Measurement-based Rotational Hand Gesture Recognition Using mmWave MIMO FMCW Radar

Jelle Wilbrink

***Abstract*—Contact-free gesture recognition can be used to convey commands to computers, allowing for more user-friendly Human-Computer Interactions than traditional input devices, like remote controllers. This paper proposes methods to estimate the radius and period of periodic rotational gestures measured using millimeter-wave (mmWave) radar based on the Doppler-Time Map (DTM), Range-Time Map (RTM) and a combination of the DTM and RTM. Following the literature, a 3D Fast Fourier Transform is applied to the radar data, yielding the DTM and RTM. These are then further processed using existing techniques, like subtracting the mean over time from the RTM to remove static objects and constant false alarm rate detection to find the approximate distance between the center of the gesture and the radar. The accuracy and robustness of these methods, that combine existing processing steps in a novel way, are determined using a set of measurements of various radii and periods of the hand's circular trajectory at different distances and angles with respect to the radar. Virtual Reality glasses are used to establish a ground truth for the measurements. It is found that the combined method using both the DTM and RTM is most robust for radius recognition with 4.4cm average error. Besides, the estimate for period recognition is found accurate in 83% of the scenarios for all methods.**

I. Introduction

SMART objects are seemingly everywhere nowadays. These objects, ranging from curtains to light bulbs, have a computer in them, which needs to be controllable by the users. Traditional mobile input devices like remote controllers are not ideal for this, as they have to be picked up and touched in order to convey information to a computer. An alternative is provided by contact-free interaction methods, such as voice commands or gesture recognition, where users interact with a fixed controller at a distance [1]. Cameras or radio devices, like radar systems, can be used for gesture recognition. Cameras have high recognition accuracy, but they depend on lighting and atmospheric conditions. They may also raise privacy concerns [2]. Radar systems do not have these drawbacks, but generally have worse resolution than cameras.

In [1], [3], [4], [5] and [6], mmWave radar is used for gesture recognition, focusing on distinguishing between different gestures from a set of both linear and rotational motions. This paper will look into the recognition of variations in radius and period of a single periodic hand gesture, since such variations can be used to control computers, but have not yet been investigated. A periodic circular hand gesture of the hand in the elevation-range plane, as seen from the radar, is chosen, because a circle performs a sinusoidal motion over time along both axes of this plane. So, if a circle can be characterized, the same algorithm is expected to work for other periodic motions, like moving back and forth a line.

A. mmWave radar theory

A radar works by first transmitting a signal, which is then partially scattered back by objects in space. This back scattering is received by another antenna on the radar.

This paper will use millimeter-wave (mmWave) radar, specifically the TI IWR1843BOOST radar in combination with the DCA1000 data capture adapter [7]. This radar has a carrier frequency f_c of 77GHz and a maximum bandwidth of 4GHz. It has 4 receivers (Rx) and 3 transmitters (Tx), forming an array of Multiple Input and Multiple Output (MIMO) antennas, as shown in Fig. 1.

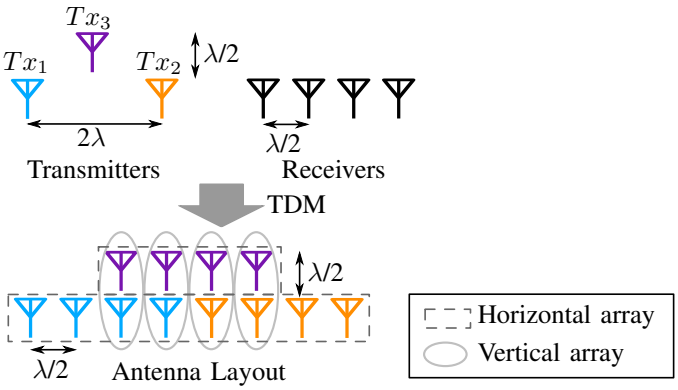

Fig. 1: Antenna configuration of TI IWR1843BOOST and antenna layout after TDM for $N_{Tx} \times N_{Rx} = 3 \times 4$. Adapted from [8].

The radar used in this paper is a so-called Frequency Modulated Continuous Wave (FMCW) radar, indicating that it sends out chirps. These are signals that linearly increase their frequency over time during period T_c. By mixing the transmitted and received chirps together, the beat frequency f_{beat} is obtained [1]. Each chirp is sampled with a number of N_{ADC} samples. N_c chirps of a single transmitter together form a frame with duration T_f and each measurement contains N_f frames.

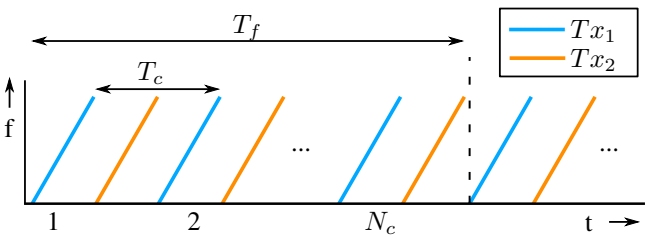

Fig. 2: Chirps transmitted by the radar.

By Time-Disivion Multiplexing (TDM) of the Tx antennas (Fig. 2), antenna layouts can be formed with combinations of N_{Tx} transmitters and N_{Rx} receivers in the following set: $N_{Tx} \times N_{Rx} = \{1 \times 4, 2 \times 4, 3 \times 4\}$ (Fig. 1). It will be assumed that the different antennas are receiving their signals simultaneously. However, as Fig. 2 shows the chirps are separated by the time T_c between two consecutive chirps from the same transmitter. The target should be approximately static for time T_c for this assumption to hold. This is the case for the measurements in this paper, as the change in velocity, distance and angle within the period of one chirp is much smaller than the resolution that can be detected.

Using the data from the radar, the radial distance, radial velocity and angle with respect to the radar can be measured. Below their relation to the physical parameters of the radar [9] will be given, because they play a role in the estimation of the radius of the gesture. The range resolution and maximum range are given in (1) and (2), respectively. In these equations, c denotes the speed of light, B the bandwidth in Hz, F_s the ADC sample frequency and S the slope of the chirps in Hz/s.

$$d_{res} = \frac{c}{2B} \tag{1}$$

$$d_{max} = \frac{f_s c}{2S} \tag{2}$$

$$v_{res} = \frac{\lambda}{2T_f} \tag{3}$$

$$v_{max} = \frac{\lambda}{4T_c} \tag{4}$$

For velocity, the resolution and maximum are given by (3) and (4). Here, the wavelength is denoted by $\lambda = \frac{c}{f_c}$ [10].

$$\theta_{res} = \frac{\lambda}{Nd\cos(\theta)} \tag{5}$$

$$\theta_{max} = \sin^{-1}\left(\frac{\lambda}{2d}\right) \tag{6}$$

The maximum measurable angle is given in (6). N denotes the number of antennas $N_{Tx} \times N_{Rx}$ and d the distance between them. The antennas of the IWR1843BOOST are spaced by $d = \frac{\lambda}{2}$ (Fig. 1), yielding a field of view of $\pm 90\,\mathrm{deg}$. Using (5), this results in an optimal angular resolution of 14.2° at $\theta = 0^\circ$, but the resolution gets worse with an increasing angle.

So far, factors for the radius recognition have been discussed, but the frequency of the gesture will also be estimated. According to the Nyquist sampling rate, the frame frequency (i.e. $\frac{1}{T_f}$) should be at least twice the frequency of the gesture [11].

B. Processing

The processing of the measured radar data can be divided in three stages: sensor data representation, feature extraction and detection [12]. These will be discussed next.

1) Sensor data representation: The first processing steps are the same in [1]-[6], [10] and [12]. As mentioned, the data measured by the radar is grouped in ADC samples, chirps and frames for each antenna. For each measurement this results in a data matrix with dimensions $(N_{Rx} \cdot N_{Tx}, N_{ADC}, N_c, N_f)$. Performing Fast Fourier Transforms (FFTs) along the ADC sample-, chirp- and antenna-dimensions of the matrix yields distance, velocity (i.e. (micro-)Doppler) and angle information, respectively. The frame dimension is unaltered and contains time information.

Now, the points in the matrix represent the intensity of the back scattered signal. However, the corresponding axes are in terms of f_{beat} or phase shift ω, so (7), (8) and (9) [9] are used to obtain the intensity of distance, velocity and angle, respectively:

$$d = \frac{f_{beat} c}{2S} \tag{7}$$

$$v = \frac{\lambda \omega}{4\pi T_c} \tag{8}$$

$$\theta = \sin^{-1}\left(\frac{\lambda\omega}{2\pi d}\right) \tag{9}$$

After the 3D-FFT has been performed on the raw data, the complex-valued data is converted to magnitude in dB. Then it is passed on to the feature extraction stage.

2) Feature extraction: Unlike the similarity in sensor data representation, a variety of different features is extracted in literature. For example, slices from the data matrix are directly used as features, like a 2D-matrix of Doppler or angle against range for each frame in time [6]. The so-called Range-Doppler Map (RDM) is also used in [1] and [2]. Range-Time and Doppler-Time Maps (RTM and DTM, respectively) can also be used as features [12]. Additionally, a number of features are listed in [13], subdividing them in three categories: explicit tracking of scattering centers, low level descriptors of the radio frequency environment and data-centric features.

3) Detection: Various types of Artificial Intelligence (AI) are used to detect gestures in [1]-[6], [10] and [12], like deep learning [3] or support vector machines [5].

Properties of periodic gestures, like the one in this paper, can be extracted using discrete signal processing techniques and without AI. Since the trajectory of the hand is a circle in the elevation-range plane, the radius r_g and period $T_g = \frac{1}{f_g}$ of the gesture can directly be estimated using distance, velocity or angle information over time, which are all obtained from the the signal processing stage.

The resolution $r_{res,d}$ for estimating the radius of the gesture using distance information is equivalent to d_{res} in (1), since both are distances along the same axis. Velocity is the time derivative of distance, so scaling v_{res} from (3) yields the resolution $r_{res,v}$ of the gesture radius that can be detected using velocity information:

$$r_{res,v} = \frac{1}{2\pi f_g}\frac{\lambda}{2T_f} = \frac{\lambda}{4\pi T_f f_g} \tag{10}$$

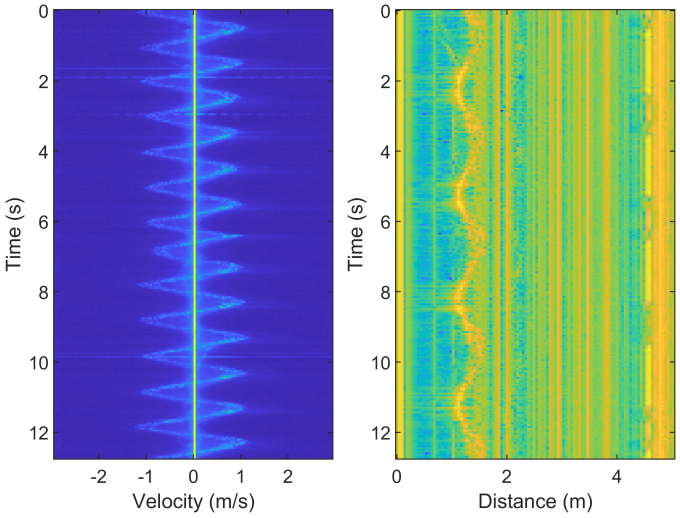

Fig. 3: Example of DTM (a) and RTM (b).

$$r_{res,\theta} = 2d \sin\left(\frac{\theta_{res}}{2}\right) \tag{11}$$

Eq. (11) is used to convert the angle to the radius of a gesture. In this equation, $r_{res,\theta}$ is the smallest detectable radius and d denotes the distance to the radar.

C. Contribution of this paper

The radius resolution $r_{res,\theta}$ in (11) is worse than $r_{res,d}$ and $r_{res,v}$. Even though this could be improved using algorithms like MUSIC [2], the accuracy of detecting radius of a gesture is inversely proportional to d, whereas $r_{res,d}$ and $r_{res,v}$ do not have this dependence on distance. Hence the processing methods in this paper focus on distance and velocity over time, i.e. using the RTM and DTM (Fig. 3).

No AI is used, unlike in [1]-[6], [10] and [12], since the desired characteristics can be obtained using signal processing. Contrary to AI, the proposed recognition method is not a black box and hence might be more insightful. Besides, the computational complexity and the amount of data to be processed might be lower. For instance, a 2D-matrix is fed to the recognition algorithm for each frame in [6], whereas the current approach will only need a 1D-vector for each frame. Also, this approach does not require the large set of training data the needed for AI.

This paper will first design post processing algorithms to recognize the radius and period of the specified hand gesture in Section II. Then, a measurement is set up in Section III to test the algorithms. The results of these experiments are shown and discussed in Sections IV and V. Finally, a conclusion will be drawn comparing the robustness and error of the proposed processing methods.

II. Methodology

As mentioned in the introduction, methods for estimating the radius and period of a gesture from the RTM and DTM will be designed. The outcomes of the two methods will also be combined. Fig. 4 shows a block diagram of the proposed processing methods. Between the blocks in the diagram filtering steps take place.

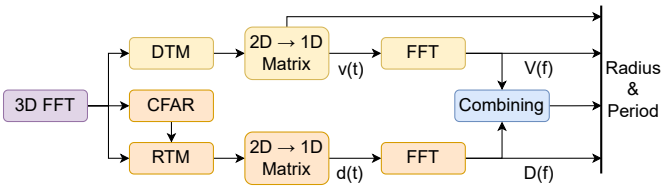

Fig. 4: Block diagram of the proposed processing methods to estimate a hand gesture's radius and period: DTM method (yellow), RTM method (orange) and combined method (blue).

A. DTM-based method

After the data matrix in dB is obtained by applying a 3D-FFT, the data matrix is summed in the range and angle dimensions in order to summarize the data along these dimensions in a single value. This step reduces the 4D-matrix to a 2D-matrix, i.e. the DTM. For the angle dimension, only the angles θ around $0°$ are summed, because the gesture is not covering the entire field of view of the radar.

In the DTM, all static objects are concentrated at $0m/s$, resulting in a high peak (Fig. 3a), which is removed by taking the the mean over time for velocities near $0m/s$.

Next, 75th percentile highest intensity value is set as a lower bound for the intensity, to which all values below this threshold are clipped. This removes low intensity noise and increases the average intensity of the DTM, for example around $v = 0m/s$

The next step is the application of a Gaussian filter, which applies a weighted average to each cell in the matrix, where the weights are a Gaussian distribution with standard deviation σ [14]. This filter reduces noise caused by single data points with high intensity.

Following Fig. 4, the DTM is now reduced to the 1D signal $v(t)$, by only keeping the velocity corresponding to the maximum value from each intensity vector along the time axis of the DTM. Ideally, $v(t)$ is a sinusoid representing the hand's velocity over time, but in reality, this is not the case. Hence, more filtering steps are applied, like a low-pass filter with cutoff frequency f_{cut} to filter out high frequency noise and subtracting the mean over time to remove the DC-component.

Next, an FFT is applied to $v(t)$, yielding $V(f)$. Now, the only step left is estimating the radius and period of the gesture. The estimated frequency $f_{g,DTM} = \frac{1}{T_{g,DTM}}$ is found by taking the frequency of the highest peak of $V(f)$, because that peak corresponds to the hand gesture.

$$r_{DTM,timedomain} = \frac{1}{2\pi f_{g,DTM}} \frac{max(v_g(t)) + min(v_g(t))}{2} \tag{12}$$

$$r_{DTM,fft} = \frac{1}{2\pi f_{g,DTM}} max(V(f)) \tag{13}$$

The radius can be estimated in two ways, using either $v(t)$ or $V(f)$, as shown in (12) and (13), respectively. Eq. (13) is expected to give a low estimate of the radius, because ideally all signal energy is in a single peak, but in reality the energy is spread out over a wider frequency band. In contrast, (12) is expected to yield a higher estimate, because only the very highest and lowest peak are considered. However, in order for (12) to work, there cannot be any outliers left in $v(t)$ after the filtering steps. This requirement is met, so (12) is used to estimate the radius of the gesture for the DTM method.

B. RTM-based method

Similar processing steps are taken for the RTM method as for the DTM method (Fig. 4), but there are some differences.

The main difference is that the RTM may include multipath propagations which make the gesture appear to happen at multiple distances, due to different delays of back scattered signals that take a longer path. This would cause detected 1D distance over time $d(t)$ to sometimes pick up the real gesture and sometimes detect the back scattering a few meters further, corrupting the extracted data. To prevent this, 2D Constant False Alarm Rate (CFAR) detection is used to detect the center distance of the gesture $d_{g,center}$, before the RTM is filtered and used to estimate the radius and period of the gesture.

$$T_h = \beta P_n \tag{14}$$

2D-CFAR looks at each cell in a matrix and compares it to a threshold value T_h. If the cell is above the threshold, a detection is made. The threshold depends on the noise power P_n, which is scaled with a factor β as shown in (14). The CFAR detector automatically adjusts β to achieve the desired Probability of False Alarm (PFA). The CFAR detector estimates P_n from a band of N_{tb} nearby cells called training band. These cells are separated from the cell under test by N_{gb} cells, the guard band [15].

To implement the CFAR detector, first an unfiltered RTM is constructed by summing along the velocity and angle dimensions of the data matrix. The RTM is then fed to the CFAR algorithm, yielding a 2D-matrix with value 1 for each cell of the RTM that exceeds threshold T_h and value 0 for other cells. This matrix is summed along the time axis to get the number of detections for each distance $N_{det}(d)$. After applying a moving average filter, the distance corresponding to the highest peak of $N_{det}(d)$ is taken as $d_{g,center}$. The RTM processing below will only look at distances $d_{g,center} \pm r_{g,max}$. Here, $r_{g,max}$ is the maximum detectable radius for the RTM method, which will be chosen later.

Next, the RTM is constructed (Fig. 4) similar to the DTM, apart from the range restrictions. However, the filtering steps are not all the same as for the DTM. Next to the Gaussian filter, the mean over time is subtracted from the RTM in order to remove static objects [8]. Examples of static objects are visible as vertical lines in Fig. 3b.

For estimating the period $T_{g,RTM}$ of the gesture, the same steps are followed as for the DTM: extracting peaks to get a 1D signal $d(t)$ of distance of the hand over time, low-pass filtering, DC-component removal and an FFT yielding $D(f)$. However, because the signal has too much outliers to use (12), the radius is determined similar to (13):

$$r_{RTM} = max(D(f)) \tag{15}$$

C. Combined method

Having two separate methods to determine the radius and period of a gesture allows us to combine them. For the radius the average of the DTM and RTM estimates is taken:

$$r_{combi} = \frac{r_{DTM,timedomain} + r_{RTM}}{2} \tag{16}$$

This can result in an overall more accurate and robust outcome, since the DTM method is expected to give a high and the RTM method a low estimate, relative to each other.

$$T_{combi} = \begin{cases} \frac{T_{g,DTM}+T_{g,RTM}}{2}, & if\ |\frac{1}{T_{g,DTM}} - \frac{1}{T_{g,RTM}}| < 0.25Hz \\ T_{g,DTM}, & otherwise \end{cases} \tag{17}$$

Eq. (17) shows the combined period T_{combi}. Early testing has shown that the RTM method produces significantly more outliers than the DTM methods, so the RTM estimate is only used if it is close the DTM estimate. For this, a maximum of 0.25Hz deviation is chosen, because that is the smallest frequency step that will be measured in this paper.

III. Measurement campaign

A set of experiments is designed to test the processing algorithms developed in the previous section. First, the experiment setup will be described, followed by the variations of the parameters of the experiment.

A. Measurement setup

This paper is only interested in angle measurements of the elevation dimension in 3D-space. Hence, the radar will be rotated 90°, such that the array of 8 antenna elements from Tx_1 and Tx_2 (Fig. 1) are directed along the elevation axis. This is also why Tx_3 (Fig. 1) will not be used in this paper.

For the radar configuration, a rate of 20 frame/s is chosen, well above the Nyquist sampling rate [11] for the fastest gesture frequency in the measurements in this paper. A total of 256 frames for each measurement is chosen, yielding a total measurement duration of 12.8s. A slope of $60MHz/\mu s$ is used to keep T_c low and thus v_{res} optimal. Besides, ramp a start- and endtime of $6\mu s$ and $66\mu s$ are chosen for each chirp, respectively. This leads to a bandwidth of 3.96GHz, which is close to the radar's maximum of 4GHz and thus close to the optimal d_{res}. Lastly, a sample rate of 10Msps, $N_{ADC} = 512$ and $N_c = 128$ are chosen.

Using equations (1) and (10), the range and velocity resolution of the radar for the chosen settings are calculated (Table I). Besides, because the radar captures discrete data points, the axes of velocity and distance are divided in N_c and N_f equal intervals, respectively. This gives these axes a finite resolution, which will be referred to as "discrete resolution" (Table I). The discrete resolution is slightly worse than the

radar resolution for both $r_{res,d}$ and $r_{res,v}$. This was chosen on purpose, because doubling the number of samples would double the amount of data, but only yield a slight improvement of the resolution, since the radar resolution would then become the limiting factor.

TABLE I: Range and velocity resolution.

	$r_{res,d}$ (m)	$r_{res,v}$ (m)
Radar resolution	0.038	0.025
Discrete resolution	0.049	0.029

Besides choosing parameters for the radar, a number of parameters need to be chosen for the estimation algorithm proposed in Section II. These are shown in Table II, along with a short motivation. The parameters for CFAR are chosen by sweeping over the possible values and choosing the combination with the minimum square error between the detected center distance and ground truth distance.

TABLE II: Algorithm parameters including motivation.

Parameter	Value	Motivation
Gaussian filter σ	0.5	Default value in Matlab [14]
$f_{cut,LPF}$	2Hz	Target cannot comfortably exceed this
$r_{g,max}$	$\pm 0.5m$	Target cannot comfortably exceed this
CFAR PFA	0.36	This yields the minimum square error
CFAR N_{tb}	9 cells	This yields the minimum square error
CFAR N_{gb}	3 cells	This yields the minimum square error

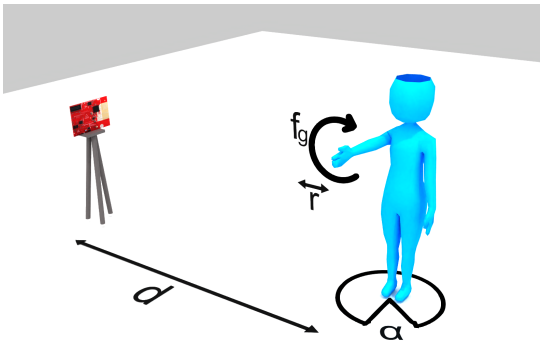

Fig. 5: Experiment setup with variable parameters indicated.

To make the experiments more accurate and repeatable, a ground truth is needed, since one can only move his hand with $\pm 5cm$ accuracy when guided. For this purpose the Virtual Reality (VR) app shown in Fig. 6 is built and a Meta Quest 2 [16] is used to run the app. All one has to do is face in the direction of the desired angle and follow the target disk with his hand. To verify whether the accuracy goal of $\pm 5cm$ is reached, the screen of the VR glasses is recorded.

Besides the VR glasses, a fencing glove that contains a metal mesh is worn in order to increase the reflectivity of the hand during the measurements. The complete setup is shown in Fig. 7.

B. Measurement scenarios

For the experiment a human subject will perform the specified gesture in front of the radar. The radius r and period

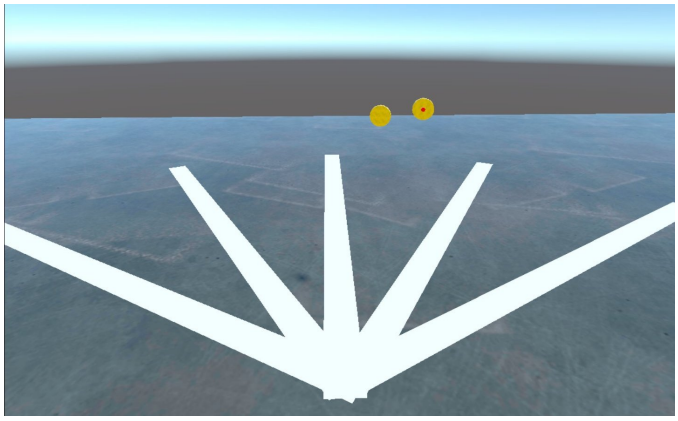

Fig. 6: Screenshot of virtual reality app, where the yellow disk with the red dot rotates around the other yellow disk at a configurable speed and radius, demonstrating the gesture. The white lines indicate $\alpha = \{-60, -30, 0, 30, 60\}°$.

Fig. 7: Measurement setup.

T of the hand gesture will be measured for various distances d between the center of human gravity and the radar and various angles α of the human torso with respect to the radar (Fig. 5). All variations that will be measured are listed in appendix A.

IV. Results

A. Overall performance

After processing the measured data, the overall performance of the estimation methods proposed in Section II is evaluated. For this purpose, the robustness and error for each method are computed. The robustness is defined as the number of measurements that are recognized within acceptable deviation divided by the total number of relevant measurements. For the radius, the threshold is an absolute error of 5cm and the threshold for the frequency is 0.25Hz, which corresponds to the maximum gesture period of 4s. Only the measurements for $\alpha = 0°$ are included for the robustness of the radius detection, because other angles are not expected to be accurate. This yields a total of 20 measurements for the detected radius and 40 measurements for the period. To get more insight, the average absolute error is also computed. Both these measures are shown in Table III.

The table shows that the combined method is the most robust and has the lowest error for radius detection, whereas only 50% of the detections are found accurate for the DTM and RTM methods. Note however, that the average error is only slightly above the threshold of 5cm, so many data points are almost accurate.

Looking at the robustness of the period estimation, all 3 methods perform equally good. However, the DTM has a lower average error, meaning it is more accurate than the other methods.

TABLE III: Robustness and error for each method.

	DTM	RTM	Both
Radius robustness (%)	50	50	75
Radius error (m)	0.061	0.051	0.044
Period robustness (%)	83	83	83
Period error (s)	0.068	1.12	0.075

Table III shows an exceptionally high error for the period estimate of the RTM method compared to the errors of the other methods. Besides, the individual data points of the RTM method have some outliers. When investigating this, it is found that the CFAR detector sometimes finds the wrong center distance for the gesture. This causes the actual gesture to be outside of the distance window of the RTM. Most measurements are detected correctly for $d = 1m$, whereas all measurements for $d = 2m$ and $d = 3m$ are misdetected. This is found to be the cause of the large errors at $\alpha = \pm 60°$ in Fig. 13. It also explains and invalidates the datapoints for the RTM (and thus the combined method) in Fig. 9 and 12. It appears that the CFAR parameters found in Table II are perhaps not optimal. A reason for this could be that the parameters with the lowest average error are selected. However, the data set contains more measurements for $d = 1m$ than the other distances, biasing the CFAR detector towards $d = 1m$.

B. Gesture radius

Fig. 8 shows the detected radius for all three methods discussed in the previous section. It shows that both the DTM and RTM methods are able to detect the radius, but that the RTM method is less accurate. In line with the different peak detection for the DTM and RTM methods, as discussed in Section II, the RTM method overall estimates a lower radius than the DTM method.

The relation between the detected radius and the distance d between the center of human gravity and the radar is shown in Fig. 9. The error of the DTM method increases with the distance. This might be due to propagation loss, causing the signal intensity to decrease at larger distances and making it harder to distinguish the gesture from noise, like in-phase back scattered signals (e.g. Fig. 14) that have high intensity.

The relation between the detected radius the angle α with respect to the radar is also investigated, as shown in Fig. 10. For reference, the theoretical radius that should be detected based on radial distance to the radar is also displayed. The detected radius follows the shape of the theoretical curve, but has a higher error. This shape shows that the error increases

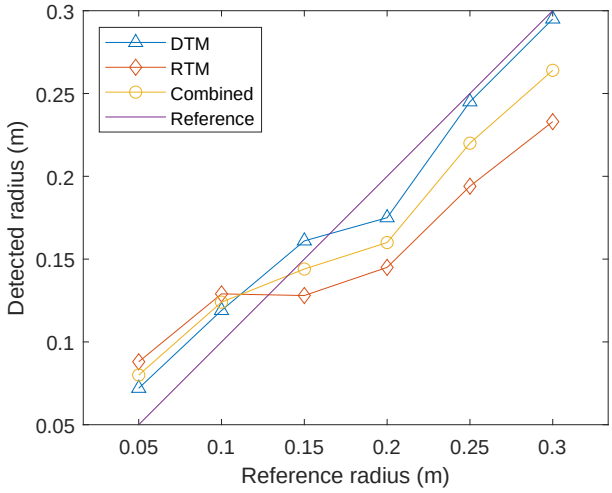

Fig. 8: Detected radius against actual radius for different detection methods, where $d = 1m$, $T = 1s$ and $\alpha = 0°$.

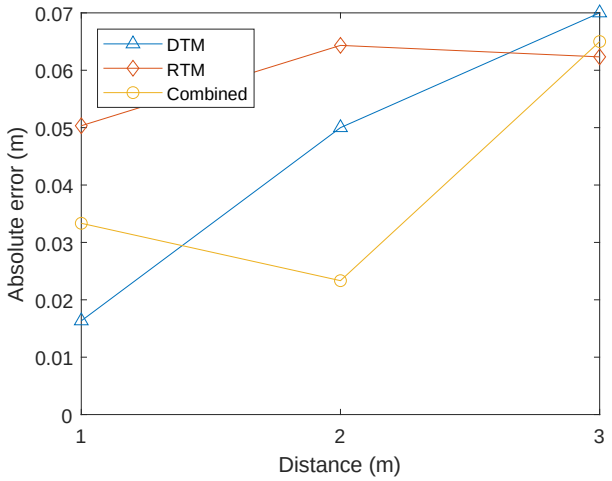

Fig. 9: Average error $|r_{detected} - r_{ideal}|$ for various distances d, where $T = 1s$ and $\alpha = 0°$.

with an increase in angle, so at some angle α the estimation will not be sufficiently accurate anymore. However, more measurements would be required to determine the maximum angle for which the proposed methods can accurately be used to detect the radius of the specified gesture.

C. Gesture period

The detected period is shown in Fig. 11. For all these measurements, both the DTM and RTM methods detected the same period. These measurements all have the minimum possible error, since their detected period corresponds to the discrete frequency closest to the reference frequency. The largest error is about 7% (Fig. 11) and could be improved by measuring for a longer time.

Fig. 12 shows the error for period detection for different distances d. It follows that the DTM method can accurately recognize the period of a gesture independent of the distance d. Similarly, from Fig. 13 we learn that period estimation using

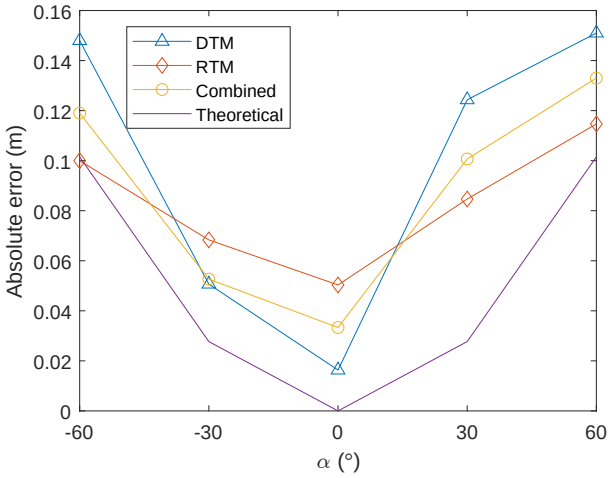

Fig. 10: Average error $|r_{detected} - r_{ideal}|$ for various angles α, where $d = 1m$ and $T = 1s$.

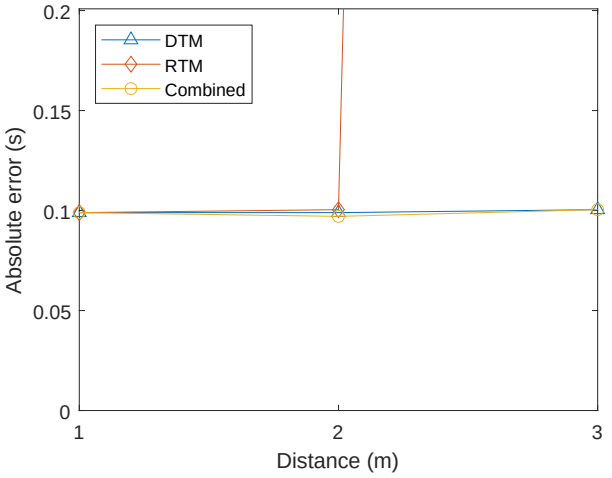

Fig. 12: Average error $|T_{detected} - T_{ideal}|$ for various distances d, where $r = 0.2m$ and $\alpha = 0°$.

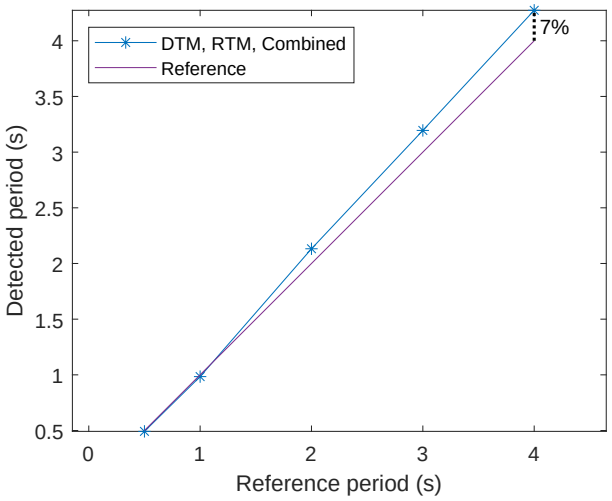

Fig. 11: Detected period against actual period for different detection methods, where $d = 1m$, $r = 0.2m$ and $\alpha = 0°$.

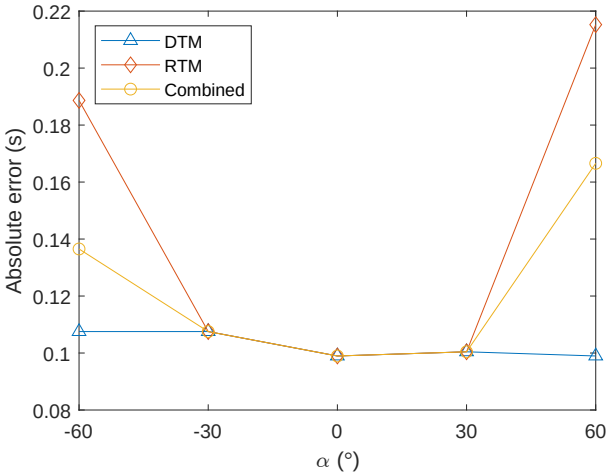

Fig. 13: Average error $|T_{detected} - T_{ideal}|$ for various angles α, where $d = 1m$ and $r = 0.2m$.

the DTM is independent of angle α within the measured range of $|\alpha| < 60°$.

V. Discussion and Recommendations

From the findings of this research a number of recommendations can be made. Firstly, as mentioned in Section IV, the CFAR detector did not detect the correct center distance of the gesture for $d > 1m$. This leaves the performance of the RTM method with respect to distance inconclusive, so future work could investigate this using different CFAR parameters. Besides, a detection method that does not depend on angle α could be investigated. If the θ_{res} can be made sufficient, a method based on angle θ can be used for this.

Furthermore, the accuracy of the ground truth of the measurements can be improved. During this research, it was assumed that an accuracy of $\pm 5cm$ can be reached for the ground truth, if the human subject is presented with a target to follow. Even though this method of obtaining a ground truth was sufficiently accurate for the measurements in this paper, capturing the coordinates of the VR controller would allow for a more quantitative comparison between the data from the radar and the ground truth. Lastly, The accuracy and repeatability of the gesture itself could be improved by using a humanoid robot.

VI. Conclusion

In this paper, processing methods were designed and evaluated to recognize the radius and period of a periodic rotational hand gesture using signal characteristics of mmWave radar. A range based method, velocity based method and a combination of the other two were constructed. Also, the dependence of the estimated period and radius on the distance d and the angle α were tested. For $d = 1m$ and $\alpha = 0°$, both the RTM and DTM methods were found accurate. However, when evaluating over more measurements with different parameters, their robustness dropped to 50% for the detected radius. The combined method shows an improved robustness of 75%. For larger distance d, the error of the detected radius increased for the DTM method. The RTM method was invalidated for $d > 1m$, because of

misdetections of the CFAR algorithm used to find the center distance of the gesture. Besides, the error of the detected radius was found to increase, when the angle α was increased.

The period estimates for all three methods were found to be accurate in 83% of the cases. Hence overall, the combined method is found the most robust of the three proposed methods for recognizing the radius and period of periodic rotational gestures.

VII. Acknowledgements

The author wants to thank his supervisor, dr. Y. Miao for her excellent input, guidance and feedback. Also thanks to M. Sellés Valls for sharing his knowledge on the radar.

References

[1] S. Hazra and A. Santra, "Robust Gesture Recognition Using Millimetric-Wave Radar System," *IEEE Sensors Letters*, vol. 2, no. 4, pp. 1–4, 11 2018.

[2] Y. Wang, A. Ren, M. Zhou, W. Wang, and X. Yang, "A novel detection and recognition method for continuous hand gesture using fmcw radar," *IEEE Access*, vol. 8, pp. 167 264–167 275, 2020.

[3] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, "Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum," *UIST 2016 - Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 851–860, 10 2016. [Online]. Available: http://dx.doi.org/10.1145/2984511.2984565

[4] S. Palipana, D. Salami, L. A. Leiva, and S. Sigg, "Pantomime," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, 3 2021. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3448110

[5] C. Liu, Y. Li, D. Ao, and H. Tian, "Spectrum-Based Hand Gesture Recognition Using Millimeter-Wave Radar Parameter Measurements," *IEEE Access*, vol. 7, pp. 1–13, 2019.

[6] J. T. Yu, L. Yen, and P. H. Tseng, "MmWave Radar-based Hand Gesture Recognition using Range-Angle Image," *IEEE Vehicular Technology Conference*, vol. 2020-May, 5 2020.

[7] Texas instruments, "xWR1843 Evaluation Module User's Guide," no. December 2018, pp. 1–20, 2019. [Online]. Available: https://www.ti.com/lit/pdf/spruim4

[8] M. Sellés Valls, "LOCALIZATION AND TRACKING USING A MONOSTATIC MM-WAVE INTERNSHIP REPORT," 2022.

[9] S. Rao and Texas Instruments, "Introduction to mmwave Sensing: FMCW Radars."

[10] X. Li, X. Wang, Q. Yang, and S. Fu, "Signal Processing for TDM MIMO FMCW Millimeter-Wave Radar Sensors," *IEEE Access*, vol. 9, pp. 167 959–167 971, 2021.

[11] B. P. Lathi and R. Green, "SAMPLING: THE BRIDGE FROM CONTINUOUS TO DISCRETE," in *Linear systems and signals*, 3rd ed., 2018, ch. 8.

[12] S. Ahmed, K. D. Kallu, S. Ahmed, and S. H. Cho, "Hand Gestures Recognition Using Radar Sensors for Human-Computer-Interaction: A Review," *Remote Sensing 2021, Vol. 13, Page 527*, vol. 13, no. 3, p. 527, 2 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/3/527/htmhttps://www.mdpi.com/2072-4292/13/3/527

[13] 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 Transactions on Graphics*, vol. 35, no. 4, 7 2016.

[14] Mathworks Inc., "Matlab imgaussfilt." [Online]. Available: https://nl.mathworks.com/help/images/ref/imgaussfilt.html

[15] ——, "Constant False Alarm Rate (CFAR) Detection," 2017. [Online]. Available: https://nl.mathworks.com/help/phased/ug/constant-false-alarm-rate-cfar-detection.html

[16] Meta, "Meta Quest 2." [Online]. Available: https://store.facebook.com/nl/quest/products/quest-2/

Appendix A
Measurement variations

TABLE IV: All variations of parameters for which a measurement is conducted.

d (m)	α(°)	T (s)	r (m)
1	0	1	0.05
1	0	1	0.1
1	0	1	0.15
1	0	1	0.2
1	0	1	0.25
1	0	1	0.3
1	0	0.5	0.2
1	0	2	0.2
1	0	3	0.2
1	0	4	0.2
1	-60	1	0.1
1	-30	1	0.1
1	30	1	0.1
1	60	1	0.1
1	-60	1	0.2
1	-30	1	0.2
1	30	1	0.2
1	60	1	0.2
1	-60	1	0.3
1	-30	1	0.3
1	30	1	0.3
1	60	1	0.3
1	-60	0.5	0.2
1	-30	0.5	0.2
1	30	0.5	0.2
1	60	0.5	0.2
1	-60	4	0.2
1	-30	4	0.2
1	30	4	0.2
1	60	4	0.2
2	0	0.5	0.2
2	0	1	0.2
2	0	4	0.2
2	0	1	0.1
2	0	1	0.3
3	0	0.5	0.2
3	0	1	0.2
3	0	4	0.2
3	0	1	0.1
3	0	1	0.3

Appendix B
Additional figures

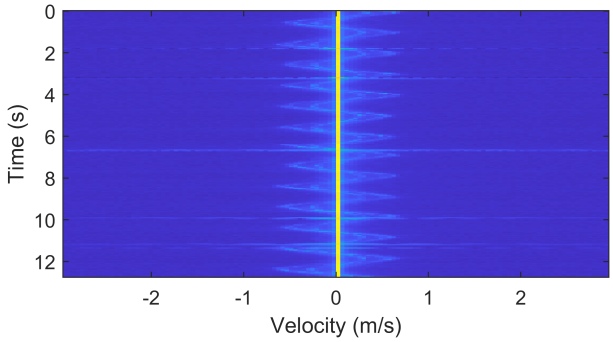

Fig. 14: DTM for $d = 2m$, $\alpha = 0°$, $T = 1s$ and $r = 0.1m$. Decreased intensity compared to $d = 1m$ and a smaller sinusoid visible in the larger one.