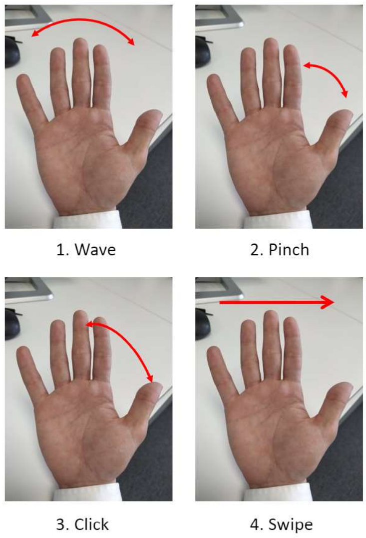

Fig. 3: Gestures 1. Wave; 2. Pinch; 3. Click; 4. Swipe

a participant who was sitting in front of the device. The 4 gestures that were recorded, which matched those used in the FMCW dataset, were a Wave (3 waves of the arm and hand), Click (a single action between thumb and 2nd finger), Swipe (between thumb and index finger) and Pinch (using the whole arm and hand); seen in Fig. 3. A few key features about the gestures are the following. The waving gesture has an oscillatory shape and longer duration. The click gesture happens over the shortest time frame (as a click is only a short sharp action). Then the pinch and swipe actions do show some level of similarity which could make them challenging for a classifier.

B. Classification of signals

The database generated initially included 4 x 80 separate files from all the participants and all the repetitions. These participants are be labelled as A to D from now on. Each file was processed to produce a raw beat frequency signal, see Fig. 4.

This beat frequency signal was then processed in order to extract the required features in order to successfully classify the different actions. Five features were extracted from each gesture examples. The features are: Standard Deviation, Entropy, Median, Maximum and Mean of the Summed Singular Values. The summed signal values were obtain via Singular Value Decomposition (SVD), of the input signal. This method has been proven to be effective on micro-Doppler signal classification and is defined in detail within [8], [9].

The classification processing applied to the obtained features was to use four separate classifiers. The classifiers selected for this paper were Linear and Quadrature Discriminate Analysis (DA) methods, Bagged-Tree mechanism and a Quadratic Support Vector Machine (SVM). These were selected as from a quick look classification check using the MATLAB Classification Learner utility. They are also fairly

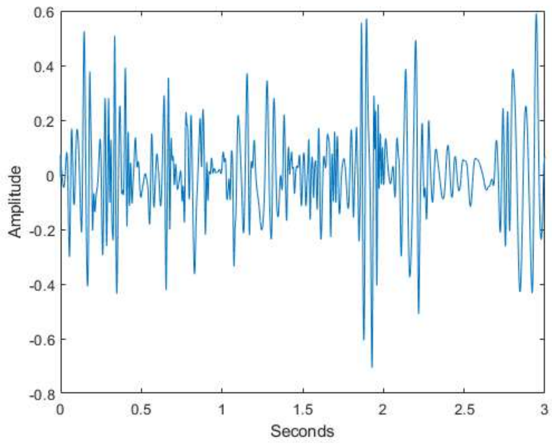

Fig. 4: Raw Beat Frequency Signal

simple classifiers that do not require extensive training or computational load.

Each classifier was trained using a 90% random subset of the features provided and then tested on the remaining 10% of data excluding this training set. The classification process was repeated 100 times to produce an average result. The best classification result of 84.1% from the single range bin data was generated by the co-pol data with the Quadratic SVM classifier.

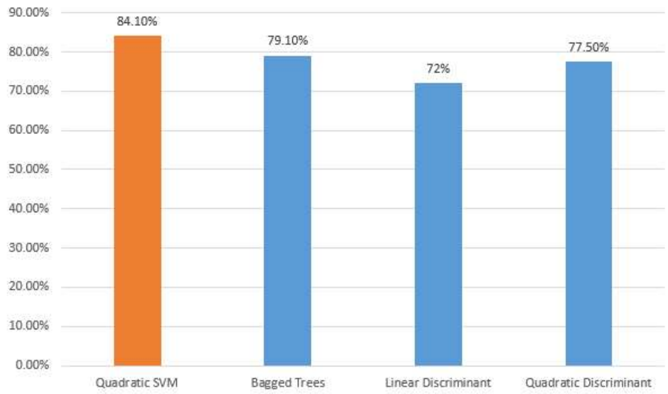

Fig. 5: Results from CW radar gesture classification using 4 different classifiers

C. Comparison of previously generated FMCW data

Prior research used an FMCW Radar to capture hand gestures, see II-B for details. Each gesture was made in the same way as with the CW radar. The raw data capture was processed into individual time windows for feature extraction as completed for the CW dataset. These individual gesture actions have varying matrices sizes hence a cell data format was used to create a ragged data cube. The data that has been shared as part of this challenge was created by the following flow of pre-processing:

- Divide vector of raw samples into a 2D matrix of chirp vs. time.
- FFT samples to convert to the range domain. Resulting in a Range vs. Time matrix (RTI)
- Apply a Moving Target Indicator (MTI) Filter signal to suppress static targets
- Extract range bins within the MTI data that contain the gesture movement and coherently sum these.
- Generate a Doppler vs. Time 2D matrix by using a Short Time Fourier Transform on the vector of selected samples.
- Store the complex samples of the Doppler vs. Time matrix within a larger cell array which is a data cube of the N repeats of the 4 gestures from each person.

The FMCW dataset of was analysed using a series of different classifiers, the results are shown within Fig. 6. The best result that was obtained using the FMCW radar was 81.4% from the single range bin data was generated by the co-pol data with a Bagged Trees classifier [8].

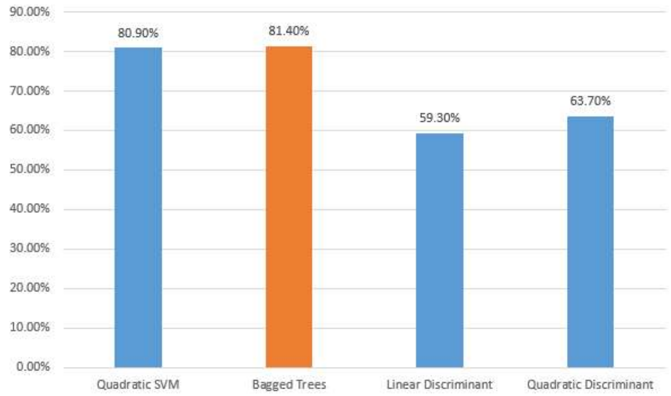

Fig. 6: Results from FMCW radar gesture classification using 4 different classifiers

It is very clear that the classification using a CW radar was slightly more accurate in the best case by approximately 3%. Additional benefits include the simple hardware and signal processing flow compared to the FMCW methods.

We believe the fundamental advantages of CW systems over FMCW systems in this application are threefold: decreased cost, decreased processing overhead, and a large capture envelope not limited to a particular set of range bins.

IV. Conclusions and Future work

This paper has shown the integration of a low cost COTS CW radar module into a gesture recognition sensor. It was hypothesised that if the gesture was the dominant signal measured by the sensor that comparable performance could be achieved between CW and FMCW sensors. The initial results here show this is valid for a selection of 4 key gestures. This is significant as one of the key advantages of CW modules is their simplicity and extremely low cost.

Future work will look to expand this into a real-time classifier that aims to constantly evaluate the raw beat frequency signal in front of the radar and then take simple features from the signal to continually update the predicted gestures that are occurring directly above the RF sensor. This will be achieved by using a Raspberry Pi module with external sound card to sample the signals.

In addition a complex (I/Q) CW module that provides direction and speed will be use for comparison to the non-complex module used to gather the data shown here. It is hoped that this additional degree of freedom within the data will provide more information about the Doppler signal, which we hypothesise will improve the classification accuracy, especially in the case of blind classification of new signals.

ACKNOWLEDGEMENTS

The authors would like the thank Colin Horne, Dr. Riccardo Palama', Alvaro Arenas Pingarron and Folker Hoffman for their support in data collection for this research.

REFERENCES

[1] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graph.*, vol. 35, no. 4, Jul. 2016. [Online]. Available: https://doi.org/10.1145/2897824.2925953

[2] V. Chen and H. Wechsler, "Micro-doppler effect in radar: phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, 1 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1603402

[3] D. Tahmoush and J. Silvious, "Radar Micro-Doppler for Long Range Front-View Gait Recognition," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, 2009.

[4] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Personnel Recognition and Gait Classification Based on Multistatic Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 669–673, 2018.

[5] T. Hall *et al.*, "Non-contact sensor for long-term continuous vital signs monitoring: A review on intelligent phased-array doppler sensor design," *Sensors (Basel, Switzerland)*, vol. 17, no. 11, p. 2632, Nov 2017.

[6] B. Dekker, S. Jacobs, A. S. Kossen, M. C. Kruithof, A. G. Huizing, and M. Geurts, "Gesture recognition with a low power FMCW radar and a deep convolutional neural network," *European Microwave Week 2017: "A Prime Year for a Prime Event", EuMW 2017 - Conference Proceedings; 14th European Microwave Conference, EURAD 2017*, vol. 2018-Janua, pp. 163–166, 2018.

[7] E. Klinefelter and J. A. Nanzer, "Interferometric radar for spatially-persistent gesture recognition in human-computer interaction," *2019 IEEE Radar Conference, RadarConf 2019*, pp. 0–4, 2019.

[8] M. Ritchie, A. Jones, J. Brown, and H. D. Griffiths, "Hand gesture classification using 24 GHz FMCW dual polarised radar," *International Radar Conference*, vol. 2017, no. CP728, pp. 1–6, 2017.

[9] M. Ritchie and A. Jones, "Micro-Doppler Gesture Recognition using Doppler, Time and Range Based Features," in *IEEE Radarcon*, Boston, 2019.

[10] M. Ritchie, R. Capraru, and F. Fioranelli, "Dopnet: A micro-doppler radar data challenge," *Institution of Engineering and Technology, Electronics Letters*, February 2020.

[11] F. Alimenti *et al.*, "A low-cost 24GHz Doppler radar sensor for traffic monitoring implemented in standard discrete-component technology," *Proceedings of the 2007 European Microwave Conference*, pp. 1441–1444, 2007.