Department of Electrical, Computer, and Software Engineering

Part IV Research Project

Literature Review and

Statement of Research Intent

Project Number: 77

mmWave based human

activity recognition

Beck Busch

Sam Mason

Kevin I-Kai Wang

28/04/2023

Declaration of Originality

This report is my own unaided work and was not copied from nor written in collaboration with any other person.

B Busch

Name: Beck Busch

ABSTRACT: Human activity recognition is an essential area of study, as it allows for many advances in how technology can interact with the people that use it. A fluent understanding of human actions and conditions would allow our technology to make decisions about various aspects of human life and work alongside us closer than ever before. Millimeter wave radar (mmWave) is an up-and-coming solution to this problem as it has several advantages over other technologies. Unlike optical sensors, mmWave is not reliant on external lighting conditions and has vastly reduced privacy concerns. Privacy is an essential component of these systems, as some of the best use cases for HAR are intimate and personal environments, such as bedrooms and living spaces. To expand the usage of this technology, we aim to develop a new algorithm for interpreting mmWave data that can precisely and accurately identify various human actions with minimal delay. Current mmWave solutions can only identify actions after they have been completed. Our approach would be able to identify an action as it takes place, making predictions that progressively get more accurate. We are calling this approach smooth human activity recognition.

1. Literature Review

Analysis of prior works is vital for undertaking a project of this scope. With this literature review, we aimed to expand our knowledge of the field while also identifying an area of research that would be beneficial. A decent amount of research exists in this field, with various papers focusing on different aspects of the solution, such as recognition technology, recording and segmentation methods, and especially the algorithms that analyze and classify the data. We reviewed a selection of these research outputs focusing on real-time classification and detection. Below in Table 1 is an index of the papers I will reference, with relevant information on how they conducted their research.

Table 1: Index of literature review source material

Paper	Activities	Segmentation Method	Technology	Real-time Implications
Full-body Activity Recognition				
Chengxi Yu et al. [1]	Walking, falling, sitting, standing, including transitions	Sliding window	mmWave Radar DVCNN	Real-time operation possible; detection not attempted.
Jun-Huai Li et al. [2]	Walking, sitting, laying, including stairs and transitions	Sliding window	Physical sensors HMM	No attempt at real-time detection
Geethika Bhavanasi et al. [3]	Walking, falling, sitting, laying, rolling, including chair, bed, and transitions	Recording sample	mmWave Radar CNN with various inputs	No attempt at real-time detection
Renyuan Zhang, Siyang Cao [4]	Walking, sitting, hand waving, including transitions	No mention	mmWave Radar CNN with Micro-Doppler	Real-time classification possible; no results provided

Xiaoyan Cheng et al. [5]	Walking, sitting, squatting, falling, boxing, pushing, waving	Recording sample	Channel State Information GMM-HMM	No attempt at real-time detection
Jiahui Huang et al. [6]	Walking, jogging, standing, sitting, ascending stairs, descending stairs	Sliding window	Wearable sensors. Two-stage end-to-end CNN.	No attempt at real-time detection
Peijun Zhao et al. [7]	Various motions	Recording sample	CubeLearn, CNN&LSTM	An attempt made at real-time implementation.
Arm Gesture Recognition				
Jih-Tsun Yu et al. [8]	Various hand motions, including grabbing, tilting, turning, and swiping.	Recording sample	CNN&LSTM with Range-angle image	Real-time classification was not attempted.
Alexandros Ninos et al. [9]	Swiping, rotating, pattern, and random	Recording samples with event detection	MLP and custom algorithm	Capable of real-time recognition.
Haihua Xie et al. [10]	Lifting, lowering, pushing, pulling, circle motions, and combinations	Recording sample	PointNet classifier	Lightweight classifier for real-time operation.
Arindam Sengupta et al. [11]	Walking, swinging arms	Individual frame analysis	CNN with point cloud input	Capable of real-time operation.

1.1. HAR Technologies

There are several different base technologies at the forefront of HAR. Optical methods like visual and infrared cameras excel at pattern recognition and motion tracking, but this approach has several concerns. The most apparent concern is that placing a constantly recording camera in an area with frequent human activity can be an invasion of privacy. Consumers will never be comfortable with third-party camera systems installed in private places like hospital rooms or houses. Hence, an effective HAR system needs to operate without the ability to record such detailed information. Another problem with optical sensors is environmental lighting. While infrared cameras can circumvent this with infrared LEDs, regular cameras rely on the lighting conditions of the environment they are monitoring to record quality data. This leads us to non-optical measurements such as accelerometers, channel state information, and mmWave radar.

1.1.1. Analysis of physical sensors and CSI

Physical sensors like accelerometers and gyroscopes offer several advantages over vision-based sensors, but they come with their own limitations. While physical sensors are often able to capture cleaner data with much less noise, [6] they require that the sensor itself be attached to the subject. This works well for situations where the activity monitoring is temporary and involved, such as physical exercise. However, in more passive monitoring environments, such as a bedroom or office, wearable devices are an unwanted burden. Channel State Information monitoring refers to the process of observing the changes in wireless signals created by objects moving in an environment. As human bodies move between a signal source and receiver, they create variations in how the signal propagates through the environment. Analysis of this change can be accurate enough to identify the posture and motion of a human form, allowing for a non-invasive form of HAR. One of the most significant advantages of CSI detection is that wifi signals can be used as the signal source. Since wifi signals are already highly prevalent in our everyday lives, signal sources are already readily available. The downsides to wifi CSI are related to the quality and accuracy of the received information. Due to the wavelength of the wifi signal, the data lacks granularity, making small-scale motion challenging to detect. On top of this, CSI experiments require precisely placed signal sources and receivers that cannot be used for regular data transmission to achieve the best result [5]. This contradicts the notion that CSI is advantageous since wifi signals are already available in the environment.

1.1.2. Advantages of mmWave FMCW radar

FMCW radar is a technique that can identify moving objects in the environment by analyzing their effect on an emitted signal. Signal bursts with a constant linear change in frequency, known as chirps, are emitted by the radar, then the reflection of this signal is received. When the signal reflects off an object, the object's velocity creates variations in the signal that can be analyzed through comparison to the initially emitted chirp. This technique allows us to gather more information points than CSI without optical sensors. The larger wavelengths of FMCW radar also provide more accurate information due to the granularity of the received data, [8] making it our preferred technology for HAR.

1.2. Data recording

There are several different methods for recording and interpreting the FMCW chirp, each with its own advantages and disadvantages. The two most commonly used methods are range-doppler image and range-angle image. Range-doppler imaging works by interpreting the variations in frequency created by the signal reflecting off a moving object. The Fourier transform of this received signal can be used to estimate the velocity of a moving object, providing data for activity recognition. Range-angle imaging compares the reflected signal received by two antennas. It uses the phase difference along with the distance between antennas to calculate the range and angle of objects in the environment. While range-doppler imaging is often preferred for its fast generation time and depth of information [8], it can often fall short when movement occurs parallel to the radar transmitter [9]. Horizontal movements are still detectable by range-doppler image due to the change of velocity of the object, but since the distance from the object to the radar (range) stays relatively constant, RDI can struggle to define the gesture. Range-angle imaging does not suffer from these problems and generally performs better at detecting motion than RDI, especially horizontal motion. For these reasons, most approaches utilize a fusion of RDI and RAI as the inputs to their systems, which achieve excellent experimental results compared to only one data system [8].

1.2.1. Alternative pre-processing

In [7], Peijun Zhao et al. investigate a new preprocessing technique they call CubeLearn. This novel approach to FMCW radar interpretation aims to train a machine learning model based on conventional Fourier transform techniques to achieve a similar processed signal without the costly Fourier transform operations. This results in accuracies of up to 95% with a processing time of only 25ms.

1.3. Data segmentation

When analyzing radar data for classification of events, data segmentation is necessary to reveal the relevant parts of the data that contain movements or behavior that we aim to detect. This can be performed in several ways, with varying results from each method. Several studies manually segment the data into the separate activities and transitions before beginning analysis. This can be fine when researching a smaller part of the HAR pipeline, but for our purposes, it is unsatisfactory. Another common approach is the usage of sliding windows. ‘Windows’ of a fixed length are slid over the data recording, and the data within the window is analyzed. Based on this analysis, the section of the recording containing the most useful information can be identified. This approach can cause issues when the fixed window size is either too small or too large for the relevant action, so the sliding window process can be performed multiple times with windows of varying lengths [2]. Alexandros Ninos et al. [9] proposed a novel solution to data segmentation that uses an event detection algorithm to identify events taking place in real-time. By analyzing the radar image for a velocity change above a certain threshold, they are able to identify the start of an action. Although this approach still relies on analyzing a data segment of a fixed length, the event detection allows it to start the analysis in real-time instead of investigating pre-recorded data.

When identifying an activity is not dependent on a period of time, individual frame analysis can be used to classify data. Arindam Sengupta et al. [11] created a CNN model to estimate the posture of the human frame. Since their desired outcome does not involve multiple data frames at once, they can analyze each frame to identify human posture. In this way, they do not need to utilize data segmentation.

1.4. Classification algorithms

The core functionality of a HAR system is the algorithm that can receive the radar input and determine which action is being performed. Best results are often achieved through the use of a machine learning model such as convolutional neural networks, sometimes in tandem with long short-term memory networks. In [3], Geethika Bhavanasi et al. perform a comparison of various ML models for mmWave FMCW HAR. They compared a CNN model to LSTM, CNN&LSTM, and other models. They found that the CNN model performed the best at 97% accuracy in a clinical environment and still performed the best in other environments. In [1], Chengxi Yu et al. propose a novel HAR pipeline involving a dual-view convolutional neural network. They found that this approach greatly outperforms other similar models, with an accuracy of 97%, compared to the similar models with a max accuracy of 70%.

1.5. Real-time applicability

Our research has been focused on an HAR system that can function in real-time. To this end, we have investigated the various experimental attempts at behavior classification in real-time. In [10], Haihua Xie et al. devised a lightweight classification algorithm to maximize the running speed of the system. They did this through the use of a PointNet algorithm to efficiently identify spatial features in the data. PointNet is a deep net architecture that allows for the processing of point clouds, a format of radar data that is less refined than micro-doppler formats like range-doppler images. By working directly with point clouds, the system can avoid the costly data calculations that are required to generate range-doppler or range-angle images. This is a great advantage for real-time systems, as the PointNet classifier was able to achieve higher accuracy than conventional algorithms such as LSTM while only taking 81 milliseconds to complete.

In [4], Renyuan Zhang and Siyang Cao design an HAR system that uses a constant false alarm rate (CFAR) algorithm running on the radar board to retrieve the relevant data points from the radar, and construct a point cloud. This removes unnecessary information, allowing for faster data transmission. The point cloud data is then processed by the host computer into a micro-doppler image for the classification algorithm. The faster transmission rates allow for data to be processed in real-time. To allow the algorithm to detect human behavior in real-time, they create a constantly updated buffer of micro-doppler frames to be fed into the algorithm. This allows the algorithm to detect sequential changes in motion and identify behavior.

2. Research Intent

We are looking to develop a system for real-time human activity recognition that is responsive to changes in behavior. Our system will be able to detect and classify human behavior accurately, within a response time of 500 milliseconds. Classification predictions should begin as soon as the behavior is detected, with an accurate classification made within half the time the action takes to occur. If an action is periodic, the classification time should be half of a single repetition. Our system will also be able to rapidly detect a change in behavior and report this within an adequate timeframe. We are interested in this system since existing real-time approaches can only classify an action after it has been completed, and systems capable of detecting a change in behavior rely on the analysis of pre-recorded data. This development will be beneficial to the field of HAR since it will allow for more adaptive and responsive implementations of behavior recognition that can be utilized in real-time as opposed to identifying behavior after it has been recorded. This project will be quite a significant undertaking, so planning will be crucial to our success.

2.1. Project outline

1: Confirm system architecture. The start of the project will be focused on making design decisions for all aspects of the solution. It is important that we make these decisions at the start since each part of the design will affect the others. We aim to have this completed by week 10 of semester 1.

2: Configure mmWave radar module. This will involve researching and purchasing an adequate radar board and configuring it until we can view the resulting range-doppler graphs on a Windows computer. This should be the easiest part of our project, taking no more than two weeks, not including shipping disparity. Other aspects of the project can be progressed in tandem with this objective. We aim to have this completed by week 11 of semester 1.

3: Construct pre-processing pipeline. We will need to research and implement the system required to process the radar signal into an input for a CNN or similar machine-learning algorithm. This stage will involve testing various approaches and researching their effectiveness for a real-time implementation of HAR. We aim to have this completed before the start of semester 2.

4: Collect training data. We will investigate the practicalities of recording our own training data and research existing mmWave datasets. Once we have made several key decisions, such as the behaviors we will be detecting and the methods we will use to record the data, we can proceed with the collection of our dataset. This step is under a time constraint since the use of third-party volunteers will require a completed ethics proposal. The ethics proposal needs to be submitted by week 8 of semester 1. We aim to have this step completed by week 2 of semester 2.

5: Create the machine learning algorithm. The core of our project will be the ML model that classifies our radar data to detect human activity. We will investigate the effectiveness of various algorithms and pipelines before completing our own approach. This will include the creation and testing of our machine-learning model and any algorithmic components that will run alongside it. We aim to have this step completed by week 8 of semester 2.

6: Complete the HAR system. Once the individual components of our system are functional, the entire pipeline will be merged into a single executable program with a user display to confirm our detection capabilities and experimental results. Included in this objective is carrying out experimental research to determine the accuracy and reliability of our system. We aim to have this step completed by week 10 of semester 2

Acknowledgments

I would like to thank my supervisor Kevin I-Kai Wang for his support and guidance during this project. His input has been invaluable for understanding the research process and is critical to making this project successful. I would also like to thank my project partner Sam Mason for their commitment to teamwork and guidance in research.

References

[1] C. Yu, Z. Xu, K. Yan, Y. -R. Chien, S. -H. Fang and H. -C. Wu, "Noninvasive Human Activity Recognition Using Millimeter-Wave Radar," in IEEE Systems Journal, vol. 16, no. 2, pp. 3036-3047, June 2022, doi: 10.1109/JSYST.2022.3140546.

[2] J. -H. Li, L. Tian, H. Wang, Y. An, K. Wang and L. Yu, "Segmentation and Recognition of Basic and Transitional Activities for Continuous Physical Human Activity," in IEEE Access, vol. 7, pp. 42565-42576, 2019, doi: 10.1109/ACCESS.2019.2905575.

[3] Bhavanasi, G., Werthen-Brabants, L., Dhaene, T. et al. Patient activity recognition using radar sensors and machine learning. Neural Comput & Applic 34, 16033–16048 (2022).

[4] R. Zhang and S. Cao, "Real-Time Human Motion Behavior Detection via CNN Using mmWave Radar," in IEEE Sensors Letters, vol. 3, no. 2, pp. 1-4, Feb. 2019, Art no. 3500104, doi: 10.1109/LSENS.2018.2889060.

[5] X. Cheng, B. Huang and J. Zong, "Device-Free Human Activity Recognition Based on GMM-HMM Using Channel State Information," in IEEE Access, vol. 9, pp. 76592-76601, 2021, doi: 10.1109/ACCESS.2021.3082627.

[6] J. Huang, S. Lin, N. Wang, G. Dai, Y. Xie and J. Zhou, "TSE-CNN: A Two-Stage End-to-End CNN for Human Activity Recognition," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 1, pp. 292-299, Jan. 2020, doi: 10.1109/JBHI.2019.2909688.

[7] P. Zhao, C. X. Lu, B. Wang, N. Trigoni and A. Markham, "CubeLearn: End-to-end Learning for Human Motion Recognition from Raw mmWave Radar Signals," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2023.3237494.

[8] J. -T. Yu, L. Yen and P. -H. Tseng, "mmWave Radar-based Hand Gesture Recognition using Range-Angle Image," 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 2020, pp. 1-5, doi: 10.1109/VTC2020-Spring48590.2020.9128573.

[9] A. Ninos, J. Hasch and T. Zwick, "Real-Time Macro Gesture Recognition Using Efficient Empirical Feature Extraction With Millimeter-Wave Technology," in IEEE Sensors Journal, vol. 21, no. 13, pp. 15161-15170, 1 July1, 2021, doi: 10.1109/JSEN.2021.3072680.

[10] H. Xie, P. Han, C. Li, Y. Chen and S. Zeng, "Lightweight Midrange Arm-Gesture Recognition System From mmWave Radar Point Clouds," in IEEE Sensors Journal, vol. 23, no. 2, pp. 1261-1270, 15 Jan.15, 2023, doi: 10.1109/JSEN.2022.3216676.

[11] A. Sengupta, F. Jin, R. Zhang and S. Cao, "mm-Pose: Real-Time Human Skeletal Posture Estimation Using mmWave Radars and CNNs," in IEEE Sensors Journal, vol. 20, no. 17, pp. 10032-10044, 1 Sept.1, 2020, doi: 10.1109/JSEN.2020.2991741.