

## Pulse Doppler Radar Target Recognition using a Two-Stage SVM Procedure

# Correspondence

It is possible to detect and classify moving and stationary targets using ground surveillance pulse-Doppler radars (PDRs). A two-stage support vector machine (SVM) based target classification scheme is described here. The first stage tries to estimate the most descriptive temporal segment of the radar echo signal and the target signal is classified using the selected temporal segment in the second stage. Mel-frequency cepstral coefficients of radar echo signals are used as feature vectors in both stages. The proposed system is compared with the covariance and Gaussian mixture model (GMM) based classifiers. The effects of the window duration and number of feature parameters over classification performance are also investigated. Experimental results are presented.

### I. INTRODUCTION

Detection and classification of ground moving and stationary targets are among the main functions of ground surveillance pulse-Doppler radars (PDRs). Human operators take an essential part in the target classification based on Doppler frequencies of an object. Trained operators can classify a target with a reasonable degree of accuracy by listening to the audio tone of the target. However, this audio-based classification scheme increases the work load of the operator. As a result he or she may not properly execute other radar functions. In addition, operators have to be trained to recognize PDR echo sounds [1].

Therefore, an automatic classification system will be an important improvement and will provide valuable support for ground surveillance pulse-Doppler operators [2, 3]. In [2] preliminary results of radar target recognition using speech recognition based methods [4] are reported. In [3] Doppler signatures and cepstrum feature parameters are classified using hidden Markov models (HMMs) and Gaussian mixture models (GMMs) and an average recognition rate of 88% is achieved. It is reported that a neural network classifier performs much worse than the HMM-GMM based classifier [3]. This is probably due to the nonstationary nature of the spectrum of target signal. Recently, another HMM-based approach presenting a PDR target recognition method was

Manuscript received October 24, 2009; revised March 29, 2010; released for publication July 7, 2010.

IEEE Log No. T-AES/47/2/940856.

Refereeing of this contribution was handled by C. Baker.

---

0018-9251/11/\$26.00 © 2011 IEEE

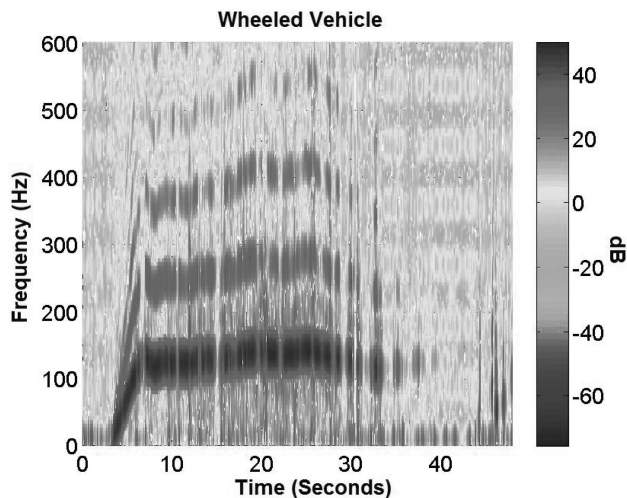


Fig. 1. Spectrogram of PDR echo signal of wheeled vehicle.

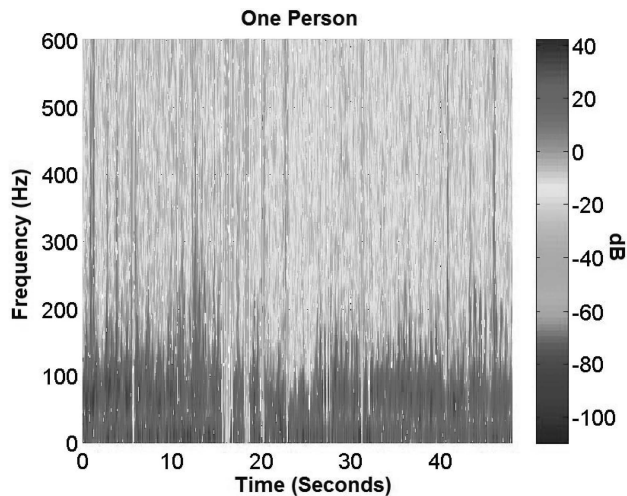


Fig. 3. Spectrogram of PDR echo signal of single person.

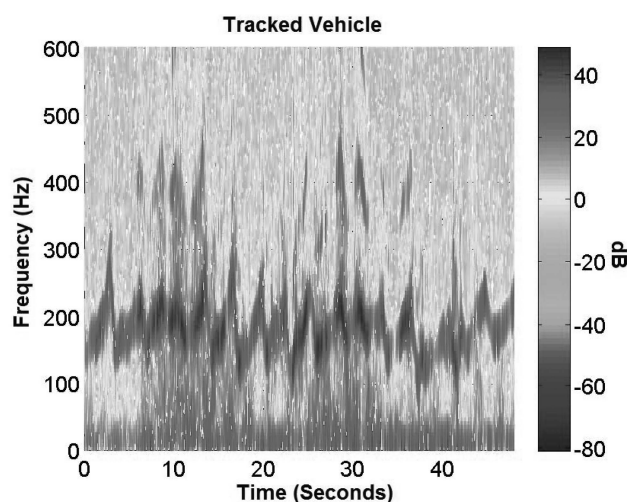


Fig. 2. Spectrogram of PDR echo signal of tracked vehicle.

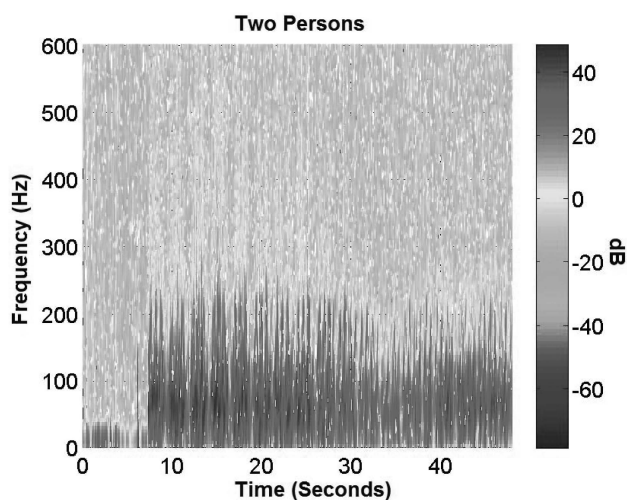


Fig. 4. Spectrogram of PDR echo signal of two persons.

reported in [5]. This recognition method relies on the track data for determining the target dynamics, and uses the mel-scale cepstrum for feature extraction from the recorded data.

Support vector machines (SVMs) were successfully used in many recognition and classification problems [6]. In this article we propose the use of a two-stage SVM-based approach. In the first stage, a set of SVMs tries to estimate the most descriptive temporal segment of the radar echo signal and in the second stage another set of SVMs classifies the target using the selected temporal segment. It is experimentally shown that the SVM based approach provides superior classification accuracies compared with other recent methods in a publicly available dataset [7]. Mel-frequency cepstrum coefficients (MFCCs) are used as feature parameters. They are extensively used in speech recognition and sound classification [4, 8–12]. The proposed method is compared with the GMM-based classifier and a recently introduced covariance matrix based approach, which is used in object detection in images [13].

This paper is organized as follows. In Section II A, MFCC feature extraction from PDR signals is described. In Section III, our novel two-stage SVM approach is introduced and presented. In Section IV the covariance matrix method is explained. In Section V experimental classification results are presented. Finally, concluding remarks are presented in Section VI.

## II. PDR ECHO SIGNALS AND CEPSTRAL FEATURE EXTRACTION FROM PDR SIGNALS

In this section, properties of PDR echo signals are reviewed and the cepstral domain feature extraction process from the radar echo signals is described.

The PDR echo signals used in this paper are collected by a 9 GHz ground surveillance radar [7]. The radar has 3 MHz bandwidth, 12  $\mu$ s pulsewidth, 125 m range resolution and 4 deg azimuth resolution. Signals are recorded with a sampling frequency of 5.682 KHz. Spectrograms of some PDR echo signals are shown in Figs. 1–5. The recording procedure was

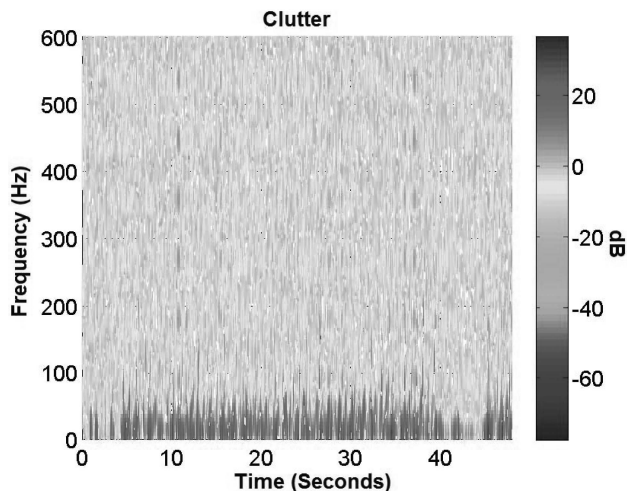


Fig. 5. Spectrogram of PDR echo signal of clutter.

explained in [7] and [14] in detail. Targets from the following categories were recorded in the publicly available database [7]:

- a wheeled vehicle,
- a tracked vehicle,
- a single person,
- two persons,
- the vegetation clutter.

Data collection is done in controlled environment and conditions. The recorded targets were within the line-of-sight of the radar in an open field, in the presence of ground clutter with low vegetation and without any interference. One target at a time was recorded in each case. Target motions are controlled and they were close to the radar to obtain high signal-to-noise ratio (SNR) PDR signals. All targets were 200–600 m away from the radar. For each case, only one target was recorded at a time [14]. During recording each target was detected and tracked automatically by the radar, allowing continuous target echo records. Target signals were obtained at different speeds (slow (10 to 20 km/hr), normal (20 to 30 km/hr), and fast (30 to 90 km/hr) for vehicles and 2 to 9 km/hr for people) and various bearing angles (0, 15, 30, 45, 60 deg) towards the radar. Target people data were also obtained at (0, 15, 30, 45, 60 deg) bearing angles towards the radar.

Target signal spectrograms of a wheeled vehicle, a tracked vehicle, one and two persons, and the clutter are shown in Figs. 1, 2, 3, 4, and 5, respectively. These plots indicate that targets have (nonstationary) time-varying spectrograms. This is because targets are moving in this study. Therefore a single feature vector is not sufficient to represent a given target signal. Similar to the speech and audio analysis, the target signal should be divided into short time windows and a representative feature vector should be obtained for each window of the PDR signal of a given target. It is assumed that the signal is stationary in each short

time window as in speech analysis. As a result, a sequence of MFCC feature vectors representing the recorded target echo signal is obtained. The nature of the target can be determined from the sequence of MFCC vectors.

#### A. Cepstral Feature Extraction

In this paper, MFCCs are used as feature parameters to represent radar target echo signals as in [2] and [3]. In this section the real cepstrum is reviewed first. Afterwards, the MFCC is reviewed. As pointed out above the MFCC is the most widely used feature parameter in speech and speaker recognition.

In [15] the one-dimensional (1-D) cepstrum is introduced as the inverse Fourier transform of the log magnitude spectrum of a signal. This is also called as real cepstrum. The cepstrum  $\hat{x}[n]$  of a discrete-time signal  $x[n]$  is defined as follows:

$$\hat{x}[n] = F^{-1}(\log(|X(e^{j\omega})|)) \quad (1)$$

where  $(|X(e^{j\omega})|)$  is the magnitude of the discrete-time Fourier transform (DTFT) of the signal  $x[n]$ .

The MFCC method introduced by Davis and Mermelstein [8] is basically a variant of cepstrum representation. Instead of the linearly-spaced frequency values used in the normal cepstrum, logarithmically-spaced bands are used in mel-frequency cepstrum, which approximates the human auditory system's response more closely than the real cepstrum. This frequency warping allows a better representation of speech signals than the linearly-spaced frequency domain bands. The logarithmic spacing of bands also provides a more descriptive representation of PDR target echo signals because most of signal energy is in low frequency bands as shown in spectrogram plots (Figs. 1–5). Computation of MFCCs is carried out using the discrete cosine transform (DCT) of a log power spectrum on a nonlinear mel-scale of frequency domain [9] as follows:

$$\tilde{x}[i] = \sum_{k=1}^N X_k \cos \left[ i \left( k - \frac{1}{2} \right) \frac{\pi}{20} \right], \quad i = 1, 2, \dots, M \quad (2)$$

where  $M$  is the number of MFCC coefficients and  $X_k$ ,  $k = 1, 2, \dots, N$  represents the log-energy output of the  $k$ th bandpass filter, which divide the spectrum in a logarithmic manner [8, 10, 11]. As described above the target echo signal is divided into short time windows of 50 ms duration as in speech and for each time window of the PDR target echo signal a mel-cepstral vector containing the MFCC coefficients  $\tilde{x}[i]$  is obtained. In speech processing, a sequence of  $M$ -dimensional MFCC vectors obtained from overlapping time windows is computed to represent a given word or an utterance, and these

MFCC vectors are successfully used in modeling various unstationary audio and sound signals including impact acoustical signals [10–12]. Similarly, a sequence of MFCC vectors is obtained for a given PDR signal. Classification of the target echo sound will be based on the sequence of MFCC vectors.

### III. SUPPORT VECTOR MACHINE BASED TARGET CLASSIFICATION APPROACH

Recently, SVMs have been successfully used in a wide range of classification problems [6, 16–18]. In this section, a two-stage SVM-based classification method employing MFCC feature vectors is developed for PDR target echo signal classification as shown in Fig. 6.

An ordinary SVM [6] is a binary decision engine providing a decision for a given input feature vector according to decision surfaces in an  $M$ -dimensional vector space of feature vectors. Decision boundaries are determined during the training phase of the SVM. Decision boundaries are hyperplanes in  $M$ -dimensional spaces. During the recognition phase, the goal is to decide which class a new feature vector will belong to. A decision about a test input is reached by comparing the test vector to the decision hyperplane. Therefore, SVMs are not computationally costly decision engines during the recognition process.

Direct application of SVMs yields poor performance on speaker identification and speech recognition, as indicated in [17]. This is because a single  $M$ -dimensional MFCC feature vector is not enough to represent a given word or utterance as the speech signal is nonstationary and SVMs do not have a built-in structure to handle time-varying signals and problems. Similarly, PDR signals are not stationary signals, either. Therefore, the temporal behavior of the signal has to be incorporated into the classification process. The proposed two-stage SVM classification system provides a solution to the classification of time-varying radar echo signals.

Although the overall system can be trained with long duration echo signals, only a short duration echo signal may be available during the recognition phase. Therefore a matching temporal region has to be determined for decision making. The goal of the first stage of our system is to determine the most representative temporal region of the given target signal. The target signal is classified based on the selected temporal region in the second stage.

The proposed system uses the “simulated probability” concept in SVMs [18]. Ordinary SVM classifiers simply determine where the given feature vector is in the feature space and provide the corresponding class label (approximate target value). They do not provide any probability information. Probabilities can be assigned to feature vectors in the SVM framework according to the distance of the feature vector to the decision surface [18]. If the

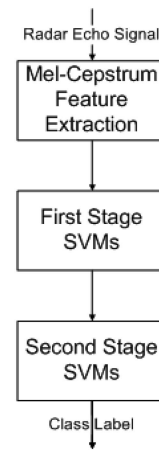


Fig. 6. Two-stage PDR echo signal classification scheme based on SVMs.

feature vector is close to (far away from) the decision surface a low (high) probability value is assigned. Probability estimates, which provide information about the confidence of the decision, are used to select a representative time category for a given target in the proposed classification method.

In the next subsections, the training and testing procedures of the two-stage SVM classification system are described.

#### A. Training and Testing of the First Stage SVMs

The aim of this stage is to match the PDR echo signal of a target to a corresponding time category in the training database. For example, the first 5 s and the last 20 s of the wheeled vehicle sound have different frequency characteristics compared with the time window of [5, 20] seconds as shown in Fig. 1. This is because the target PDR signal shows different Doppler characteristics when the target enters and leaves the beam of the radar. Similarly, the initial and the last portions of all target signals are different from the middle part the signal. After examining various target signals we decided to divide the recorded target PDR signals into four categories: the initial segment, the first middle segment, the second middle segment, and the last segment. This is similar to the state concept in HMMs. During the training phase MFCC vectors extracted from short time segments are used to train SVMs corresponding to four time categories (or time states). The first stage SVMs determine a matching time category among four possible cases for a given 5 s or shorter duration radar echo signal and this decision is used by the second stage SVMs instead of using a time-averaged representation of the target MFCC vectors during classification.

As described in the previous section the decision space of SVMs are formed from the MFCC feature vectors, which are computed from 50 ms long short time windows. Since the duration of the recorded PDR signal is in the order of seconds, a sequence of

MFCC feature vectors is computed for a given echo signal.

During the recognition phase, the target echo signal is assumed to be shorter than 5 s (if not, it can be divided into 5-s long segments and the following analysis can be carried out for each 5-s long segment). The MFCC vectors obtained from 50 ms short time windows of the PDR signal are computed. The first stage SVMs produce a probability estimate for each MFCC feature vector. For each time category, an average probability value is obtained as follows

$$p_l = \frac{1}{L} \sum_{i=1}^L p_{i,l}, \quad l = 1, 2, 3, 4 \quad (3)$$

where  $L$  is the number of MFCC vectors extracted from the test signal,  $p_{i,l}$  is the probability estimate for the  $i$ th MFCC vector by the  $l$ th SVM corresponding to the  $l$ th time category. The time state of the input test signal is simply selected according to the maximum  $p_l$  value.

The above framework can determine a time category even for a single 50 ms duration PDR echo signal. However, this will not be reliable. At least a 1-s long time segment ( $L = 20$ ) is necessary to get reasonable probability estimates in (3). When we have a 5-s long target echo data,  $L = 100$ , which provides reliable probability estimates in (3).

In one-against-all classification, the binary SVM of each category separates members of that category from members of other classes. In this approach, there may be a gray area in the decision space in a multi-class classification problem, when only binary decisions are allowed. Since we use probability estimates, it is highly unlikely that the probability estimates  $p_l$  in (3) will become exactly equal to each other. We have also not encountered any classification problems in our experimental studies.

#### B. Training and Testing of the Second Stage SVMs

In an ordinary SVM-based classification scheme, a single SVM for each target class is designed and tested according to one-against-all strategy to determine the type of the target. In this paper, an SVM for each target class is designed for each time category. Since the nature of a typical radar echo signal is different for each time category, it is better to design different SVMs for each time category for a given target class. To make a final decision, the SVM of the selected time category is used. The decision space of second stage SVMs is also formed from the MFCC feature vectors, which are computed from 50-ms long time windows as in the first stage SVMs.

During the recognition phase, it is assumed that the time category of the test signal is determined in the first stage. MFCC vectors of the input test signal are processed by the target class SVMs of the current time category, which produce a probability estimate

for each feature vector. For the  $k$ th target class, an average probability value is obtained as follows

$$q_k = \frac{1}{L} \sum_{i=1}^L q_{i,k} \quad (4)$$

where  $q_{i,k}$  is the probability estimate for the  $i$ th MFCC vector and  $L$  is the number of MFCC vectors extracted from the test signal. The target class is determined according to the maximum  $q_k$ .

If the duration of the PDR signal is longer than 5 s, then the probability values  $q_k$  are computed for each 5-s long segment and the target category is determined according to the highest  $q_k$  value of all segments.

We compared the proposed two-stage SVM procedure to the GMM method and the covariance matrix based method and experimentally observed that the proposed two-stage approach provides higher classification rates. We present our experimental results in Section V. We do not describe the well-known GMM method here but briefly review the recently introduced covariance matrix method in the next section.

#### IV. THE COVARIANCE MATRIX BASED PDR SIGNAL CLASSIFICATION

Porikli et al. introduced the covariance matrix method as a new image region descriptor, and experimentally showed that the covariance method is superior to other image texture classification methods [13]. A multiplier-less operator is also introduced in [19] approximating the covariance matrix. This multiplier-less operator reduces the computational cost and increases the performance of the covariance matrix based method in some image processing problems. It is experimentally observed that the covariance method provides the second-best PDR signal classification results after the two-stage SVM process described in the previous section.

We construct a covariance matrix from MFCC coefficients instead of the actual signal samples. We divide the PDR signal into short time windows of duration 50 ms and compute an MFCC vector in each window as in the SVM-based method introduced in the previous section. Let  $\mathbf{x}$  be a  $d$ -dimensional MFCC feature vector for a short time window. Let us assume that there are  $N$  time segments in a given PDR echo signal. As a result we have  $N$   $d$ -dimensional feature vectors  $(\mathbf{x}_k)_{k=1,\dots,N}$ .

The covariance matrix of a given PDR signal is defined as follows

$$\Sigma = \frac{1}{N-1} \sum_{k=1}^N (\mathbf{x}_k - \mu)(\mathbf{x}_k - \mu)^T \quad (5)$$

where  $\mu$  is the mean vector of the feature vectors.

Euclidean distance measure is not a good distance measure for covariance matrices [20]. Since many machine learning methods operate on Euclidean spaces, they cannot be used in a straightforward manner for covariance matrix parameters [13]. The nearest neighbor (NN) algorithm is used as the classifier. A generalized eigenvalue based distance metric is used to compare covariance matrices in the NN algorithm. This distance measure was introduced in [20], and used as a part of the NN method:

$$D(\Sigma_1, \Sigma_2) = \sqrt{\sum_{k=1}^d \log^2 \lambda_k(\Sigma_1, \Sigma_2)} \quad (6)$$

where  $\lambda_i(\Sigma_1, \Sigma_2)$  are the generalized eigenvalues of covariance matrices  $\Sigma_1$  and  $\Sigma_2$ . During the training phase covariance matrices of the training data are computed and they are stored. During the recognition phase distances between the instance covariance matrix to be classified and the covariance matrices in the training database are computed. Then, the test instance is assigned to the class of its NN. In the K-NN approach the test instance is compared with the K NNs in the training set using (6).

The covariance matrix combines multiple features which may be correlated. The diagonal entries of the covariance matrix reflect the variance of each feature and the nondiagonal entries reflect the correlations. For radar signals, correlation is an important property to be exploited since consecutive signal segments include information about the same target. Furthermore, the averaging operation in the covariance computation filters out the noise which corrupts the signal [13].

Recognition results of the covariance matrix method are close to the two-stage SVM procedure in the Ben-Gurion University [7] database. It may even provide the best results in other databases.

## V. EXPERIMENTAL RESULTS

In this section the classification results for a single-stage SVM method, the proposed two-stage SVM described in Section III, the covariance method described in the previous section, and the well-known GMM method are presented. As mentioned in Section IIA there are five PDR signal classes corresponding to wheeled vehicles, tracked vehicles, a single person, two persons, and the clutter. As pointed out in Section II, the PDR data collection process is described in detail in the document located at (<http://www.ee.bgu.ac.il/~testproj/ReadMe.doc>).

MFCC vectors are used as feature parameters in all methods. It is experimentally observed that MFCC is superior to ordinary real cepstrum as in speech recognition and speaker identification problems. A total of 1496 PDR MFCC vectors for each category from the database located at Ben-Gurion University

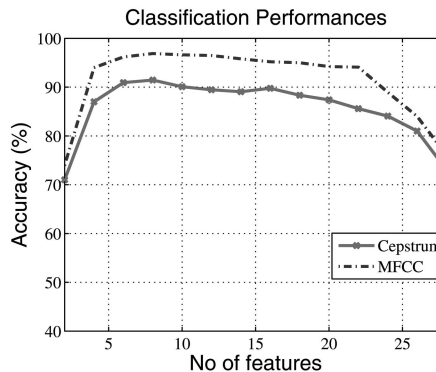


Fig. 7. Classification accuracy of GMM method using cepstrum and MFCC features for five-class PDR classification problem as function of number of feature vector entries. Model order is 10 for all classes.

TABLE I  
Confusion Matrix of GMM Classifier with MFCC Coefficients in Five-Class Problem

	Wheeled	Tracked	One Person	Two Persons	Clutter
Wheeled	92.1	7.0	0	0	0.9
Tracked	4.1	95.9	0	0	0
One Person	0	0	95.0	5.0	0
Two Persons	0	0	0.3	99.7	0
Clutter	0	0	0	2.8	97.2

[7] are used for training and 1056 test vectors are used for testing purposes. MFCC vectors are extracted from 50-ms long short time windows. Since a human operator classifies targets by listening to the PDR echo sounds, the short time window duration is selected as the same as the window duration in speech recognition applications. The data used in training is not included in the test set. As can be seen from Figs. 1, 2, 3, 4, and 5, the PDR signal can be assumed to be stationary in short-time windows. Therefore, an MFCC vector extracted from a 50-ms long window will represent that window in an accurate manner.

The classification performance of the GMM-based classifier with both cepstrum and MFCC coefficients is shown in Fig. 7. Similar results are obtained in all of the classification schemes. This is essentially because most of the signal energy lies in low frequency bands in PDR signals as in speech. As a result MFCC vectors produce better results because more emphasis is given to low-frequency bands during the MFCC computation compared with regular cepstrum which gives equal emphasis to all frequency values. Table I presents the confusion matrix of the GMM-based classifier using the MFCC vector as the feature vector. The experiments are done with a mixture model order of 10 and the feature vector length is also 10 for all classes. The GMM-based approach achieved an average classification accuracy

TABLE II  
Confusion Matrix of Covariance Approach with MFCC Feature Vectors in the Five-Class Problem

	Wheeled	Tracked	One Person	Two Persons	Clutter
Wheeled	95.3	4.7	0	0	0
Tracked	2.8	97.2	0	0	0
One Person	0	0	97.8	2.2	0
Two Persons	0	0	0	100.0	0
Clutter	0	0	0	1.1	98.9

TABLE III  
Confusion Matrix of Single Stage SVM Approach with MFCC Vectors in the Five-Class Problem

	Wheeled	Tracked	One Person	Two Persons	Clutter	None
Wheeled	89.4	7.6	0	0	0	3.0
Tracked	6.4	89.2	1.5	0	0	2.9
One Person	0	0	91.4	8.6	0	0
Two Persons	0	0.9	3.8	94.1	0	1.2
Clutter	1.7	0	1.8	2.50	93.1	0.9

TABLE IV  
Confusion Matrix of Two-Stage SVM Approach with MFCC Coefficients in Five-Class Problem

	Wheeled	Tracked	One Person	Two Persons	Clutter	None
Wheeled	96.4	2.6	0	0	0	1.0
Tracked	1.2	97.8	0	0	0	1.0
One Person	0	0	97.4	1.8	0	0.8
Two Persons	0	0	0	100.0	0	0
Clutter	0	0.9	0	0	99.1	0

of 96.0%. Longer MFCC vectors do not improve the classification accuracy.

Table II presents the confusion matrix of the covariance-based approach when 22 MFCCs are used. The covariance approach achieves a classification accuracy of 97.8%, which is higher than the best result obtained by the GMM-based approach.

Single-stage SVM with a radial bases function kernel has an average classification accuracy of only 91.4% as shown in Table III. SVMs are essentially binary classifiers. Therefore, for multiple classes (more than two classes), a single SVM is constructed for each class using the one-against-all strategy. The LIBSVM software [18] is used to train the SVMs, which implements the sequential minimal optimization (SMO) algorithm [18] and provides an estimated probability value for each MFCC vector of the test PDR echo signal. The average value of the estimated probabilities are computed using the entire PDR signal. The class with the highest probability is selected as the classification result.

The proposed two-stage SVM approach described in Section III makes use of temporal characteristics of signal, which is not the case for traditional

single-stage SVM classification. In this approach, the average classification accuracy is 98.1%. Table IV shows the confusion matrix with an MFCC vector of size 10. Although the computation load is higher than single-stage classification methods, it provides the highest classification accuracy. However, this recognition accuracy is very close to the covariance matrix based approach which may produce better results in other data sets obtained under different PDRs or under different conditions.

## VI. CONCLUSION

In this paper, a two-stage SVM-based target classification scheme is presented. In the first stage the most descriptive temporal segment of a given radar echo signal is determined and the target is classified using the selected temporal segment in the second stage. Mel-frequency cepstral coefficients of radar echo signals are used as feature vectors in both stages. Experimental results indicate that the proposed two-stage SVM and the covariance-based classification methods outperform the GMM-based classification method. The main advantage of the two-stage SVM method over the ordinary single-stage SVM, the covariance matrix method, and the GMM-based classification is that it reaches a decision about a test signal according to the temporal behavior of the PDR signal which is nonstationary in nature. In ordinary SVMs and GMMs temporal information is not used.

**ABDULKADIR ERYILDIRIM**  
**IBRAHIM ONARAN**  
**A. ENIS ÇETIN**  
Dept. of Electrical Engineering  
Bilkent University  
Ankara, Turkey 06800  
E-mail: (onaran@ee.bilkent.edu.tr)

## REFERENCES

- [1] Stove, A. G. and Sykes, S. R.  
A Doppler-based automatic target classifier for a battlefield surveillance radar.  
In *Proceedings of RADAR 2002*, 419–423.
- [2] Hughes, E. J. and Lewis, M.  
The application of speech recognition techniques to radar target Doppler recognition: A case study.  
In *IET Seminar on High Resolution Imaging and Target Classification*, 2006, 145–152.
- [3] Jahangir, M., Pointing, K. M., and O’Loughlen, J. W.  
A robust Doppler classification technique based on hidden Markov models.  
In *Proceedings of IEEE Radar Conference*, Oct. 2002, 162–168.
- [4] Reynolds, D. A. and Rose, R. C.  
Robust text-independent speaker identification using Gaussian mixture speaker models.  
*IEEE Transactions on Speech and Audio Processing*, 3 (1995), 72–83.
- [5] Kouemou, G., Neumann, C., and Opitz, F.  
Sound and dynamics of targets—Fusion technologies in radar target classification.  
In *Proceedings of the 11th International Conference on Information Fusion*, June 2008, 1–7.

- [6] Vapnik, V.  
*Statistical Learning Theory*.  
Hoboken, NJ: Wiley, 1998.
- [7] Bilik, I., Tabrikian, J., and Lapidus, V.  
The database of radar echoes from various targets, 2001.  
<http://www.ee.bgu.ac.il/~spl/database/radar.htm>.
- [8] Davis, S. B. and Mermerlstein, P.  
Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences.  
*IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-28* (Aug. 1980), 357–366.
- [9] Zheng, F., Zhang, G., and Song, Z.  
Comparison of different implementations of MFCC.  
*Journal of Computer Science and Technology*, **16** (2001), 582–589.
- [10] Pearson, T. C., Cetin, A. E., and Tewfik, A. H.  
Detection of insect damaged wheat kernels by impact acoustics.  
*In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, 2005, vol. 5, 649–652.
- [11] Jabloun, F., Cetin, A. E., and Erzin, E.  
Teager energy based feature parameters for speech recognition in car noise.  
*IEEE Signal Processing Letters*, 1999.
- [12] Erzin, E., Cetin, A. E., and Yardimci, Y.  
Subband analysis for robust speech recognition in the presence of car noise.  
*In Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'96)*, 1996.
- [13] Tuzel, O., Porikli, F., and Meer, P.  
Region covariance: A fast descriptor for detection and classification.  
*In Proceedings of European Conference on Computer Vision (ECCV)*, vol. 3984, May 2006, 589–600.
- [14] Bilik, I., Tabrikian, J., and Cohen, A.  
GMM-based target classification for ground surveillance Doppler radar.  
*IEEE Transactions on Aerospace and Electronic Systems*, **42** (2006), 267–278.
- [15] Bogert, B. P., Healy, M. J. R., and Tukey, J. W.  
The quefrency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphé cracking.  
*In Proceedings of the Symposium on Time Series Analysis*, 1963, 209–243.
- [16] Krebel, U.  
*Pairwise Classification and Support Vector Machines (Advances in Kernel Methods: Support Vector Learning)*.  
Cambridge, MA: MIT Press, 1990, 255–268.
- [17] Raghavan, S., Lazarou, G., and Picone, J.  
Speaker verification using support vector machines.  
*In Proceedings of the IEEE Southeast Conference on Signal Processing*, 2006, 188–191.
- [18] Chang, C. C. and Lin, C. J.  
LIBSVM : A library for support vector machines, 2001.  
<http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [19] Tuna, H., Onaran, I., and Cetin, A. E.  
Image description using a multiplier-less operator.  
*IEEE Signal Processing Letters*, **16** (Sept. 2009), 751–753.
- [20] Forstner, W. and Moonen, B.  
A metric for covariance matrices.  
Dept. of Geodesy and Geoinformatics, Stuttgart University, Stuttgart, Germany, Technical Report, 1999.

## Adaptive Clutter Measurement Density Estimation for Improved Target Tracking

**In a surveillance situation the origin of each measurement is uncertain. Each measurement may be a false (clutter) measurement, or it may be a target detection. Probabilistic methods are usually used to discriminate between the clutter and the target measurements. Clutter measurement density is an important parameter in this process. The values of the clutter measurement density in the surveillance space are rarely known a priori, and are usually estimated using sensor data and track information. A novel approach is presented and evaluated for estimating the values of clutter measurement density, which significantly enhances target tracking. Simulation results validate this approach.**

### I. INTRODUCTION

The surveillance systems usually amplify received signals and take detections whenever the amplitude of the received signal is greater than the detection threshold. The surveillance space projections of these detections are input to the target tracking filter employed, and in this paper these projections are referred to as the (target tracking) measurements. Measurements may also have additional components or features [1–3], including the amplitude or the Doppler velocity. Proposed clutter measurement density estimator does not preclude measurement features from being used in the usual manner.

If the targets are present, they are detected in every scan with a probability of detection  $P_D < 1$ . Additionally, detections may be generated by other random phenomena, including the thermal noise effects and random objects in the surveillance space. In this paper these nontarget random detections are termed clutter. At each measurement time, the target tracking filter is presented with a set of measurements, without prior information on the origin of each measurement.

Manuscript received October 21, 2009; revised April 1, 2010; released for publication August 19, 2010.

IEEE Log No. T-AES/47/2/940857.

Refereeing of this contribution was handled by T. Luginbuhl.

This work was supported by the Defense Acquisition Program Administration and Agency for Defense Development (Republic of Korea) under the International Cooperative Research Project ADD-09-70-01-03.

0018-9251/11/\$26.00 © 2011 IEEE