THROUGH-THE-WALL RADAR SIGNAL CLASSIFICATION USING DISCRIMINATIVE DICTIONARY LEARNING

Abdesselam Bouzerdoum, Fok Hing Chi Tivive, and Jia Fei

School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Northfields Avenue, Wollongong NSW 2522, Australia e-mail: {a.bouzerdoum@uow.edu.au, tivive@uow.edu.au, jf981@uowmail.edu.au}

ABSTRACT

Through-the-wall radar imaging is an electromagnetic wave sensing technology capable of detecting targets behind walls, doors, and opaque obstacles. Identification of stationary targets is often achieved by first forming an image of the scene, and then segmenting and classifying the targets of interest. In order to provide prompt and reliable situational awareness, this paper proposes a radar signal classification approach that does not rely on image formation. Here, a dictionary learning based method is employed to classify targets behind a wall using the signals received from individual antennas. The cepstrum coefficients of the high resolution range profile are first extracted as features. Then, the latent consistent K-SVD algorithm is used to learn a discriminative dictionary and a linear classifier simultaneously. Experimental results show that the proposed method can classify individual radar signals with high accuracy, without having recourse to image formation.

Index Terms— Through-the-wall radar imaging, signal classification, dictionary learning, LC-KSVD

1. INTRODUCTION

Through-the-wall radar imaging (TWRI) technology applies electromagnetic (EM) waves to image targets behind opaque obstacles, such as walls or doors, and inside an enclosed building structure. It can produce an image of the interrogated scene and provide vital intelligence information to law enforcement officers, soldiers or search-and-rescue personnel [1]. However, to develop a reliable TWRI system, several challenging tasks have to be addressed, one of which is target classification. For moving target classification, several methods based on micro-Doppler signature have been proposed [2–5]. Change detection and micro-Doppler based techniques are not applicable to stationary target classification. Thus, alternative classification methods are required to identify stationary targets.

Several approaches have been proposed for classifying stationary targets in TWRI [6–9]. Some of the classification methods extract features directly from the radar signal, but the majority of the methods perform image formation and target segmentation before feature extraction [7–9]. Ho *et al.* applied Prony's algorithm and singular value decomposition to estimate the poles of the received signal for different target materials and then used the parameters of the estimated poles as features [6]. However, the estimation of the poles becomes unstable at low signal-to-noise ratio. Balthasar *et al.* employed a compressive classification method using smashed filter to discriminate the number of targets and the target types in TWRI [10]. One issue with the smashed filter is that its performance depends on the stability of the dimensionality reduction of the data manifold [11].

Instead of processing the radar signal for feature extraction, Debes et al extracted statistical and geometrical features from a three-dimensional (3D) beamformed image to classify targets from clutter [7]. They first used the iterated conditional modes or the level set method to segment the image. Then, they estimated the parameters of the Weibull distribution as statistical features and the parameters of the superellipsoids fitted to the segmented voxels as geometric features. Smith and Mobasseri extracted features from a 3D beamformed image and used a naive Bayes classifier to discriminate four different types of targets [13]. The image was segmented by selecting the voxels at 3 dB in intensity within the 3D iso-surface. For each target voxel position, high range resolution profiles (HRRPs) in all three dimensions were extracted as features. The HRRPs were then compressed using principal component analysis to create a feature vector [13]. These image-based classification methods require the formation of a 3D image for feature extraction, which can be a time consuming process.

In this paper, we propose a signal classification method for TWRI based on dictionary learning and sparse representation. The proposed method avoids the image formation step by learning a set of data-driven bases from the received signals to classify different stationary targets behind the wall. Instead of applying a predefined orthoprojector to select a set of measurements or features as described in [10], the radar signal is represented by a set of sparse coefficients, where the bases (or atoms) of the dictionary are learnt from the data. Furthermore, the proposed method is independent of the cross-range resolution since it classifies the received signal at each individual antenna separately.

The remainder of the paper is organized as follows. Section 2 presents the through-the-wall radar signal model. Section 3 describes the dictionary learning based signal classification method, followed by experimental results in Section 4. Finally, Section 5 presents concluding remarks.

2. THROUGH-THE-WALL RADAR SIGNAL MODEL

This section presents the signal model of a through-the-wall monostatic stepped-frequency radar system. Consider a linear array of N antenna elements placed at a certain standoff distance from the wall, where each antenna transmits and receives a wideband steppedfrequency signal comprising M frequencies, ω_m ($m = 0, \ldots, M -$ 1). Assuming that there are P_t point targets behind the wall, the radar signal received by the n-th antenna at the m-th frequency can be expressed as

$$y_{m,n} = \sum_{k=1}^{P_{w}} \sigma_{n}^{w} A_{k} e^{-j\omega_{m}\tau_{n,w}^{k}} + \sum_{i=1}^{P_{t}} \sigma_{i}^{t} e^{-j\omega_{m}\tau_{n,i}} + \epsilon_{m,n}, \quad (1)$$

where σ_n^w is the complex reflectivity of the wall, $\tau_{n,w}^1$ represents the propagation delay associated with the direct return from the wall, $\tau_{n,w}^k$ (k > 1) is the propagation delay associated with the k-th wall reverberation, A_k represents the path loss factor of the k-th wall return, P_w is the number of wall reverberations, σ_i^t is the complex reflectivity of the *i*-th target, $\tau_{n,i}$ is the two-way propagation delay between the *n*-th antenna and the *i*-th target, and $\epsilon_{m,n}$ is the measurement noise. Since the wall returns are relatively stronger than the target echoes, they need to be removed, or at least significantly attenuated, before target classification. Here, the background subtraction technique was used to remove the direct wall returns and wall reverberations from each antenna signal.

3. SIGNAL CLASSIFICATION USING LC-KSVD

After the removal of the wall returns, signal classification is achieved by learning a dictionary over features extracted from individual antenna signals. The following subsections describe in more detail feature extraction, dictionary learning, and classification of radar signals.

3.1. Feature Extraction

Feature extraction is a key step for improving the accuracy and robustness of the classification method. In radar target classification, the received signal is often represented as a HRRP, which depicts the target strength (magnitude) as a function of location (range). Suppose that the maximum unambiguous range R is divided into M bins, with a range resolution Δr , $R = M \Delta r$. Let $\hat{y}_{m,n}$ denote the n-th antenna signal at the m-th frequency after wall clutter mitigation and $z_{k,n}$ be the k-th range bin of the HRRP at the n-th antenna location. The signals $z_{k,n}$ and $\hat{y}_{m,n}$ ($k, m = 0, \ldots, M - 1$) are related by the discrete Fourier transform:

$$z_{k,n} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{y}_{m,n} e^{j\omega_m (2k\Delta r)/c} \quad k = 0, \dots, M-1, \quad (2)$$

where c is the speed of light. After obtaining the HRRP from the radar signal, we can use its magnitude as a feature for classification. In this paper, two other types of features are investigated, namely magnitude spectrum and the cepstrum coefficient of the HRRP itself. The cepstrum coefficient is obtained by computing the inverse discrete Fourier transform of the log-magnitude Fourier spectrum of the HRRP. Mathematically, the cepstrum of the *n*-th HRRP can be written as

$$\boldsymbol{v}_n = |\mathcal{F}^{-1}\{\log(|\mathcal{F}\{|\boldsymbol{z}_n|\}| + 1)\}|, \qquad (3)$$

where $|\cdot|$ denotes the magnitude operator, $z_n = [z_{0,n}, \ldots, z_{M-1,n}]^T$ is the HRRP of the *n*-th antenna signal, and \mathcal{F} and \mathcal{F}^{-1} denote, respectively, the discrete Fourier transform (DFT) and inverse DFT operators.

Due to changes in aspect angle and the position of the target with respect to the antenna location, not all the received signals contain sufficient target information for classification. Therefore, we only classify those antenna signals whose energies are above a certain threshold. Here, the threshold η is computed as the average of the energies of all the HRRPs obtained from the antenna array,

$$\eta = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=0}^{M-1} |z_{k,n}|^2.$$
(4)

3.2. Dictionary Learning

Dictionary learning methods have been employed in various applications, including face recognition [14], speech emotion recognition [15], and EEG signal classification [16]. One of the popular techniques for learning a dictionary from a given data is K-SVD, proposed by Aharon *et al.* [17]. The K-SVD method solves an optimization problem by alternately finding sparse representations given a dictionary and then updating the dictionary. To improve the K-SVD method, Jiang *et al.* incorporated the label information of the data into the optimization problem and proposed the label consistent K-SVD (LC-KSVD) [18]. These two dictionary learning methods are briefly described in the following subsections.

3.2.1. K-SVD

Let $\mathbf{Y} \in \mathbb{R}^{M \times P}$ denote a set of *M*-dimensional feature vectors extracted from the HRRPs, where *P* is the number of feature vectors. To learn a dictionary $\mathbf{D} \in \mathbb{R}^{L \times G}$, G > M, the K-SVD algorithm solves the following optimization problem:

$$\langle \mathbf{D}^*, \mathbf{X}^* \rangle = \operatorname*{arg\,min}_{\mathbf{D}, \mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2$$

s.t. $\|\boldsymbol{x}_i\|_0 \le K, \quad \forall i \quad (5)$

where $\mathbf{X} \in \mathbb{R}^{G \times P}$ is a sparse matrix in which each column \boldsymbol{x}_i contains no more than K non-zero entries. After learning the dictionary \mathbf{D}^* and computing the sparse representations of the radar signals \mathbf{X}^* , we can train a linear classifier by solving the following linear ridge regression problem:

$$\mathbf{W}^* = \operatorname*{arg\,min}_{\mathbf{W}} \|\mathbf{H} - \mathbf{W}\mathbf{X}^*\|_F^2 + \lambda \,\|\mathbf{W}\|_F^2, \qquad (6)$$

where $\mathbf{H} \in \mathbb{R}^{C \times P}$ denotes the label matrix with C classes and λ is a regularization term.

3.2.2. LC-KSVD

The principle of LC-KSVD method is to learn a discriminative dictionary and train a linear classifier simultaneously. In [18], the dictionary learning and classifier training are formulated as a constrained optimization problem, which incorporates a discriminative sparse-code error term to encourage similarity among sparse representations of the signals of the same class. The constrained optimization problem is given by

$$\langle \mathbf{D}^*, \mathbf{W}^*, \mathbf{A}^*, \mathbf{X}^* \rangle = \underset{\mathbf{D}, \mathbf{W}, \mathbf{A}, \mathbf{X}}{\arg \min} \| \mathbf{Y} - \mathbf{D} \mathbf{X} \|_F^2 + \alpha \| \mathbf{Q} - \mathbf{A} \mathbf{X} \|_F^2$$

+ $\beta \| \mathbf{H} - \mathbf{W} \mathbf{X} \|_F^2 \quad \text{s.t.} \quad \| \boldsymbol{x}_i \|_0 \le K, \quad \forall i \quad (7)$

where $\mathbf{Q} \in \mathbb{R}^{G \times P}$ is the discriminative sparse-code matrix, $\mathbf{A} \in \mathbb{R}^{G \times G}$ is a linear transformation, and α and β are the regularization parameters that control the relative contribution of the discriminative sparse-code error and the classification error terms, respectively.

Once the dictionary D^* and the matrix weight W^* of the linear classifier are obtained using LC-KSVD, the classification of a test sample y is performed as follows. First, the sparse representation of the test sample is calculated by solving

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \mathbf{D}^* \boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_0 \le K.$$
(8)

Let $\mathbf{W}^* = [w_{ij}^*]$ and $\boldsymbol{x}^* = [x_1^*, \dots, x_G^*]^T$. Then, the class label of the test sample is determined as

$$l = \arg\max_{i} \sum_{j=1}^{G} w_{ij}^{*} x_{j}^{*}, \qquad i = 1, \dots, C.$$
(9)

4. EXPERIMENT RESULTS

The proposed radar signal classification approach is tested on synthetic data obtained with EM numerical simulation using XFDTD software. The experimental setup and data collection, including the imaged scene, the radar parameters, and the targets, are described in the next section. This is followed by an evaluation of the proposed classification method.

4.1. Experimental Setup and Data Collection

A 41-element antenna array covering a horizontal scanning distance of 2.4 m was placed at a standoff distance of 1.0 m in front of a homogeneous wall. The imaged scene depicted in Fig. 1(a) was a 3 m \times 4 m area (i.e., 3 m down-range and 4 m cross-range) behind the wall of thickness 0.15 m and a dielectric constant of 7.76. A steppedfrequency signal of 1.1 GHz bandwidth centered at 2.75 GHz with a step size of 5 MHz was used to interrogate the scene behind the wall. Five metallic objects with similar cross-section-a dihedral, a trihedral, a square plate, a cylinder, and a sphere-were placed behind the wall. Each target was placed at five different positions in the scene: top-left (P1), top-right (P2), center (P3), bottom-left (P4), and bottom-right (P5). Figure 1 depicts a schematic diagram of the TWRI scene along with the scene layout showing the target locations. For each target type, a set of 41 signals was collected from the antenna array at each target position (P1-P5). Five-fold crossvalidation was used to estimate the classification rate (CR) in all experiments. In each validation fold, all the radar signals at four target positions were used for training and the radar signals whose HRRP energies were greater than a given threshold at the remaining target position were selected for estimating the CR. This was repeated five times for different test sets and the final CR was computed as the percentage of correctly classified samples, which were aggregated across all the validation folds.

4.2. Performance of Feature Extraction Methods

Three types of features are used to form the discriminative dictionary: the magnitude of the HRRP, the magnitude of the Fourier coefficient of the HRRP, and the cepstrum coefficient of the HRRP. The parameters of LC-KSVD are set as follows: the dictionary size G = 700, the sparsity level K = 25, and the regularization parameters $\alpha = 4$ and $\beta = 2$. The CR for each feature type is obtained by using five-fold cross-validation. Table 1 presents the CRs of LC-KSVD using the three different types of features considered here. Among the three types of features, the cepstrum coefficient gives the highest CR, followed by the magnitude of the Fourier coefficient. Learning the dictionary using only the magnitude of the HRRP achieves the lowest CR of 87.9%.



Fig. 1. (a) Schematic diagram of the simulated TWRI scene, and (b) scene layout and target location.

 Table 1. Classification rate of LC-KSVD using different types of features.

Feature	CR
Cepstrum coefficient of the HRRP	96.0%
Magnitude Fourier coefficient of the HRRP	93.9%
Magnitude of the HRRP	87.9%

4.3. Effects of Dictionary Size and Sparsity Level on Classification Rate

The dictionary size and sparsity level both have an influence on the CR of LC-KSVD. Several simulations are performed by varying the number of atoms in the dictionary from 200 to 1400, with a step size of 100 and the sparsity level from 10 to 100, with a step size of 10. Figure 2 illustrates the evolution of the CR of LC-KSVD as a function of the dictionary size for a fixed sparsity level of 20. The highest CR is achieved when the number of atoms reaches 800. Further increasing the number of atoms in the dictionary barely improves the CR of LC-KSVD. Figure 3, on the other hand, depicts the CR of LC-KSVD as a function of the sparsity level with the dictionary size fixed to 800. From this figure, it is clear that increasing the sparsity level beyond 20 decreases the CR of the dictionary learning method. In the following experiments, the dictionary size and the sparsity level are set to 800 and 20, respectively.



Fig. 2. Classification rate of LC-KSVD with K = 20.



Fig. 3. Classification rate of LC-KSVD with G = 800.

4.4. Effect of Target Position on CR

The aim here is to investigate the influence of target location on the CR of LC-KSVD. The dictionary and classifier are trained with sig-

nals from 4 of the 5 target positions, and they are tested on signals from the remaining target position. Here, the cepstrum coefficients are used for training the discriminative dictionary. Table 2 presents the CR at the five target locations: P1, P2, P3, P4 and P5. The dictionary learning method achieves 100% CR at 4 target locations.

 Table 2. CRs of LC-KSVD when a stationary target is placed at different positions behind the wall.

	Target position						
	P1	P2	P3	P4	P5		
CR	100%	100%	93.0%	100%	100%		

In terms of classification of different types of targets, the confusion matrix of the dictionary learning method is shown in Table 3. The main diagonal of the matrix lists the correct CR for each target type. The off-diagonal entries indicate misclassification rates. The entry at (i, j) gives the CR of the target known to be in class *i* but classified as class *j*. For instance, the cylinder target is misclassified as a plate with a rate of 2.6%. Other targets such as the plate and sphere targets are classified perfectly. The overall CR of the system is 98.1%.

 Table 3. Confusion matrix of LC-KSVD for different types of stationary targets.

		Predicted					
		Plate	Cylinder	Dihedral	Sphere	Trihedral	
Actual	Plate	100%	0%	0%	0%	0%	
	Cylinder	2.6%	97.4%	0%	0%	0%	
	Dihedral	5.3%	0%	94.7%	0%	0%	
	Sphere	0%	0%	0%	100%	0%	
	Trihedral	0%	0%	1.7%	0%	98.3%	

5. CONCLUSION

This paper proposed a technique for through-the-wall radar signal classification, which does not require the image formation. The proposed method employs dictionary learning and sparse representation to classify individual antenna signals. Features extracted from HRRP, such as cepstrum coefficients, are used to jointly learn a discriminative dictionary and train a linear classifier. Experiments based on simulated data show that it is possible to classify the behind-the wall targets from the individual radar signals with high classification rates.

Acknowledgment

This work is supported by a grant from the Australian Research Council (ARC).

6. REFERENCES

- M. G. Amin (Ed.), *Through-the-wall radar imaging*, CRC Press, Boca Raton, FL, USA, 2010.
- [2] 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.
- [3] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 2015–2023, 2010.
- [4] F. H. C. Tivive, S. L. Phung, and A. Bouzerdoum, "Classification of micro-doppler signatures of human motions using loggabor filters," *IET Radar, Sonar and Navigation*, vol. 9, no. 9, pp. 1188-1195, 2015.
- [5] M. G. Amin, F. Ahmad, Y. D. Zhang, and B. Boashash, "Human gait recognition with cane assistive device using quadratic timefrequency distributions," *IET Radar, Sonar and Navigation*, vol. 9, no. 9, pp. 1224-1230, 2015.
- [6] Anthony T. S. Ho, W. H. Tham, and K. S. Low, "Improving classification accuracy in through-wall radar imaging using hybrid prony's and singular value decomposition method," *IEEE International Geoscience and Remote Sensing Symposium*, pp. 4267–4270, 2005.
- [7] C. Debes, J. Hahn, A. M. Zoubir, and M. G. Amin, "Target discrimination and classification in through-the-wall radar imaging," *IEEE Transaction on Signal Processing*, vol. 59, no. 10, pp. 4664–4676, 2011.
- [8] A.A. Mostafa, C. Debes, and A. M. Zoubir, "Segmentation by classification for through-the-wall radar imaging using polarization signatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 9, pp. 3425–3439, 2012.
- [9] V. Kilaru, M. Amin, F. Ahmad, P. Sevigny, and D. DiFilippo, "Gaussian mixture modeling approach for stationary human identification in through-the-wall radar imagery," *Journal of Electronic Imaging*, vol. 24, no. 1, pp. 013028 1–013028 11, 2015.
- [10] M. R. Balthasar, M. Leigsnering, and A. M. Zoubir, "Compressive classification for through-the-wall radar imaging," *Proc. of the 20th European Signal Processing Conference*, pp. 2288– 2292, Bucharest, Romania, 2012.
- [11] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," *Proc. SPIE, Computational Imaging V*, vol. 6498, pp. 64980H 1– 64980H 12, California, USA, 2007.
- [12] C. Debes, M. G. Amin, and A. M. Zoubir, "Target detection in single- and multiple-view through-the-wall radar imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1349–1361, 2009.
- [13] G. E. Smith and B. G. Mobasseri, "Robust through-the-wall radar image classification using a target-model alignment procedure," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 754–767, 2012.
- [14] Q. Zhang and B. Li, "Discriminative k-svd for dictionary learning in face recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2691–2698, California, USA, 2010.

- [15] M. J. Gangeh, P. Fewzee, A. Ghodsi, M. S. Kamel, and F. Karray, "Multiview supervised dictionary learning in speech emotion recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 6, pp. 1056–1068, 2014.
- [16] W. Zhou, Ya Yang, and Z. Yu, "Discriminative dictionary learning for eeg signal classification in brain-computer interface," *IEEE 12th International Conference on Control, Automation, Robotics and Vision*, pp. 1582–1585, Guangzhou, China, 2012.
- [17] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [18] Zhuolin Jiang, Zhe Lin, and Larry S. Davis, "Label consistent k-svd: learning a discriminative dictionary for recognition," *IEEE Transactions on Pattern Analysis and Machine Intelli*gence, vol. 35, no. 11, pp. 2651–2664, 2013.