Joint Fall And Aspect Angle Recognition Using Fine-Grained Micro-Doppler Classification

Qingchao Chen[1], Matthew Ritchie[2], Yang Liu[3], Kevin Chetty[1], Karl Woodbridge[2]

[1] Department of Security and Crime Science, University College London, London, U.K.
[2] Department of Electronic and Electrical Engineering, University College London, London, U.K.
[3] Computer Laboratory, University of Cambridge, Cambridge, U.K.
{Qingchao.chen.13, m.ritchie, k.chetty, k.woodbridge}@ucl.ac.uk, yl504@cam.ac.uk

***Abstract*—Activity recognition and monitoring are finding important applications in ambient assisted living healthcare. Among the various types of motions which researchers are attempting to detect and recognize, fall detection has gained significant interest. In this paper, we investigate the application of high-frequency (24 GHz) FMCW radar for multi-perspective micro-Doppler (μ-D) activity recognition. Data from two different types of motion; falling and picking-up an object, were collected from three aspect angles and put through a fine-grained classifier to not only differentiate the motions, but to also identify their aspect towards the radar receivers. A key novel component of this work is the application of the fine-grained classification task, where a label discriminate sparse representation classifier is proposed to improve recognition performance over very similar μ-D signatures. This is achieved by learning a discriminate dictionary constrained by the label information and meanwhile preventing the overfitting problem. The greatest increase in classification performance was found to be of the order of 8 %.**

Keywords—Fall Recognition, Fine-Grained Micro-Doppler Classification, Aspect Angles, Sparse Representation Classifier, Label Discriminate.

I. INTRODUCTION

Driven by demands from an ageing population, Ambient Assistant Living (AAL) [1] is becoming increasingly important to promote wellbeing, and reduce the high care costs from healthcare providers. Within the AAL framework, activity recognition has been an important research topic to facilitate enhanced situational awareness. Large amount of activity monitoring technologies have been investigated to detect, recognize and monitor human activity. Among these, sensors embedded in wearable and mobile phones such as accelerometers and gyroscopes are able to provide physical information about subjects, but suffer from low movement update rates of typically less than 5 Hz [1]. In addition, people may forget to wear or even lose their on-body sensors due to the physical discomfort. Passive infrared (PIR) sensors are able to only provide the coarse-grained room level existence [2], while RFID based devices employ complex transmitters and receivers, and require pre-planning in order to optimally site the positions of the nodes [3]. Similar to on-body sensors RFID tags or transmitters can also be easily damaged, lost or forgotten [3]. Video system such as MS Kinect and Intel RealSense might provide very detailed information in home monitoring, however, it requires optimal lighting conditions and the acceptability of deploying video cameras in home environments raises many privacy issues.

Activity recognition using micro-Doppler (μ-D) signatures from radar sensors have been thoroughly investigated for a range of different applications [4]. μ-D signatures can be utilized to differentiate between different types of target and activity, especially human motions [5], and for distinguishing various types mechanical motions such as those associated with wind turbines and aircraft propeller blades [6], or between bird and drones [7]. A wide range of features have been adopted into the μ-D classifications. Empirical features, such as the maximum frequency, Radar Cross Section (RCS) amplitude and time duration have been investigated, however, these features require very accurate detection and typically vary between experimental scenarios. Singular Value Decomposition (SVD) has been utilized to extract the abstract and discriminate features based on the projected eigenvector in time-frequency space [8]. However, the final feature selection is empirical also. Principle Component Analysis (PCA) based features are able to reduce the dimension of the whole spectrogram data but the method is limited by training sample numbers.

Deep Neural Networks (DNN) [9] is a feature plus classification method that is gaining popularity. However, DNN typically requires large volumes of data in the training stages, and these networks are prone to be over-fitting due to less training classes than the parameters of networks. The classifiers that have been investigated include linear classifiers, linear discriminate classifiers, random forest classifier, Support Vector Machine (SVM) and the Sparse Representation Classifier (SRC), which differ in performance capability across various applications.

For fall recognition based on radar μ-D signature classification, Jocanovic et al. [10] first proposed to apply DNN for differentiating falling down from three other motion types using their VNA-based Continuous Wave (CW) radar operating at 6 GHz center frequency. Although demonstrated to have good recognition performance, the motions most similar to the falling action was (i) sitting down from a standing position, (ii) and bending over. The authors considered bending then straightening up as a whole signature, which makes the whole recognition task easier. Moreover, although the μ-D signature depends a lot on the aspect angle and the antenna geometry in an indoor environment, the exact geometry is unclear besides that the antenna feed point is at the height of 1m above the ground. We have recently demonstrated activity recognition results [11] using a 2.4 GHz Wi-Fi passive radar[12], and used SRC to differentiate six motions, including falling. However, in the experiments, the effect of aspect angle was not considered and both false positives and negatives in the confusion matrix arose with the action of picking up objects from the floor.

An important requirement for healthcare monitoring devices is to have very low false alarms associated with health critical situations such as falling. Such a system may be linked to a local healthcare service that in an emergency would respond directly. This is a very challenging requirement because as a plethora of different motions undertaken by a person on daily basis could potentially be misrecognized as a fall. To further develop a fall recognition system with low false alarm rate, we have made the following contributions:

Firstly, we argue that an optimal geometry to detect and differentiate falls is such that the receiver antenna is placed on the ground pointing upwards. Next, we observe that within the indoor environment and common daily life, the most similar motion to a fall is the action of picking-up items from the ground while bending over. Moreover, fall motion depends a lot on the effect of the aspect angle any the experimental geometry.

Based on the above, we carried out a series of experiments consisting of a person either falling or bending over to pick up objects, and these were taken from three different aspect angles. We propose a challenging problem termed fine-grained μ-D classification, which is aimed to not only to differentiate falling from other very similar signatures, but to identify the aspect angle from which the measurement originates. Finally, to tackle the fine-grained μ-D classification task, we proposed the Label-Discriminative SRC (LD-SRC), integrating the ground-truth label as discriminate constraints and overfitting prevention constraints into the optimization procedure.

II. Fine-grained Micro-Doppler Classification Using Sparse Representation Classifier and Label Discrimination

Feature extraction is an essential step in μ-D classification and can be regarded as a special representation or transformation of the original signal data. Among the techniques, sparse coding is a special representation where the data is utilized to represent themselves with some additional sparsity constraints. In this section, the sparse representation classifier (SRC) and the label discrimination SRC are introduced for feature learning and classification.

A. Sparse Representation Classifier

First we provide some theoretical background. The dictionary, D, the training datasets Y_t and test datasets, Y_s are represented as:

$$D = [d_1 \quad d_2 \quad \dots \quad d_n \quad \dots \quad d_{N_D}] \in R^{M \times N_D}, \tag{1}$$

$$Y_t = [y_1 \quad y_2 \quad \dots \quad y_{n_T} \quad \dots \quad y_{N_T}] \in R^{M \times N_T}, \tag{2}$$

$$Y = [y_1 \quad y_2 \quad \dots \quad y_{n_S} \quad \dots \quad y_{N_S}] \in R^{M \times N_S}, \tag{3}$$

where M is the dimension of each data sample and N_D, N_T and N_S are number of data samples in the dictionary, training and test datasets respectively. It is worth noting that the dictionary size and the training sample size do not need to be the same and the dictionary can be either the exact training dataset or some learned new dataset. We could further suppose that there are P classes and the dictionary, training and test data labels are $L_D \in R^{N_D \times 1}$, $L_T \in R^{N_T \times 1}$ and $L_S \in R^{N_S \times 1}$ respectively.

In general, SRC consists of two stages: sparse coding and classification. More specifically, sparse coding of test samples using a training dataset can be solved by the following optimization function [13]:

$$x = \arg\min_x \|y - Dx\|_2^2 + \alpha \|x\|_1, \tag{4}$$

where y is one of the test sample, x is the sparse code vector and α is the sparsity level penalty parameter. The first term in equation 4 is aimed at reconstructing the test sample with minimum residuals using a linear combination of the training samples, while the one-norm constraint on the coding vector x is aimed to ensure the sparsity of x.

The classification task is operated based on the reconstruction error using the i^{th} class samples. More specifically, we wish to determine the predicted label based on the following:

$$\arg\min_i \|y - D_i x_i\|_2, \tag{5}$$

where x_i is the sparse coding weight from i^{th} class and D_i is the training samples from the i^{th} class.

B. Dictionary Learning based on Label Discrimination for fine-grained Micro-Doppler classification

The effectiveness of the SRC depends on an important fact that correlations between the inter-class training samples should be small. For the fine-grained μ-D classification problem which requires differentiating between very similar signatures, data samples in the dictionary D will potentially have large correlation coefficients, which could degrade the SRC performance. Therefore, we adopt a Label-Consistent K-SVD (LC-KSVD) dictionary learning method [14, 15], known for its capability to embed the label discrimination and to optimize the linear classifiers jointly in the training procedure. However, in spite of the advantages of LC-KSVD, its learning stage is very sensitive to the training data and prone to be overfitting, especially for the μ-D dataset consisting of only six classes at most but with thousands of pixels in the whole spectrogram. To tackle the fine-grained fall recognition task using label discrimination, and to also overcome the overfitting problem, we modify LC-KSVD and propose the new method, called Label-Discriminative SRC (LD-SRC), which solves the following optimization problem:

$$\begin{aligned} \langle X, D, A, W \rangle = \arg\min_{X,A,D,W} & \alpha \|Q - AX\|_2^2 + \beta \|H - WX\|_2^2 \\ & + \gamma_1 \|A^T A - I\|_2^2 + \gamma_2 \|W^T W - I\|_2^2 + w_1 \|W\|_2^2 + w_2 \|A\|_2^2 \\ & , s.t. \mu \|Y - DX\|_2^2 + (1-\mu) \|X\|_1 \end{aligned} \tag{6}$$

where X is the sparse code, W is the linear classifier weighting matrix transforming sparse code to the label matrix H and A is the matrix transforming the sparse code to the label-discriminative matrix Q. Label matrix H is defined as: $H = [h_1 \dots h_{N_T}]^T \in R^{P \times N_T}$ and $h_i = [0,0 \dots 1 \dots 0,0]^T \in R^P$ is the label vector where the non-zero position indicates the class of a training sample. α, β, γ_1, γ_2 and μ are parameters used to control the relative contribution of each constraint. The matrix Q is defined as: $Q = [q_1 \dots q_{N_T}]^T \in R^{N_D \times N_T}$ and the elements in $q_i \in R^{N_D \times 1}$ should either be '1' or '0'. The use of matrix Q integrates ground-truth label information to learning the dictionary so that the distances of intra-class sparse codes are minimized while those of inter-class sparse codes are maximized. More specifically, an example dataset of six training samples consisting of three classes and an eight-sample dictionary is shown in Fig.1. In this scenario, the sparse codes from three classes of the six training samples using D are represented in three different colors. The indices of '1' in Q represent locations where corresponding dictionary atoms should share the same label with input sample. In other words, those dictionary atoms should have larger weights than others in the sparse coding stage.

Our newly proposed constraints are as follows:

$$\gamma_1 \left\| A^T A - I \right\|_2^2 + \gamma_2 \left\| W^T W - I \right\|_2^2 . \tag{7}$$

The purpose of equation (7) is to avoid the overfitting problem in the optimization. By constraining A and W to be close to orthogonal matrices, i.e. they should only rotate the vector x from a geometrical point of view, overfitting is prevented. The advantage of this optimization is that the implementation of the sparse coding, as shown in equation 6 is regarded as the implicit constraints which would guarantee the conformity of sparse coding between the test and training stages.

$$X = \arg\min_X \mu \left\| Y - DX \right\|_2^2 + (1-\mu) \left\| X \right\|_1 \tag{8}$$

$$H_{predicted} = WX \tag{9}$$

As D and W have been learned in the test stage, the equation 8 is used to sparse coding for the test samples and the predicted label matrix H can be obtained by equation (9). The algorithm is implemented using the stochastic gradient decent method and the Lasso solver[13].

III. Experimental Setup and Implementation

The radar used for experiments was the 24 GHz Ancortek SDK FMCW radar, which was setup according to the operational parameters described in Table 1, as well as the geometry illustrated in Fig. 2. The radar was deployed on the ground approximately 1.5 m from the target with the antennas angled +45° in elevation, in-order to have the torso of the individual centrally within the beam of the system. Future work will further investigate different geometries to observe how the classification results depend on the radar geometries.

Two different actions were performed; falling down and picking up an item while bending over. Each of these actions is very similar in the downward motion of the main body, and could easily confuse a simple Doppler classifier. Additionally, as shown in Fig. 3, each action was recorded at three aspect angles with respect to the antenna pointing direction, -45°, 0° and 45°. This was done to evaluate how easily one action in a given geometry could be confused with another. The dataset generated contained 52 repeats of each action at each angle, providing a total of 312 repeats in total (156 for each action).

Table 1 Specification of the Ancortek SDK FMCW Radar

Center Frequency	24GHz
Chirp Bandwidth	500MHz
Chirp Period	1ms
Baseband Sampling Rate	512KS/s

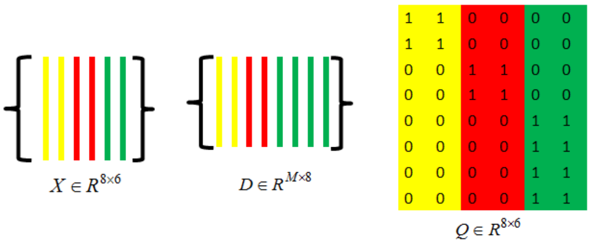

Fig. 1 The Q Matrix Generated From Ground Truth Label

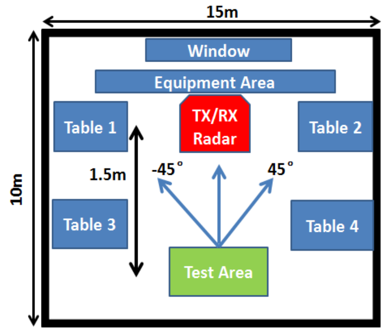

Fig. 2 Experiment Scenarios

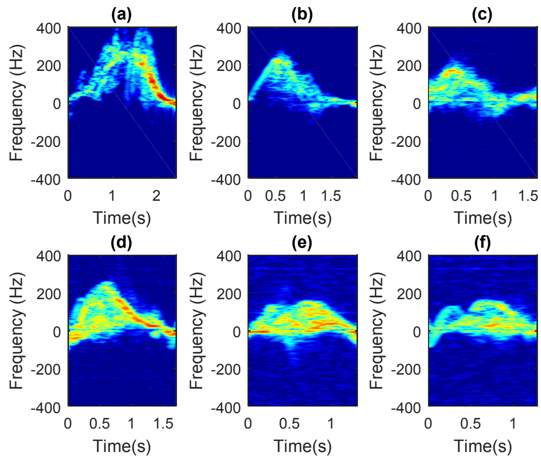

Fig. 3 Falling and pick-up while bending down from three aspect angles. Falling down from (a) 0°, (b) +45°, (c) -45° and bending down from (d) 0°, (e) +45°, (f) -45°. Note: all the figures are normalized to [-60 0] in a dB scale.

For the μ-D extraction, we first perform Fast Fourier Transform (FFT) to the baseband signal and apply the Moving Target Indicator (MTI) filter to remove the DC frequency component. Next, the averaged range-time bins of the moving target are selected and we perform the Short Time Fourier Transform (STFT) to obtain the spectrogram with a 0.2 second Hanning window and 0.19 second overlapping window. After the spectrogram is obtained, we cut time duration of the data to isolate just the section when the person target was moving.

To train the dictionary and fine tune the classifiers based on the method shown in Section 2, a pre-processing method onto the spectrogram matrix is necessary to unify the dimension of the feature vector. When we obtain the spectrogram matrix, due to the fact that the time duration of each μ-D signature might be different, a bi-cubic interpolation method [16] was used to map different spectrogram matrices to the same dimensionality. Finally, we stack each column (Doppler profile) of spectrogram matrix into a single column vector. It is worth noting that, although the 2-D matrix structure is transformed into a single column, the relevant data is still present but just relative positions of the pixels have changed into a 1-D plane. In addition, data interpolation is able to remain the original structure and feature, which is similar to the fact that expanding or shrinking an image won't change its local features and data representations.

An example μ-D data capture from the radar of a straight on fall, and a bend down pickup action from six different aspect angles can be seen in Fig. 3(a) to (f) respectively. The action of falling appeared to occur over a time window of around 1.5 to 2 seconds, and had a maximum Doppler frequency shift of 400 Hz (velocity around 2.5 ms^{-1}) compared with around 1.3-1.8 seconds and around 250 Hz Doppler frequency shift (1.6 ms^{-1}) of the bending down movement from the zero degree.

As shown in Fig. 3(a), (b) and (c), when the μ-D signature of the falling targets were measured from three perspectives, a different maximum velocity and power distributions in the time-frequency space is exhibited. Maximum velocities in Fig.3 (b) and (c) are similar and both are smaller than that shown in Fig. 3(a). The same is true for the results shown in Fig. 3(d), (e) and (f). It is worth mentioning that if we focus on differentiating the μ-D signature in Fig. 3(d), where the target is bending down at 0 degrees and the one in Fig.3(b) or (c), where they are falling down from +45˚ and -45˚, the task would be more difficult, as they present very similar signatures.

IV. CLASSIFICATION RESULTS AND ANALYSIS

Three different types of classifier challenges are presented here with increasingly fine-grained properties. This was done to evaluate the least to most complex classification task and learn when each method was most effective.

The first classifier challenge (EXP1) considered this as a two-class classification problem, where the variation in aspect angle is ignored and all data from each motion is considered as a single class. The second classifier challenge (EXP2) considered a four-class classification problem, where we regard the falling or bending down from 45˚ and -45˚ as a single class. To summarize, we try to differentiate not only the two motions, but also their angular perspective. The third classifier challenge (EXP3) considered this as a six-class classification problem, where our aim is to differentiate not only two motions but also from 0˚, -45˚ or +45˚.

As shown in Table 2, three classifiers are compared facing three different tasks, as shown in the previous sections. It seems that the LD-SRC outperforms SRC and SVM around 5% in the recognition rate and SRC outperforms SVM around 4% in recognition rate in EXP2 and EXP3 but stays the same as SVM in the EXP1 recognition result.

Table 2 Recognition Results of Three Experiments

	SVM	SRC	LD-SRC
EXP1	90%	90%	93.43%
EXP2	70%	84%	89.5%
EXP3	57.29%	74%	82%

V. CONCLUSIONS

Micro-Doppler Data from a real radar system was recorded to generate a 336 sample dataset of various activities. A newly proposed classification technique based on sparse representations and label discriminants was successfully applied to this data and show 8% increase in performance relative to other techniques. Future work will look to evaluate this LD-SRC classifier across a many class problem, and analyze the effect of the number of training data and across different geometries.

REFERENCES

[1] M. Chan, D. Estève, C. Escriba, and E. Campo, "A review of smart homes—Present state and future challenges," Computer methods and programs in biomedicine, vol. 91, pp. 55-81, 2008.

[2] T. Yamazaki, "The ubiquitous home," International Journal of Smart Home, vol. 1, pp. 17-22, 2007.

[3] K. Kalimeri, A. Matic, and A. Cappelletti, "RFID: Recognizing failures in dressing activity," in 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, 2010, pp. 1-4.

[4] V. C. Chen, The micro-Doppler effect in radar: Artech House, 2011.

[5] S. Sundar Ram and L. Hao, "Simulation of human microDopplers using computer animation data," in Radar Conference, 2008. RADAR '08. IEEE, 2008, pp. 1-6.

[6] A. Naqvi, S.-T. Yang, and H. Ling, "Investigation of Doppler features from wind turbine scattering," Antennas and Wireless Propagation Letters, IEEE, vol. 9, pp. 485-488, 2010.

[7] P. Molchanov, R. I. Harmanny, J. J. de Wit, K. Egiazarian, and J. Astola, "Classification of small UAVs and birds by micro-Doppler signatures," International Journal of Microwave and Wireless Technologies, vol. 6, pp. 435-444, 2014.

[8] F. Fioranelli, M. Ritchie, and H. Griffiths, "Personnel recognition based on multistatic micro-Doppler and singular value decomposition features," Electronic Letters, 2015.

[9] Y. Bengio, "Learning Deep Architectures for AI," Found. Trends Mach. Learn., vol. 2, pp. 1-127, 2009.

[10] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in 2016 IEEE Radar Conference (RadarConf), 2016, pp. 1-6.

[11] Q. Chen, B. Tan, K. Chetty, and K. Woodbridge, "Activity Recognition Based on Micro-Doppler Signature with In-Home Wi-Fi," arXiv preprint arXiv:1611.01801, 2016.

[12] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-Wall Sensing of Personnel Using Passive Bistatic WiFi Radar at Standoff Distances," Geoscience and Remote Sensing, IEEE Transactions on, vol. 50, pp. 1218-1226, 2012.

[13] J. Mairal and B. Yu, "Complexity analysis of the lasso regularization path," arXiv preprint arXiv:1205.0079, 2012.

[14] J. Zhuolin, L. Zhe, and L. S. Davis, "Label Consistent K-SVD: Learning a Discriminative Dictionary for Recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, pp. 2651-2664, 2013.

[15] F. K. Coutts, D. Gaglione, C. Clemente, G. Li, I. K. Proudler, and J. J. Soraghan, "Label Consistent K-SVD for sparse micro-Doppler classification," in 2015 IEEE International Conference on Digital Signal Processing (DSP), 2015, pp. 90-94.

[16] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," SIAM Journal on Numerical Analysis, vol. 17, pp. 238-246, 1980.