Dictionary Learning for Classification of Indoor Micro-Doppler Signatures Across Multiple Carriers

Shelly Vishwakarma and Shobha Sundar Ram
Dept. of Electronics and Communication Engineering
Indraprastha Institute of Information Technology Delhi
New Delhi, India 110020
shellyv@iiitd.ac.in and shobha@iiitd.ac.in

Abstract—**Micro-Doppler signatures of dynamic targets such as humans, animals and vehicles are very effective feature vectors for classification based on machine learning algorithms. In the existing works, the test data have been measured in nearly identical operating conditions to the training data that were gathered for the classifiers. However, this assumption may be violated in real life scenarios. In this work, we demonstrate that classification based on sparsity based dictionary learning can overcome this limitation. Here, we learn unique target class dictionaries from micro-Dopplers gathered at multiple carriers. Then we test the classifier using data gathered at another carrier (distinct from those used for training). We test the performance of the classification algorithm for both simulation and measurement data. Our results show a classification accuracy of** 99% **and** 89% **for simulated and measurement data respectively.**

Keywords–micro-Dopplers, classification and, dictionary learning

I. INTRODUCTION

Continuous wave radar scatterings from the motions of non-rigid bodies exhibit micro-Doppler phenomena in the joint time-frequency space [1]–[3]. These micro-Doppler signatures have been extensively researched for target recognition and classification purposes for applications such as military surveillance, law enforcement, bio-medical studies and in search and rescue operations. In [4], the authors exploited time domain information embedded in the micro-Doppler signal in a dynamic time warping algorithm to discriminate between multiple targets. On the other hand, in [5], spectral information was used to discriminate micro-Dopplers. Authors have employed both heuristic methods in [6], [7] and more sophisticated methods such as principle component analysis (PCA), independent component analysis (ICA) and empirical mode decomposition to extract distinctive features from time-frequency spectrograms for classification [8]–[10]. All of these works require domain knowledge to extract the most useful and discriminative features from micro-Doppler data. Alternately, in [11], the authors used an unsupervised feature learning method- deep convolution neural network (DCNN), for human activity classification.

The fundamental assumption underlying all of these works is that the training data, required for machine learning purposes, and the test data must be gathered in similar operating conditions. These conditions include radar parameters such as the carrier frequency, transmitted waveform, power and sensor location as well as target parameters such as the motion trajectory of the target. However, this assumption may be violated frequently in real life conditions. For instance, a particular radar carrier may be unsuitable in an urban environment due to the presence of interference sources such as WiFi. If the radar is operated in a non-line-of-sight environment (through-wall, foliage penetration) the dispersive propagation channel may support a particular carrier over others. Therefore, there is a tremendous advantage in allowing for a reconfigurable radar such as [12], [13] where the radar parameters, such as the carrier frequency, can be easily modified using software during actual deployment. The reconfigurable radar must be accompanied by suitable classification algorithms which can tolerate the diversity in radar parameters and resulting micro-Doppler signatures.

In this work, we propose a dictionary learning based classifier that is trained and tested with distinct carrier frequencies. Dictionary learning is a popular method that has been recently introduced in [14], [15]. It has been used extensively in image processing, audio and video processing, denoising and source separation applications such as energy disaggregation. We have previously demonstrated the advantages of dictionary learning for representing underlying micro-Doppler signals in a sparser fashion compared to fixed data independent transforms like DCT, wavelet or Fourier in [16], [17]. In this work, we use training micro-Doppler data captured at multiple carriers to learn an exclusive model or dictionary for each moving target class. Once learned, these discriminative dictionaries can directly be used to classify test signal measured at either the same or different carrier than that being used at the time of training the corresponding dictionaries.

We demonstrate the performance of this algorithm with both simulation and measurement data. For the measurement data, we gathered micro-Doppler data from 50 individuals (men and women) for three different classes- two humans walking together, a boxing human, single human walking with a stick and a rotating table fan. We gathered the data at five different carrier frequencies - 2.5GHz, 3GHz, 3.5GHz, 4GHz and 4.5GHz. Choice of multiple carriers can be adapted to different radar operating scenarios. We learned the dictionaries from the micro-Doppler data corresponding to four of the five carriers and tested the classifier with data from the fifth carrier. We used sparse representation based classifier (SRC)

for classification of micro-Dopplers from multiple target types. We performed five-fold cross validation to measure the classification accuracy of the proposed technique. Our results show that the classifier uses the learned dictionaries to accurately identify the target returns at different carrier frequencies in 99% of the test cases using simulated data and 89% of the test cases using measured data.

Our paper is organized as follows. In section II, we briefly explain dictionary learning and the classification algorithm used in this paper. In section III, we describe the simulation set up and the results. In section IV, we describe the experimental set up for measurement data collection and the results from the measurement data. Section V concludes the paper.

II. THEORY

Traditional transforms such as Fourier and wavelets have been widely used to represent micro-Doppler signatures for classification purposes [7], [18]. However the selection of optimum parameters for these data independent dictionaries, becomes challenging when multiple targets are present. For instance, the choice of the short time window (dwell time) duration, in the case of short time Fourier transform, is critical for successfully representing the micro-Doppler signature of a dynamic target. However, the dwell time can vary significantly between different targets (say humans and fans). Dictionary learning provides an alternative where unique dictionaries can be learned for each target class from the training data. As a result, the dictionaries represent the signals sparsely and capture their uniqueness. Secondly, since these dictionaries are data dependent, they can be used to learn the features from multiple signatures gathered at different carriers. This flexibility is not available with the data independent transforms.

A. Dictionary Learning Framework

Dictionary learning aims at learning a pool of signal vectors also known as basis or atoms B, which can be tuned to the underlying signal Y. Hence, they yield slightly sparser representations A, of the signal than off-the-shelf fixed dictionaries as shown in (1).

$$Y = BA \qquad (1)$$

This learning problem fundamentally involves minimizing the objective function $J(B, A)$ as shown in (2).

$$J(B, A) = \min_{B, A} \|Y - BA\|_F^2 \ s.t. \|A\|_0 \leq \tau \qquad (2)$$

Here A is τ sparse and $\|.\|_0$ is the l_0 norm of the vector that counts the number of non-zero elements in that vector to ensure sparse representation of the signal Y. Since l_0-minimization is NP-hard [19], it is relaxed to convex l_1-minimization problem shown in (3).

$$J(B, A) = \min_{B, A} \|Y - BA\|_F^2 + \lambda \|A\|_1 \qquad (3)$$

This formulation is simply a constrained Euclidean cost function that measures the quality of signal representation. $\lambda \in \Re$ is the regularization parameter that controls the tradeoff between sparsity and data fitting error.

First consider P dimensional R training signals forming a training matrix Y such that $Y \in \Re^{P \times R}$. These training signals comprise of micro-Doppler data gathered at multiple carrier frequencies. The dictionary, $B \in \Re^{P \times Q}$, is undercomplete if $P > Q$ or overcomplete if $P < Q$. We choose undercomplete representations for each class. We solve for B and A using a two staged iterative procedure where atoms of B and coefficients, A, are updated alternatively as discussed below.

For $k = 1$ stage, B^1 is initialized with randomly selected signal vectors from the data itself. Subsequently, for every k^{th} stage, the coefficients A^k are learned using sparse coding as shown in (4).

$$\{A^k\} = \min_{A^k} \|Y - B^k A^k\|_F^2 + \lambda \|A^k\|_1 \qquad (4)$$

Equation (4) is known as Least Angle Shrinkage and Selection operator (LASSO) which can be solved using a number of basis pursuit algorithms. In this paper, we solve this using the Iterative soft thresholding algorithm (ISTA) suggested by [20].

Once sparse coefficients A^k are estimated, obtaining B^{k+1} becomes a least squares problem [21] which can be solved using the formulation given in (5).

$$\begin{aligned} \{B^{k+1}\} \quad = \quad & \min_{B^{k+1}} \|Y - B^{k+1} A^k\|_F^2 \\ & s.t. \|b_p\|_2^2 \leq 1 \forall p = 1, 2, ...P \end{aligned} \qquad (5)$$

The columns of dictionary are normalized to have l_2-norm less than unity. This two staged process is iterated until the objective function $J(B, A)$ converges or reaches a very low tolerance level. The algorithm, thus, automatically learns discriminative features and classification boundaries from the training dataset.

If there are N target classes, the corresponding dictionaries, B_n can be learned for each nth class using this procedure. Once learned, these dictionaries can be used directly for the task of classification using the sparse representation based classifier (SRC) which is discussed next.

B. Sparse Representation Based Classification

Sparse representation based classifier combines the dictionaries, $B_n, \forall n = 1, 2, 3, ...N$, from all N target classes to form B as shown in (6).

$$B = [B_1 \, B_2 \, B_3 \, ..B_N] \in \Re^{P \times N} \qquad (6)$$

Although we considered undercomplete representation of the dictionaries at the time of training but, the overall dictionary B shown in (6) results into an overcomplete one. We can now classify any test signal Y_{test} by first finding its sparse representation, A_{test}, over B using any l_1-minimization technique as shown in (7).

$$\{A_{test}\} = \min_{A_{test}} \|Y - BA_{test}\|_F^2 + \lambda \|A_{test}\|_1 \qquad (7)$$

Here, the test signal comprises of micro-Doppler data gathered at a distinct carrier frequency from those used in the training stage. In ideal, noise less conditions, the entries of A_{test}, that correspond to the class which the test sample Y_{test} belongs to, will be nonzero while all the other entries (corresponding

to other classes) are zero. Realistically, this implies that the test sample is assigned to the class having the least error of representation amongst all classes as shown in (8) and (9).

$$r_n(Y_{test}) = \|Y_{test} - B_n A_n\|_2^2 \quad \forall n = 1, 2, 3, ...N \quad (8)$$

$$\hat{k} = \min_n r_n(Y_{test}) \quad (9)$$

Here, $\hat{k}$ is the class assigned to the target based on the minimum residual error, $r_n(Y_{test})$, criterion.

III. SIMULATION

In this section, we examine the performance of the classifier when training and test data are gathered at different carrier frequencies. We simulate the time domain micro-Doppler returns from three classes of moving targets- human walking towards radar (FH), human walking away from radar (BH) and a rotating ceiling fan (CF). We have only considered these simple human motions over more complicated motions such as crawling or boxing due to the unavailability of suitable kinematic models. In the past, the radar returns of these motions have been simulated with motion capture data from existing animation databases [22]. But in those cases, the animation data were captured for only a very limited number of test subjects. On the other hand, the popular Boulic model, that is derived from biomechanical experimental data, can be used to simulate a wide variety of human walking motions [23]. In this model, the dynamics of 17 reference points on the human body (head, the shoulders, knees, the neck, spine base, elbows, hips, ankles and toes) are described by 12 time dependent trajectories which are functions of three parameters - the height and the relative velocity of the human and the starting phase of his stride. Thus by varying these parameters, we can describe a corresponding variety of human walking motions. We assume that each target is composed of a few discrete primitives (spheres, plates, cylinders, ellipsoids) with corresponding scattering points and the superposition of the returns from these points give rise to the aggregate reflected signal [3]. We have modeled the ceiling fan using a 3-point scatterer model, where each point scatterer is located midway on three flat rectangular blades. The radar backscattered returns is a function of three parameters- the angular velocity of rotational motion, the blade length and blade width.

A. Simulation Set Up

We assume a monostatic continuous wave radar configuration located at $[10, 0, 0]m$, with three target classes, as shown in Fig. 1.

We assume that the propagation channel consists of only a single moving target at a time. Here, the human is upright along the Z axis and walks along the XY plane either towards the radar or away from the radar. The fan rotates about the Z axis in the XY plane. We simulate the radar returns from the three target classes at multiple carrier frequencies- 2.5GHz, 3GHz, 3GHz, 4GHz and 4.5GHz for a duration of 1s with a sampling frequency of 1KHz.
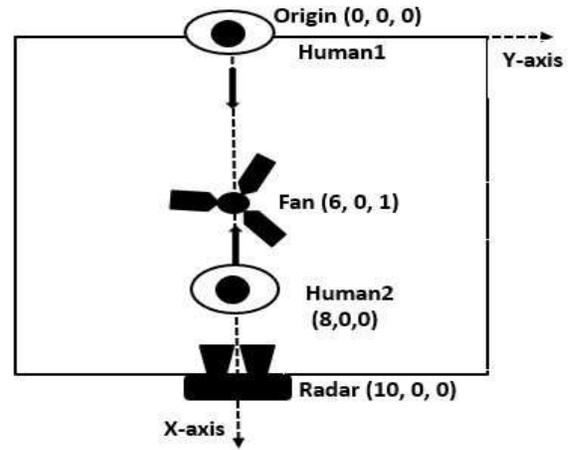

Fig. 1. Room Geometry

In Fig.2, we show the spectrograms generated for simulated human walking towards radar (FH), human walking away from radar (BH) and a ceiling fan (CF) at 4GHz. In the case of the human, the Dopplers are mostly positive when the target is approaching the radar. The strongest returns arise from the torso while the weaker returns are generated from the motions of the arms and legs. Some of the negative Dopplers arise from the backswing of the arms and legs. The spectrogram of the human walking away from the radar is very similar, except for the sign reversal of the Doppler values. The spectrogram of the fan shows very different pattern. The motion of the three blades give rise to three sinusoidal curves. The amplitude of these curves is a function of the product of the blade length and the angular velocity of the fan. As the Dopplers are directly a function of the carrier frequency, there are two direct consequences to the micro-Doppler signatures. Provided the short time duration is kept fixed, the frequency resolution and the Doppler extent of the signatures will be lowered when the carrier frequency is reduced and higher when increased. Second order effects such as scattering, absorption and multiple interactions between the different body parts are not considered in the simulation data.

As mentioned earlier, the position, size and motion of the three target classes are varied to generate a variety of training and test data. We have varied the human height from 1.5m to 1.8m and its relative velocity from $1(H_t/s)$ to $2(H_t/s)$ where H_t is the height from toe to hip as described in [23] and derived from the height of the human. Similarly we simulated multiple distinct cases for rotating fan by varying its blade length from 0.20m to 0.40m, the width of blades from 0.14m to 0.17m and the angular velocity of rotation between 200RPM to 400RPM. Overall, 360 distinct human and fan cases were generated for each particular carrier frequency. Thus for 5 different frequencies we simulated a total of 1800 cases out of which 80% (data from 4 carriers) were used for training and 20% (data from the fifth, distinct carrier) for testing purposes.
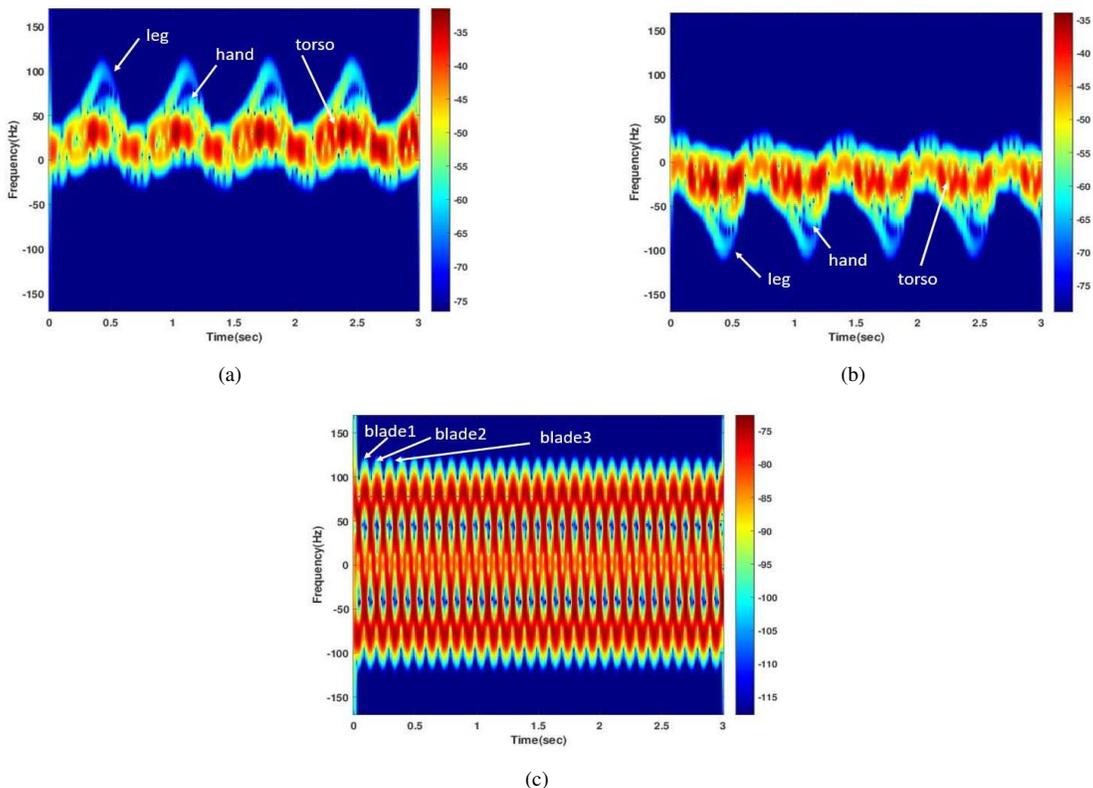
Fig. 2. Spectrogram generated using STFT of a simulated (a) human walking towards radar, (b) a human walking away from the radar and (c) a rotating ceiling fan at 4GHz

B. Results

The training matrix of each target class was of size $[1000 \times 1440]$ with 1400 distinct training cases (80% of total 1800 cases) each with 1000 time domain samples (sampling frequency of 1KHz of 1 second data). We used the two staged dictionary learning technique described earlier to learn human and fan undercomplete dictionaries of size $[1000 \times 500]$ each. We tested our algorithm with test data comprising of returns from a single target simulated at a carrier frequency different from those used while gathering the training data. We performed five-fold cross validation to measure the accuracy of classification using the SRC algorithm. Each test sample was assigned to the class having least residual error computed using (9). The algorithm used a regularization parameter of value 0.01 and ran in MATLAB on a 2.4GHz Intel processor.

Resulting classification accuracies are presented in TABLE I for each fold. The results clearly show that the SRC classifier correctly identifies the different target classes more 99% of the time using their corresponding dictionaries even if both training and testing were performed using distinct datasets extracted at different carrier frequencies.

IV. MEASUREMENTS

We examine the performance of the proposed algorithm in real world scenarios using radar measurement data. Compared to the simulation setup, we have considered more complex

TABLE I
CLASSIFICATION ACCURACY FOR EACH FOLD USING SIMULATED
MICRO-DOPPLER DATA CAPTURED AT MULTIPLE CARRIERS

Cases	Fold 1	Fold 2	Fold 3	Fold 4	Fold5	Average
FH	100	100	100	100	100	100
BH	100	100	100	100	100	100
CF	100	98.33	100	100	96.66	98.99

target motion classes - a boxing human (HB), human walking holding a stick (HHS), two humans walking in opposite directions (TH), and a table fan (TF).

A. Set Up

A monostatic configuration of continuous wave Doppler radar was configured with two linearly polarized double-ridged horn antennas (HF907) separated by a distance of 30cm, and a N9926A FieldFox vector network analyzer (VNA) as shown in Fig. 3. We captured time domain back-scattered returns from the targets using S_{21} measurements of the VNA at multiple carrier frequencies -2.5GHz, 3GHz, 3.5GHz, 4GHz and 4.5GHz. Again, we considered only a single dynamic target class in the propagation channel. The measurements were made over a duration of 2.7s with 1000 samples in indoor line of sight (LOS) conditions. The lower sampling frequency in the measurement data (compared to the simulationd data) is due to the limitations of the time-domain measurement capabilities of the VNA.
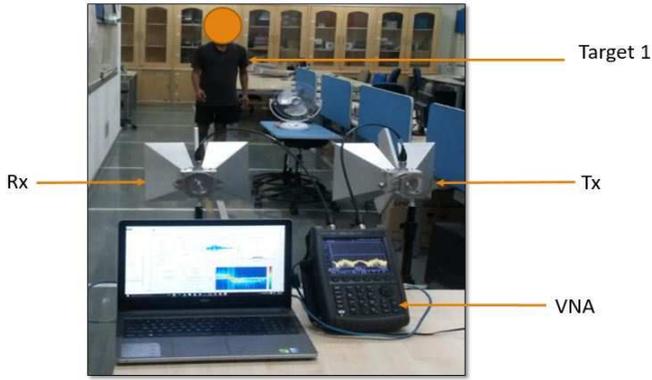
Fig. 3. Experimental setup using continuous wave monostatic radar with two linearly polarized horn antennas (HF907) and a VNA

T/P	TH	HB	HHS	TF
TH	100	0	0	0
HB	0	100	0	0
HHS	15	10	75	0
TF	0	0	0	100

Cases	Fold 1	Fold 2	Fold 3	Fold 4	Fold5	Average
TH	95	100	75	85	80	87
HB	85	100	85	100	100	94
HHS	85	75	85	80	100	85
TF	90	100	90	85	90	91

The spectrograms of all four target classes are shown in Fig. 4. Fig. 4(a), shows the spectrogram generated from the micro-Doppler returns from a boxing human. Here, there is no translational motion of the human body due to which the torso Doppler is at 0Hz. Likewise, there are very weak returns from the legs. The Dopplers, from the hands, follow a periodic pattern corresponding to hand movement in a boxing activity. Doppler are positive when hand moves in the direction of radar and negative when moves in opposite direction. Fig. 4(b) corresponds to spectrogram of a human walking with a stick. The pattern is similar to that shown earlier in Fig.2(a). However, now we see an additional Doppler component arising from the movement of the metallic stick (almost like a third limb motion). Fig. 4(c) captures the motion of two humans walking simultaneously. The one who is approaching the radar has positive Doppler frequencies and the who walks away has negative Dopplers. Fig. 4(d) shows the micro-Doppler from the table fan. This is quite different from the spectrogram from simulated data seen in Fig.2(c). The difference arises due to the lower sampling frequency of the measurement radar that gives rise to aliasing of the signals.

We gathered experimental data from 50 different human subjects, both men and women, (with distinct gaits, heights, velocities and undergoing different activities) and a table fan (with varying angular velocities, distances and orientations with respect to the radar). The human subjects moved between 1m and 9m in front of the radar. We measured in total 100 unique cases for each target class at 5 different carrier frequencies. 80% of the data (corresponding to 4 carrier frequencies) were used for training purposes and 20% (corresponding to the distinct carrier frequency) for testing the performance of algorithm.

B. Results

We applied dictionary learning to measured data from different target classes in a similar manner to the simulation data. We present the confusion matrix of the results of the classification algorithm in TABLE II. In this particular case,

the classifier is trained with data from the following carrier frequencies- 2.5GHz, 3GHz, 3.5GHz, 4GHz while the test data is measured at 4.5GHz. The results show that the human walking with a stick (HHS) and boxing human (HB) are confused in 10% of the cases. This can be attributed to the fact that Dopplers are directly proportional to the carrier frequency. As a result, HHS at lower carrier is getting confused with HB at higher carrier frequencies. Also HHS gets confused with the two humans (TH) class in 15% of the cases. This is because, in the two human case, there are some time instants when the signal from the second target is very weak and therefore, the radar incorrectly assumes that there is only a single mover (usually the target closer to the radar) in the channel. In all the other cases, the algorithm correctly classified each target type with an accuracy of 100%. We performed a five-fold cross validation and observed that the average classification accuracy for TH is 87%, HHS is 85% while HB recorded 94% classification accuracy highest amongst all the classes as shown in TABLE III. Average accuracy across five-folds for TF was 91%.

V. CONCLUSION

In this paper, we used a sparse coding based dictionary learning method to represent the micro-Doppler data from indoor moving targets at multiple carriers for classification purposes. We evaluated the performance using both simulated data and measurement data and demonstrated a high classification accuracy. The primary advantage of the proposed method is that it is flexible and can be used to learn features under diverse radar operating conditions. This provides a tremendous advantage for deploying a reconfigurable radar that can be adapted to challenges in real life scenarios.
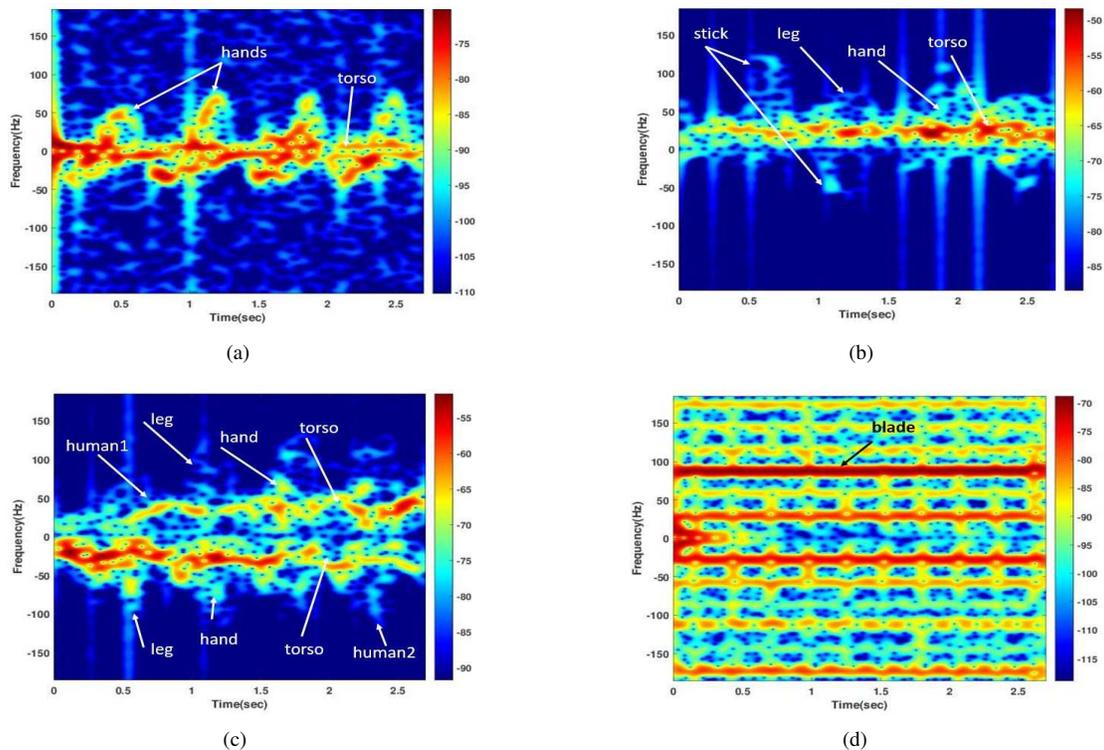
ACKNOWLEDGMENT

Fig. 4. Spectrogram using STFT of (a) a boxing human, (b) a human walking while holding a stick, (c) two walking humans-one walking towards radar and other walking away from radar and (d) a rotating table fan at 4.0GHz

REFERENCES

[1] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Analysis of micro-doppler signatures," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–6, 2003.

[2] V. C. Chen, "Analysis of radar micro-doppler with time-frequency transform," in *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on*. IEEE, 2000, pp. 463–466.

[3] V. Chen, "Doppler signatures of radar backscattering from objects with micro-motions," *IET Signal Processing*, vol. 2, no. 3, pp. 291–300, 2008.

[4] G. E. Smith, K. Woodbridge, and C. J. Baker, "Template based micro-doppler signature classification," in *High Resolution Imaging and Target Classification, 2006. The Institution of Engineering and Technology Seminar on*. IET, 2006, pp. 127–144.

[5] A. Stove and S. Sykes, "A doppler-based automatic target classifier for a battlefield surveillance radar," in *RADAR 2002*. IET, 2002, pp. 419–423.

[6] T. Thayaparan, L. Stanković, and I. Djurović, "Micro-doppler-based target detection and feature extraction in indoor and outdoor environments," *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 700–722, 2008.

[7] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.

[8] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. J. Soraghan, and S. Marshall, "Robust pca for micro-doppler classification using svm on embedded systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2304–2310, 2014.

[9] F. H. C. Tivive, S. L. Phung, and A. Bouzerdoum, "Classification of micro-doppler signatures of human motions using log-gabor filters," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1188–1195, 2015.

[10] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 8, no. 5, pp. 425–434, 2014.

[11] Y. Kim and T. Moon, "Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, 2016.

[12] Y. Wang, Y. Yang, and A. E. Fathy, "A reconfigurable uwb system for real-time through wall imaging applications," in *2010 IEEE Radio and Wireless Symposium (RWS)*. IEEE, 2010, pp. 633–636.

[13] S. Costanzo, F. Spadafora, A. Borgia, H. Moreno, A. Costanzo, and G. DiMassa, "High resolution software defined radar system for target detection," *Journal of Electrical and Computer Engineering*, vol. 2013, p. 7, 2013.

[14] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, "Supervised dictionary learning," in *Advances in neural information processing systems*, 2009, pp. 1033–1040.

[15] Q. Zhang and B. Li, "Discriminative k-svd for dictionary learning in face recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2691–2698.

[16] S. Vishwakarma and S. S Ram, "Classification of multiple targets based on disaggregation of micro-doppler signatures," in *Proceedings of the Asia-Pacific Microwave Conference*. IEEE, 2016.

[17] ——, "Detection of multiple movers based on single channel source separation of their micro-dopplers," *submitted to IEEE Transactions on Aerospace and Electronic Systems*, 2016.

[18] T. Thayaparan, S. Abrol, E. Riseborough, L. Stankovic, D. Lamothe, and G. Duff, "Analysis of radar micro-doppler signatures from experimental helicopter and human data," *IET Radar, Sonar & Navigation*, vol. 1, no. 4, pp. 289–299, 2007.

[19] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[20] I. W. Selesnick and I. Bayram, "Sparse signal estimation by maximally sparse convex optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1078–1092, 2014.

[21] C. M. Bishop, "Pattern recognition," *Machine Learning*, vol. 128, 2006.

[22] B. Erol, C. Karabacak, S. Z. Gürbüz, and A. C. Gürbüz, "Simulation of human micro-doppler signatures with kinect sensor," in *2014 IEEE Radar Conference*. IEEE, 2014, pp. 0863–0868.

[23] R. Boulic, N. M. Thalmann, and D. Thalmann, "A global human walking model with real-time kinematic personification," *The visual computer*, vol. 6, no. 6, pp. 344–358, 1990.