

RESEARCH

Open Access

# Classification of ground moving targets using bicepstrum-based features extracted from Micro-Doppler radar signatures

Pavlo O Molchanov<sup>1\*</sup>, Jaakko T Astola<sup>1</sup>, Karen O Egiazarian<sup>1</sup> and Alexander V Totsky<sup>2</sup>

## Abstract

In this article, a novel bicepstrum-based approach is suggested for ground moving radar target classification. Distinctive classification features were extracted from short-time backscattering bispectrum estimates of the micro-Doppler signature. Real radar data were obtained using surveillance Doppler microwave radar operating at 34 GHz. Classifier performance was studied in detail using the Gaussian Mixture Mode and Maximum Likelihood decision making rule, and the results were verified on a multilayer perceptron and Support Vector Machine. Experimental real radar measurements demonstrated that it is quite feasible to discern three classes of humans (single, two and three persons) walking in a vegetation cluttered environment using proposed bicepstrum-based classification features. Sophisticated bispectrum-based signal processing provides the extraction of new classification features in the form of phase relationships in the radar data. It provides a novel insight into moving radar target classification compared to the commonly used energy-based strategy.

## 1 Introduction

In recent years, radar analysis of human motion using measurements of evolutionary Doppler frequency variations has been under intensive study [1-9]. Recognition, identification and classification of persons moving in a vegetation cluttered environment using ground surveillance Doppler radar systems have a number of applications including security, military intelligence and battlefield purposes. One of the particular and effective discriminative features for the classification of moving persons is the micro-Doppler (m-D) contributions contained in the backscattering radar signature [4,5].

The m-D signature of a target is a time-varying frequency modulation contribution arising in radar backscattering and caused by the movement of separate parts of the target. Joint time-frequency (TF) analysis is the basis of most of the existing methods used to extract m-D features [1-3]. The time-varying trajectories of the different instantaneous m-D frequencies mapped into the TF domain are robust discriminative features belonging

to a person or a group. However, it should be stressed that the problem of recognizing single, two, three or more moving persons using their m-D radar signatures is one of the most difficult problems to solve.

Recently, approaches exploiting the m-D radar signatures for moving person classification have been reported in the literature [2,4,5,8]. Most approaches deal with quadratic (spectrogram-based) TF analysis of non-stationary and multi-component backscattered radar signals. According to these approaches, discriminative features are extracted from the energy-based TF distributions, i.e., estimates of backscattered signal energy distribution per unit time per unit frequency. Unfortunately, phase and frequency relationships between certain Doppler spectral components in radar returns that contain important information about phase- and frequency-coupled m-D contributions are irretrievably lost in the energy-based TF radar signatures. Therefore, a common drawback of the energy-based TF analysis is the impossibility of retrieving additional particular information concerning frequency- and phase-coupling of instantaneous frequencies contained in the radar backscattering. Phase coupling contained in radar backscattering carries important information about individual target properties.

\*Correspondence: pavlo.molchanov@tut.fi

<sup>1</sup>Department of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 10, FI-33720, Tampere, Finland

Full list of author information is available at the end of the article

Extraction of the latter phase relationships from radar backscattering could provide additional insight into radar moving target classification and improve the classification probability rate compared with commonly used energy-based classification features such as power spectrum and cepstrum features.

It has been shown in our previous articles [6,7,10,11] that sophisticated bispectrum-based signal processing permits the extraction and use for radar target recognition of the discrimination and classification of phase coupled harmonics in raw radar backscattering contaminated by additive white Gaussian noise (AWGN) and vegetation clutter. The distinctive benefits of the bispectrum-based radar signature compared to a common energy-based spectrogram can be characterized as follows. First, the possibility of retrieving the phase coupled contribution which common energy-based techniques are simply unable to provide. Second, the bispectrum tends to zero for a stationary zero-mean AWGN, which means that there are no phase coupled frequencies in a linear Gaussian process. Therefore, bispectrum-based signal processing provides suppression of the AWGN contribution in the TF radar signature. As a result, the powerful, i.e., the phase coupled contributions unambiguously belonging to a moving target are emphasized, and the weak, i.e., the phase independent spectral contributions belonging to vegetation clutter and AWGN, are diminished. Therefore, the bispectrum-based approach allows improved signal-to-noise ratio (SNR) in collected radar signatures and, hence, provides robustness in its discriminative features.

In this article, novel discriminative features computed in the form of bicepstral coefficients extracted by using bispectral estimation from radar echo-signals are suggested and studied. The performance of the suggested bicepstrum-based classifier is examined using experimental radar data processing for solving one of the most important and difficult problems in radar automatic target recognition (ATR) systems, which deals with discrimination and classification of a single walking person and group of walking persons in a vegetation clutter and AWGN environment.

The objective of this article is a comparative study of ATR system performances evaluated by using the common spectrogram-based and suggested novel bicepstrum-based approaches. Classification features based on the bispectrum estimate have been proposed earlier [11,12]. In this article, we extend the comparative analysis of the proposed features to estimate their advantages and disadvantages. In addition, a multilayer perceptron and support vector machine are used as additional classifiers to compare the results of the different classification schemes.

The article is organized as follows. First, in Section 2 the theoretical background of the solution is considered

and new feature extraction techniques are proposed. Next, in Section 3 the description of experimental data for classification is given. Then, the proposed feature extraction techniques are evaluated and compared in Section 4. Finally, conclusions are provided.

## 2 Theoretical background

The idea for the suggested approach deals with the well-known properties of the bispectral estimation method described in detail in [13]. Bispectral signal processing allows the assessment of the magnitude and phase of the correlation relationships between different harmonics. When a phase relationship exists, the phase-coupled harmonics contribute considerably to the bispectrum estimate in the form of corresponding peaks arising in the bifrequency plane. On the other hand, the bispectrum is identically zero for a stationary zero-mean AWGN. Therefore, unlike the energy spectrum, the bimagnitude, i.e., magnitude bispectrum estimate, contains the peaks in the bifrequency domain caused only by coherent contributions in the signal under study.

We have demonstrated in our previous study [6,7] that the swinging legs and arms of a walking person are not independent mechanical sources provoking time-varying instantaneous frequencies (IF) in the m-D spectrum content, but are related to each other via the “common basis” or “common carrier” which is the translating and swaying human torso.

The evident presence of phase-coupled harmonics retrieved from real radar measurements performed by a ground surveillance microwave radar has been shown in [6]. It has also been demonstrated [6] that multi-component and chirp-like returns collected by surveillance radar contain the contributions of a number of correlated scattering centers spatially distributed on the surface of the moving human body. Extraction of these bicoherent dependences and studying their evolutionary behavior enables the acquisition of a new class of information features for solving the tasks of radar target recognition, identification and classification.

The Doppler frequency shift  $\Delta f_D$  observed in a radar signal backscattered by a moving person is equal to  $\Delta f_D = 2v/\lambda$ , where  $v$  is the velocity of target and  $\lambda$  is the radar emitted wavelength. For a person moving with normal speed of motion equal to 3–5 km/h, the Doppler frequency shift caused by translating the locomotion of the human torso, is equal to  $\Delta f_D = 190 - 316$  Hz. The latter frequencies are within the audio signal band. Though swinging human arms provoke larger Doppler frequency shift values, they are also within the audio frequency band. Therefore, the received Doppler signal caused by backscattering from a moving person can be related to an audio signal. From this point of view, it is reasonable

to consider two bispectrum-based algorithms that find applications in the analysis and recognition of human speech [14,15].

The main concept of radar data processing implemented in this article is shown in Figure 1, where  $1, \dots, M$  are the  $M$  input segments. In this manner, an entire non-stationary received signal  $s$  can be divided onto a series of quasi stationary segments  $x_1, \dots, x_M$  using a sliding window function  $\psi$ . The segment  $x_m$  is assessed as  $x_m(k) = \psi(k) \cdot s((m-1)L+k)$ , where  $L$  is the length of the window function expressed in the number of temporal samples. The features are then extracted and conditional posterior probabilities are computed for those segments.

Each segment is assumed to be independent from the others. Therefore, the conditional probabilities for each class of the entire sequence are equal to the product of the conditional posterior probabilities of each segment. The decision is made using the maximum likelihood rule.

The integrated bispectrum (IB) proposed in [14] is defined as

$$IB(f) = \frac{1}{L-f} \sum_{u=1}^{L-f} B(f, u), \quad (1)$$

where  $f = 1, \dots, K-1$  is the frequency index;  $B(f, u) = X(f)X(u)X^*(f+u-1)$  is the signal bispectrum;  $X$  is the

Fourier transform of the signal  $x$ ;  $L$  is the width of the window function;  $K$  is the maximum frequency index; and  $*$  denotes the complex conjugation.

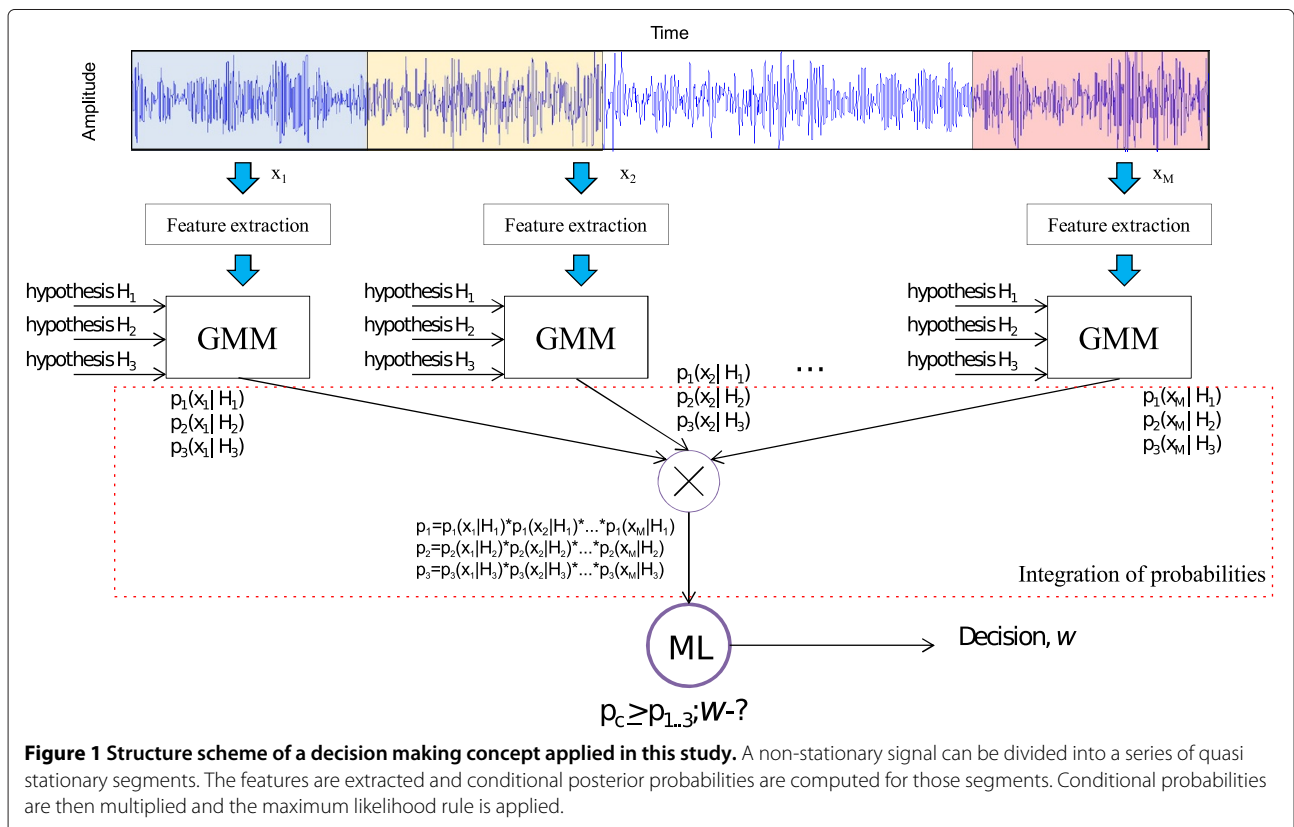
Another bispectrum-based algorithm designed for speech recognition purposes is described in [15]. This algorithm, referred to as the DFB, deals with the averaging of the bimagnitude samples in the bifrequency domain along the fixed frequency direction  $f_3$  such as  $f_1 + f_2 = f_3$ :

$$DFB(f) = \frac{1}{K} \sum_{f_2=1}^K (B(f-f_2, f_2))^{\frac{1}{3}}. \quad (2)$$

A bicepstrum is the result of taking the inverse Fourier transform of the logarithm of the bispectrum. In this article, the following bicepstral coefficients denoted below by  $CIB(f)$  and  $CDFB(f)$  are computed using the bispectral data  $IB$  (1) and  $DFB$  (2). These bicepstral values are exploited as the discriminative classification features as:

$$CIB(f) = \left| \frac{1}{K} \sum_{j=1}^K \log(|IB(j)|) e^{i2\pi f j / K} \right|, \quad (3)$$

$$CDFB(f) = \left| \frac{1}{K} \sum_{j=1}^K \log(|DFB(j)|) e^{i2\pi f j / K} \right|. \quad (4)$$



**Figure 1 Structure scheme of a decision making concept applied in this study.** A non-stationary signal can be divided into a series of quasi stationary segments. The features are extracted and conditional posterior probabilities are computed for those segments. Conditional probabilities are then multiplied and the maximum likelihood rule is applied.

To evaluate and compare correctly the performance of the suggested bicepstrum-based classifier with the common power cepstrum-based classifier, the following power cepstrum coefficients  $C(f)$  are considered and computed in our survey as:

$$C(f) = \left| \frac{1}{K} \sum_{j=1}^K \log(|X(j)|^2) e^{i2\pi f j / K} \right|^2. \quad (5)$$

A human operator can distinguish different targets listening to the baseband version of the received continuous wave radar signal [4,16]. The cepstrum coefficients (5) are commonly used as features for speech and audio recognition [17,18]. Taking into account the similarity in the signals used for speech recognition and ground moving target classification [16], the selection of cepstrum coefficients for the comparison seems reasonable. Moreover, the cepstrum coefficients as a feature for ground target classification are of great interest to other researchers [4], and comparison with their research is sensible.

From the various existing approaches to radar target recognition and classification, the maximum likelihood (ML) rule and the Gaussian mixture model (GMM) are selected to evaluate the bispectrum-based classifier performance. In our opinion, the GMM is a good strategy for the unknown probability density function (pdf) approximation [19].

In the general case, the GMM [19] approximates the probability density function of a feature vector  $y$  under hypothesis  $H$  as:

$$p(x|H) = \sum_{n=1}^N r_n \phi(y|\theta_H^{(n)}), \quad (6)$$

where  $N$  is the number of mixture components called the GMM order;  $r_n$  is the mixture weight of the  $n$ th component, such that  $\sum r_n = 1$ ;  $\theta_H^{(n)}$  is the parameter referred to the distribution of component  $n$  under the class hypothesis  $H$ ; and  $\phi(x|\theta_H^{(n)})$  is the probability distribution of  $x$  parameterized by  $\theta_H^{(n)}$ .

The probability distribution of each component given in (6) can be written as

$$\phi(x|\theta_H^{(n)}) = \frac{(2\pi)^{-\frac{D}{2}}}{|\Gamma_H^{(n)}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu_H^{(n)})^T \cdot (\Gamma_H^{(n)})^{-1} (x - \mu_H^{(n)})\right) \quad (7)$$

where  $D$  is the dimension of the feature vector;  $\Gamma_H^{(n)}$  is the covariance matrix of component  $n$  for the hypothesis  $H$ ; and  $\mu_H^{(n)}$  is the vector of the mean values of component  $n$  for the hypothesis  $H$ . In this article, a full covariance matrix is used [20].

The posterior class conditional probability for the entire received signal  $s$  is the product of the posterior class conditional probabilities  $p(x_m|w)$ :

$$p(s|w) = \prod_{m=1}^M p(x_m|w), \quad (8)$$

where  $w$  is the class hypothesis.

The decision-making rule exploited in the ATR system using the maximum likelihood (ML) method can be defined as follows:

$$\hat{w} = \arg \max_{w=1 \dots W} p(s|w), \quad (9)$$

where  $p(s|w)$  is a likelihood function conforming to the signal  $s$  referred to the classification hypothesis  $w$ .

### 3 Discussion of experimental results

Real radar data were collected by experimental measurement performed with a ground surveillance Doppler homodyne, monostatic, polarimetric and continuous wave radar.

The radar backscattering data relevant to three pedestrian classes were accumulated and recorded. The following scenarios were considered both in vegetation clutter and open space environments. (1) Single moving person: Person walking towards or away from the radar at a velocity of 3–5 km/h. (2) Group of moving persons: Two persons walking towards or away from the radar at a velocity of 3–5 km/h. (3) Group of moving persons: Three persons walking towards or away from the radar at a normal velocity of 3–5 km/h and either synchronously or asynchronously.

The ground surveillance radar system is shown in Figure 2. The parameters of the radar are: wavelength—8.8 mm; emitted radar microwave power—15 mW; receiving/transmitting antenna beam width in both E and H planes—60°; level of side lobes in the horn antenna pattern—24 dB; cross-polarization level lower than—30 dB; receiver noise figure—20.2 dB; two-channel 16 bit ADC; and a sampling rate in the digital records—8 KHz. The averaged signal-to-noise ratio (SNR) values are equal to 4 dB, 6 dB, and 11 dB for single, two and three moving persons, respectively.

The total length of all recorded wave-files is more than 23 minutes. The measurements were performed during the autumn period. Despite the radar being able to operate in both vertical and horizontal polarization modes, only the horizontal mode was considered for the classification.

Collection of the data for the dataset was performed as follows. The initial position of a target was fixed at a few meters from the radar. The person started to walk away from the radar for approximately 40 s, stopped for 2 s, turned around and came back, stopped for about 2 s and repeated the motion several times. Each considered



**Figure 2** Ground surveillance radar system used for experimental data collection. Wavelength—8.8 mm, emitted radar power—15 mW; receiving/transmitting antenna beam width in both E and H planes—60°; level of side lobes in horn antenna pattern—24 dB; cross-polarization level  $\leq -30$  dB; receiver noise figure—20.2 dB; two-channel 16 bit ADC; and sampling rate in digital records—8 KHz.

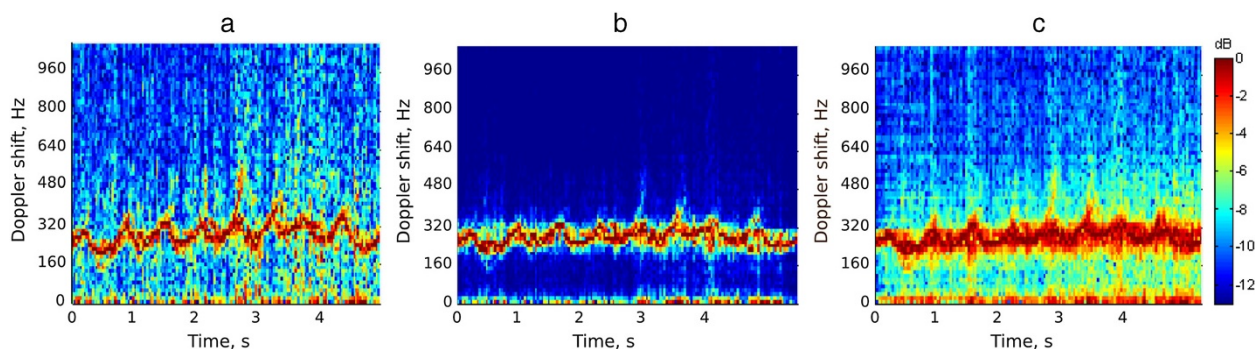
class contained six sets of experiments performed with a person walking away from the radar and five sets of experiments with a person walking towards the radar. The same persons have participated in all experiments.

Examples of time-frequency radar signatures of a single person moving in vegetation clutter are shown in Figure 3. Three types of TF distributions are shown: a spectrogram computed in the form of the amplitude of the Short Time Fourier Transform (STFT), and bispectrum-based radar signatures computed by IB (1) and DFB (2). The time-frequency distributions are computed with a Hamming window of length  $L = 64$  ms, without overlap.

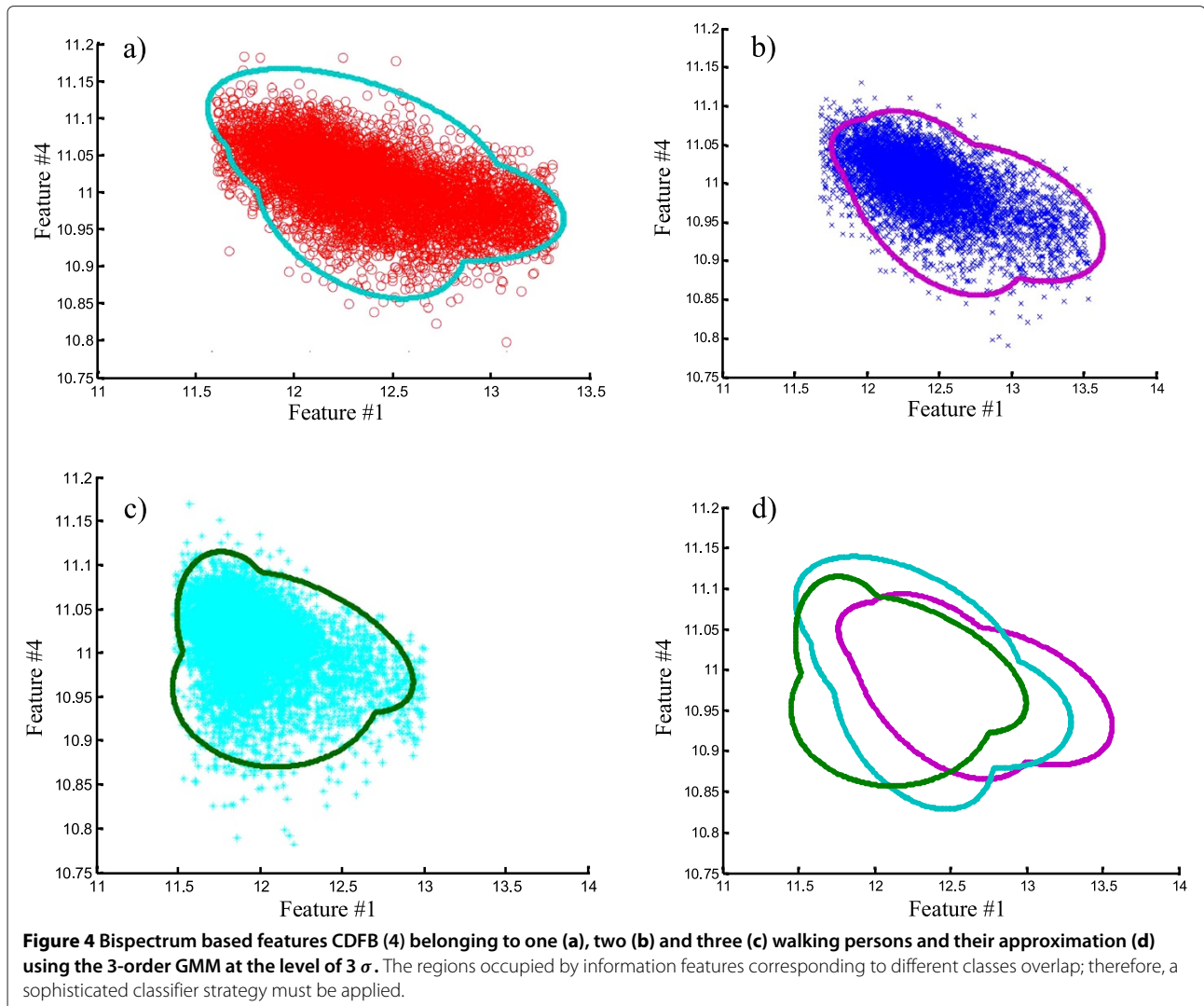
It can be seen from Figure 3 that AWGN is suppressed better in the bicepstrum-based radar signatures plotted in both Figure 3b,c compared to the spectrogram represented in Figure 3a.

It can be seen from the Figure 3 that the analyzed signal does not contain frequencies higher than 700 Hz, therefore, a sampling frequency for the ADC equal to 8 KHz is a reasonable choice.

Dependencies between the values conforming to the first and fourth bicepstrum coefficients given in CDFB (4) and their GMM approximation represented by a 3-order model at a level of  $3\sigma$  are illustrated in Figure 4. As can



**Figure 3** Time-frequency radar signatures measured in vegetation clutter and represented by: (a) spectrogram; (b) bicepstrum-based features (1); (c) bicepstrum-based features (2). The spectrogram is computed in the form of the amplitude of the short time Fourier transform (STFT), and the bicepstrum-based features are computed using IB (1) and DFB (2).



be seen from Figure 4, the regions occupied by information features corresponding to different classes overlap. Therefore, a feature space with a higher dimensionality is necessary to discriminate the classes.

Histograms illustrating the distribution laws for the second cepstral and bicepstral coefficients computed by using (5) and (4), respectively, are represented in Figure 5. The histograms show that it is difficult to discern a single walking person from two or three persons using just this one feature. Both histograms in Figure 5a,b referred to a single walking person but contain the overlapping domains corresponding to the histograms obtained for both two and three persons. However, classes belonging to two and three walking persons are more separated in the histograms plotted using bicepstral coefficients (4).

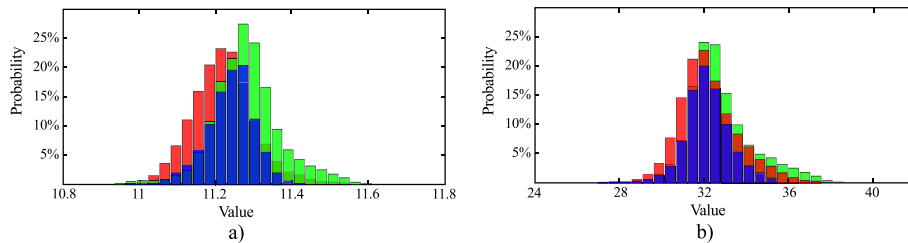
Classifier performance can be achieved using the features with lower inter-class similarity, i.e., when the same classifier but different feature vectors are used. To estimate inter-class similarity, the Euclidean metric has been

computed for the sampled cross-correlation function. The similarity measure (SM) is evaluated as follows:

$$SM(j) = \frac{1}{3K} \sum_{i=1}^j \sum_{\substack{k=1,2,3 \\ l=2,3,1}} ||XCF\{Y_{k,i}, Y_{l,i}\}||, \quad (10)$$

where  $j$  is the number of used cepstral or bicepstral coefficients; XCF is the cross-correlation function;  $k, l$  are the indexes belonging to three classes;  $i$  is the cepstral/bicepstral coefficient number; and  $Y_{k,i}$  is the set of cepstral/bicepstral coefficients number  $i$  belonging to the class  $k$ .

Dependencies of SM on the number of first cepstral/bicepstral coefficients are illustrated in Figure 6. One can see the benefit of using the bispectrum-based strategy (see the straight curve in Figure 6) compared to the power spectrum-based technique (dashed curve in Figure 6). This benefit can be assessed by comparing the values belonging to the straight and dashed curves in Figure 6.



**Figure 5** Histograms of the second bicepstrum/cepstrum coefficients related to the single (blue), two (red) and three (green) walking persons and computed using: (a) (4) and (b) (5). The histograms show that it is difficult to discern a single walking person from two or three persons by using only this one feature.

It is clearly seen that the correlation values are smaller for the bispectrum-based feature extraction technique, which means the latter technique possesses better orthogonality of its features. Therefore, a better classifier performance should be achieved for the bispectrum-based technique.

#### 4 Analysis of classifier performance

##### 4.1 Data separation

Commonly [3,4], to evaluate classifier performance, the classification dataset is divided into two subsets of the same size. One subset is used as a training dataset and the other as a testing dataset. The disadvantage of such an approach is that the classification probability rates might vary if the small original dataset is split in a different manner. To obtain more accurate and reliable classification results, the  $K = 11$  cross-validation technique is applied. The initial data under analysis are split into  $K$  subsets of

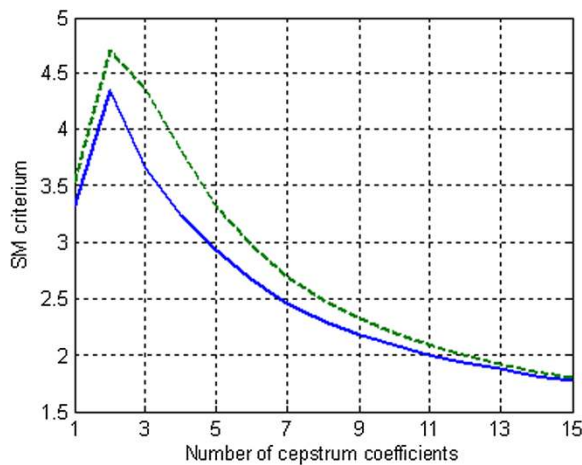
the same length, and  $K - 1$  subsets are used as a training dataset, and the remaining one as a testing dataset. The cross-validation process is repeated  $K - 1$  times ( $K - 1$  folds) with each of the  $K$  subsets used as a testing dataset. The  $K$  results from the folds are averaged to evaluate a single estimation. The most important benefit of the  $K$ -fold cross-validation strategy is that all measured data are distributed somewhat uniformly within both training and testing operations.

Eleven diverse experiments were performed for each of three radar target classes and the 11-fold cross-validation technique was exploited for target classification. This implies that the features have been extracted from the measured radar data ten times for the training dataset and once for the testing dataset during each experiment, i.e., 0.91 part of the data collected is used as a training dataset, and the remaining 0.09 part as the testing (validation) dataset for each fold.

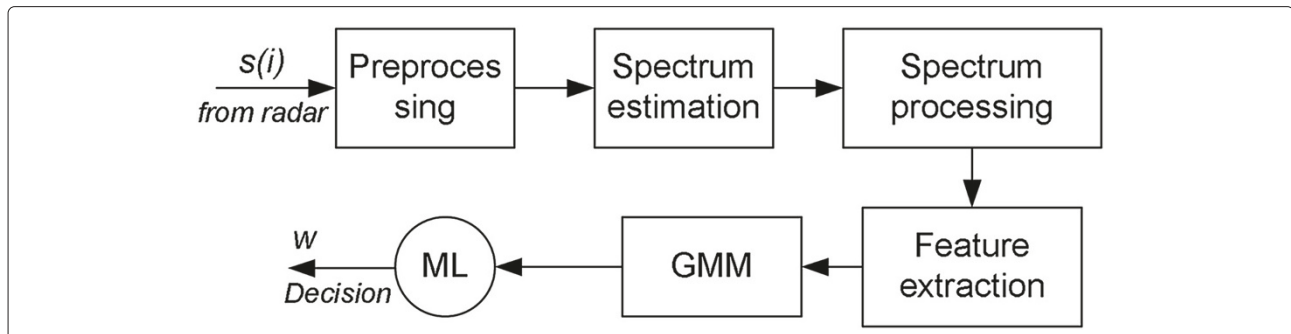
##### 4.2 Classification scheme

A scheme for the proposed classifier is shown in Figure 7. The preprocessing block partitions the input signal into a series of frames of  $L$  samples length. The spectrum estimation block computes the spectrum of each frame using a Hamming window. The spectrum contains the frequencies higher than can be provoked by human gait, i.e., those frequencies that are higher than 900 Hz. Therefore, in the next block denoted as “Spectrum processing” they are removed by an ideal low pass filter. It can be seen from Figure 3 that the maximum frequency in the signal under consideration is near 640 Hz, therefore, higher frequencies could be removed. Next, features are extracted from the spectrum at the block denoted as “Feature extraction”. It could be one of the above-mentioned techniques C (5), CIB (3), or CDFB (4). The conditional posterior probabilities are computed at the block denoted as “GMM”, and the decision is made using the ML rule (9).

Some parameters such as the length of segment  $L$ , the number of GMM components and the number of used classification features must be defined a priori. The experimental system illustrated in Figure 8 is presented for this



**Figure 6** Inter-class similarity computed for the feature vector C (5) (dashed curve) and CIB (3) (straight curve). The correlation values are smaller for the bispectrum-based feature extraction technique, which means the latter technique possesses better orthogonality in its features.



**Figure 7 Scheme of the proposed classifier.** The main steps necessary to refer a signal  $s(i)$  observed at the output of the radar to one of the possible classes are illustrated.

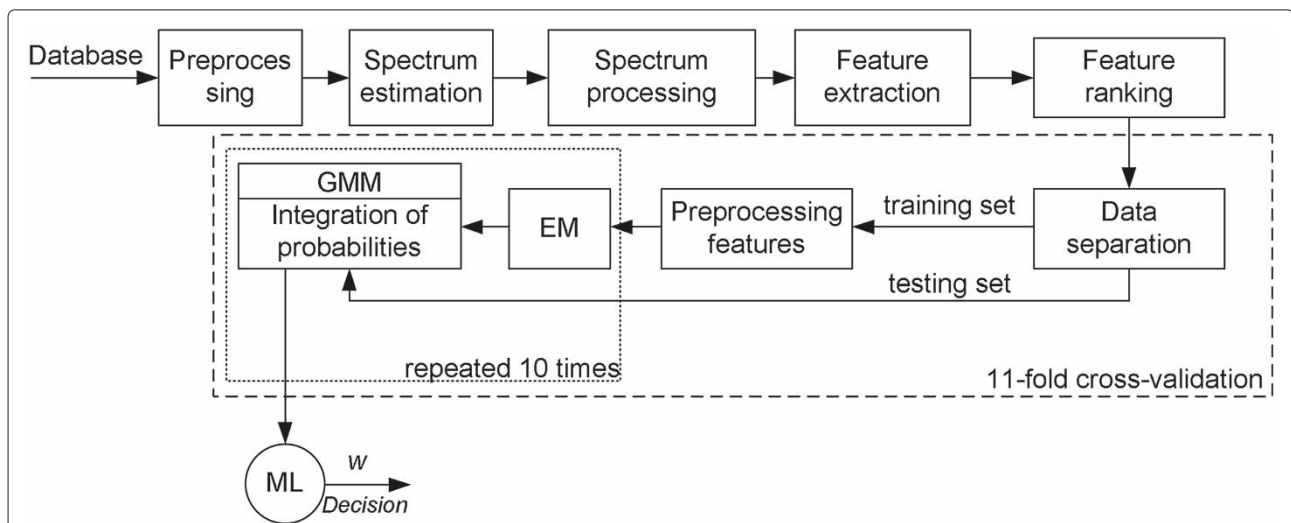
purpose. An 11-fold cross-validation is applied for the performance evaluation of the ATR system, where data mining is carried out and the optimal parameters are estimated. The concept “optimal” means the parameters with which the best classification performance is obtained.

The scheme illustrated in Figure 8 is used for performance evaluation. The block denoted “preprocessing features” removes the outliers from the training set, discarding 1% of the highest and lowest values. Next, the parameters  $\theta$  of GMM are estimated using the Expectation Maximization algorithm [21]. The initial estimate of the parameters is obtained by the k-means algorithm, and for statistical stability the results of 10 GMMs are averaged. Posterior class conditional probabilities extracted from the segments are multiplied to obtain the posterior class conditional probability of the entire received signal. The latter operation is performed in the block “Integration of probabilities”.

The optimal number of both power cepstrum (5) and bicepstral coefficients given by (3) and (4) has been

estimated. From one side, if a small number of them is selected compared with the information containing in other coefficients, worse probabilities will be obtained. From the other side, the so called “curse of dimensionality” can arise if a large number of coefficients have been selected. We compute the probabilities of classification only for the training set changing the number of coefficients from 1 to 50% of their maximum number. Then, the number of coefficients is selected according to the maximum value of classification probability. The estimated number of used coefficients depends on the segment length, and decision-making time. Therefore, it is unique for a fixed set of parameters. We consider 50% features at most because of their symmetry (see (4), (3), and (5)). The maximum number of features is  $z = L/9$ .

Empirically it was established that by varying the GMM order, the classification probability rates do not depend significantly on the GMM order. However, with increasing feature vector dimensionality, the GMM order should decrease. When only a few feature vectors are



**Figure 8 Scheme of parameter estimation.** An 11-fold cross-validation is applied for the performance evaluation of the ATR system, where a data mining is carried out and the optimal parameters are estimated.



available GMM requires more components to achieve good approximation of the probability density function. However, when the ATR system operates with many feature vectors, a few components will be adequate. In the considered case, we deal with a few feature vectors. The processing data length is equal to 2 s (18 feature vectors are used for each class within one realization of cross-validation), and a large quantity of data are used when the data length is equal to 16 ms (2468 feature vectors for each class are used within one realization of cross-validation). The GMM order was defined empirically. When the processing data lengths were equal to 2 s, 1 s and 512 ms, the GMM order was selected to be equal to five. For the lengths of 256 ms, 128 ms and 64 ms, the GMM order was equal to four and, finally, for 32 ms and 16 ms it was equal to three.

### 4.3 Feature ranking

Feature ranking is an important operation contained in classification algorithms. It ranks features according to certain information criteria and only the most informative features are used for classification. The concept for the criteria used in our feature ranking procedure is based on Information Theory and came from the article [22]. Assume that we have a feature vector  $y$  with available values  $\{y_1, \dots, y_J\}$  and a class label vector  $z$  with the values of  $\{z_1, \dots, z_W\}$ . The following conditional entropy  $H(z|y)$  can then be computed as:

$$H(z|y) = \sum_{j=1 \dots J} p(y_j)H(z|y = y_j), \quad (11)$$

where  $p(y_j)$  is the probability of  $y$  taking the state  $y_j$ ; and  $H(z) = \sum_{w=1 \dots W} p(z_w) \log_2 p(z_w)$  is the entropy of  $z$ .

The conditional entropy (11) indicates how much entropy is left if the state of the feature  $y$  is known. The information gain (IG) indicating the amount of additional

information about the class provided by the feature  $y$  is given as:

$$IG(z|y) = H(z) - H(z|y). \quad (12)$$

The features are sorted according to their IG value. We will vary the number of used features and those having a higher IG will be used first in the classification.

### 4.4 Performance evaluation

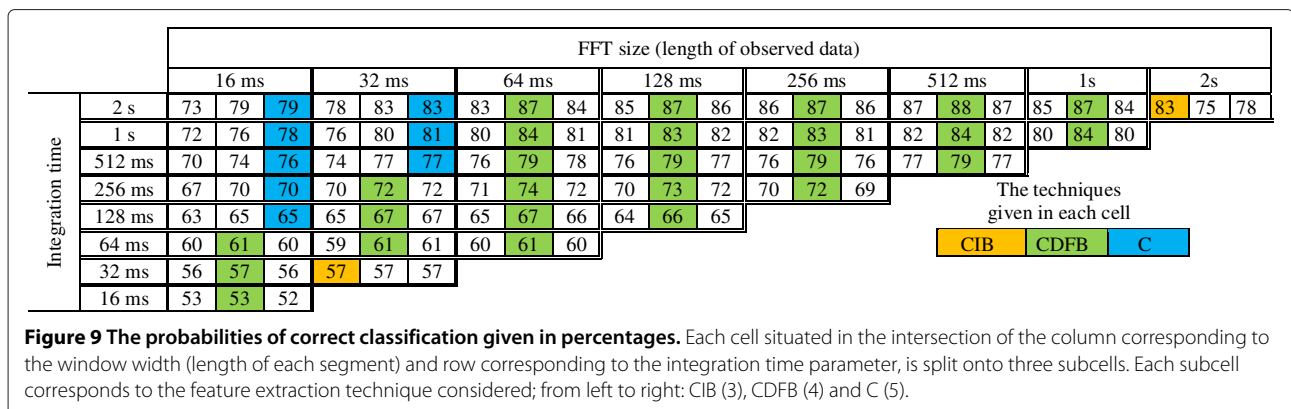
The proposed classification scheme uses the integration of probabilities by dividing a non-stationary signal of length  $N$  into  $M$  segments of length  $L$ . As a result, a sequence of  $M$  quasi stationary signals is obtained.

The probability of correct classification is computed as:

$$P = \sum_{w=1}^W \frac{U_{\text{cor}}(w)}{U_{\text{total}}(w)} P_a(w), \quad (13)$$

where  $W$  is the total number of available classes;  $U_{\text{cor}}(w)$  is the number of correctly classified instances related to the class  $w$ ;  $U_{\text{total}}(w)$  is the total number of classified instances related to the class  $w$ ; and  $P_a(w)$  is a priori probability related to the class  $w$ . Unfortunately, a priori probability is impossible to estimate using the available experimental data. Because of this, we assume that a priori probability related to each separate class is of the same value and equal to  $P_a(w) = \frac{1}{W} \forall w$ .

The probabilities of correct classification depending on the length of input signals and integration time are given in percentage terms in Figure 9. Each cell in Figure 9 situated at the intersection of the column corresponding to the window width (length of each segment) and row corresponding to the integration time parameter, is split onto three subcells. Each subcell corresponds to the feature extraction technique considered. The left subcell shows the probability obtained using the proposed bicepstrum-based technique CIB (3). The



middle subcell corresponds to the suggested bicepstrum-based technique CDFB (4). The right subcell contains data obtained using the common cepstrum-based technique C (5). Subcells containing the maximum value of probability obtained using different techniques with the same parameters are highlighted according to the technique used.

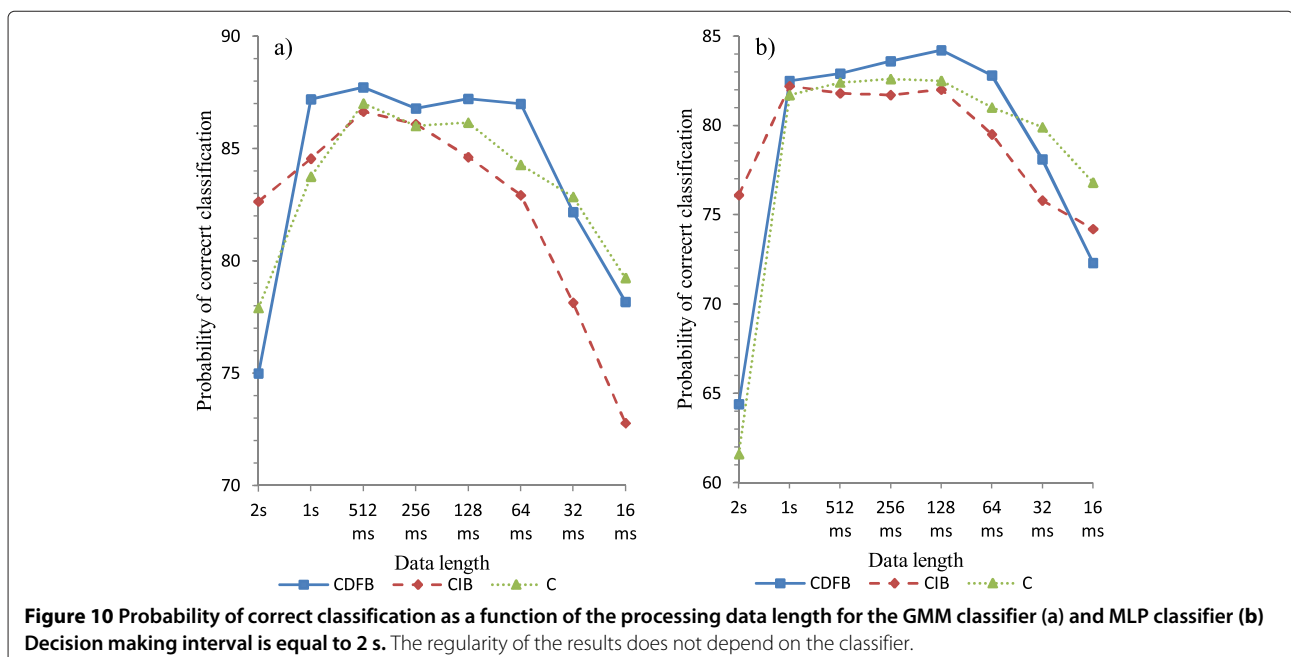
The comparative analysis data represented in Figure 9 demonstrate the benefits of the bicepstrum-based techniques compared with the common cepstrum-based technique. The bicepstrum-based techniques provide better results with data lengths equal to or more than 64 ms. The conventional cepstrum-based technique outperforms the suggested techniques only when the data lengths are less than 64 ms and the integration time is more than 128 ms. However, the difference between the techniques under comparison is not very significant. The worst performance of the bicepstrum-based classifier is caused by low frequency resolution depending on the window width exploited in the STFT. The considered non-parametric estimation provides a frequency resolution equal to 63 Hz and 125 Hz for data lengths of 32 ms and 16 ms, respectively. To improve the frequency resolution in ATR systems, the parametric bispectrum-based techniques [10] can be used.

It is well-known that the performances of ATR systems depend on the classifier, and may vary with different types of classifier. All the above-mentioned results were obtained using the statistical-based classifier, GMM with the ML decision rule. Recently, the popularity of Neuron networks (NN) and multi layer perception (MLP) has

increased in ATR systems. Therefore, it is reasonable to compare the obtained results with those obtained using MLP. The MLP is selected to be a feed-forward back-propagation Artificial Neural Network (ANN) with two hidden layers (ten neurons are contained in each hidden layer). Their transfer function is selected to be the tan-sigmoid. The output layer contains three output nodes with a purely linear transfer function. The mean squared error performance function is selected to estimate the ANN performance. The MLP is trained to estimate the class conditional posterior probability of the feature vector, and this is archived using three output nodes. The ANN is generated using standard Matlab functions.

Figure 10 shows the dependences of the correct classification probabilities on processing the data length for a decision making interval equal to 2 s for the two classifiers. The following peculiarities should be emphasized in the comparison of the results presented in Figure 10:

- For both considered bicepstral classifiers, the CDFB technique (4) provides the best results with data lengths larger than 64 ms and less than 1 s;
- The common cepstrum-based technique C (5) outperforms the suggested techniques only when the data length is less than 64 ms;
- The best probability of correct classification for a data length of 2 s is obtained using the bicepstral CIB technique (3);
- Regularity of the results does not depend on the classifier.

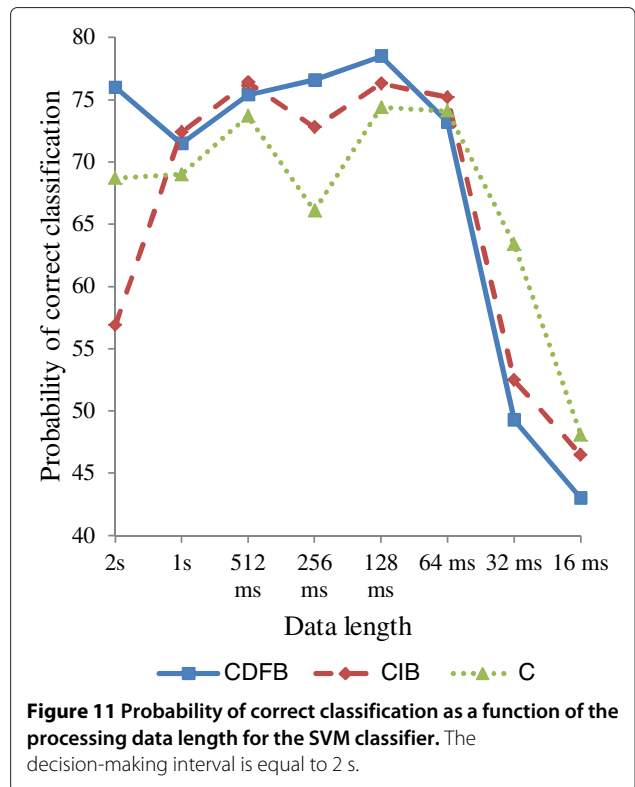


Confusion matrices computed with a data length of 512 ms and decision making time of 2 s for all considered techniques are listed in Table 1:

- The best classification performance for a single walking person is achieved using the CDFB features with 90% of the correct classifications outperforming the other features by 2%;
- The class of two walking persons is the most complicated for all considered feature extraction techniques. The highest probability of correct classification, equal to 83%, is provided using the CIB features. The probabilities of correct classification equal to 82% and 81% are provided using the CDFB and C features, respectively;
- The last considered class of three walking persons is classified with a probability of correct classification at a level of 89% using the CDFB and C features. The CIB features provide 88% of the correct classifications.

Next, classification is performed using the Support Vector Machine (SVM) with a linear kernel. It is a non-probabilistic classifier, therefore, “integration of the probabilities” (8) could not be performed. This step is replaced by a majority voting method. The probabilities of correct classification are computed using 2-fold cross-validation.

The classification results computed by SVM are shown in Figure 11. The results are similar to those obtained



**Figure 11** Probability of correct classification as a function of the processing data length for the SVM classifier. The decision-making interval is equal to 2 s.

**Table 1** Confusion matrices for the considered techniques with a data length of 512 ms and decision making time of 2 s

CDFB features			
	One person	Two persons	Three persons
One person	90	4	6
Two persons	6	82	12
Three persons	6	5	89

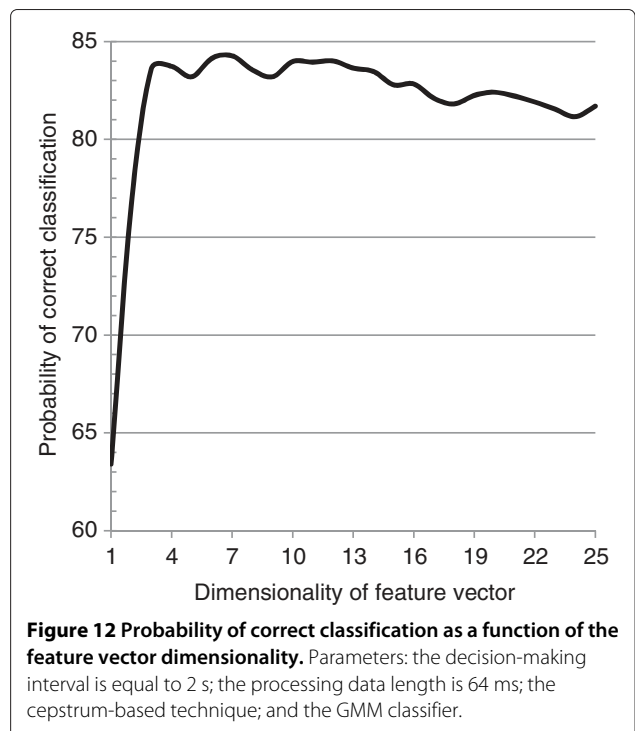
  

CIB features			
	One person	Two persons	Three persons
One person	88	8	4
Two persons	5	83	12
Three persons	4	8	88

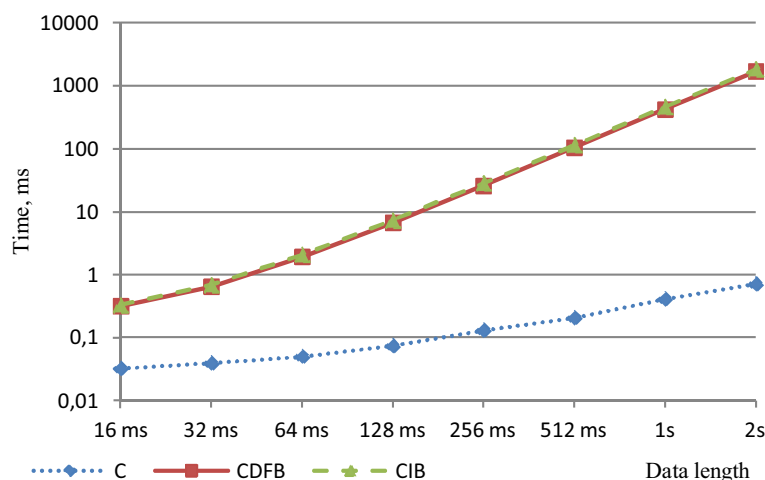
  

C features			
	One person	Two persons	Three persons
One person	88	7	5
Two persons	6	81	13
Three persons	4	7	89

The confusion matrix shows the probability of declaring a class  $w_j$  if the feature set from class  $w_i$  was at the input of the classifier. The three classes examined are listed in the left column and the declared classes are listed in the top row. The values within the main diagonal illustrate the probability of the correct classification when the input class  $w_i$  is correctly declared as  $w_i$ .



**Figure 12** Probability of correct classification as a function of the feature vector dimensionality. Parameters: the decision-making interval is equal to 2 s; the processing data length is 64 ms; the cepstrum-based technique; and the GMM classifier.



**Figure 13** Elapsed time for feature extraction depending on the processing data length and method used. The processing time required for bispectrum-based techniques is significantly larger than for the cepstrum-based technique. Fortunately, real-time implementation is possible, because the processing time is smaller than the data length.

earlier. The cepstrum-based technique (C) outperforms other techniques when the data length is less than 64 ms. The bispectrum-based techniques outperform the cepstrum-based technique when the data length is higher than 64 ms, by 2–4%.

The probability of correct classification depending on the feature vector dimensionality is shown in Figure 12. The curve is calculated for the cepstrum-based features and GMM classifier, with a processing data length of 64 ms and decision making time of 2 s. Before calculating the probabilities, the features are sorted according to the IG criterion and therefore, more of the informative features are used first. It can be seen in Figure 12 that the function rapidly rises when there are fewer than four features. The function has a peak at number of features equal to seven, and corresponded features are selected to provide the classification result. The function then decreases with the increase in feature vector dimensionality.

It should be noted that bispectrum-based techniques require larger digital signal processing times because of additional computation for the 3-D-valued bispectral density. Therefore, it is of practical interest to estimate the computational time and compare it for all the techniques considered. Computations were performed by a computer with the following parameters: Intel Core 2 DUO CPU 3 GHz, 3.2 Gb RAM, operation system Windows XP SP 2, and Matlab R2010a.

Figure 13 illustrates the time required for feature extraction using the three techniques considered. For all techniques and all available values of data length, the processing time is smaller than the data length. Therefore, a real-time implementation of all algorithms is possible. The processing time required for the bispectrum-based techniques is significantly larger than for cepstrum-based

techniques. This is the cost of better classification performance. Fortunately, the signal processing time required for bispectrum-based techniques can be optimized using the symmetry properties of bispectra [13].

## 5 Conclusions

This article proposed bispectrum-based feature extraction from micro-Doppler radar signatures to classify moving radar targets. Data were collected using ground surveillance Doppler radar for one, two and three moving persons. Pattern features were extracted from integrated and averaged short-time bispectrum estimates of transient Doppler radar signals in the form of two types of bicepstral coefficients. Diverse scenarios were considered and the 11-fold cross-validation test was employed to improve the classification accuracy. Experimental results demonstrate that it is quite feasible to recognize three classes of persons moving in a vegetation cluttered environment using the proposed bispectrum-based features extracted from micro-Doppler radar backscattering. Bispectrum-based pattern features extraction from radar backscattering provides additional insight into moving target radar classification that is superior to the commonly used energy-based information features. The experimental results obtained are useful from the point of view of practical recommendations for security and military ATR systems and open new possibilities for ground moving target recognition and classification.

### Competing interests

The authors declare that they have no competing interests.

### Acknowledgements

The study was supported by Graduate School in Electronics, Telecommunications and Automation, GETA, Finland, and Tampere International Center for Signal Processing, TICSP, Finland.

#### Author details

<sup>1</sup>Department of Signal Processing, Tampere University of Technology, Korkeakoulunkatu 10, FI-33720, Tampere, Finland. <sup>2</sup>Department of Transmitters, Receivers and Signal Processing, National Aerospace University, Chkalova Str. 17, 61070, Kharkiv, Ukraine.

Received: 25 July 2012 Accepted: 9 February 2013

Published: 27 March 2013

#### References

1. P van Dorp, F Groen, Feature-based human motion parameter estimation with radar. *IET Radar Sonar Navig.* **2**(2), 135–145 (2008)
2. V Chen, Doppler signatures of radar backscattering from objects with micro-motions. *IET Signal Process.* **2**(3), 291–300 (2008)
3. J Geisheimer, W Marshall, E Grenaker, in *Conference on Signals, Systems and Computers*, vol. 1. A continuous-wave (CW) radar for gait analysis (Pacific Grove, USA, 2001), pp. 834–838
4. I Bilik, J Tabrikian, A Cohen, GMM-based target classification for ground surveillance Doppler radar. *IEEE Trans. Aerosp. Electron. Syst.* **42**, 267–278 (2006)
5. G Smith, K Woodbridge, C Baker, Radar micro-Doppler signature classification using dynamic time warping. *IEEE Trans. Aerosp. Electron. Syst.* **46**(3), 1078–1096 (2010)
6. J Astola, K Egiazarian, G Khlopov, S Khomenko, I Kurbatov, V Morozov, A Totsky, Application of bispectrum estimation for time-frequency analysis of ground surveillance Doppler radar echo signals. *IEEE Trans. Instrument. Meas.* **57**(9), 1949–1957 (2008)
7. J Astola, K Egiazarian, G Khlopov, S Khomenko, I Kurbatov, V Morozov, A Teplyuk, A Totsky, in *Proceedings of International Radar Symposium*, vol. 1. Time-frequency analysis of ground surveillance Doppler radar echo signals by using short-time cross-bispectrum estimates (Cologne Germany, 2007), pp. 805–808
8. G Smith, K Woodbridge, C Baker, H Griffiths, Multistatic micro-Doppler radar signatures of personnel targets. *IET Signal Process.* **4**(3), 224–233 (2010)
9. A Balleri, K Chetty, K Woodbridge, Classification of personnel targets by acoustic micro-Doppler signatures. *IET Radar Sonar Navig.* **5**(9), 943–951 (2011)
10. P Molchanov, J Astola, K Egiazarian, G Khlopov, V Morozov, B Pospelov, A Totsky, in *11th International Radar Symposium*. Object recognition in ground surveillance Doppler radar by using bispectrum-based time-frequency distributions, (2010), pp. 1–4
11. P Molchanov, J Astola, K Egiazarian, A Totsky, in *2011 European Radar Conference (EuRAD)*. Moving target classification in ground surveillance radar ATR system by using novel bicepstral-based information features, (2011), pp. 194–197
12. P Molchanov, J Astola, K Egiazarian, A Totsky, in *2011 Proceedings International Radar Symposium (IRS)*. Target classification by using pattern features extracted from bispectrum-based radar Doppler signatures, (2011), pp. 791–796
13. C Nikiyas, M Raghuvier, Bispectrum estimation: a digital signal processing framework. *Proc. IEEE.* **75**(7), 869–891 (1987)
14. A Moreno, M Rutllan, in *Fourth International Conference on Spoken Language*, vol. 2. Integrated polispectrum on speech recognition (Philadelphia, USA, 1996), pp. 1281–1284
15. K Onoe, S Sato, S Homma, A Kobayashi, T Imai, T Takagi, Bi-spectral acoustic features for robust speech recognition. *IEICE Trans. Inf. Syst.* **E91-D**(3), 631–634 (2008)
16. T McConaghy, H Leung, E Bosse, V Varadan, Classification of audio radar signals using radial basis function neural networks. *IEEE Trans. Instrument. Meas.* **52**(6), 1771–1779 (2003)
17. S Davis, P Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans. Acoustics Speech Signal Process.* **28**(4), 357–366 (1980)
18. D Dimitriadis, P Maragos, A Potamianos, Robust AM-FM features for speech recognition. *IEEE Signal Process. Lett.* **12**(9), 621–624 (2005)
19. R Duda, P Hart, D Stork, *Pattern Classification*, 2 edn (Wiley, New York, 2001)

20. P Bell, Full Covariance Modelling for Speech Recognition. *PhD thesis*, The University of Edinburgh (2010)
21. A Dempster, N Laird, D Rubin, Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Stat. Soc.* **39**, 1–38 (1977)
22. J Deng, C Simmermacher, S Cranefield, A study on feature analysis for musical instrument classification. *IEEE Trans. Syst. Man Cybern. Part B: Cybern.* **38**(2), 429–438 (2008)

doi:10.1186/1687-6180-2013-61

**Cite this article as:** Molchanov et al.: Classification of ground moving targets using bicepstrum-based features extracted from Micro-Doppler radar signatures. *EURASIP Journal on Advances in Signal Processing* 2013 **2013**:61.

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)