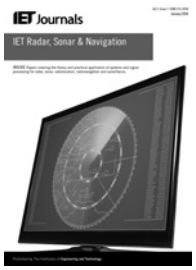
Published in IET Radar, Sonar and Navigation
Received on 16th March 2011
doi: 10.1049/iet-rsn.2011.0087



ISSN 1751-8784

Classification of personnel targets by acoustic micro-Doppler signatures

A. Balleri[1] *K. Chetty*[2] *K. Woodbridge*[1]

[1]*Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*
[2]*UCL Jill Dando Institute of Security and Crime Science, UCL Department of Security and Crime Science, 35 Tavistock Square, London, WC1H 9EZ*
E-mail: a.balleri@ee.ucl.ac.uk

Abstract: Classification of targets using their micro-Doppler signatures has attracted a growing interest in recent years. In addition to their main bulk translation, targets may exhibit additional motions, such as vibrations and rotations, which generate Doppler modulations in the echo that contain unique target features and thus can be used to perform target recognition. Although target classification by micro-Doppler signatures has been exploited in the radio frequency regime for radar systems, much less work has been done in acoustic. In this work, an ultrasound radar operating at 80 kHz has been developed to gather micro-Doppler signatures of personnel targets performing various actions. The performance of a range of classifiers and feature extraction algorithms in distinguishing between these micro-Doppler signatures is presented.

1 Introduction

The echo produced by a moving target that is illuminated by a radar or ensonified by a sonar system contains frequency modulations caused by the time-varying delay occurring between the target and the sensor. The main bulk translation of the target towards or away from the sensor induces a frequency or Doppler shift of the echo as a result of the well-known Doppler effect [1]. Additional movements of small parts of the target contribute with frequency modulations around the main Doppler shift and these are commonly called micro-Doppler modulations. Micro-Doppler modulations contain a signature of the target that can be used for target recognition.

Classification of targets by micro-Doppler signatures has been widely investigated for the radio frequency (RF) regime and in particular for applications related to radar systems. Chen *et al.* [2–4] have modelled the radar micro-Doppler phenomenon and simulated micro-Doppler signatures for various targets, such as rotating cylinders, vibrating scatterers and personnel targets. The authors also show that a time–frequency analysis of the radar return can be used to extract micro-Doppler signatures from the received signal, leading to additional information on the target that can be used for classification and recognition. An easy way to perceive target micro-Doppler modulations is to listen to an appropriate audible version of the radar return. Indeed, this method has been used to perform classification of targets but it has the disadvantage of requiring the presence of a human operators. For these reasons, the most recent challenge that has to be addressed is to develop micro-Doppler classification methods that allow automatic target recognition. Various studies have looked at classifying targets automatically through the analysis of their micro-Doppler signatures. For example, classification by micro-Doppler signatures of a tracked vehicle, a wheeled vehicle and a personnel target was performed in [5–7]. Here, classification performance was analysed via a K-nearest neighbourhood (K-NN) classifier and a Naïve Bayesian classifier by testing the main features of the micro-Doppler signatures that were extracted with both the principal component analysis (PCA) and the dynamic time warping algorithms. The use of micro-Doppler signatures for biometric purposes, and more specifically to identify humans, has been at the centre of a number of works in the last few years. In particular, there has been attempts to relate specific components of micro-Doppler gait signatures to parts of the body for identification purposes [8]. In [9] a technique introduced for visual pattern recognition is applied to human gait signatures to distinguish between different human motions. The exploitation of micro-Doppler information for classification has been recently extended to the field of radar imaging and in particular to inverse synthetic aperture radar (ISAR) [10–12] and there has also been an attempt to exploit signatures of moving personnel targets through multi-static radar measurements [13].

Although target classification by micro-Doppler signatures has been widely exploited in the RF regime, there has been rather little research done on the same topic in the acoustic regime. An acoustic Doppler sensor operating at 40 kHz was developed in 2007 to characterise a person's gait with the final goal to perform human recognition [14]. The system was deployed to measure Doppler signatures of men and women walking towards and away from the sensor which were used to assess classification performance of a Bayesian classifier. An additional active acoustic sensing

system, still operating at 40 kHz, was developed at Johns Hopkins University [15, 16]. The system was used to gather micro-Doppler signatures of humans and four-legged animals that were used to perform target classification [17].

It is desirable for an active surveillance system to operate indiscreetly, ideally be undetectable at all times and have minimal impact on the environmental surroundings including on any animals other than humans. This requires an acoustic radar to operate at frequencies outside the hearing sensitivity of, at least, most common non-human targets. Although humans can hear up to about 20 kHz, common pets' hearing system is sensitive up to much higher frequencies; dogs, for example, can hear up to about 40 kHz and cats can hear up to 60 kHz. For these reasons, it may be an operational requirement that an acoustic radar operates at frequencies above 40 kHz. The use of higher frequencies leads to an additional advantage; as the Doppler shift is directly proportional to the carrier frequency, given a certain frequency resolution, at higher frequencies small movements induce a wider frequency shift and therefore a clearer contribution to the micro-Doppler signature. On the other hand, the increased acoustic attenuation at higher frequencies lowers the operational range of the system.

In this paper a fully coherent acoustic radar operating at 80 kHz is described together with a set of experimental data, containing micro-Doppler signatures of personnel targets performing various actions. The data were collected at University College London between 2010 and 2011. Classification performance of a K-NN classifier and a Naïve Bayesian classifier trained to distinguish between different human motions and different personnel targets are then presented.

2 Micro-Doppler theory

The analytic signal of a pure tone $s(t)$ is defined as the signal $\hat{s}(t)$, such that $s(t) = \text{Real}\{\hat{s}(t)\}$, and is generally expressed in polar format as

$$\hat{s}(t) = \mathrm{e}^{\mathrm{j}2\pi f_0 t} \tag{1}$$

The Doppler shift induced by a target moving with a constant radial velocity v with respect to a radar or sonar system is

$$f_{\mathrm{D}} = 2f_0 \frac{v}{c} \tag{2}$$

where f_0 is the carrier frequency of the active sensor and c is the speed of propagation of the transmitted signal in a given medium. In the presence of a complex target, which is composed of number of parts N each one moving with a velocity $v_i(t)$, the resulting Doppler shift is given by the sum of each single Doppler shift as

$$f_{\mathrm{D}}(t) = \sum_{i=1}^{N} 2f_0 \frac{v_i(t)}{c} \tag{3}$$

The analytic signal of the echo return from such a target is given by

$$\hat{s}_{\mathrm{R}}(t) = \mathrm{e}^{\mathrm{j}2\pi f_0 t} \mathrm{e}^{\mathrm{j}2\pi f_{\mathrm{D}}(t)t} \tag{4}$$

Mixing the received signal $\hat{s}_{\mathrm{R}}(t)$ with the transmitted signal $\hat{s}(t)$ as follows

$$\hat{s}_{\mathrm{R}}(t)\hat{s}(t)^* = \mathrm{e}^{\mathrm{j}2\pi f_{\mathrm{D}} t} \tag{5}$$

allows extraction of the Doppler signature from the data. This is the component of the signal that contains the micro-Doppler information on the target and that can be used for target recognition and classification. The bandwidth of the resulting signal is normally much smaller than the carrier frequency with the information contained in the lower frequencies. The micro-Doppler signature can be represented in the two-dimensional (2D) time–frequency space by a short time Fourier transform. This is computed by calculating the Fourier transform of a series of 50% overlapping sliding windows $x_i(n)$ of a given length N. The ith window is defined as $x_i(n) = \hat{s}_{\mathrm{R}}(k)w(n)$, where $w(n)$ is a suitable weighting function and $k = n + i(N/2)$ for $n = 0, \ldots, N-1$. Under these circumstances the short time Fourier transform of the sequence $\hat{s}_{\mathrm{R}}(k)$ is given by

$$\mathrm{STFT}(i, K) = \sum_{n=0}^{N-1} x_i(n)\, \mathrm{e}^{-\mathrm{j}2\pi(nK/N)}, \qquad K = 0, \ldots, N-1 \tag{6}$$

The frequency resolution can be approximated as the inverse of the duration of window $T_{\mathrm{w}} = N/f_{\mathrm{s}}$, where f_{s} is the sampling rate, and therefore only Doppler shifts that are greater than $1/T_{\mathrm{w}}$, corresponding to velocities

$$v > \frac{c}{2f_0 T_{\mathrm{w}}} \tag{7}$$

will be clearly visible. Equation (7) shows that the the use of higher frequency has the additional advantage to induce a wider micro-Doppler bandwidth and that, given a certain frequency resolution, small movements are more easy to detect when the carrier frequency is higher.

3 Description of the acoustic radar

The acoustic radar developed is composed of a signal generator, a loudspeaker, a microphone, two pre-amplifiers that amplify the signals before transmission and after recording and a Data AcQuisition (DAQ) card that digitises the recorded signal. The waveforms are generated by a National Instrument NI PXI-6733 card capable of generating up to 500 MS amples/s (32 bit resolution) per channel. Each card can transmit on eight channels and therefore eight different waveforms could potentially be transmitted simultaneously. The generated waveform is given as in input to the preamplifier (Ultra Sound Advice S55A) capable of generating a maximum output of 140 V peak to peak. The loudspeaker provides a monitor output which is 1/100 of the actual output of the amplifier. This allows monitoring of the output and of the gain of the amplifier. The signal is transmitted by a loudspeaker (Ultra Sound Advice) which can nominally operate between 20 kHz and over 200 kHz. It can generate an output level greater than 105 dB SPL (sound pressure level) between 20 and 50 kHz (measured at 0.25 m) and greater than 85 dB SPL up to 150 kHz. The round-shaped active area of the loudspeaker has got a 50 mm diameter [18]. In the receive section, echoes are captured by a microphone capsule (Ultra Sound Advice UM3 capsule). The sensitive

doi: 10.1049/iet-rsn.2011.0087

element of the capsule is a 1.8 cm diameter disk made of a very thin aluminised polyester film. Nominally, its sensitivity is better than −57 dB from 20 to 120 kHz and better than −70 dB at 180 kHz. The maximum gain obtainable by the amplifier is nominally ×220 at 100 kHz [19]. The amplified received signal is digitised by a DAQ card (NI PXI-6133) which is capable of receiving eight channels simultaneously with a sampling frequency of up to 500 kHz per channel. The digitised data are stored in a file and the signal processing is performed offline with Matlab.

4 Classification of personnel targets

4.1 Description of the experiment

The data containing micro-Doppler signatures were gathered at UCL between 2010 and 2011. The microphone and the loudspeaker were arranged in a pseudo mono-static configuration next to each other as shown in Fig. 1. Their active areas were contained on the same plane perpendicularly to the ground floor with their phase centres positioned 12 cm far from each other and both at a height of 105 cm with respect to the ground floor. During the experiments the personnel targets were facing the microphone and the loudspeaker at a distance of 2 m. A 10 s long CW tone at 80 kHz with 0.3 V amplitude was generated with the NI PXI-6733 card to ensonify the targets. The gain of the loudspeaker pre-amplifier, that was measured by using the monitor output, was about 20 dB. Echoes were recorded with the microphone and then sampled at a rate of 500 kHz. Micro-Doppler signatures were gathered for three different personnel targets undertaking various actions

- walking
- running
- walking while carrying an object in one hand
- walking while carrying an object with both hands
- walking with a heavy backpack on the shoulders

The physical characteristics of the targets are given in Fig. 2. When walking or running the targets were moving on a Pro Fitness manual treadmill and were ensonified from the back to avoid any masking effects. This was deployed to remove the main Doppler shift contribution from the Doppler signatures, and to keep the signal-to-noise ratio of the received signal as constant as possible during the recording. Also the treadmill removed any movement constraint on the targets and therefore allowed longer recordings. A measure of the background noise and background clutter were taken.

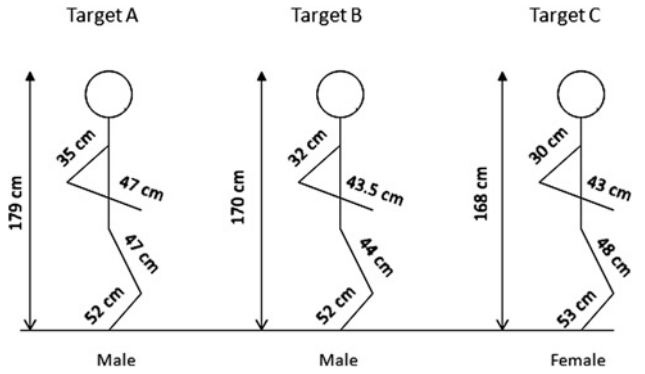


Fig. 2 *Physical characteristics of the three personnel targets, a female and two males, used for the experiments*

4.2 Data pre-processing

A recording of the background clutter was removed from the raw data for each target recording. The analytic signal of the difference signal was mixed with the complex conjugate of the analytic signal of the transmitted tone to obtain the baseband signal containing the micro-Doppler signatures. The analytic signals were obtained by applying the Matlab *Hilbert.m* function to the original real signals. The mixer output was filtered with an ideal low-pass filter cutting off all the frequencies over 2.5 kHz. It was experimentally observed, in fact, that the micro-Doppler signatures were contained within the frequencies below 2.5 kHz. The resulting signal was down-sampled of a factor 50 leading to a final sampling frequency of 10 kHz. This was done to reduce the load of the data without altering the information contained in the frequencies below 2.5 kHz. The short time Fourier transform was applied to all the recordings to visualise the recorded micro-Doppler signatures. This was calculated by using 30 ms long 50% overlapping windows weighted with a Hanning function and by computing their respective Fourier transform as described in Section 2. The measured micro-Doppler signatures for one of the two targets taking all the action listed above are given in Figs. 3–5.

4.3 Classification performance analysis

The baseband data obtained after pre-processing containing the micro-Doppler signatures were used to assess and compare the performance of a Bayesian classifier and a K-NN classifier. In

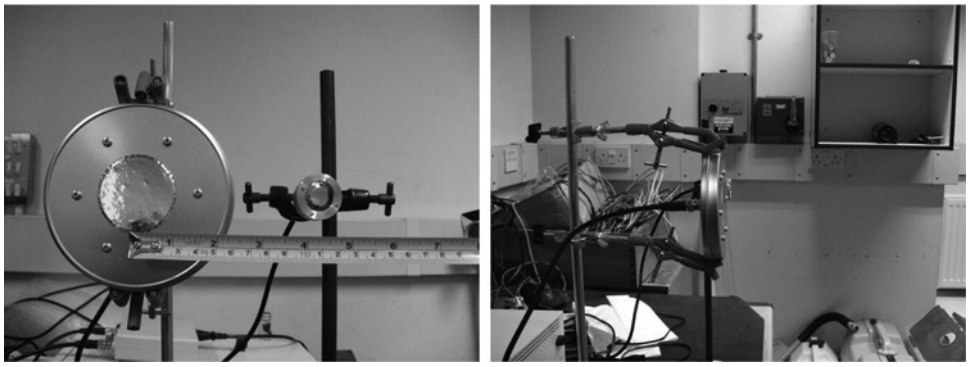


Fig. 1 *Frontal and lateral view of the space arrangement of the microphone and the loudspeaker during the experiment*

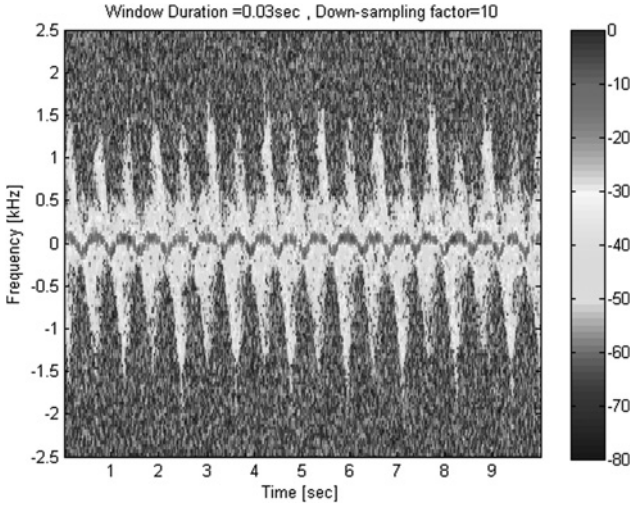


Fig. 3 *Micro-Doppler signature of Target A walking on a treadmill that was facing the acoustic radar at a distance of 2 m*

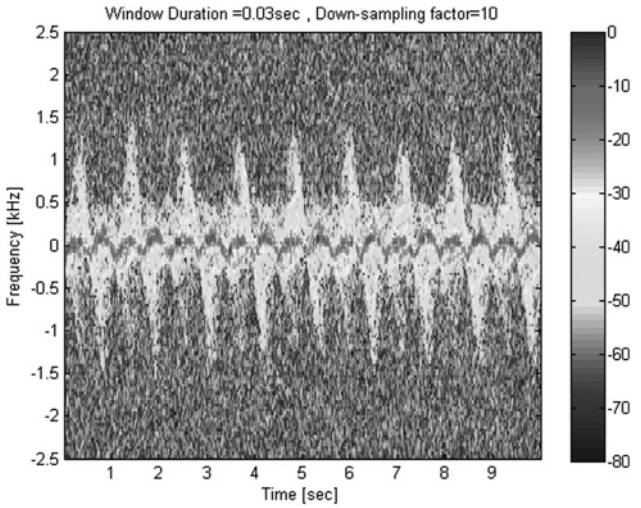


Fig. 4 *Micro-Doppler signature of Target A walking on a treadmill while carrying an object with the left hand*

Dimension of the book were such to obstruct the left arm from swinging

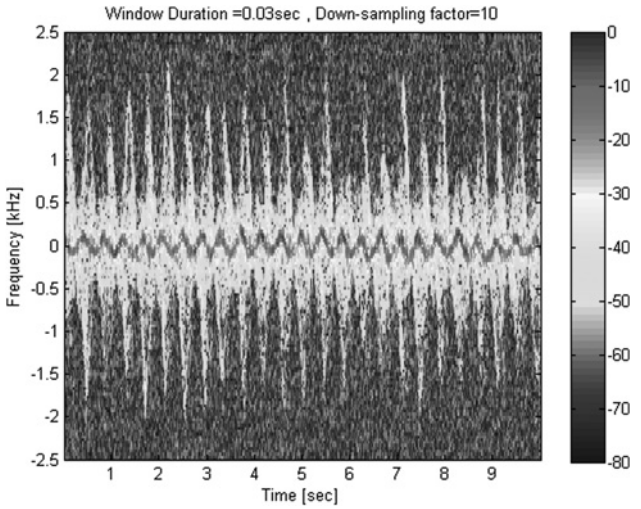


Fig. 5 *Micro-Doppler signature of Target A running on a treadmill*

Table 1 Total number of training windows N_{Train} and test windows N_{Test} used to assess the performance of the classifier for each window duration T_w

T_w	N_w	N_{Train} (1.5 s)	N_{Test}
30 ms	3330	99	3231
100 ms	995	29	966
150 ms	660	19	641

particular, our main focus was to distinguish between different personnel targets undertaking the same action or between different actions undertaken by the same target. The time sequences corresponding to each class were divided in 50% overlapping windows of duration T_w. The analysis was repeated for three different window durations; $T_w = 30$ ms, $T_w = 100$ ms, $T_w = 150$ ms. This resulted in a total number N_w of windows for each class as indicated in Table 1. The content of each window was organised in column vectors $\boldsymbol{x}_j^i$ where j indicates the jth window belonging to the ith class. The first N_{Train} windows for each class, corresponding to the first 1.5 s of the recorded signal, were used to train the classifiers, so that the information used for training remained the same in all three cases. They formed the matrix $\boldsymbol{X} = (\boldsymbol{X}^1\boldsymbol{X}^2 \ldots \boldsymbol{X}^C)$ representing the training set, where the jth columns of each sub-matrix $\boldsymbol{X}^i$ is

$$\boldsymbol{X}^i(j) = \boldsymbol{x}_j^i, \quad j = 1 \ldots N_{\text{Train}} \tag{8}$$

and C being the number of classes under test. The remaining N_{Test} windows were used to form the test set $\boldsymbol{Y} = (\boldsymbol{Y}^1\boldsymbol{Y}^2 \ldots \boldsymbol{Y}^C)$ where similarly the jth columns of each sub-matrix $\boldsymbol{Y}^i$ is

$$\boldsymbol{Y}^i(j) = \boldsymbol{x}_{j+N_{\text{Train}}}^i, \quad j = 1 \ldots N_{\text{Test}} \tag{9}$$

The PCA, Cepstrum and Mel-Cepstrum algorithms were used to extract N_f features from each window to reduce the dimensionality of the data and reduce the computational load of the classifiers [20, 21].

4.4 Feature extraction

1. *PCA:* The training matrix $\boldsymbol{X}$ was used to calculate the cross correlation matrix $\boldsymbol{S}$ defined as

$$\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^{\text{H}} \tag{10}$$

where $\boldsymbol{X}^{\mathbf{H}}$ is the transposed and conjugate of matrix $\boldsymbol{X}$. The N_f eigenvector of $\boldsymbol{S}$ corresponding to its N_f largest eigenvalues were used to form the column of a matrix $\boldsymbol{A}$. This was used to perform the dimensionality reduction as

$$\boldsymbol{D} = \boldsymbol{A}^{\text{H}}\boldsymbol{Y} \tag{11}$$

and

$$\boldsymbol{E} = \boldsymbol{A}^{\text{H}}\boldsymbol{X} \tag{12}$$

$\boldsymbol{D}$ and $\boldsymbol{E}$ are the training and test sets after dimensionality reduction, respectively. Each column of $\boldsymbol{D}$ will be referred to as a training feature vector and likewise each column of $\boldsymbol{E}$ represents a test feature vector.

doi: 10.1049/iet-rsn.2011.0087

2. *Cepstrum:* The Cepstrum of a signal $s(t)$ is defined as the inverse Fourier transform of the quantity $C(f) = \log_{10}|S(f))|$ where $S(f)$ is the Fourier transform of the signal $s(t)$ [21]. The Cepstrum feature vector $\boldsymbol{c}_j^i$, for each window $\boldsymbol{x}_j^i$, was calculated as the inverse Fourier transform of the Cepstrum C_j^i as

$$\boldsymbol{c}_j^i(n) = \frac{1}{N}\sum_{K=0}^{N-1} C_j^i(K)\mathrm{e}^{\mathrm{j}2\pi(nK/N)}, \quad n = 0, \ldots, N_f - 1 \tag{13}$$

with

$$C_j^i(K) = \log_{10}\left|\sum_{n=0}^{N-1} x_j^i(n)\mathrm{e}^{-\mathrm{j}2\pi(nK/N)}\right| \quad K = 0, \ldots, N-1 \tag{14}$$

The feature vector corresponding to the N_w training windows of each class were used to form the training test $\boldsymbol{D}$ as

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{c}_1^1 & \ldots & \boldsymbol{c}_{\mathrm{N_{Train}}}^1 & \ldots & \boldsymbol{c}_1^{\mathrm{C}} & \ldots & \boldsymbol{c}_{\mathrm{N_{Train}}}^{\mathrm{C}} \end{pmatrix} \tag{15}$$

and the remaining feature vectors were used to form the test set $\boldsymbol{E}$ as

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{c}_{\mathrm{N_{Train}}+1}^1 & \ldots & \boldsymbol{c}_{\mathrm{N_w}}^1 & \ldots & \boldsymbol{c}_{\mathrm{N_{Train}}+1}^{\mathrm{C}} & \ldots & \boldsymbol{c}_{\mathrm{N_w}+1}^{\mathrm{C}} \end{pmatrix} \tag{16}$$

3. *Mel-Cepstrum:* To obtain the Mel-Cepstrum feature vectors, the Cepstrum of each window $C_j^i(K)$ was filtered with a bank of non-overlapping triangular filter $H_p(K)$, all characterised by the same the same bandwidth $B_{mel} = 100$ MEL, to generate a sequence $C_{\mathrm{mel}_j}^i(K)$ defined as

$$C_{\mathrm{mel}_j}^i(K) = \begin{cases} \sum_{\hat{K}=1}^{N} C_j^i(\hat{K})H_p(\hat{K}) & K = K_p \\ 0 & K \neq K_p \end{cases} \tag{17}$$

In the equation K_p corresponds to the frequency bin containing the central frequency of the filter $H_p(K)$. The Mel-Cepstrum coefficient were calculated as the inverse Fourier transform of the sequence $C_{\mathrm{mel}_j}^i$ as

$$c_{\mathrm{mel}_j}^i(n) = \frac{1}{N}\sum_{K=0}^{N-1} C_{\mathrm{mel}_j}^i(K)\mathrm{e}^{\mathrm{j}2\pi(nK/N)}, \quad n = 0, \ldots, N_f - 1 \tag{18}$$

The first Mel-Cepstrum coefficient $c_j^i(0)$, which has been regarded as unstable in previous work, was not used to perform classification [22, 23]. To allow a fair comparison, this was also done for the Cepstrum features. The Mel-Cepstrum feature vectors corresponding to the N_w training windows of each class were used to form the training test $\boldsymbol{D}$ as

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{c}_{\mathrm{mel}_1}^1 & \ldots & \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_{Train}}}}^1 & \ldots & \boldsymbol{c}_{\mathrm{mel}_1}^{\mathrm{C}} & \ldots & \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_{Train}}}}^{\mathrm{C}} \end{pmatrix} \tag{19}$$

and those remaining were used to form the test set $\boldsymbol{E}$ as

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_{Train}}+1}}^1 & \ldots & \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_w}}}^1 & \ldots & \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_{Train}}+1}}^{\mathrm{C}} & \ldots & \boldsymbol{c}_{\mathrm{mel}_{\mathrm{N_w}}}^{\mathrm{C}} \end{pmatrix} \tag{20}$$

4.5 Classifiers

1. *Naïve Bayesian classifier:* The Naïve Bayesian classifier assumes that the feature vector is a statistical process whose element are all Gaussian distributed and statistically independent. Under this assumption the statistical distribution of the feature vector is given by the product of the Gaussian distributions of each of its elements. The training set D was used to estimate the mean value η_i and the variance σ_i^2 of the in-phase and in-quadrature components of each element for each class as

$$\eta_i = \frac{1}{N_{\mathrm{Test}}}\sum_{n=1}^{N_{\mathrm{Test}}} d(i, n) \tag{21}$$

and

$$\sigma_i^2 = \frac{1}{N_{\mathrm{Test}}}\sum_{n=1}^{N_{\mathrm{Test}}} (d(i, n) - \eta_i)^2 \tag{22}$$

where $d(i, j)$ is either the in-phase or in-quadrature component of the element (i, j) of D. The in-phase and in-quadrature components were also assumed to be statically independent. These resulted in a number C of probability distribution function (pdf), each one associated to a class. The value

Table 2 Confusion matrices reporting the performance of the Naïve Bayesian classifier and the K-NN classifier ($k = 3$) testing the walking gait of Target A and that of Target C for all feature extraction algorithms ($T_w = 30$ ms)

	PCA		Cepstrum		Mel-Cepstrum	
	Target A	Target C	Target A	Target C	Target A	Target C
Naïve Bayesian						
Target A, %	97	3	91.7	8.3	85	15
Target C, %	11.9	88.1	27.6	72.4	24	76
	$P_{cc} = 92.57\%$		$P_{cc} = 82.03\%$		$P_{cc} = 80.52\%$	
K-NN						
Target A, %	47.2	89.8	10.2	52.8	80.3	19.7
Target C, %	1	18.8	81.2	99	12.3	87.7
	$P_{cc} = 73.1\%$		$P_{cc} = 85.52\%$		$P_{cc} = 84\%$	

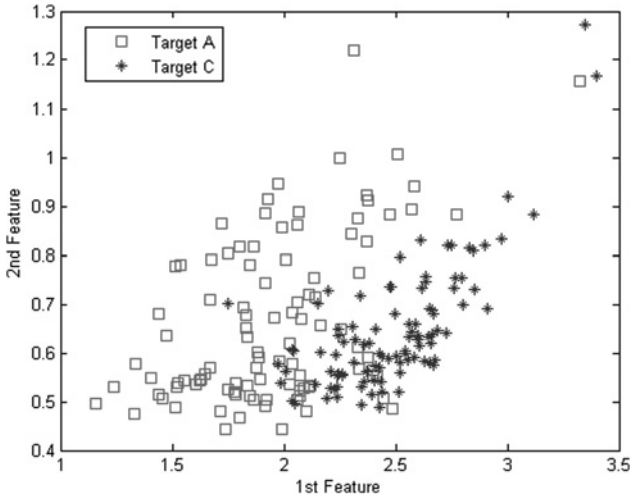


Fig. 6 *Plot of the first two Mel-Cepstrum features in the 2D plane for Target A and Target C*

of each distribution was calculated for each feature vector in the test set which was then assigned to the class presenting the largest value of the density function. This was repeated for all test feature vectors to estimate classification performance [21].

2. *The K-NN classifier:* The Euclidean distance between the test feature vector, that is, a column of matrix $\boldsymbol{E}$, and all the training vectors was calculated. These resulted in a number N_{Train} of distances for each class. The $k = 3$ or $k = 5$ lowest distances, depending on the number of classes under test, were selected and the test vector was assigned to the class that owned the highest number of training vector between those which generated the selected k lowest distances. This was repeated for all test feature vectors to assess classification performance [21].

4.6 Results

Table 2 shows the classification performance that was obtained from the comparison between the walking gait of Target A and Target C for $T_w = 30$ ms. This is divided in a sub-table per classifier, each one containing the three confusion matrices related to the three feature extraction algorithms that were used. This allows a straightforward comparison of the results obtained with different feature extraction algorithms given a specific classifier. In each confusion matrix, the percentage number is the ratio between the number of assignments and the total number of windows under test, given the same class. The averaged probability of correct classification (P_{cc}) is given below each confusion matrix. This was calculated under the assumption that each class had the same a priori probability. The parameter k of the K-NN classifier was set equal to 3. Results show that high-level classification performance can be obtained. In particular, for this case, the combination PCA plus Naïve Bayesian classifier leads to a rate of correct classification over 90%. The lowest rate, instead, is given by the PCA followed by the K-NN classifier presenting a high number of wrong decisions

Table 3 Confusion matrices reporting the performance of the Naïve Bayesian classifier and the K-NN classifier ($k = 5$) testing the walking gait of Target A, Target B and Target C for all feature extraction algorithms ($T_w = 30$ ms)

Naïve Bayesian	PCA			Cepstrum		
	Target A	Target B	Target C	Target A	Target B	Target C
Target A, %	85.2	14.2	0.5	89	6	5.1
Target B, %	15.7	39.8	44.5	36	29.8	34.1
Target C, %	3	30	67	23.7	28.5	47.8
		$P_{cc} = 64.04\%$			$P_{cc} = \%$	

Naïve Bayesian	Mel-Cepstrum		
	Target A	Target B	Target C
Target A, %	80.7	12.2	7.2
Target B, %	32.5	23.8	43.8
Target C, %	21.8	21.3	57
		$P_{cc} = 53.78\%$	

K-NN	PCA			Cepstrum		
	Target A	Target B	Target C	Target A	Target B	Target C
Target A, %	30	24.6	27.6	78.2	9	4
Target B, %	1.3	19.4	74	21	30.1	36.6
Target C, %	0.2	10.6	87	11.6	20	57.4
		$P_{cc} = 44.80\%$			$P_{cc} = 55.23\%$	

K-NN	Mel-Cepstrum		
	Target A	Target B	Target C
Target A, %	70.5	12.7	8
Target B, %	13.6	31.1	40.2
Target C, %	5.3	26.5	54
		$P_{cc} = 51.86\%$	

doi: 10.1049/iet-rsn.2011.0087

when Target A is given. Overall, in all other cases, correct classification performance is above 80%. Fig. 6 shows the magnitude of the first two Mel-Cepstrum features of each class, in the 2D plane, giving a visual idea of the degree of separability between the two classes [21].

The confusion matrices obtained from the comparison of the walking gaits of all three targets, for $T_w = 30$ ms, are given in Table 3. The parameter k of the K-NN classifier was set to 5 and all cases that could not univocally be assigned to any of the three classes were treated as unknowns and counted to average performance. Classification performance drops for both the classifiers and all feature extraction algorithms, all presenting a high rate of wrong decisions between Target B and Target C. The highest rate of correct classification is achieved by the PCA algorithm followed by the Naïve Bayesian classifier and this is equal to 64%. The drop in performance may be due to differences in the gait of Target B and Target C that cannot be resolved by the feature extraction algorithms. This

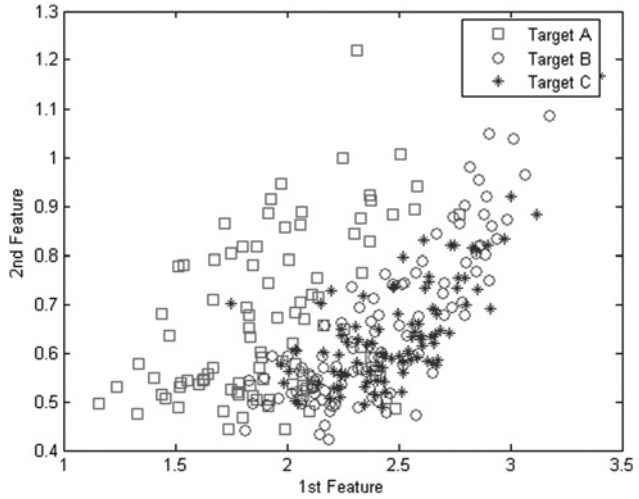


Fig. 7 *Plot of the first two Mel-Cepstrum features in the 2D plane for Target A, Target B and Target C ($T_W = 30$ ms)*

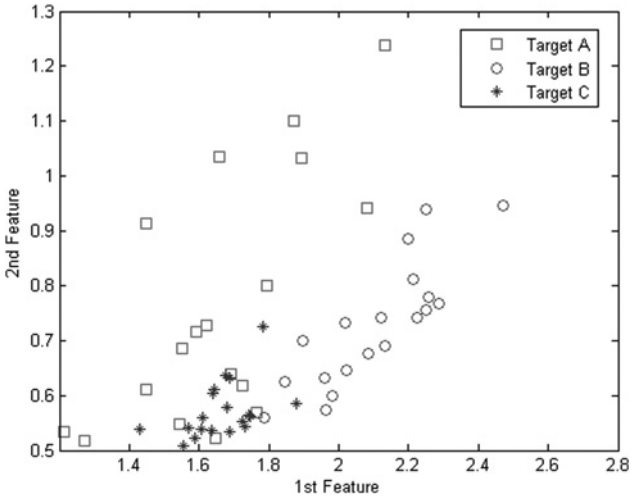


Fig. 8 *Plot of the first two Mel-Cepstrum features in the 2D plane for Target A, Target B and Target C ($T_W = 150$ ms)*

appears clear by looking at Fig. 7 where the magnitude of the first two Mel-Cepstrum features of each class are plot against each other on the 2D plane. A substantial overlapping between the features related to Target B and those related to Target C can be observed. Classification performance improves when the duration of the windows increases. Table 4 reports the results for the same case but with the window duration set to $T_w = 150$ ms. The probability of correct classification on the average improves of 20% up to a correct rate of 71% given by the combination of the PCA algorithm followed by the Naïve Bayesian classifier. The plot of the magnitude of the first two Mel-Cepstrum features, given in Fig. 8, shows a higher degree of separability between the three classes that clearly leads to the improvement in performance. The same behaviour is observed for all feature extraction algorithms.

Table 5 reports the results obtained when the classifiers were trained to distinguish between the same personnel target, Target B, undertaking two different actions for

Table 4 Confusion matrices reporting the performance of the Naïve Bayesian classifier and the K-NN classifier ($k = 5$) testing the walking gait of Target A, Target B and Target C for all feature extraction algorithms ($T_w = 150$ ms)

Naïve Bayesian	PCA			Cepstrum		
	Target A	Target B	Target C	Target A	Target B	Target C
Target A, %	81.4	16.3	2.2	84.3	4	11.9
Target B, %	8.9	49.3	41.8	3	58.8	38.2
Target C, %	2.2	13.6	84.3	2	1.9	96.3
		$P_{cc} = 71.66\%$			$P_{cc} = 70.77\%$	

Naïve Bayesian	Mel-Cepstrum		
	Target A	Target B	Target C
Target A, %	62.7	11.2	26
Target B, %	11.6	54.1	34.3
Target C, %	3.4	1.3	95.3
		$P_{cc} = 70.7\%$	

K-NN	PCA			Cepstrum			Mel-Cepstrum		
	Target A	Target B	Target C	Target A	Target B	Target C	Target A	Target B	Target C
Target A, %	0	15.9	83.5	57.3	12.5	25.6	63.2	11.6	21.1
Target B, %	0	61.8	38.2	3	52.7	41.2	5.6	59.4	28.7
Target C, %	0	1.4	98.6	0.2	0.9	98.1	9.2	1.2	88.8
		$P_{cc} = \%$			$P_{cc} = 69.37\%$			$P_{cc} = 70.05\%$	

Table 5 Confusion matrices reporting the performance of the Naïve Bayesian classifier and the K-NN classifier ($k = 5$) testing the micro-Doppler signatures of Target B walking normally and Target B walking with a heavy weight on the shoulders. Results are given for all feature extraction algorithms ($T_w = 30$ ms)

	PCA		Cepstrum		Mel-Cepstrum	
	B-WN	B-WBP	B-WN	B-WBP	B-WN	B-WBP
Naïve Bayesian						
B-WN, %	87.5	12.5	60	40	70	30
B-WBP, %	23.4	76.6	9.6	90.4	14.6	85.4
	$P_{cc} = 82.05\%$		$P_{cc} = 75.21\%$		$P_{cc} = 77.7\%$	
K-NN	B-WN	B-WBP	B-WN	B-WBP	B-WN	B-WBP
B-WN, %	96.5	3.5	57.8	42.2	77.2	22.8
B-WBP, %	46.5	53.5	7	93	19	81
	$P_{cc} = 75\%$		$P_{cc} = 75.38\%$		$P_{cc} = 79.12\%$	

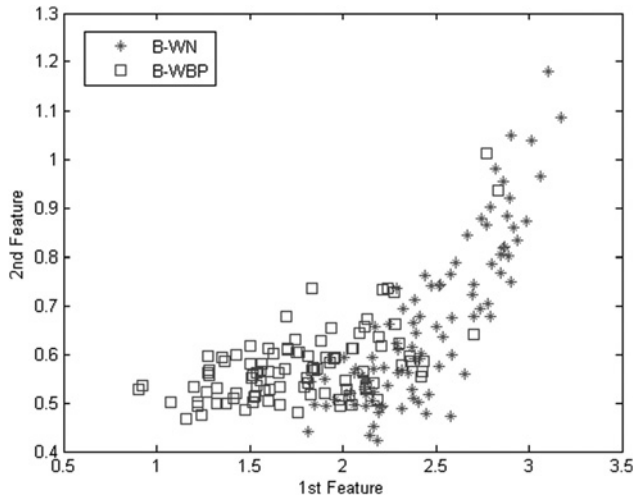


Fig. 9 *Plot of the first two Mel-Cepstrum features in the 2D plane for Target B walking normally (Class B-WN, green) and Target B walking with a heavy backpack on the shoulders (Class B-WBP, brown)*

$T_w = 30$ ms. In particular, results are related to the comparison between Walking Normally (Class B-WN) and Walking with a BackPack (Class B-WBP). The parameter k of the K-NN classifier was set to 3. Results show that it is possible to achieve correct classification performance over 80%. The highest level of performance is achieved by the PCA algorithms and the Mel-Cepstrum algorithm for the Naïve Bayesian classifier. Fig. 9 shows the magnitude of the first two Mel-Cepstrum features for each class in the 2D plane.

Overall performance varies depending on the target and the type of actions under test. In particular, as expected, results corroborate that the closest the classes under test, the highest the drop in performance. Classification of a target walking normally against the same target walking with an object in a hand proved to be highly challenging and showed, on the average, a drop of the performance to about 60% for $T_w = 30$ ms. On the average, classification performance tends to increase as the window duration T_w becomes longer.

These results show that acoustic radar micro-Doppler signatures can be easily obtained and give valuable information on target characteristics. Further research is needed to determine exactly the observation time that is required to achieve robust classification performance and whether other classification methods and additional features can be used to aid target classification.

5 Conclusions and future work

An acoustic radar operating at 80 kHz was developed for collection of micro-Doppler signatures of moving targets, doubling the frequency used in previous works. The acoustic radar was deployed in a set of experimental trials in which micro-Doppler signatures of various personnel targets undertaking a number of actions were collected in a highly indoor cluttered environment. These were used to test a K-NN and a Naïve Bayesian classifier trained to distinguish between walking gaits of different targets or between different action undertaken by the same targets. Results show that the acoustic radar can be successfully deployed at short ranges to collect micro-Doppler signatures of moving targets. Classification performance results show that the information contained in these acoustic signatures can be used to perform identification and recognition of personnel targets. Data have been gathered and analysed for a range of personnel target motions and good levels of classification have been achieved. Although the acoustic radar can only survey short ranges because of the high attenuation of sound waves in air, it can be successfully deployed in indoor environments, such as airports, to monitor, for example, the flow of passengers through doors or security checks with the potential to identify suspicious behaviour. In addition to this, the acoustic radar offers an easy and inexpensive way to collect micro-Doppler signatures of a wide range of targets, often difficult to obtain, that can be used to analyse human walking gaits or other types of behaviour in a number of applications.

Future work will look at understanding how additional features such as geometrical features (i.e. the signature period or the signatures maximum Doppler shift) or how the history of the current features (i.e. the time variation of the Cepstrum or MEL-Cepstrum coefficients) can be deployed to enhance classification performance. Extending the work presented in this paper to the underwater regime is also under consideration.

6 Acknowledgment

This work is part of a project that has been funded by the UK MoD University Defence Research Centre (UDRC) in signal processing. The authors would like to thank the UK EPSRC Engineering Instrument Pool [24].

doi: 10.1049/iet-rsn.2011.0087

7 References

1 Gill, T.P.: 'The Doppler effect, an introduction to the theory of the effect' (Logos Press, Limited, 1965)
2 Chen, V., Li, F., Ho, S.-S., Wechsler, H.: 'Analysis of micro-doppler signatures', *IEE Proc. Radar Sonar Navig.*, 2003, **150**, (4), pp. 271–276
3 Chen, V., Li, F., Ho, S.-S., Wechsler, H.: 'Micro-doppler effect in radar: phenomenon, model, and simulation study', *IEEE Trans. Aerosp. Electron. Syst.*, 2006, **42**, (1), pp. 2–21
4 Chen, V.: 'Analysis of radar micro-doppler with time-frequency transform'. Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing, 2000
5 Smith, G.E., Woodbridge, K., Baker, C.J.: 'Template based micro-doppler signature classification'. The Institution of Engineering and Technology Seminar on High Resolution Imaging and Target Classification, 21 November 2006, pp. 127–144
6 Smith, G., Woodbridge, K., Baker, C.: 'Micro-doppler signature classification'. Int. Conf. Radar 2006, CIE '06, 16–19 October, 2006, pp. 1–4
7 Smith, G., Woodbridge, K., Baker, C.: 'Naive Bayesian radar micro-doppler recognition'. Int. Conf. Radar, 2008, 2–5 September, 2008, pp. 111–116
8 Tahmoush, D., Silvious, J.: 'Radar micro-doppler for long range front-view gait recognition'. IEEE Third Int. Conf. Biometrics: Theory, Applications, and Systems, 2009, BTAS '09, 28–30 September, 2009, pp. 1–6
9 Tivive, F., Bouzerdoum, A., Amin, M.: 'Automatic human motion classification from doppler spectrograms'. Second Int. Workshop Cognitive Information Processing (CIP), 2010, pp. 237–242
10 Vignaud, L., Ghaleb, A., Le Kernec, J., Nicolas, J.-M.: 'Radar high resolution range & micro-doppler analysis of human motions'. Radar Conf. – Surveillance for a Safer World, 12–16 October, 2009, pp. 1–6
11 Ghaleb, A., Vignaud, L., Nicolas, J.: 'Micro-doppler analysis of wheels and pedestrians in isar imaging', *IET Signal Process.*, 2008, **2**, (3), pp. 301–311
12 Chen, V., Miceli, W., Himed, B.: 'Micro-doppler analysis in ISAR – review and perspectives'. Int. Radar Conf. – Surveillance for a Safer World, 2009, pp. 1–6
13 Sammartino, P., Fortuny-Guash, J.: 'Space and frequency diversity for moving personnel spectrogram estimation'. Conf. Proc. Radar 2010, 2010
14 Kalgaonkar, K., Raj, B.: 'Acoustic doppler sonar for gait recogination'. IEEE Conf. on Advanced Video and Signal Based Surveillance, 2007, AVSS 2007, 2007, pp. 27–32
15 Zhang, Z., Pouliquen, P.O., Waxman, A., Andreou, A.G.: 'Acoustic micro-Doppler radar for human gait imaging', *J. Acoust. Soc. Am. Express Lett.*, 2007, **121**, (3), pp. 110–113
16 Zhang, Z., Pouliquen, P., Waxman, A., Andreou, A.: 'Acoustic micro-doppler gait signatures of humans and animals'. 41st Annual Conf. on Information Sciences and Systems, 2007, CISS '07, 2007, pp. 627–630
17 Zhang, Z., Andreou, A.: 'Human identification experiments using acoustic micro-doppler signatures'. Argentine School of Micro-Nanoelectronics, 2008, Technology and Applications, 2008, EAMTA 2008, pp. 81–86
18 Ultra Sound Advice: 'Instruction manual for ultrasound loudspeaker system', www.ultrasoundadvice.co.uk
19 Ultra Sound Advice: 'SP3 PSU Operating instruction'. Ultra Sound Advice, www.ultrasoundadvice.co.uk
20 Duda, R., Hart, P., Stork, D.: 'Pattern classification' (John Wiley and Sons, 2001)
21 Theodoridis, S., Koutroumbas, K.: 'Pattern recognition – Fourth edition' (Academic Press, 2009)
22 Zheng, F., Zhang, G., Song, Z.: 'Comparison of different implementations of MFCC', *J. Comput. Sci. Technol.*, 2001, **16**, (6), pp. 582–581
23 Piccone, J.: 'Signal modeling techniques in speech recognition', *Proc. IEEE*, 1993, **81**, (9), pp. 1215–1247
24 Balleri, A., Chetty, K., Woodbridge, K.: 'Frequency-agile non-coherent ultrasound radar for collection of micro-Doppler signatures'. IEEE Radar Conf. 2011, Kansas City, USA, May 2011