2015

Classification of micro-Doppler signatures of human motions using log-Gabor filters

Fok Hing Chi Tivive
University of Wollongong, tivive@uow.edu.au

Son Lam Phung
University of Wollongong, phung@uow.edu.au

Abdesselam Bouzerdoum
University of Wollongong, bouzer@uow.edu.au

Publication Details

F. Tivive , S. Lam. Phung & A. Bouzerdoum, "Classification of micro-Doppler signatures of human motions using log-Gabor filters," IET Radar, Sonar and Navigation, vol. 9, (9) pp. 1188-1195, 2015.

Classification of micro-Doppler signatures of human motions using log-Gabor filters

Abstract
In recent years, Doppler radar has been used as a sensing modality for human gait recognition, due to its ability to operate in adverse weather and penetrate opaque obstacles. Doppler radar captures not only the speed of the target, but also the micro-motions of its moving parts. These micro-motions induce frequency modulations that can be used to characterise the target movements. However, a major challenge in Doppler signal processing is to extract discriminative features from the radar returns for target classification. This study presents a feature extraction method for classification of human motions from the micro-Doppler radar signal. The proposed method applies the log-Gabor filters at multiple spatial frequencies and orientations on a joint time-frequency representation. To achieve invariance to the target speed, features are extracted from local patches along the torso Doppler shift. Then, the (2D)2PCA (two-directional two-dimensional principal component analysis) method is applied to create a compact feature vector. Experimental results based on real radar data obtained from multiple human subjects demonstrate the effectiveness of the proposed approach in classifying arm motions.

Keywords
gabor, signatures, filters, doppler, micro, log, classification, motions, human

Disciplines
Engineering | Science and Technology Studies

Publication Details
F. Tivive , S. Lam. Phung & A. Bouzerdoum, "Classification of micro-Doppler signatures of human motions using log-Gabor filters," IET Radar, Sonar and Navigation, vol. 9, (9) pp. 1188-1195, 2015.

Classification of micro-Doppler Signatures of Human Motions using Log-Gabor Filters

Fok Hing Chi Tivive, Son Lam Phung, and Abdesselam Bouzerdoum

***Abstract*—In recent years, Doppler radar has been used as a sensing modality for human gait recognition, due to its ability to operate in adverse weather and penetrate opaque obstacles. Doppler radar captures not only the speed of the target, but also the micro-motions of its moving parts. These micro-motions induce frequency modulations that can be used to characterize the target movements. However, a major challenge in Doppler signal processing is to extract discriminative features from the radar returns for target classification. This study presents a feature extraction method for classification of human motions from the micro-Doppler radar signal. The proposed method applies the log-Gabor filters at multiple spatial frequencies and orientations on a joint time-frequency representation. To achieve invariance to the target speed, features are extracted from local patches along the torso Doppler shift. Then, the (2D)^{2}PCA method is applied to create a compact feature vector. Experimental results based on real radar data obtained from multiple human subjects demonstrate the effectiveness of the proposed approach in classifying arm motions.**

***Index Terms*—Human micro-Doppler signature classification, feature extraction, S-method, Log-Gabor filters.**

I. Introduction

With the advances in radar technology, there is an increasing interest in using Doppler radar for human gait recognition and activity monitoring. A modern Doppler radar detects not only the velocity of a target but also the local dynamics of its moving parts. The micro-movements induce frequency modulations around the main Doppler shift are commonly known as micro-Doppler (μ-D) effects. Several studies have been conducted to analyze μ-D signatures of moving targets [1]–[5]. An early study on μ-D effects investigated the jet engine modulation of radar returns for target identification [1]. Later, Chen *et al.* developed mathematical models and performed simulations to analyze μ-D effects of targets under translation, rotation, and vibration [2]. Other researchers conducted numerical simulations and real experiments, which demonstrated that the μ-D signature represents the kinetic motions of an object and provides a viable means for object identification [3]–[5]. Micro-Doppler signals have been used for classifying rigid targets, such as helicopters and aircrafts [6], and wheel and track vehicles [7]. They have also been used to differentiate rigid and non-rigid targets, e.g. humans and vehicles [8]–[10]. In recent years, the research on μ-D signals has been focused on analyzing human movements [11]–[21]. In these applications, one common challenge is to extract discriminative features from the radar returns for classification.

The authors are with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia. E-mails: {tivive, phung, a.bouzerdoum}@uow.edu.au.

In this paper, we propose a μ-D feature extraction method for classifying human movements. First, the μ-D radar signature is obtained using time-frequency (T-F) analysis. Then, two-dimensional (2D) filters are applied on the T-F representation. Since the μ-D modulations induced by the arm and leg motions appear close to the main Doppler frequency, a local T-F patch centered on the torso frequency shift is located for feature extraction. This step makes the proposed method stable against variations in the target speed. Then, log-Gabor filters are used to extract features from the T-F patch. This type of filters has neither DC component nor bandwidth limitation. Therefore, a small set of filters is adequate to cover the desired frequency spectrum. The remainder of the paper is organized as follows. Section II presents a brief description of the μ-D signal and the related work on human μ-D radar signature classification. Section III describes the proposed feature extraction method. Section IV presents the experimental results, and finally, Section V gives the concluding remarks.

II. Related work

In this section, a brief description of μ-D radar signal is given in Section II-A, followed by a review of existing μ-D radar signature classification approaches in Section II-B. Then, three T-F analysis techniques used to depict μ-D radar signature are presented in Section II-C.

A. Micro-Doppler radar signal

When a radar signal is backscattered from a target moving at a constant radial velocity, the carrier frequency of the radar signal is shifted according to the target velocity. If the target is a point scatterer, the received signal can be expressed as

$$x(t) = A(t)\exp\{j\ \phi_d(t)\}, \tag{1}$$

where $A(t)$ is the time-varying amplitude and $\phi_d(t)$ is the instantaneous Doppler phase [11]. Let λ be the transmitted signal wavelength and $v(t)$ be the target velocity. The instantaneous Doppler phase is given by

$$\phi_d(t) = \frac{4\pi}{\lambda}\int_0^t v(\tau)\ d\tau. \tag{2}$$

A complex target such as a human can be represented as a set of point scatterers. For example, Bilik and Tabrikian modeled a human body as a set of K segments; each segment moves at its own velocity and has M points [11]. The signal received by the antenna can be written as

$$x(t) = \sum_{k=1}^{K}\sum_{m=1}^{M} A_{k,m}(t)e^{j\frac{4\pi l_{k,m}}{\lambda}\int_0^t \hat{\beta}_k(\tau)\cos\beta_k(\tau)\ d\tau}, \tag{3}$$

where $A_{k,m}$ is the amplitude for the m-th point-target on the k-th segment, $\beta_k(t)$ is the instantaneous angle from the zenith of the segment, $\hat{\beta}_k(t)$ is the instantaneous angular velocity, and $l_{k,m}$ is the distance of the m-th point along the segment. In practice, the human locomotion is much more complex than the model given in (3). Therefore, advanced modeling tools, e.g. electromagnetic wave scattering model, have been employed to simulate the μ-D signals exhibited by different parts of a moving person [22].

B. Existing micro-Doppler classification approaches for human motions

The various μ-D classification approaches require first the extraction of salient features from the radar signal. The features are calculated from the *time*, *frequency* or *joint time-frequency* domain. In the time domain, Fairchild and Narayanan employed the empirical mode decomposition (EMD) algorithm to decompose the radar signal into a set of intrinsic mode functions (IMFs) and computed the IMF energies as features [12]. In the frequency domain, Bilik and Khomchuk applied speech and audio processing techniques to extract three types of features: real cepstrum, linear predictive coding coefficients, and Mel-frequency cepstrum coefficients (MFCC) [13]. Molchanov *et al.* computed discrete cosine transform features from μ-D signals for target classification [14]. In the joint T-F domain, Kim and Ling defined six features from the spectrogram of a moving person: 1) the torso Doppler frequency, 2) the total bandwidth of the Doppler signal, 3) the offset of the total Doppler, 4) the bandwidth without micro-Doppler, 5) the normalized standard deviation of the Doppler signal strength, and 6) the period of the limb motion [15]. Orovic *et al.* applied the Hermite S-method to convert the radar signal into a T-F representation and developed an envelope detection method to capture the evolution of the arm swing [16]. Mobasseri and Amin extracted features from the spectrogram by applying principal component analysis (PCA) [17]. Li *et al.* employed the matrix-based PCA technique on spectrogram to classify different arm motions of a walking person [18]. Bjorklund *et al.* computed the cadence velocity diagram (CVD) and extracted the cadence frequencies and velocity profiles as μ-D features [19]. Tivive *et al.* proposed a machine learning method to extract μ-D features from the spectrogram [20]. Groot *et al.* proposed to use particle filters to differentiate walking from running motions and estimate the person's height from the spectrogram [21]. Given the importance of converting the μ-D signal into a joint T-F representation for classification, the next section describes three T-F analysis methods.

C. Time-frequency representations

We review three main T-F analysis representations for depicting μ-D radar signature: the short-time Fourier transform (STFT), the pseudo Wigner-Ville distribution (PWD), and the S-method. Consider a time-domain signal $x(t)$. The STFT of the signal is given by

$$X(t,\omega) = \int_{-\infty}^{\infty} x(t+\tau)w(\tau)e^{-j\omega\tau}\ d\tau, \tag{4}$$

where $w(\tau)$ is a time window. The spectrogram (SP) is the squared magnitude of the STFT:

$$X_{\mathrm{SP}}(t,\omega) = |X(t,\omega)|^2. \tag{5}$$

The STFT has a simple implementation, but it generally provides a low resolution. In comparison, the PWD produces a high resolution T-F representation, and is related to the STFT as

$$X_{\mathrm{PWD}}(t,\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} X(t,\omega+\theta)X^*(t,\omega-\theta)\ d\theta. \tag{6}$$

However, the PWD produces cross-terms which may cause difficulties in interpreting the T-F distribution.

By contrast, the S-method achieves similar auto-term concentration, but it does not suffer from the cross-terms. The T-F distribution of the S-method (SM) is given by

$$X_{\mathrm{SM}}(t,\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} Q(\theta)X(t,\omega+\theta)X^*(t,\omega-\theta)\ d\theta, \tag{7}$$

where $Q(\theta)$ is a finite frequency domain window. The S-method behaves as the STFT when $Q(\theta) = \pi\delta(\theta)$, and as the PWD when $Q(\theta) = 1$. The discrete form of the S-method can be written as follows:

$$X_{\mathrm{SM}}(n,k) = \sum_{i=-N/2}^{N/2} Q(i)\ X(n,k+i)\ X^*(n,k-i), \tag{8}$$

where n is the discrete time index, k is the discrete frequency index, and N is the number of frequency samples. For a rectangular window, i.e., $Q(i) = 1$ for $|i| \leq J$ and zero otherwise, the S-method with J terms can be written as

$$X_{\mathrm{SM}}(n,k) = \sum_{i=-J}^{J} X(n,k+i)\ X^*(n,k-i). \tag{9}$$

The parameter J is usually set to a small value ($J \in [3,10]$) since most of the auto-term energy is located around the maximum value of the auto-term [23].

As an illustration, the above three T-F analysis techniques are applied to a μ-D signal; the respective T-F representations are shown in Fig. 1. The Doppler frequency shift induced by the torso is around 200 Hz. The main peak represents the leg swing. Although the PWD generates a high resolution T-F representation, shown in Fig. 1(b), it is hard to discern the μ-D modulations, not to mention the cross-terms. Therefore, in this paper only the STFT and S-method will be investigated to generate the joint T-F representation.

III. Micro-Doppler feature extraction

In this section, we describe the proposed approach for signal classification using μ-D radar signatures. Figure 2 shows the main steps of the feature extraction, followed by the classification step. After the T-F analysis, the μ-D radar signature is converted into a low-dimensional feature vector by using log-Gabor filters and a dimensionality reduction technique.

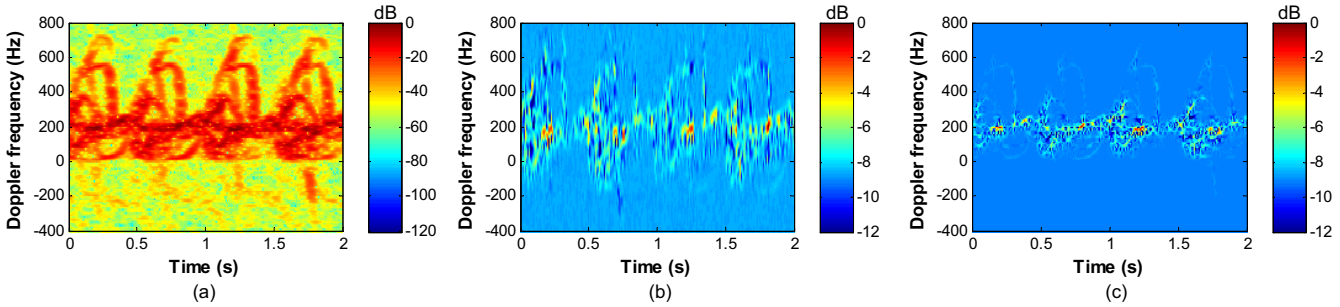


Fig. 1. Time-frequency representations (logarithmic scale) for a μ-D signal of a person walking towards the radar with both arms swinging: (a) STFT, (b) PWD, and (c) S-method.

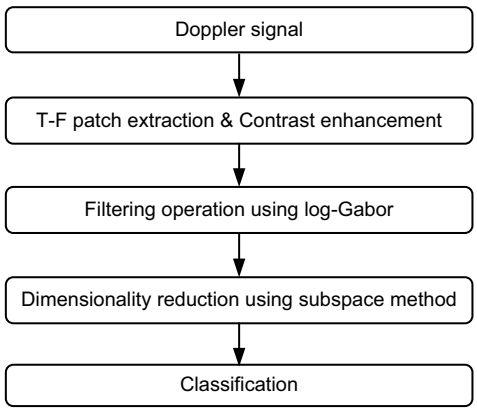


Fig. 2. A schematic diagram of the proposed μ-D feature extraction method.

A. Time-frequency patch and contrast enhancement

When a person is walking, the μ-D modulations induced by the arm and leg motions occur mainly around the human torso frequency. Therefore, we extract local patches centered on the torso frequency instead of the entire T-F representation. This strategy improves invariance to translational speed of the target. The torso frequency can be easily determined by locating the main peak in the frequency profile. From the joint T-F representation $X(n,k)$, the frequency profile can expressed as

$$F_p(k) = \sum_{n=1}^{M} X(n,k), \tag{10}$$

where M is the number of time samples. The frequency profile is passed to a median filter and then normalized by dividing by the maximum value to obtain $\widetilde{F}_p(k)$. The location of the torso frequency is computed as

$$k_t = \arg\max_k \widetilde{F}_p(k). \tag{11}$$

The patch height is determined relative to the height of the main peak, whereas the patch width is determined based on the task at hand; here, it is chosen so as to maximize the classification rate on individual patches. Assuming the person is inducing positive micro-Doppler modulations, the height of the main peak is estimated by finding the smallest frequency index k_o that satisfies the following condition:

$$\frac{\sum_{k=k_t}^{k_o} \widetilde{F}_p(k)}{\sum_{k=k_t}^{N/2} \widetilde{F}_p(k)} \geq \eta. \tag{12}$$

where N is the number of frequency samples and η is a fixed threshold ($0 < \eta < 1$). The threshold η is chosen based on the T-F representation noise level. When the T-F representation is noisy, the threshold is set to a low value and vice versa. In this paper, it is set to 0.99. The vertical span of the patch is given by the frequency interval $[k_t - k_o, k_t + k_o]$. Since different individuals swing their legs at different speeds, the patch height is fixed to a predefined value N_y through down-sampling or up-sampling operation on the columns of the patch.

The patch width is determined by the length of the input signal and the length and the step of the window in the T-F representation. Furthermore, the patch is aligned with respect to the main peak. Let T_i denote the time of the i-th main peak, ΔT denote the time duration between two consecutive main peaks, and N_x be the length of the input signal in samples. The horizontal span of the patch is given by the time interval $[T_i + \Delta T/2, T_i + \Delta T/2 + N_x - 1]$.

Next, contrast enhancement is performed on the patch. We use the Naka-Rushton equation [24] because it not only enhances the weak μ-D amplitudes but also suppresses the small amplitudes, which usually represent noise. Furthermore, it constrains the input to the range $[0, 1)$, facilitating the subsequent feature extraction. Let $W(i,j)$ denote the patch containing the absolute value of the T-F representation, where $i \in [1, N_y]$ and $j \in [1, N_x]$. The contrast-enhanced patch is computed as

$$\widetilde{W}(i,j) = \frac{W(i,j)^r}{W(i,j)^r + \mu^r}, \tag{13}$$

where μ is the mean value of the patch. In this paper, the parameter r is set to 1 ($r = 1$). Figure 3 presents the patches extracted from the spectrogram that are contrast-enhanced by the logarithmic scale and Naka-Rushton equation. The difference between Figs. 3(b) and (c) shows that the Naka-Rushton equation is better than the logarithmic scale at enhancing weak micro-Doppler modulations and suppressing clutter.

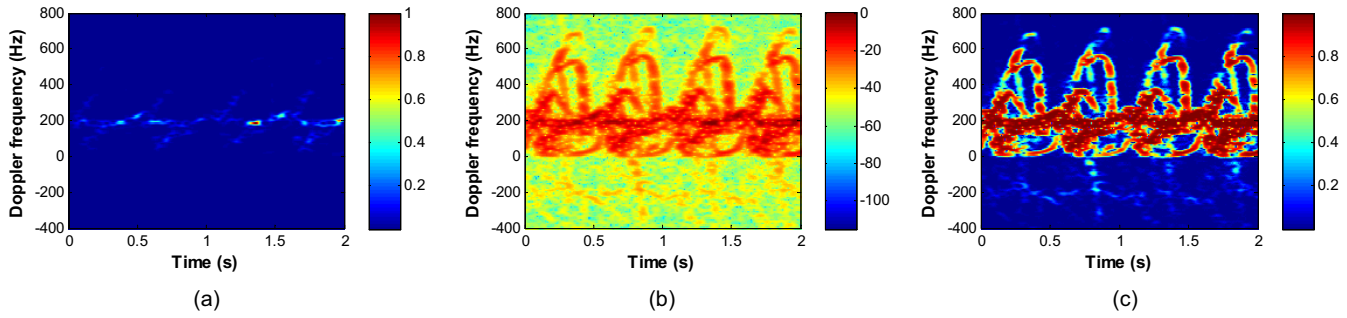


Fig. 3. T-F patch extracted from the spectrogram before and after contrast enhancement: (a) input T-F patch, (b) T-F patch enhanced by the logarithmic scaling, and (c) T-F patch enhanced by the Naka-Rushton equation.

B. Log-Gabor filtering

Log-Gabor filters, which were proposed by Field [25], are used to analyze the pre-processed T-F patch. A log-Gabor filter has similar shape to a Gabor filter on the logarithmic frequency scale with an extended tail in the high-frequency region. Due to the singularity at the origin, a Log-Gabor filter is designed in the frequency domain and is computed as

$$G_{k,l}(f,\theta) = \exp\left\{-\frac{[\ln(f/f_k)]^2}{2[\ln(\beta)]^2}\right\}\exp\left\{-\frac{(\theta-\theta_l)^2}{2\sigma_\theta^2}\right\}, \quad k = 1,\ldots,N_f, \quad l = 1,\ldots,N_\theta, \tag{14}$$

where f_k is the k-th center frequency of the filter, θ_l is the l-th orientation angle, N_f is the number of scales, N_θ is the number of orientations, β is the bandwidth of the filter, and σ_θ is the angular bandwidth. Here, the bandwidth of the filter is set to two octaves ($\beta = 0.55$), and the angular bandwidth σ_θ is set to 1.5 for even spectrum coverage. Figure 4 shows examples of the log-Gabor filter.

The filtering operation is performed in the frequency domain. Let $\widetilde{W}_F$ denote the 2D Fourier transform of the normalized patch $\widetilde{W}$, see (13). The output of the (k,l)-th log-Gabor filter is computed as

$$Z_{k,l} = |\mathrm{IFFT2}(\widetilde{W}_F\ G_{k,l})|, \tag{15}$$

where IFFT2 denotes the 2D inverse Fast Fourier transform. The output map is further normalized as

$$\widetilde{Z}_{k,l}(i,j) = \frac{Z_{k,l}(i,j)}{\sum_{l=1}^{N_\theta} Z_{k,l}(i,j)}. \tag{16}$$

This normalization step provides some degree of intensity invariance. Next, each output map is partitioned into $R = d_1 \times d_2$ non-overlapping sub-regions, and the means of all the sub-regions are concatenated to form a mean vector $\boldsymbol{\mu}_{k,l} = [\mu_1,\ldots,\mu_R]^T$. Then, the mean values of the L output maps (where $L = N_f\ N_\theta$) are arranged into a matrix $A \in \mathbb{R}^{R\times L}$:

$$A = [\boldsymbol{\mu}_{1,1},\ldots,\boldsymbol{\mu}_{N_f,N_\theta}]. \tag{17}$$

Finally, the feature matrix A is normalized to the range $[0,1]$ before applying a dimensionality reduction technique to generate a compact feature vector.

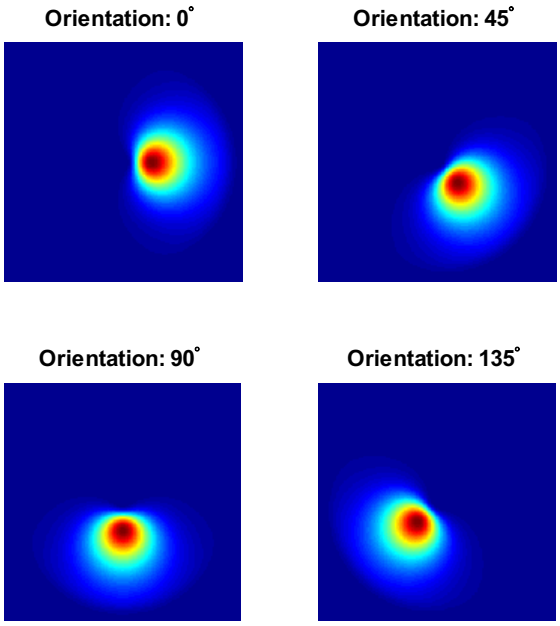


Fig. 4. Examples of Log-Gabor filters with 4 orientations at a normalized center frequency of 0.12 in the 2D frequency domain.

C. Dimensionality reduction

Two different matrix-based subspace techniques are considered to reduce the size of the feature matrix A: two-directional two-dimensional principal component analysis (2D)^{2}PCA [26] and two-directional two-dimensional linear discriminant analysis (2D)^{2}LDA [27]. In comparison to standard PCA and LDA, the matrix-based methods do not require a matrix-to-vector conversion to compute the image covariance matrices, thereby reducing the computational cost significantly. Unlike LDA, (2D)^{2}LDA does not suffer from the singularity problem for small training sets. Both (2D)^{2}PCA and (2D)^{2}LDA will be investigated in the proposed feature extraction. A brief description of each method is given in the following subsections.

1) (2D)^{2}PCA: This technique generates two projection matrices to reduce the number of rows and columns of an image simultaneously. The compressed map $D \in \mathbb{R}^{m_r\times m_c}$, $(m_r \le R,\ \ m_c \le L)$, can be rewritten as

$$D = \Phi_r^T A \Phi_c, \tag{18}$$

where Φ_r and Φ_c are the projection matrices with orthonormal components. Then, the columns of matrix D are concatenated to form a feature vector for classification. Let $Y = A\Phi_c$. The projection matrix Φ_c can be determined by maximizing the following criterion [26]:

$$\begin{aligned}\Omega(\Phi_c) &= \mathrm{trace}\{\mathrm{E}[(Y-\mathrm{E}(Y))(Y-\mathrm{E}(Y))^T]\} \\ &= \mathrm{trace}\{\mathrm{E}[(A\Phi_c-\mathrm{E}(A\Phi_c))(A\Phi_c-\mathrm{E}(A\Phi_c))^T]\} \\ &= \mathrm{trace}\{\Phi_c^T\mathrm{E}[(A-\mathrm{E}(A))^T(A-\mathrm{E}(A))]\Phi_c\}. \quad (19)\end{aligned}$$

The image covariance matrix can be defined as $G_c = \mathrm{E}[\{A-\mathrm{E}(A)\}^T\{A-\mathrm{E}(A)\}]$, which is an L-by-L nonnegative definite matrix. Suppose that the training set comprises P patches $\{A_1, \cdots, A_P\}$. The image covariance matrix G_c can be computed as

$$G_c = \frac{1}{P}\sum_{i=1}^{P}(A_i-\bar{A})^T(A_i-\bar{A}), \quad (20)$$

where $\bar{A}$ is the global mean given by

$$\bar{A} = \frac{1}{P}\sum_{i=1}^{P}A_i. \quad (21)$$

The criterion in (19) can be expressed as

$$\Omega(\Phi_c) = \mathrm{trace}(\Phi_c^T G_c \Phi_c). \quad (22)$$

The criterion function $\Omega(\Phi_c)$ is maximized when Φ_c is composed of the m_c most dominant eigenvectors of G_c: $\Phi_c = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{m_c}]$. The number of eigenvectors m_c is determined using the following condition:

$$\frac{\sum_{i=1}^{m_c}\lambda_i}{\sum_{i=1}^{L}\lambda_i} \geq \gamma, \quad (23)$$

where λ_i denotes the i-th eigenvalue and γ is a fixed threshold.

Similarly, the other projection matrix Φ_r contains the m_r most dominant eigenvectors of the image covariance matrix G_r, which is given by

$$G_r = \frac{1}{P}\sum_{i=1}^{P}(A_i-\bar{A})(A_i-\bar{A})^T. \quad (24)$$

The number of eigenvectors m_r in the projection matrix Φ_r is estimated using a similar condition to (23).

2) (2D)^{2}LDA: The LDA technique considers the class information when forming the projection matrices. Its principle is to find a linear transformation that maximizes the between-class scatter and minimizes the within-class scatter of the training set. For (2D)^{2}LDA, the optimization of the between-class and within-class scatters is performed in both row and column directions simultaneous as

$$D = \Psi_r^T A \Psi_c, \quad (25)$$

where Ψ_r and Ψ_c are projection matrices. Let C be the number of classes and N_i be the number of training samples in the i-th class. Let P be the total number of training samples, $P = \sum_{i=1}^{C}N_i$. Let A_p^i denote the p-th sample and $\bar{A}^i$ be the mean of all samples of the i-th class, $\bar{A}^i = \frac{1}{N_i}\sum_{j=1}^{N_i}A_j^i$. The between-class and within-class scatter matrices for the row direction are given respectively by

$$G_{\mathrm{bc}} = \frac{1}{P}\sum_{i=1}^{C}N_i(\bar{A}^i-\bar{A})^T(\bar{A}^i-\bar{A}), \quad (26)$$

and

$$G_{\mathrm{wc}} = \frac{1}{P}\sum_{i=1}^{C}\sum_{j=1}^{N_i}(A_j^i-\bar{A}^i)^T(A_j^i-\bar{A}^i), \quad (27)$$

where $\bar{A}$ is the global mean given in (21). Similarly, the between-class and within-class scatter matrices for designing the projection matrix Ψ_r are, respectively,

$$G_{\mathrm{br}} = \frac{1}{P}\sum_{i=1}^{C}N_i(\bar{A}^i-\bar{A})(\bar{A}^i-\bar{A})^T, \quad (28)$$

and

$$G_{\mathrm{wr}} = \frac{1}{P}\sum_{i=1}^{C}\sum_{j=1}^{N_i}(A_j^i-\bar{A}^i)(A_j^i-\bar{A}^i)^T. \quad (29)$$

The projection matrices Ψ_r and Ψ_c are obtained by maximizing the following Fisher criteria:

$$J(\Psi_r) = \mathrm{trace}\left(\frac{\Psi_r^T\, G_{\mathrm{br}}\, \Psi_r}{\Psi_r\, G_{\mathrm{wr}}\, \Psi_r}\right) \quad (30)$$

and

$$J(\Psi_c) = \mathrm{trace}\left(\frac{\Psi_c^T\, G_{\mathrm{bc}}\, \Psi_c}{\Psi_c\, G_{\mathrm{wc}}\, \Psi_c}\right). \quad (31)$$

The discrimination vectors in the projection matrices Ψ_r and Ψ_c are the eigenvectors of $G_{\mathrm{wr}}^{-1}G_{\mathrm{br}}$ and $G_{\mathrm{wc}}^{-1}G_{\mathrm{bc}}$, and the number of eigenvectors in these matrices is estimated using a similar condition to (23).

D. Classification stage

Support vector machines (SVMs) are used as a classifier because they possess good generalization capability. The key concept of SVMs is to determine an optimal hyperplane that separates two different classes. A hyperplane is optimal if it maximizes the margin between the two classes, where the margin is defined as the distance between the hyperplane and a closest training vector. Consider a training set $\{\boldsymbol{x}_i,\ y_i\}_{i=1}^{P}$, where $\boldsymbol{x}_i \in \mathbb{R}^n$ is the i-th input vector and $y_i \in \{1, -1\}$ is the corresponding class label. Training an SVM classifier involves solving the following optimization problem:

$$\begin{aligned}&\min_{\boldsymbol{w},b,\xi}\left\{\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{P}\xi_i\right\}, \quad \text{subject to} \\ &\quad y_i(\boldsymbol{w}^T\phi(\boldsymbol{x}_i)+b) \geq 1-\xi_i,\ \xi_i \geq 0,\ i = 1, \cdots, P. \quad (32)\end{aligned}$$

Here, $\phi(\boldsymbol{x}_i)$ maps an input vector $\boldsymbol{x}_i$ into a higher-dimensional space so that the classification problem becomes simpler (i.e. linearly separable). Parameter C is a positive regularization constant to control the trade-off between the margin and the

misclassification rate. A dual optimization problem is given by

$$\max_{\boldsymbol{\alpha}} \left\{ \sum_{i=1}^{P} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{P} \alpha_i \alpha_j y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \right\} \text{ subject to}$$

$$0 \leq \alpha_i \leq C, \quad \sum_{i=1}^{P} \alpha_i y_i = 0, \quad (33)$$

where $K(\boldsymbol{x}_i, \boldsymbol{x}_j) \equiv \phi(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j)$ is the kernel function. Vector $\boldsymbol{w}$ in the primal problem is related to variable α_i, $i \in [1, P]$ in the dual problem as

$$\boldsymbol{w} = \sum_{i=1}^{P} y_i \alpha_i \phi(\boldsymbol{x}_i). \quad (34)$$

Once α_i has been calculated, the decision function is given by

$$y_p(\boldsymbol{x}) = \text{sgn} \left\{ \sum_{i=1}^{P} y_i \alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b \right\} = \text{sgn} \left\{ \boldsymbol{w}^T \phi(\boldsymbol{x}) + b \right\}. \quad (35)$$

where

$$b = \frac{1}{N_s} \sum_{i=1}^{N_s} \boldsymbol{w}^T \boldsymbol{x}_i - y_i, \quad (36)$$

and N_s is the number of support vectors. In this paper, the linear kernel, $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T \boldsymbol{x}_j$, is used and parameter C is obtained from the training set via a cross-validation procedure.

IV. Experimental results

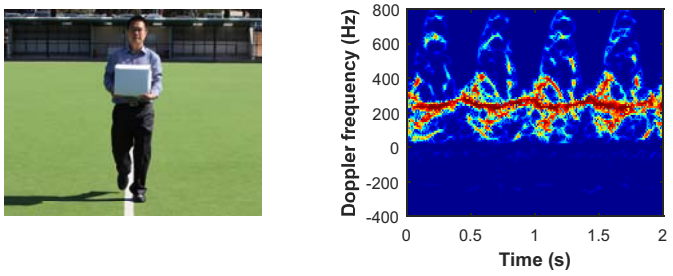
In this section, we first describe the experimental setup, and then investigate the effects of different steps in feature extraction. Finally, we compare the classification performance of the proposed method with other feature extraction methods.

A. Experimental setup

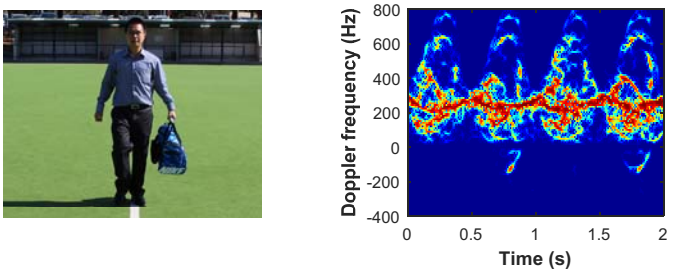
A 24-GHz frequency modulated continuous wave radar was used for data acquisition. The beam width of the radar antenna is 7° horizontal and 25° vertical. The radar was positioned at a height of 0.7 m from the ground. The Doppler data were acquired in two environments (outdoor and indoor) from 18 subjects (7 females and 11 males). Each subject performed three motion types: (i) walking with both arms swinging; (ii) walking with one arm swinging; (iii) walking with no arm swinging. Each subject walked towards the radar at azimuth angles of 0° and 3.5°, and repeated each motion type three times. The radar signal was recorded at a sampling rate of 7.812 kHz. Overall, 234 Doppler signals of length 10 s were recorded. Figure 5(a) shows images of a subject walking with different arm motions and their respective T-F representations.

B. Analysis of feature extraction steps

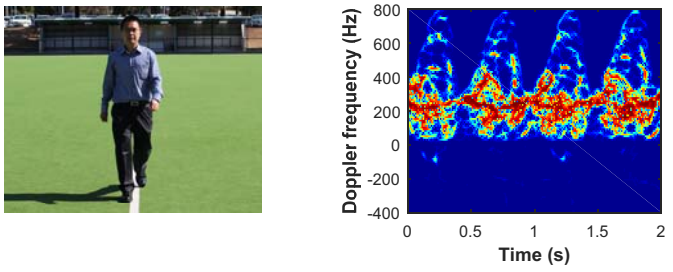
In the proposed method, several adjustable parameters need to be considered, e.g., the window length in the T-F representation, the number of scales and orientations of the log-Gabor filters, and the dimensionality reduction settings. A six-fold cross-validation is performed to investigate the effect of these parameters on the classification rate (CR). In each validation fold, five subsets are used for training, and the remaining subset is used for testing. This is repeated six times for different choices of the test subset. The final CR is computed as the percentage of correctly classified samples, which are aggregated across all the validation folds. Initially, a signal length of 1 s is used to determine the adjustable parameters of the proposed method.



(a) No-arm swing



(b) One-arm swing



(c) Two-arm swing

Fig. 5. Images of a subject walking with different arm motions towards the radar at an azimuth angle of 0° and their T-F representations produced by the S-method and enhanced by the Naka-Rushton equation.

1) Time-frequency representation: In the STFT and S-method, the choice of the window length can affect the performance significantly. In the first experiment, the number of DFT (Discrete Fourier transform) points and the overlap between consecutive windows are set to 2048 and 90%, respectively. Then a variable window length is used to form the T-F representation. When the window length is smaller than 2048 samples, the signal is padded with zeros prior to T-F analysis. The extracted patch is fixed to a height

of 128 pixels ($N_y = 128$) and contrast-enhanced by the Naka-Rushton equation. Initially, a set of 32 log-Gabor filters (4 scales and 8 orientations) are used. To compute the mean values of the filtered map as features, the height d_1 of the non-overlapping sub-region is set to 8, whereas the width d_2 is chosen to have a time duration of 125 ms. Therefore, for an input signal length of 1 s, d_2 is equal to 8. In the following two experiments, the feature vector is directly classified by the SVMs without dimensionality reduction. Figure 6 depicts the CR as a function of the window length. Both T-F analysis methods reach a peak CR with a window length of 139.3 ms. Further increasing the window length reduces the CR; this is due to poor time resolution. Both T-F representations will be used in the succeeding experiments.

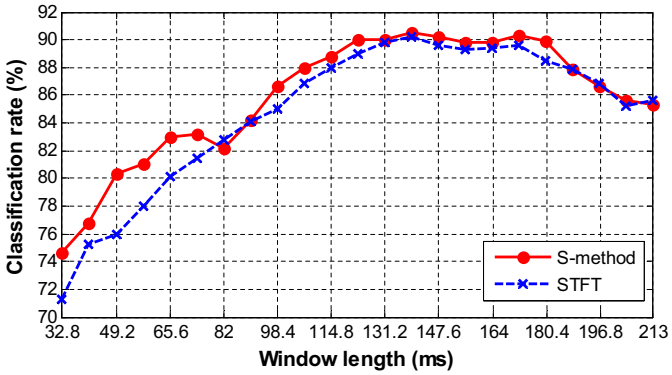
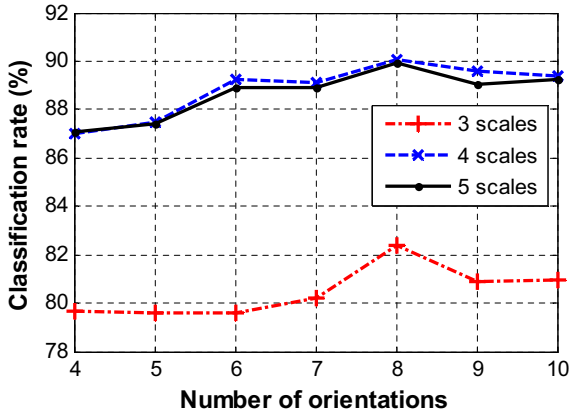


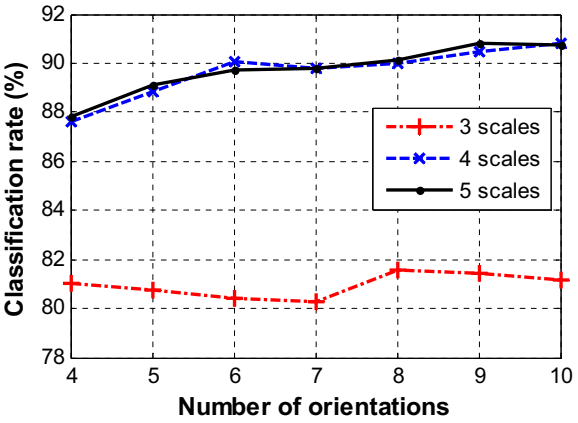
Fig. 6. Classification rates of the STFT and S-method as a function of the window length.

2) Log-Gabor filtering: Discriminative features are extracted by convolving the pre-processed patch with log-Gabor filters of different scales and orientations. To design a set of log-Gabor filters, we vary the number of scales from 3 to 5, and the number of orientations from 4 to 10. Figures 7(a) and (b) show the CR as a function of the number of scales and orientations for the STFT and S-method, respectively. The CR improves markedly when the number of scales increases from 3 to 4; it reaches a steady state when the number of scales reaches 4. With 4 scales, the STFT obtains a peak CR of 90.0% with 8 orientations, whereas the S-method achieves a CR of 90.4% with 9 orientations. Therefore, we use a set of 32 log-Gabor filters (i.e., 4 scales and 8 orientations) for the STFT, and 36 log-Gabor filters (i.e., 4 scales and 9 orientations) for the S-method.

3) Dimensionality reduction: Two subspace methods, $(2D)^2$PCA and $(2D)^2$LDA, are evaluated for feature compression. Figure 8 presents the CRs as a function of the number of features obtained by dimensionality reduction. When the number of features is small, $(2D)^2$LDA performs better than $(2D)^2$PCA. However, increasing the number of features improves the CR of $(2D)^2$PCA to the same level as $(2D)^2$LDA. For this classification problem, $(2D)^2$PCA achieves the highest CR; therefore, it will be used to reduce the number of features in the following experiments.



(a)



(b)

Fig. 7. Classification rates of the log-Gabor filter as a function of the number of scales and orientations, using (a) STFT and (b) S-method as T-F analysis.

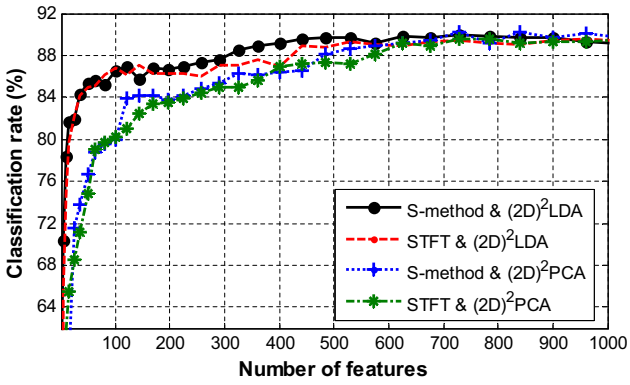


Fig. 8. Classification rates of different dimensionality reduction techniques as a function of the number of features.

4) Input signal length: The input signal length is an important factor in the classification of μ-D radar signature. A too short signal will not contain adequate cycles of the arm/leg swings to differentiate between the three human motions, whereas a too long signal will lead to data redundancy. To

investigate the effects of input signal length, we vary the signal length from 0.5 to 3.0 s with a step of 0.5 s. For each input signal length, a new SVM classifier is trained. Figure 9 presents the CRs as a function of input signal length. The CR improves when the input signal length increases from 0.5 s to 3 s and reaches a plateau for a signal length of 1.5 s for S-method and 2 s for STFT. Based on these results, we use a signal length of 2 s in the following experiments.

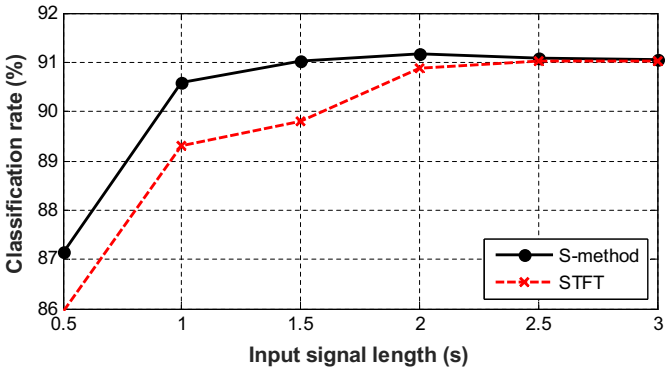


Fig. 9. Classification rates of the proposed method for different input signal lengths.

5) Effects of azimuth angle and clutter: So far in Section IV-B, the proposed method has been tested on radar data for subjects walking towards the radar at an azimuth angle of $0°$ in an outdoor environment. We also evaluate the proposed method in different configurations, especially in the presence of clutter. Here, the clutter or noise in the received signals include multi-path propagations and reflections from other stationary targets. Four data sets are used. Data sets *outdoor-0.0* and *outdoor-3.5* are acquired from subjects walking in an outdoor environment, at an azimuth angle of $0°$ and $3.5°$, respectively. Data sets *indoor-0.0* and *indoor-3.5* are acquired from subjects walking along a corridor inside a building, at azimuth angles of $0°$ and $3.5°$, respectively. Table I lists the CRs of the proposed method on different data sets using a two-fold cross-validation. The proposed method has a CR of 91.6% for the *indoor-0.0* and 91.4% for the *outdoor-0.0*. For subjects walking obliquely towards the radar at an azimuth angle of $\pm 3.5°$, the CR reduces by 0.4% for the *outdoor-3.5* and 2.5% for the *indoor-3.5*.

TABLE I
CLASSIFICATION RATES OF THE FEATURE EXTRACTION METHOD TESTED ON RADAR SIGNALS COLLECTED IN DIFFERENT ENVIRONMENTS.

Data acquisition environment	*Outdoor*		*Indoor*	
Azimuth angle	$0°$	$3.5°$	$0°$	$3.5°$
Proposed feature extraction	91.6%	91.2%	91.4%	88.9%

In summary, based on the experiments presented in Subsection IV-B, we select the following configurations for the proposed method: (i) S-method for T-F analysis; (ii) the Naka-Rushton equation for preprocessing the patch; (iii) a set of 9 orientations and 4 scales log-Gabor filters for generating the feature map; (iv) $(2D)^2$PCA for dimensionality reduction; (v) input signal length of 2 s; and (vi) SVM classifier.

C. Comparison of feature extraction methods

A six-fold cross-validation is employed to compare different feature extraction methods. The indoor radar dataset (*indoor-0.0*) is used to generate the training and test sets. The dataset is partitioned so that no human subject appears simultaneously in the training set and the test set. The radar signals are divided into segments of 2 s. Each segment is aligned so that it comprises two full gait cycles.

For comparison, four different feature extraction methods are also tested: 1) MFCC, 2) CVD, 3) EMD, and 4) Gabor filters. In the MFCC based method, 40 triangular bandpass filters are used to produce 64 mel-scale cepstral coefficients. The analysis window length and the overlap between successive windows are set to 0.5 s and 0.01 s, respectively. In the CVD based method, the first three harmonic frequencies and the velocity profile at each harmonic frequency are used to form a feature vector. Then, the feature vector is compressed using the standard PCA technique. In the EMD method, the energies of the IMFs are used as features. The Gabor filter based method employs the same number of filters and applies the same step of converting the filtered outputs into a feature vector. All the feature extraction methods employ SVMs as classifier. Furthermore, they are implemented using MATLAB software and executed on a PC with a 2.9 GHz i7-CPU.

TABLE II
CLASSIFICATION RATES OBTAINED USING SIX-FOLD CROSS-VALIDATION OF DIFFERENT FEATURE EXTRACTION METHODS.

Method	*Features*	*CR* $\pm$ *std* (%)	*Extraction time* (s)
Proposed method	3277	91.3 ± 6.9	0.2405
Gabor filters	2408	79.9 ± 7.5	0.2105
MFCC	40	72.7 ± 7.2	0.0272
CVD	1638	62.3 ± 5.1	0.0210
EMD	16	41.6 ± 3.2	8.4400

Table II presents the CRs and the processing times of different feature extraction methods. The proposed method achieves the highest CR of 91.3%, followed by the Gabor filter based method with a CR of 79.9%. The EMD method gives the lowest CR, which indicates that using only the energy of the IMF is not sufficient to discriminate the subtle differences of the arm swings. In terms of processing time, the proposed method takes about 0.2405 s on average to extract a feature vector, and its most time consuming stage is the filtering operation. We should note that no code optimization was used in the implementation of the proposed method; the processing time can be reduced by optimizing the implementation or using a different programming language. Compared to the proposed method, the Gabor filter based method is about 1.14 times faster, but it yields a CR of 11.4% lower. The MFCC based method is about 8.84 times faster than the proposed method, but its CR is 18.6% lower. The CVD based method is about 11.5 times faster than the proposed method, but its CR is 29.0% lower. The EMD method is significantly slower and less accurate compared to the proposed method. The EMD high processing time is due to the iterative technique for extracting the IMFs.

The results shown in Table II are obtained using different number of features for each method. To compare the classification performances of the different methods using the same number of features, Fig. 10 illustrates the CRs of the feature extraction methods as a function of the number of input features. Clearly, the CRs of the proposed method are still higher than those of the other methods, when using the same number of features.

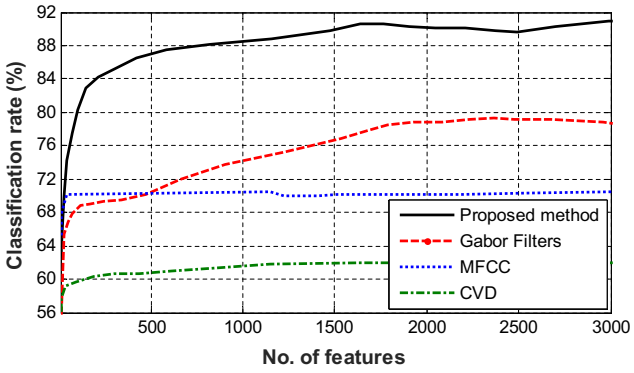


Fig. 10. Classification rates of the proposed method and other feature extraction methods as a function of the number of features.

V. Conclusion

This paper presents a 2D feature extraction method for classifying μ-D radar signature of human motions. The radar return is transformed into a T-F representation using the S-method. Instead of processing the entire T-F representation, a small patch centered on the torso frequency is automatically extracted to improve stability against the target speed. The T-F patch is then contrast-normalized to highlight the weak μ-D modulations. log-Gabor filters are employed to detect salient features at multiple scales and orientations, and $(2D)^2$PCA is applied for dimensionality reduction. The proposed method is validated using μ-D radar signals obtained from human subjects walking with various arm motions. Experimental results show that it achieves promising results in classifying different human motions.

References

[1] Bell, M. R. and Grubbs, R. A., "JEM modeling and measurement for radar target identification," *IEEE Trans. Aerospace and Electronic Systems*, vol. 29, no. 1, pp. 73–87, 1993.

[2] Chen, V. C., *The micro-Doppler effect in radar*, Boston: Artech house, 2011.

[3] Thayaparan, T., Abrol S., Riseborough, E., Stankovic, L., Lamothe, D., and Duff, G., "Analysis of radar micro-Doppler signatures from experimental helicopter and human data," *IET Radar, Sonar, & Navigation*, vol. 1, no. 4, pp. 289–299, 2004.

[4] Tahmoush, D. and Silvious, J., "Stride rate in radar micro-Doppler images," *IEEE Int. Conf. Systems, Man and Cybernetics*, pp. 4218–4223, 2009.

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

[6] Molchanov, P., Totsky, A., Astola, J., Egiazarian, K., Leshchenko, S., and Rosa-Zurera, M., "Aerial target classification by micro-Doppler signatures and bicoherence-based features," *European Radar Conf.*, pp. 214–217, 2012.

[7] Li, Y., Du, L., and Liu, H., "Hierarchical classification of moving vehicles based on empirical mode decomposition of micro-Doppler signatures," *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 3001–3013, 2013.

[8] Bilik, I., Tabrikian, J., and Cohen, A., "GMM-based target classification for ground surveillance Doppler radar," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 267–278, 2006.

[9] Ghaleb, A., Vignaud, L., and Nicolas, J. M., "Micro-Doppler analysis of wheels and pedestrians in ISAR imaging," *IET Signal Processing*, vol. 2, no. 3, pp. 301–311, 2008.

[10] Yessad, D., Amrouche, A., Debyeche, M., and Djeddou, M., "Micro-Doppler classification for ground surveillance radar using speech recognition tools," *Lecture Notes in Computer Science*, vol. 7042, pp. 280-287, 2011.

[11] Bilik, I. and Tabrikian, J., "Radar target classification using Doppler signatures of human locomotion models," *IEEE Trans. Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1510–1522, 2007.

[12] Fairchild, P. D. and Narayanan, M. R., "Classification and modeling of human activities using empirical mode decomposition with S-band and millimeter-wave micro-Doppler radars," *SPIE Radar Sensor Technology XVI*, vol. 8361, Article no. 83610X, pp. 1–15, 2012.

[13] Bilik, I. and Khomchuk, P., "Minimum divergence approaches for robust classification of ground moving targets," *IEEE Trans. Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 581–603, 2012.

[14] Molchanov, P., Astola, J., Egiazarian, K., and Totsky, A., "Ground moving target classification by using DCT coefficients extracted from micro-Doppler radar signatures and artificial neuron network," *Microwaves, Radar and Remote Sensing Symposium*, pp. 173–176, 2011.

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

[16] Orovic, I., Stankovic, S., and Amin, M., "A new approach for classification of human gait based on time-frequency feature representations," *Signal Processing*, vol. 91, no. 6, pp. 1448–1456, 2011.

[17] Mobasseri, B. G. and Amin, M. G., "A time-frequency classifier for human gait recognition," *SPIE, Optics and Photonics in Global Homeland Security V and Biometric Technology for Human Identification VI*, vol. 7306, Article no. 730628, pp. 1–9, 2009.

[18] Li, J., Phung, S. L., Tivive, F. H. C., and Bouzerdoum, A., "Automatic classification of human motions using Doppler radar," *Proc. Int. Joint Conf. on Neural Networks*, pp. 1–6, 2012.

[19] Bjorklund, S., Johansson, T., and Petersson, H., "Evaluation of a micro-Doppler classification method on mm-wave data," *Proc. IEEE Radar Conf.*, pp. 0934–0939, 2012.

[20] Tivive, F. H. C., Bouzerdoum, A., and Amin, M. G., "A human gait classification method based on radar Doppler spectrograms," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. Article ID 389716, pp. 1–12, 2010.

[21] Groot, S., Harmanny, R., Driessen, H., and Yarovoy, A., "Human motion classification using a particle filter approach: Multiple model particle filtering applied to micro-Doppler spectrum," *Int. Journal of Microwave and Wireless Technologies*, vol. 5, no. 3, pp. 391–399, 2013.

[22] Ram, S. S. amd Lin, H., "Simulation of human micro-Dopplers using computer animation data," *Proc. IEEE Radar Conf.*, pp. 1957–1966, 2008.

[23] Thayaparan, T., Stankovic, L., Djurovic, I., Penamati, S., and Venkataramaniah, K., "Intelligent target recognition using micro-Doppler radar signatures," *Proc. SPIE, Radar Sensor Technology XIII*, vol. 7308, Article no. 730817, pp. 1–11, 2009.

[24] Naka, K. I. and Rushton, W. A. H., "S-potentials from colour units in the retina of fish (cyprinidae)," *The Journal of Physiology*, vol. 185, pp. 536–555, 1966.

[25] Field, D. J., "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[26] Zhang, D. and Zhou, Z.-H., "(2D)2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition," *Journal Neurocomputing*, vol. 69, no. 1-3, pp. 224–231, 2005.

[27] Noushath, S., Kumar, G. H., and Shivakumara, P., "2D2LDA: An efficient approach for face recognition," *Pattern Recognition*, vol. 39, no. 7, pp. 1396–1400, 2006.