Micro-Doppler radar classification of humans and animals in an operational environment

W.D Van Eeden[a,*], J.P. de Villiers[a,b,**], R.J. Berndt[b], W.A.J. Nel[b], E. Blasch[c]

[a]*University of Pretoria, Pretoria, South Africa,*
[b]*Defence, Peace, Safety and Security, Council of Scientific and Industrial Research, Pretoria, South Africa*
[c]*Air Force Research Lab, Rome, NY, USA*

Abstract

A combined Gaussian mixture model and hidden Markov model (HMM) is developed to distinguish between slow moving animal and human targets using mel-cepstrum coefficients. This method is compared to the state-of-the-art in current micro-Doppler classification and an improvement in performance is demonstrated. In the proposed method, a Gaussian mixture model (GMM) provides a mixture of mel-frequency distributions while a hidden Markov model is used to characterise class specific transitions between the mel-frequency mixtures over time. A database of slow moving targets in a cluttered environment is used to evaluate the performance of the model. It is shown that the combined Gaussian mixture Hidden Markov model (GMM-HMM) approach can accurately distinguish between different classes of animals and humans walking in these environments. Results show that the classification accuracy of the model depends on the continuous observation time on target and ranges from 75% to approximately 90% for times on target between 250 ms to 1.25 s respectively. A confidence based rejection scheme is also presented to reduce false classification rates. Possible applications include border safeguarding and wildlife anti-poaching operations for species such as rhinos or elephants.

---

[*]Corresponding author
[**]Corresponding author
Email addresses: `vaneeden.willem@gmail.com` (W.D Van Eeden), `pieter.devilliers@up.ac.za` (J.P. de Villiers), `RBerndt@csir.co.za` (R.J. Berndt), `WAJNel@csir.co.za` (W.A.J. Nel), `erik.blasch.1@us.af.mil` (E. Blasch)

Highlights

- A Gaussian mixture Hidden Markov model is used to classify ground moving targets.

- An improvement in performance over popular state-of-the-art method is demonstrated.

- Targets can be accurately classified at ranges exceeding 1km.

- The model can distinguish between different classes of animals and humans.

- The accuracy ranges from 75

1. Introduction

The detection and classification of ground based targets has a number of useful applications spanning military, security and disaster management domains. There are many examples of classification studies that make use of electro-optical systems as a sensor (e.g., video surveillance). However, radar provides some unique benefits as it is a day-night, all-weather sensor and its Doppler sensing capability allows easy detection of moving objects. Accurate radar classification systems could be used to monitor borders for illegal crossings, detect wildlife poachers within nature reserves, as well as monitor farmlands for livestock theft. Over the last decade, poaching and livestock theft specifically have become considerable problems in South Africa (Mashala, 2013; Duffy, 2016). To this end, research is being conducted into ways to detect, track and classify potential livestock thieves and poachers within nature reserves and farmlands.

The terrain is typically savanna and grassland-type regions in which clutter conditions vary, and animals are prevalent. The goal is to develop a classification system that is capable of accurately distinguishing between animals and humans based on their micro-Doppler return, making the assumption that the micro-Doppler returns would be produced for some time period as the target crosses an open area. Being able to deploy all-weather radar systems to critical sites could aid conservation and security efforts. Additionally, the results can help determine the movement patterns of the animals such as when they enter or depart an area.

A variety of classification methods have been investigated in open radar literature. The most prominent among them are support vector machines (SVMs) (Li et al., 2012; Kim & Ling, 2009; Fairchild & Narayanan, 2014), Bayesian classifiers (Challa & Pulford, 2001), hidden Markov models (HMM) (Padar et al., 2016) and Gaussian mixture models (GMM) (Bilik et al., 2006; Molchanov et al., 2011). Artificial neural networks (ANN) have also been used successfully (Kim & Moon, 2016; Kim & Ling, 2008). Some of these neural networks have been trained on raw spectrogram data (Jordan, 2016). In most of the applications mentioned in the aforementioned papers, classification accuracies of 75% and greater were obtained with a time on target

(ToT) of greater than 1 second. Notably, Bilik et al. (2006) obtained a maximum classification accuracy of 96% with a ToT of 4 seconds using mel-Cepstrum coefficients and a GMM Majority voting approach. In (Molchanov et al., 2012; Boulic et al., 1990) it is shown that the accuracy of micro-Doppler based classification systems were directly influenced by the ToT. It was shown that increased time on target would improve classification accuracy; however, a time on target of greater than 1 second would not yield a significant improvement. Much of the literature does not consider animals and humans in applicable operational environments.

Many studies only consider humans, and although they show it is possible to accurately distinguish between a single human and a group of humans, their targets do not include wild animals. The studies are usually performed in the open without the presence of trees and other clutter typical of operational environments. The number of animals in a range cell also affects the classification process and most studies do not address multiple targets. Distinguishing between vehicles and humans is, in some cases, a trivial matter as the targets differ significantly in both radar cross section (RCS) and radial velocity (vehicle observations do not have visible micro-Doppler when approaching the radar site head on (Nanzer & Rogers, 2009)). Both the target types and the measurement environments are shortcomings in some of the studies in the literature (Bilik et al., 2006; Nanzer & Rogers, 2009). It is important to study classification between humans, vehicles, animals, and clutter (HVAC) in realistic and operationally applicable environments and scenarios.

Automatic target recognition (ATR) using radar has been achieved using various methods. Different statistical parameters of the micro-Doppler signal have been used to classify humans, wheeled vehicles, tracked vehicles, helicopters and unknown targets (farm animals etc.) while achieving an accuracy of greater than 90% (Bilik et al., 2006; Nebabin, 1995). Baulic walking models (Boulic et al., 1990) have been used to model and classify the human motion (Van Dorp & Groen, 2008). Image processing techniques such as pseudo-Zernike moments have also been applied in radar target classification (Clemente et al., 2015). A common theme amongst most ATR measurements is that targets are in the open and measured at high signal to noise ratios (SNRs). The focus in this paper is on targets measured in their natural environment, at realistic operational ranges and whose Doppler spectral profiles occupy similar spectral bands or Doppler returns exhibit similar spectral content.

Recent developments include the use of multistatic radar systems for classification between humans and humans carrying weapons (Fioranelli et al., 2015, 2017) and the development of transforms to obtain finer grained signatures when using ultra-wide band radar (UWB) (Qi et al., 2016). Most of this research is however still conducted at low technology readiness levels (TRLs) in controlled conditions.

In this paper, the radar return is treated in a similar way to an audio signal, and audio classification techniques are used to extract features from the data. In a past study (De Witt et al., 2012), it was shown that trained human operators can distinguish between different classes of targets when listening to the Doppler frequency characteristics of the down-sampled radar return from dismounts. Using this knowledge, it should be possible to develop an automatic scheme for classifying targets based on the audio-like characteristics of their radar micro-Doppler return. Here, a Gaussian Mixture Model (GMM) provides a mixture of mel-frequency distributions while the hidden Markov Model (HMM) is used to characterise class specific transitions between the mel-frequency mixtures over time. The GMM model combined with the HMM model is abbreviated as the GMM-HMM model. The GMM part of the GMM-HMM model uses mel-frequency cepstrum coefficients. They are frequently used as features in audio classification. They appear to produce an efficient scale that is more representative of human audio perception of the frequency spectra than other methods such as linear predictive coding (LPC) (Deng & O'Shaughnessy, 2003). The HMM part of the GMM-HMM model is also used frequently in speech recognition to model and recognise transitions between formants in speech, and has been shown to produce satisfactory performance (Rabiner, 1989). Mel-cepstrum frequency components have been shown to outperform LPC and other feature spaces both in terms of number of required dimensions and overall micro-Doppler classification performance of targets (Bilik et al., 2006).

The results obtained by the GMM-HMM model are compared to Bilik's method (Bilik et al., 2006) which combines a GMM with a voting scheme. No time varying information between consecutive bursts is considered in Bilik's method, whereas the GMM-HMM method includes information about transitions between the GMM components, attempting to extract additional information from the time varying information present in the data. The authors conjecture that this could increase the accuracy of the model, however steps should be taken to keep the model complexities between the two approaches close to optimal ranges. To measure model complexity in the case

Parameter	Value
Sweep Repetition Frequency	1.33 kHz
Range Resolution	≈ 5 m
Transmit Frequency	10.2 GHz
Beam Width	0.6°

Table 1: Radar Measurement Parameters

of the GMM-HMM is non trivial and is reserved for future work; however, different model complexities (number of model parameters) for the GMM-HMM model are investigated in this work.

2. Dataset

As far as the authors are aware, no large database of animals (specifically African animals) and humans in a natural environment is publicly available. In an effort to develop robust classifiers for ground observation radar systems, a data collection effort was conducted. The parameters of the X-band frequency modulated continuous wave (FMCW) radar system that was used to collect data is given in Table 1, a photograph of the experimental set-up is shown in Fig. 1. Human and animal targets were recorded in the wild in a staring mode. In total, approximately 1200 seconds worth of data were recorded, which included multiple target types. The human targets were willing participants and their park ranger guides (for safety), whereas the animals were targets of opportunity that happened to be moving towards or away from the radar site at the time. The recordings are not controlled in any manner as the animals do not co-operate with the measuring process nor are their angle of incidence and speed controlled. Sometimes radar range-Doppler cells contain a single animal or human, and at other times more than one of each of the same type. As such this could cause within-class variability of the micro-Doppler features in the data. Most recordings were performed on animals between 1 and 3 km from the radar site.

3. Mel-Cepstrum coefficient extraction

Before any processing is performed to extract features, a simple 10 Hz high pass filter was applied to the recorded data to remove the quasi-stationary ground clutter.

Figure 1: Radar experimental set-up used for measurements on a hilltop overlooking the operational savanna environment that can be seen in the background.

A burst is composed of N sweep repetition intervals (SRIs) which have been observed consecutively upon which a length-N fast Fourier transform (FFT) is performed, resulting in FFT energies $S(1), \ldots, S(N)$ at frequencies $f_1, \ldots, f_N$. Bursts lengths are typically chosen so that the signal can be considered stationary for the burst duration. Bursts can also be allowed to overlap. In our data, short burst lengths of less than 64 (50 ms) cause consecutive bursts to be assigned to the same component in the GMM, implying stationarity of the signal for such short burst lengths. In speech recognition 20-30 ms is considered stationary (McLoughlin, 2016), however stationarity intervals for Doppler walking motion differs from this. To determine the length that the data is stationary, the data were subjected to stationarity tests. Specifically, the null hypothesis was tested against the alternative hypothesis of a unit root (Kwiatkowski et al., 1992). To determine the average length that the data is stationary, the burst length was varied and the percentage of bursts that passed the stationarity test were recorded. The results showed that for a burst length of 128 points (100 ms) 71% of samples where considered stationary. For 64 points (50 ms), 85% of bursts would pass a stationarity test. At longer burst lengths more bursts are considered to be non-stationary than stationary. As such, the chosen burst length was 100 ms. This burst length is also used by Bilik (Bilik et al., 2006) with whom this work is compared. A Hamming window is applied to the burst to reduce side-lobes. For each burst, a set of mel-frequency cepstrum coefficients are

calculated and stored. A set of mel-frequency cepstrum coefficients arising from consecutive bursts is defined as a *frame*. A spectrogram of a moving zebra created with these parameters is shown in Fig. 2.
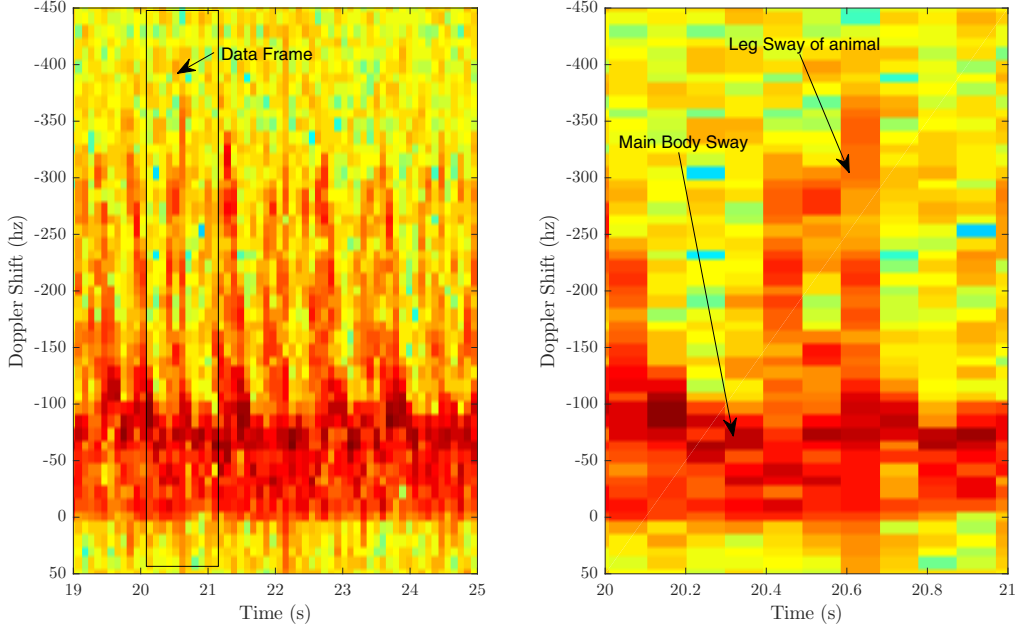

Figure 2: Each vertical line of the spectrogram corresponds to 1 burst of 128 samples upon which a length-N=128 FFT has been computed. The rectangle indicates the boundaries of a frame. Here a frame represents the chosen (simulated) time on target, were the radar to operate in scanning mode. The frame in turn, consists of a specified number of bursts. The suspected main body sway, as well as the movement of the animals legs are indicated in the spectrogram.

In order to compute the mel-frequency cepstrum coefficients the following steps are performed:

1. The time domain signal $s(n)$ is divided into frames of length L corresponding to the observation time one wishes to consider. The length of these frames are chosen based on the radar system in use and the dynamics of the targets being measured.
2. Next the frames are partitioned into B bursts of specific length N. The time domain signal comprising a frame will be referred to as $s_i(n)$, where n is the index of the nth data point in the burst and the subscript i is the index of the ith burst in the frame of length L.

3. For each burst in a frame, an FFT operation is performed, thereby obtaining the frequency domain signal $S_i(k)$:

$$S_i(k) = \sum_{n=1}^{N} s_i(n)h(n)e^{\frac{-j2\pi kn}{N}} \quad 1 \leq k \leq N, \tag{1}$$

where $h(n)$ is an N-sample-long Hamming window, n is the index over samples, j is the imaginary number $j^2 = -1$, and k is the Fourier domain index. The periodogram-based power spectral estimate for the burst $s_i(n)$ is given by:

$$P_i(k) = \frac{1}{N}|S_i(k)|^2. \tag{2}$$

4. Compute the mel-spaced filterbank $H_m(k)$. This is a set of triangular filters that are applied to the periodogram power spectrum of step 3. These cepstral wavelet functions are essentially wavelets spaced according to the mel-scale. Each individual mel-cepstrum coefficient covers a certain range of the spectrum and sums all the data in that region together to give a single value corresponding to the power in that region. The filterbank is M vectors of length N where M is equal to the chosen number of cepstral coefficients. To calculate the filterbank energies, the coefficients of each filter $H_m(k)$ are multiplied with the power spectrum $P_i(k)$ of the burst and summed to obtain a single value A_m, $m \in \{1, \ldots, M\}$ for each of the M filters, i.e.

$$A_m = \sum_{k=0}^{N} P_i(k)H_m(k). \tag{3}$$

5. The logarithm of each of the filterbank energies from step 4 is taken, i.e.

$$Y_m = \log_{10} A_m. \tag{4}$$

6. To determine the cepstrum coefficients, a discrete cosine transform (DCT) is used. The initial K coefficients y_n, $n \in \{1, \ldots, K\}$ of an inverse discrete cosine transform of the first K components of $Y_1, Y_2, \ldots Y_K$, where $K < N$ is given by

$$y_n = \sum_{k=1}^{M} Y_k \cos\left[n\left(k - \frac{1}{2}\right) * \frac{\pi}{20}\right]. \tag{5}$$

The initial value y_0 represents the average power of the signal, which can vary widely, and is usually discarded. In this paper 10 cepstral coefficients are used. If a higher degree of resolution is desired, a higher number of coefficients can be used.

o calculate the coefficients $H_m(k)$ of the mel-filterbanks from step 4, the ext four steps are followed:

1. Convert the upper and lower frequencies of the relevant frequency interval of the data to the mel scale with the equation:

$$\mathcal{M}(f) = 1125 \ln(1 + \frac{f}{700}). \tag{6}$$

where f is the frequency being converted. In this case the lower frequency limit is 15 Hz (23.85 mels) and the upper frequency limit is dependent on the sweep repetition frequency (SRF) and is usually equal to the Nyquist frequency. If the radar system has a SRF of 1.3 kHz the upper limit will be around 650 Hz (738.87 mels).

2. Next, points are linearly spaced between the lower limit and upper limit in the mel-spectrum to obtain $M+1$ mel-frequency points β_m. Convert these points back to frequency domain (from the mel-domain) with the equation:

$$h(m) = \mathcal{M}^{-1}(\beta_m) = 700(e^{\frac{\beta_m}{1125}} - 1), \tag{7}$$

where β_m indexes a frequency on the mel-scale, to obtain m points $h(m)$ in Hz.

3. Since the frequency resolution does not allow these points to be placed anywhere in space where desired, these points can be converted to frequency bin numbers based on the locations of the discrete frequency bins of the FFT, where

$$f(m) = \lfloor (Nh(m)/T_s) \rfloor, \tag{8}$$

where T_s is equal to the sample period and $\lfloor \rfloor$ is the floor of the function. This results in a sequence of bin numbers $f(m)$.

4. Now the filters $H_m(k)$ are created. Each filter will start increasing at a point $f(m-1)$, peak at $f(m)$ and decrease back to zero at $f(m+1)$.

All other points are equal to 0 according to the following equation:

$$H_m(k) = \begin{cases} 0 & k < f(m-1), \\ \frac{k-f(m-1)}{f(m)-f(m-1)} & f(m-1) \leq k \leq f(m), \\ \frac{f(m+1)-k}{f(m+1)-f(m)} & f(m) \leq k \leq f(m+1), \\ 0 & k > f(m+1). \end{cases} \tag{9}$$

Only $D = M - 1$ coefficients y_1 to y_M are used. The mel-ceptrum coefficients are chosen to occupy the frequencies where most of the target energy is concentrated. As such parts of the Doppler spectrum that do not contain energy from the target were discarded manually. In this application, the coefficients are placed between 10 Hz and 650 Hz as the data only considers targets that are moving at relatively slow, speeds typically less than 5 ms^{-1}. The mel-cepstrum coefficients are then passed to a Gaussian mixture model (GMM) for further analysis. This choice is supported by previous work from which it is known that the maximum micro-Doppler of targets typically reach at least twice this speed (van Eeden et al., 2015).

In the dataset presented here, targets exhibit either negative or positive Doppler, not both. Since the cepstral coefficients are only defined for positive spectra, the spectra of targets moving away from the radar (with negative Doppler spectra), were flipped such that every negative Doppler frequency presented as a positive Doppler frequency and *vice versa*. The implicit intuitive assumption for the sake of simplicity, although not provably accurate, is that the Doppler spectrum of a target moving away from the radar with some aspect angle, would present as the mirror image of the Doppler spectrum of a target moving towards the radar with the opposite aspect angle. As such, the radial direction of the target is not assumed to contain discriminative features for the purpose of classifying between species. Future studies with more representative data may be employed to test the assumption that closing and receding targets would exhibit mirror image spectra.

4. Gaussian Mixture Model

Once the mel-cepstrum coefficients have been determined, they are separately clustered for each class by a GMM using the iterative expectation-maximization (EM) algorithm (Bishop, 2006). The vector of observed cepstrum coefficients is given by the column vector $\mathbf{y} = [y_1, ..., y_D]^T$. A set of these vectors are then modelled as a mixture of Gaussian distributions for

each class. Consider $D \times B$ dimensional matrix $\mathbf{Y}$, where the nth column is populated element wise with the elements of column vector $\mathbf{y}_n$, where D is equal to the number of cepstrums per burst, and B is equal to the number of bursts. Similarly the corresponding latent variables will be denoted by an $K \times B$ matrix $\mathbf{Z}$ with rows populated with the elements of $\mathbf{z}_n$. Assuming that the data points are drawn independently from the distribution, then the log likelihood of the Gaussian mixture model for this data set is given by

$$\ln p(\mathbf{Y}|\kappa, \mu, \Sigma) = \sum_{n=1}^{B} \ln \left\{ \sum_{k=1}^{K} \kappa_k \mathcal{N}(\mathbf{y}_n|\mu_k, \Sigma_k) \right\}, \tag{10}$$

where $\mathcal{N}(\cdot|\mu, \Sigma)$ represents a multivariate Gaussian distribution with mean vector μ and covariance matrix Σ. This likelihood function needs to be maximised with respect to the parameters $\{\kappa_k, \mu_k, \Sigma_k\}$ for $k = 1, \ldots, N$. The expectation maximization (EM) algorithm of Dempster is used to maximizes this likelihood (Dempster et al., 1977). The first step of the EM algorithm is to initialize the means μ_k, covariances Σ_k and mixing coefficients κ_k and evaluate the initial value of the log likelihood. Then perform the **EStep** by evaluating the weights using the current parameter values

$$\gamma(z_{nk}) = \frac{\kappa_k \mathcal{N}(\mathbf{y}_n|\mu_k, \Sigma_k)}{\Sigma_j \kappa_j \mathcal{N}(\mathbf{y}_n|\mu_j, \Sigma_j)}. \tag{11}$$

Next the **MStep** is performed by re-estimating the parameters using the current weights.

$$\mu_k^{new} = \frac{1}{B_k} \sum_{n=1}^{B} \gamma(z_{nk}) \mathbf{y}_n, \tag{12}$$

$$\Sigma_k^{new} = \frac{1}{B_k} \sum_{n=1}^{B} (\mathbf{y}_n - \mu_l^{new})(\mathbf{y}_n - \mu_l^{new})^T, \tag{13}$$

$$\kappa_k^{new} = \frac{N_k}{B}, \tag{14}$$

where

$$B_k = \sum_{n=1}^{N} \gamma(z_{nk}). \tag{15}$$

Finally, the log likelihood and parameters are checked for convergence, if the values have not changed by more than 10^{-6}, the process is terminated. If the convergence criterion is not satisfied, then return to the **EStep** and continue until the algorithm converges.

The number of components of the GMM for each class is automatically computed using the Akaike's information criterion (AIC) (Ljung, 1999). The lowest AIC score determines the GMM distribution weights. For the training dataset used in this study the GMM typically has between 6 and 10 components dependent on the class, (10 for humans and walking animals, 6 for grazing animals). Finally, the states (GMM components) of the classes are combined into a single GMM so as to allow the probabilities of the different components to be related. The separate GMMs for each of the classes are merged into a single augmented GMM, in which all the components of the separate class dependent GMMs are merged in a single "augmented" space. To determine whether or not any components significantly overlap and should be merged, the closed form equation for a bi-directional version of the Kullback-Leibler (KL) divergence between the Gaussian components is used as a similarity score between components (Hershey & Olsen, 2007). The equation for the bi-directional KL divergence is give by:

$$D_{KL}(\mathcal{N}_a||\mathcal{N}_b) = 0.25(tr(\Sigma_b^{-1}\Sigma_a) + (\mu_b - \mu_a)^\top \Sigma_b^{-1}(\mu_b - \mu_a) - k + \ln \frac{\det \Sigma_b}{\det \Sigma_a})$$
$$+ 0.25(tr(\Sigma_a^{-1}\Sigma_b) + (\mu_a - \mu_b)^\top \Sigma_a^{-1}(\mu_a - \mu_b) - k + \ln \frac{\det \Sigma_a}{\det \Sigma_b}),$$
$$(16)$$

where $\mathcal{N}_a$ and $\mathcal{N}_b$ are the two components of the Gaussian mixture models being compared. The terms Σ_a and Σ_b are the covariance matrices of the Gaussian mixture components. The values μ_a and μ_b are the means of the Gaussian mixture components. In this bi-directional version, the value is not dependent on the order of $\mathcal{N}_a$ and $\mathcal{N}_b$ as is the case with the standard KL-divergence.

In this application it was found that no components ever overlap when using the KL divergence criterion, implying that the states of the different animal/human classes are well separated in the cepstrum feature space. If any were to overlap a merging strategy would have to be followed. Next a hidden Markov model is constructed to classify the targets based on their state transitions.

5. Hidden Markov Model

The HMM determines the probability of a class C given a sequence of observed cepstrum coefficients $\mathbf{y}_{0:t}$, where the subscript $0{:}t$ represents all observations from discrete time 0 to t. The corresponding sequence of hidden states is denoted by $x_{0:t}$. At any time instant t, x_t is a scalar that can take on an integer value between 1 and Q, where Q is the number of mixture components of the GMM. The probabilities of these states are inferred by the expectation maximization algorithm already described. This is the standard state space description of the hidden Markov model. Using Bayes theorem

$$P(C, x_{0:t}|\mathbf{y}_{0:t}) = \frac{P(\mathbf{y}_{0:t}|C, x_{0:t})P(C, x_{0:t})}{P(\mathbf{y}_{0:t})}. \tag{17}$$

The prior probability in the numerator can be expanded using the product rule as follows

$$P(C, x_{0:t}) = P(C)P(x_{0:t}|C). \tag{18}$$

The first-order Markov assumption present in the model results in

$$P(x_{0:t}|C) = P(x_0|C)P(x_1|x_0, C)P(x_2|x_1, C)\ldots P(x_t|x_{t-1}, C), \tag{19}$$

or

$$P(x_{0:t}|C) = P(x_0|C)\prod_{l=1}^{t} P(x_l|x_{l-1}, C). \tag{20}$$

Using the conditional independence properties of the model, i.e. that C is independent of y given x, the first term of (17) can be expanded as

$$P(\mathbf{y}_{0:t}|C, x_{0:t}) = P(\mathbf{y}_{0:t}|x_{0:t}) \tag{21}$$

$$= \prod_{j=0}^{t} P(\mathbf{y}_j|x_j). \tag{22}$$

Substituting equations (20) and (22) into equation (17) results in

$$P(C, x_{0:t}|\mathbf{y}_{0:t}) = \frac{P(C)\prod_{j=0}^{t} P(\mathbf{y}_j|x_j)P(x_0|C)\prod_{l=1}^{t} P(x_l|x_{l-1}, C)}{P(\mathbf{y}_{0:t})}. \tag{23}$$

Only the probability of the class given the observations is of interest, as such a marginalization over the states $x_{0:t}$ can be performed and the denominator can be ignored due to it only being a normalization constant:

$$P(C|\mathbf{y}_{0:t}) \propto \sum_{x_{0:t}} P(C)P(x_0|C) \prod_{l=1}^{t} P(x_l|x_{l-1},C) \prod_{j=0}^{t} P(\mathbf{y}_j|x_j). \qquad (24)$$

The conditional probability in (24) can then either be computed using full enumeration or the Viterbi sequence can be computed to approximate the probability of the classes.

$$\max_{x_{0:t}} P(C|\mathbf{y}_{0:t}) \propto P(C)P(x_0|C) \max_{x_t}(P(x_t|x_{t-1},C) \max_{x_{0:t-1}} P(y_{0:t}|x_{0:t}) \qquad (25)$$

6. Method

The data were partitioned into three different classes, namely walking humans, walking animals and grazing animals. Each class consists of roughly 6 minutes worth of recorded data, whilst utilising the radar in the staring mode.

The classes are defined as follows:

1. Animals that were measured walking approximately radially towards or from the radar, were grouped together as a class "walking animals".

2. Animals that were standing still, eating grass, raising their heads and looking around and then potentially moving slowly to another nearby location to continue grazing were grouped into the class "grazing animals"

3. Human participants were observed at a range of around 3 km walking towards the radar installation and were classed as "human targets", where no distinction is made whether 1 or 2 humans are present in a range cell.

4. A fourth "nondescript" class is added. This class is analogous to a uniform distribution over motion types in our model, and is represented by an equi-probable random walk through the hidden Markov model states (i.e. all state transitions are equally probable). It serves primarily as a method to normalize the classification results and lower the model's confidence in its own answer, owing to the fact that not all types of target motion are accounted for in the model. In practice, a

uniform background motion class has almost no effect on the classification outcomes, except for reducing confidence in the final answer if the exhibited motion is close to random walk behaviour.

The background class can significantly reduce the error rate by requiring a higher probability threshold for choosing a particular class, if a thresholding procedure is employed. The approach of defining this "nondescript" class is similar to the approach followed in (Bishop, 1994), and is referred to as the "Other" class in the results section.

The measurements of the targets were split into three sets (folds) of roughly equal size, with approximately the same number of measurements of each class in each fold. For every test performed, the classifier is trained on one fold, validated on another, and tested on the remainder.

A custom 3-fold cross validation strategy was then implemented, which is depicted in Table 2. The data fold partitioning is performed such that in any sub-fold shown in Table 2, one set is completely held out until testing. This ensures that test set results are never used to determine model order.

	Training	Model Validation	Testing
a)	Train on 1	Validate on 2	Test on 3
	Train on 2	Validate on 1	Test on 3
b)	Train on 2	Validate on 3	Test on 1
	Train on 3	Validate on 2	Test on 1
c)	Train on 1	Validate on 3	Test on 2
	Train on 3	Validate on 1	Test on 2

Table 2: This table indicates how the data was partitioned for training, validation and testing over three folds. Consider for example a "sub-partition" a). Fold number 3 is held out for a test set, and as such no test data is used to in any way determine the models resulting from the training/validation set combinations 1/2 and 2/1. The same can be said for sub-partitions b) and c). The validation results are averaged within each sub-partition (i.e. average of only two results per sub-partition), however the test results are averaged over all six test results.

Once the mel-cepstrum coefficients have been extracted and clustered appropriately, the transition probabilities for each class between the different "cepstral-states" are computed and stored through supervised learning. A frame is observed, divided into appropriately sized bursts and each burst is assigned a "state" based on its cepstrum coefficients and the GMM cluster it

is closest to. After cluster selection, the state transitions are calculated. To avoid errors based on low emission probabilities, transition probabilities that are zero are set to small non-zero values (10^{-4}). Increasing this value would result in the "Other" class to become more probable, reducing the confidence in the classification result. The converse is also true. These transition probabilities can be determined by observing how often such "between class" transitions take place in the state transitions between mixtures components of measured data, however in the case of this paper the choice was arbitrary, but resulted in acceptable performance. Along with this, the class transition probabilities for each cepstral-state is computed. It should be apparent that certain cepstral-states are more likely to occur in certain classes than others. The cepstral information is also useful and collected at the same time as it forms the basis for GMM based classification employed by Bilik et al. (2006).

In the GMM-HMM classification model, the most probable sequence through the Markov model is calculated using the Viterbi algorithm in Eq. 25. Once the probability of each class is calculated a maximum a-posteriori (MAP) approach (using the Viterbi algorithm) is used and the class with the highest probability is selected as the classification result.

Two different studies have been conducted for this research. The first focuses on distinguishing between *broad classes* (human/bipedal motion, walking animals, grazing animals) for practical purposes and the second focuses on distinguishing between two specific animals. Giraffe and zebra data from the walking animal class were selected to determine if the classifiers under consideration could distinguish between individual species of animals. These animals where observed in the greatest number and had the most data available. They also tend to have similar micro-Doppler spectral ranges. However, their relative body motions and spectrograms differ in terms of periodicity and the largest Doppler shift achieved by the legs of the animal. This difference is directly related to the physiology of their movement. To determine whether or not the classifier is merely using the relative radial velocity component produced by the main body of the target we isolate only these two similar targets to determine if the classifier can in fact distinguish based on micro-Doppler features. This does not mean that in the human animal classification it has not classified based on radial velocity. It merely shows that the classifier is capable of distinguishing based on micro-Doppler features using this scheme.

For both studies, the burst length was set to 100 ms and the overlap was set to 0. This corresponds to the approach taken by Bilik, since allowing

overlaps between bursts would cause the votes to be dependent and would consequently introduce double counting. This is not a restriction in the case of the GMM-HMM model; however, to perform a like-for-like comparison, a zero overlap was chosen for both cases. Exploratory experiments by the authors have also shown that the introduction of overlap did not seem to produce a significant improvement. The increase in resolution from a longer burst length provides a more significant advantage.

7. Results

7.1. Broad Class classification

7.1.1. Model Selection

If the AIC method is used to determine the number of parameters for the GMM part of the GMM-HMM model, the addition of the HMM state transition parameters would result on a combined model which is much too complex for this data set. This would lead to overfitting, since none of the HMM transition parameters are considered by the AIC method. The recommended total model order is 27 for the GMM part (10 for human and walking animal each, 7 for grazing animal). This implies an HMM model with 729 possible transitions while only approximately 3000 state transitions are available for each fold to train upon. An equivalent AIC which considers both the GMM and HMM parameters would need to be developed to better determine the exact complexity of the model or more data should be obtained to enable improved parameterisation of the transition probabilities. The creation of such an AIC measure would certainly be of interest, but would be beyond the scope of this work. However, the authors optimised the number of GMM parameters (and by implication the number of HMM transition probability parameters) which maximised the classification accuracy between folds using a validation set, resulting in a GMM-HMM model with the right number of parameters for this dataset. This can be seen in Table 3.

Validation is performed according to the dataset partitioning scheme detailed in Table 2, obtaining results that are shown in Table 3. The model order that had the highest accuracy (averaged between 2 sub-partition validation results) was chosen to be run on the test set. In the case of Table 3 this corresponds to the model with 14 components (5/5/4). The model would then be tested on the remaining third fold of the data to obtain the test results. The process is repeated for every fold of the data rotating the sets as per Table 2. The tests performed for the rest of this work (i.e., in

Section 7.1.2) are reported as the mean and standard deviation over six test results, with the model order for the six tests being chosen based on their specific validation results.

Model order	GMM Component Partition	Training Accuracy	Validation Accuracy
6	2/2/2	96.62% ±1.74	92.26% ±1.15
8	3/3/2	96.97% ±1.66	94.13% ±1.64
10	4/4/2	97.68% ±1.48	93.99% ±0.27
12	5/5/2	97.29% ±0.45	93.28% ±0.054
12	4/5/3	97.83% ±1.66	92.96% ±0.1
12	5/4/3	97.90% ±1.66	94.42% ±1.49
13	5/5/3	97.61% ±0.7	93.68% ±2.21
14	5/6/3	97.61% ±0.81	93.41% ±2.36
14	6/5/3	98.38% ±0.73	92.53% ±2.52
14	5/5/4	98.96% ±0.062	94.66% ±0.24
15	6/6/3	97.28% ±1.41	93.22% ±2.72
18	7/7/4	98.74% ±0.5	94.08% ±3.83
21	8/8/5	99.02% ±0.072	93.11% ±4.78
23	9/9/5	99.18% ±0.009	93.93% ±4.96

Table 3: Table of accuracy for model order of a single fold of data (average of two validation results per sub-patition). Model order is denoted as human/walking animal/grazing animal model order. There is a broad range close to the best point where model order does not affect performance significantly, i.e. performance is insensitive to choice of model order within this range. However, performance does degrade for very high and low model orders.

The greater than 90% accuracy obtained by all models seems to be a satisfactory result, however it becomes increasingly difficult to significantly improve performance for this dataset owing to the law of diminishing returns.

7.1.2. GMM-HMM method and Bilik's method comparison

The ToT was varied for the broad class datasets (human, walking animal, grazing animals) in order to show the effect of observation time on classification accuracy. The ToT is an important factor for scanning radar systems. An increased ToT will also increase the revisit time which will make accurate tracking more difficult to achieve. Here, the results of the GMM-

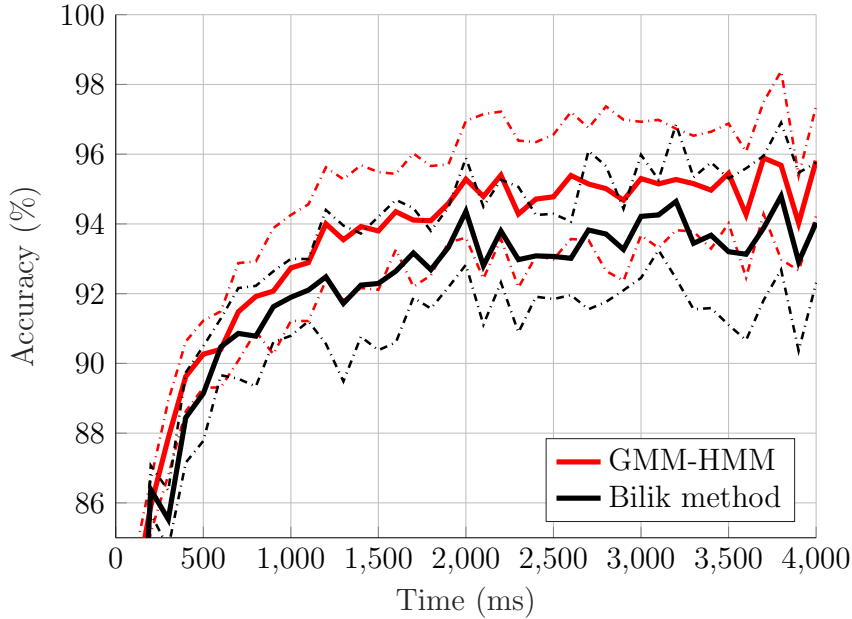HMM are compared to Bilik's method (Bilik et al., 2006) for this broad class classification problem.

Figure 3: A comparison of human classifier accuracy vs. observed time for Bilik's method and HMM-GMM. The dashed-dotted lines indicate ± single standard deviations for the different folds.

Fig. 3 shows the classification accuracy compared with that of Billik's method. In this figure targets classified as the "non-descript" or "Other" class in the GMM-HMM method are excluded from the accuracy calculations, since the data classified as "Other" look more like a random walk through the GMM space than any other class. This indicates poor data quality, and the advantage of including an "Other" class assists with discarding such poor data, i.e. data which does not contain significant discriminative information. This seems to be reflected in the figure in improved accuracy and reduced variance. Although the GMM-HMM results are not statistically significantly better than Bilik's method at any single ToT, the results are consistently better over almost all ToTs.

Considering the results in Tables 4 and 5 and Fig. 3, it is evident that the accuracy of the system is a function of the ToT. The longer the target is observed, the more accurate the classification becomes. However, this

500 ms ToT			
HMM-GMM/Bilik	**True Class**		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	85.87/84.58	3.39/3.48	11.09/9.11
Walking animal	4.33/5	93.06/94.06	1.36/0.96
Grazing animal	6.93/10.4	1.49/2.46	86.55/89.93
Other	2.87/-	2.07/-	1.0/-
1 s ToT			
HMM-GMM/Bilik	**True Class**		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	89.2/87.06	2.31/0.237	8.79/5.62
Walking animal	3.47/3.41	94.39/96.1	3.7/0.37
Grazing animal	4.93/9.54	1.65/1.53	90.48/94.01
Other	2.4/-	1.65/-	0.37/-
1.5 s ToT			
HMM-GMM/Bilik	**True Class**		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	91.04/88.39	1.49/1.7	7.14/5.5
Walking animal	3.98/3.8	96.29/96	0.55/0.8
Grazing animal	4.18/7.7	0.99/2.2	91.48/93.5
Other	0.8/-	1.24/-	0.82/-
2 s ToT			
HMM-GMM/Bilik	**True Class**		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	93.09/91.15	1.66/2.3	6.25/3
Walking animal	3.19/2.68	96.69/96.36	0/0.37
Grazing animal	2.66/6.17	0.33/1.32	93.75/96.64
Other	1.06/-	1.32/-	0/-

Table 4: 3-fold cross validation confusion matrices for broad classes using MAP approach, with no rejection present for these results. Results are reported as percentages for HMM-GMM method followed by Bilik's method.

500 ms ToT			
75%/90%	True Class		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	90.23/93.02	2.79/1.74	7.68/5.4
Walking Animal	3.41/2.73	94.94/96.7	1.01/0.8
Grazing Animal	5.08/3.41	0.96/0.73	90.7/93.22
Other	1.29/0.85	1.31/0.82	0.61/0.57
Rejection Rate	10.25/18.25		
1 s ToT			
HMM-GMM/Bilik	True Class		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	92.29/93.69	1.85/1.54	7.78/6.53
Walking Animal	2.86/2.7	95.64/96.76	0.38/0.4
Grazing Animal	3.57/2.7	1.34/0.85	91.46/92.87
Other	1.29/0	1.17/0	0.38/0
Rejection Rate	5.09/8.21		
1.5 s ToT			
HMM-GMM/Bilik	True Class		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	92.43/92.95	1.24/1.0	5.78/5.07
Walking Animal	3.68/3.53	96.52/97	0.29/0.3
Grazing Animal	3.27/2.9	1.0/1.0	93.64/94.63
Other	0.61/0.62	1.24/1	0.29/0
Rejection Rate	3.29/4.71		
2 s ToT			
HMM-GMM/Bilik	True Class		
Classifier Out	**Human**	**Walking**	**Grazing**
Human	93.28/93.73	1.67/1.68	4.56/3.86
Walking Animal	3.23/3.27	97.33/97.64	0/0
Grazing Animal	2.69/2.45	0.33/0.34	95.44/96.14
Other	0.81/0.54	0.67/0.34	0/0
Rejection Rate	2.06/3.13		

Table 5: 3-fold cross validation confusion matrix for broad classes with thresholding. Results are reported as percentages for GMM-HMM method with 75% required confidence followed by 90% required confidence

performance comes with a trade-off in a scanning radar setup. The ToT will increase classification performance, but detection and tracking of the targets will be penalised. The penalty could be reduced by increasing the number of beams on the radar system; however, this increases the power and financial cost of the radar system. Bilik et al. (2006) showed that approximately 4 seconds of observation was optimal for humans and that after this time frame, longer staring would not significantly increase the accuracy of the classification system. This observation is confirmed in this study. However, significant improvement is rarely seen in this data for ToTs greater than 1 second, with the required ToT also being dependent on the species observed. A human specifically requires a ToT of around 1 second as this is the period of their walking gait. Some animals can be observed for shorter periods to achieve the same maximum in accuracy due to the period of their movement cycle being shorter.

It should be noted that the micro-Doppler for different targets can be very similar for short ToTs, especially when considering that the target aspect angle is not controlled or recorded in our data. The ranger and researcher would often stop to look around for wild animals and the ranger would look back to see if the researcher was still following closely. As such for short periods of time human data and grazing animal data could match. This emphasizes the need for the classifier to handle multiple incidence angles and should be investigated in future work. The error could be isolated if the angle of incidence was known at all times. Likewise, the error causes some uncertainty in the model and the relatively low accuracy can be ascribed to the variability in the human behaviour just described. However, a system that is robust to such ambiguities, as well as the number of same-class targets in a range-Doppler cell, would be advantageous. Poachers or thieves will not conveniently stay together or spread out equally at all times, nor will they walk straight towards the radar presenting with a perfect human walking micro-Doppler signature.

The spike in misclassification of grazing animals for short observation periods can be explained by the manner in which animals graze. They typically stand still when grazing, lift their heads for a time and then move short distances. This short movement period closely matches other walking animals and even humans when it occurs, and is frequently misclassified. Some short portions of human data also match grazing animals. The misclassification persisting after the use of confidence metrics could indicate overfitting of the data owing to the AIC score not incorporating the added complexity of

the additional HMM parameters. However, when reducing the GMM-HMM model order below that which corresponds to the optimal AIC score these errors remain, but increases the accuracy and reduces the variance between folds are observed. Caution is required with high accuracies as that could indicate that some overfitting has occurred. However, the human/grazing-animal misclassifications do not seem to be a result of overfitting, since they are not influenced by model order.

7.1.3. Thresholding for controlling desired resultant remaining accuracy

In this section, a rejection scheme for classifying only targets that achieve a certain confidence threshold is implemented for the GMM-HMM model. The selection of a threshold is over and above the inclusion of the "Other" class already described. A rejection approach is often used in the medical and quality control industries as the cost of a false classification is much worse than the cost of simply requiring human intervention (Scheme & Englehart, 2015) which is true in the case of poacher observation as well. It is much more desirable to have the operator use secondary methods to classify a target than for the uncertain targets to be misclassified and valuable resources needlessly deployed. As such accuracies reported when thresholding is employed, are resultant accuracy after a certain number of uncertain samples have been rejected based on their confidence score (the required confidence in the threshold methods was 75%). Hence, the rejection scheme only classifies targets if their probability is higher than a certain percentage. If a low false alarm rate and high accuracy is desired, the use of confidence metrics can significantly reduce the errors obtained by the system at the cost of not classifying a given observation due to lack of confidence. Targets that do not pass this threshold are left non-classified and are not included in accuracy results. These could be seen as "missed classifications", but could in practice be left for later classification in anticipation of improved measurement opportunities.

Figure 4 shows the accuracy of the GMM-HMM classifier with varying ToTs as well as showing the effects of requiring a confidence threshold. A confidence threshold of 75% was required for the model. It can be seen that at the cost of rejecting a certain number of samples, the confidence metric will increase the accuracy. The false alarm level can be lowered to a desired or required rate with a high enough specified confidence level.
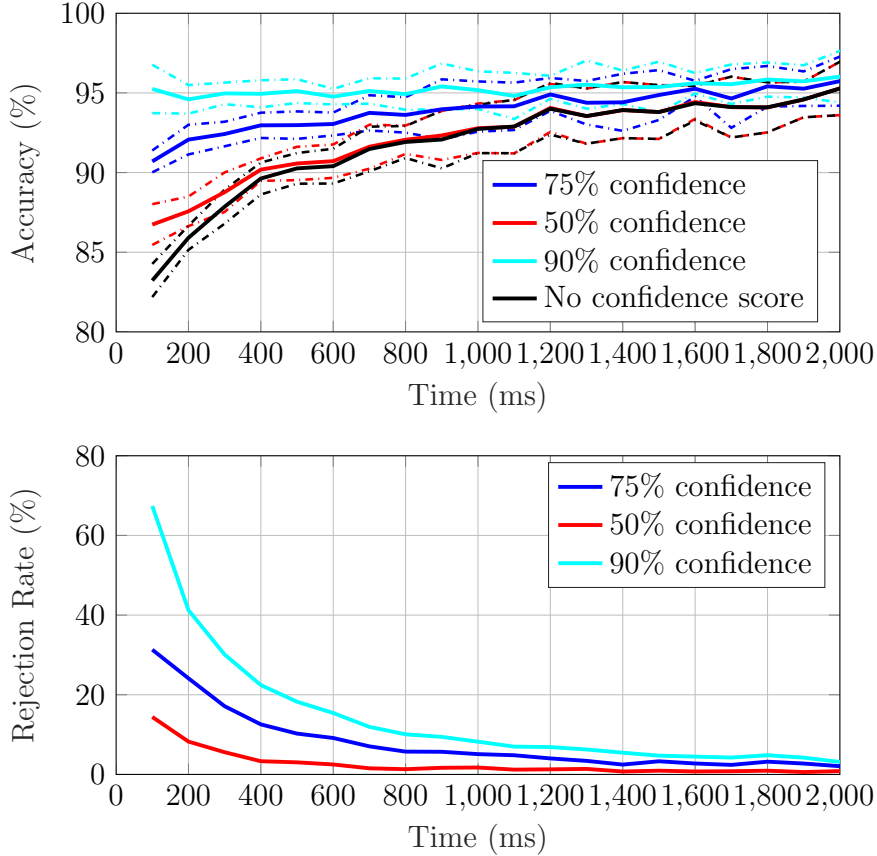
Figure 4: A comparison of classifier remaining accuracy and rejection rate vs. ToT for GMM-HMM method using confidence scores

7.2. Special case: Giraffe-Zebra Classification

In many cases, classifiers that supposedly use micro-Doppler modulation as the source of information for classification end up actually relying on the relatively strong radial velocity component produced by the main body of the target. For example, a car and a human can easily be distinguished based only on average or maximum radial velocity over time. However, in animal-human classification, many of the targets have very similar radial velocities. As such, any classification algorithm should be able to discriminate based on micro-Doppler features alone. The following example was chosen where Doppler data from two different species (giraffe and zebra) had ap-

proximately the same main body radial velocity component. In Fig. 5, a comparison of a typical giraffe spectrogram and a typical zebra spectrogram is presented. Each burst of the data roughly corresponds to a vertical line
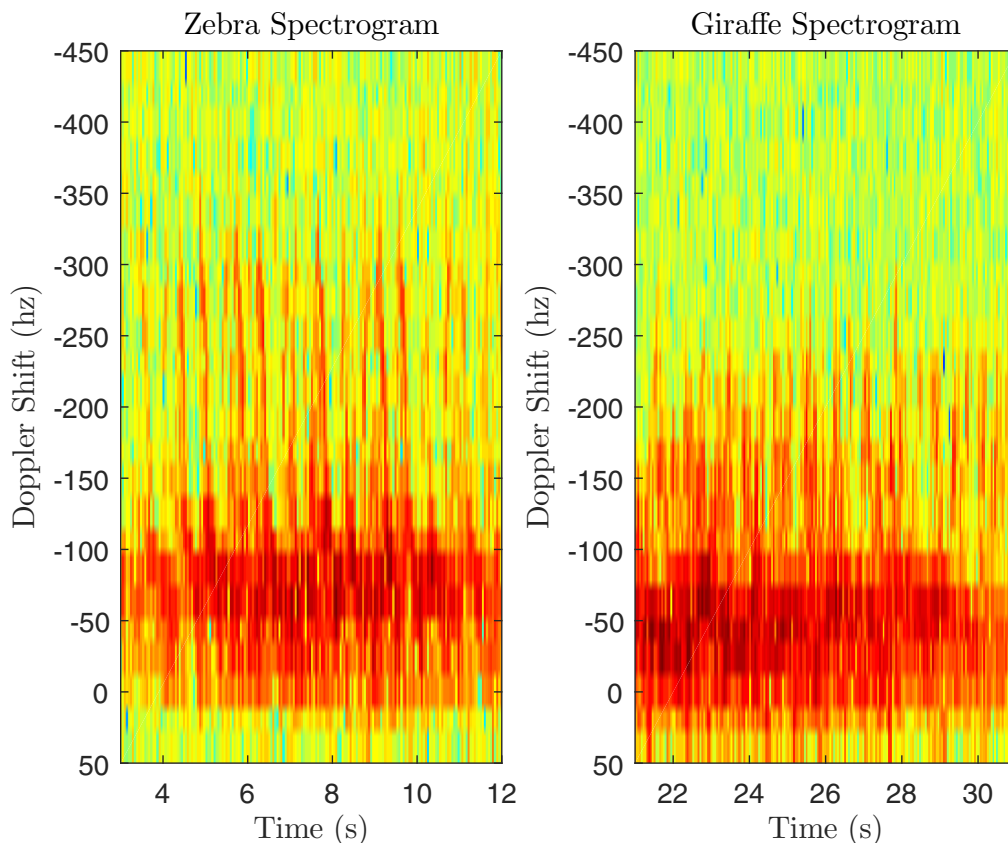

Figure 5: A comparison of typical zebra and giraffe spectrograms from the recorded dataset. There are some variations between different examples of a single species, but some features in the spectrograms are discriminative between species within this particular dataset.

within the spectrogram and over time the target's motion repeats itself in a cyclic manner. The form of the spectrogram is also expected to be dependent on what type of movement the animal is performing at a particular point in time and its aspect angle to the radar. That is, the Doppler signature of a zebra can take multiple different forms. Comparing arbitrary bursts of giraffe and zebra data, their cepstrum coefficients are often similar at comparable phases of their gait, especially when not observing a time at which

the legs of the animal are moving. However the way in which they transition through different GMM components (or "cepstral-states") is discriminative. This difference allows the zebra and giraffe to be accurately classified using the GMM-HMM. Notably the difference between a zebra and giraffe are expected to occur in the frequency of leg movements as well as the maximum frequency shift obtained from their legs. The ratio of maximum Doppler shift obtained by the legs to the Doppler shift obtained by the main body of the animal is expected to correspond to the length of the animals' legs and should be a distinguishing feature for every species. This hypothesis will have to be confirmed in future research when a more comprehensive dataset can be obtained that includes aspect angles.
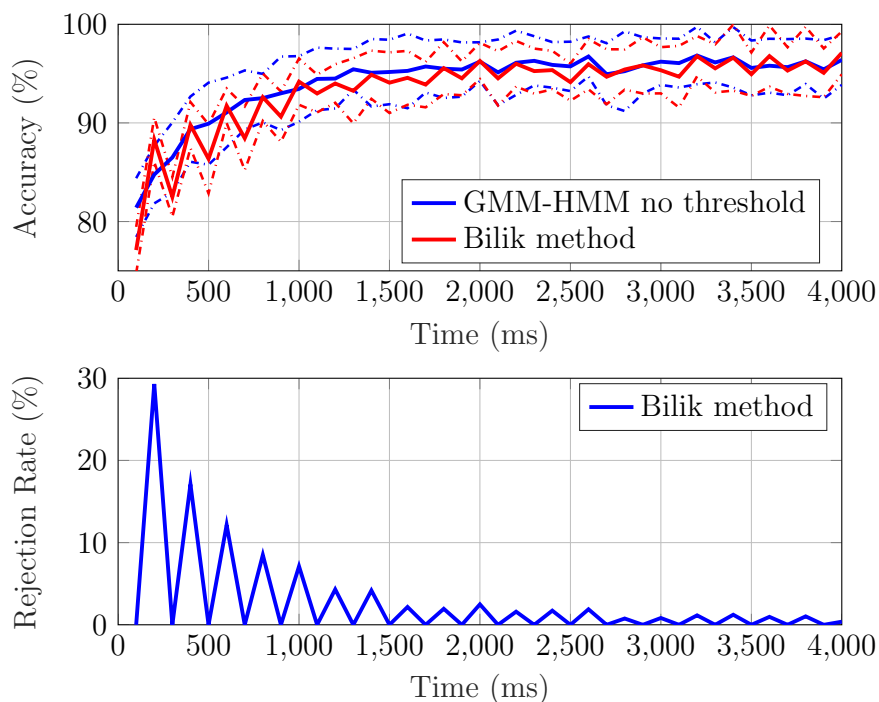

Figure 6: Comparison of classification accuracy and rejection rate vs. ToT with GMM-HMM method and Bilik's method trained for giraffe and zebra data only. The only form of rejection present in this graph is when Bilik's method obtains equal votes for both classes and from the graph it can be seen that this only happens on even intervals. This also explains the spiky nature of the graph for short ToTs, as observations with equal votes are more likely to be incorrect

500 ms ToT		
HMM-GMM/Bilik	**True Class**	
Classifier Out	**Zebra**	**Giraffe**
Zebra	86.13/78.9	7.79/9.04
Giraffe	13.87/21.1	92.21/90.96
1 s ToT		
HMM-GMM/Bilik	**True Class**	
Classifier Out	**Zebra**	**Giraffe**
Zebra	90.7/88.08	4.88/2.18
Giraffe	9.3/11.92	95.12/97.82
1.500 s ToT		
HMM-GMM/Bilik	**True Class**	
Classifier Out	**Zebra**	**Giraffe**
Zebra	93.4/89.58	3.78/3.11
Giraffe	6.6/10.42	96.22/96.89
2 s ToT		
HMM-GMM/Bilik	**True Class**	
Classifier Out	**Zebra**	**Giraffe**
Zebra	94.44/93.1	2.68/1.79
Giraffe	5.56/6.9	97.32/98.21

Table 6: 3-Fold cross validation results for giraffe and zebra classification

From Fig. 6 it can be see that both classification methods perform equally well. Based on the results, the GMM mixtures are capable of distinguishing between individual species based only on the micro-Doppler features. Specifically the higher Doppler shift in the leg movements of the zebra as well as the slight increase in the Doppler shift of the main body that is present when the animals are moving at approximately the same velocity.

The same type of 3-fold cross validation and testing was performed as described in Table 2, resulting in Table 6. It can be seen that if the classifier is trained to distinguish between two different species, and as long as the target was observed for a long enough time-frame, accurate classification is possible using any method. The longer a target is observed the more accurate the classifier. A preferable approach would be to jointly fuse the macro behavioural features (walking/running/grazing) of animals that can be determined from radar track data, as well as micro-Doppler features. This could be a topic of future research.

8. Conclusion

This paper shows that accurate human-animal classification of slow moving land based targets is possible when observed in their natural habitat. The classification performance is dependent on the ToT. It should be noted that since the diversity of the dataset on which this study is based is somewhat limited, further data is required to definitively characterise the performance of the classifier and identify possible shortcomings. Using the currently available dataset, the approach used produces promising results. It is shown that human-animal classification is possible with reasonable accuracies while the targets are in their natural environments at ranges exceeding 1 km.

The results obtained for the GMM-HMM classifier are not statistically significantly better than Billik's method for any specific ToT; however, the GMM-HMM consistently performs better over all ToTs. As mentioned in Section 7, if only the AIC score were to be used to determine model order, this dataset would not contain enough transitions per fold in order to accurately estimate the HMM transition probabilities between GMM components. This is a form of over-fitting, since in addition to the GMM parameters determined by the AIC, HMM transition parameters are added, resulting in an overly complex model. Hence, by using a validation set, the best model order was determined in Section 7.1.1. However the performance seems to be insensitive

to model selection choices over a fairly broad band close to the best model order.

The GMM-HMM model was designed based on previous work (van Eeden et al., 2015) where it was shown that the cepstrum for human targets would repeat in a periodic manner and it was believed that this periodicity would be present in many other species. Even if individual mel-cepstrum observations were similar the state transitions over time would serve as a discriminating feature. Although the HMM-GMM method seems to outperform Billik's method, the performance increase is not particularly marked. This can be better understood when the bi-directional KL-distances between the GMM components are investigated. The bi-directional KL-distance of less than one would mean that the specific components could be merged. Upon studying the KL-distances for the Giraffe Zebra GMM components it was found that the smallest KL-distance was approximately 2 and the average was approximately 50. This implies that the cepstral features for different animals do not overlap within the data set. As such the cepstral information contained in the GMM provides most of the discriminative power of the features. On the other hand since there are very few transitions that overlap between different species (classes) the transitions do not significantly contribute to the classification results. It is believed that with a more representative data, clusters would overlap more and the contribution of state transitions in the GMM-HMM model could contribute to improved performance of the GMM-HMM classifier.

Regarding future work, Dabrowski & de Villiers (2015a,b) have shown how movement models from tracking systems could be used to classify targets. This work could be adapted and included in the classification model for tracked animals. Furthermore, the effect of adding classification results into the tracking system for more accurate association and tracking should be investigated. Having accurate classification could significantly improve model selection for tracking systems and reduce the prevalence of false tracks. Accurate tracking should also allow repeated classification to take place on a target, thereby allowing for improved classification results over time. The classification method suggested in this work improves accuracy of classification for short ToTs (between 150 ms and 300 ms). However, if the target is observed for a very short time (shorter than 50ms) any temporal or spectral based methods would struggle, owing to the time scales at which human or animal motion occur. As such an approach such as that suggested in Nanzer & Rogers (2009) should be investigated.

9. Acknowledgement

References

Bilik, I., Tabrikian, J., & Cohen, A. (2006). GMM-based target classification for ground surveillance Doppler radar. *IEEE Transactions on Aerospace and Electronic Systems*, *42*, 267–278. doi:10.1109/TAES.2006.1603422.

Bishop, C. (1994). Novelty detection and neural network validation. *IEE Proceeding on Image Signal Processing*, *141*, 217–222.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*.

Boulic, R., Thalmann, N. M., & Thalmann, D. (1990). A global human walking model with real-time kinematic personification. *The visual computer*, *6*, 344–358.

Challa, S., & Pulford, G. W. (2001). Joint target tracking and classification using radar and ESM sensors. *IEEE transactions on Aerospace and Electronic Systems*, *37*, 1039–1055.

Clemente, C., Pallotta, L., Maio, A. D., Soraghan, J. J., & Farina, A. (2015). A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-Zernike polynomials. *IEEE Transactions on Aerospace and Electronic Systems*, *51*, 417–430. doi:10.1109/TAES.2014.130762.

Dabrowski, J. J., & de Villiers, J. P. (2015a). Maritime piracy situation modelling with dynamic Bayesian networks. *Information fusion*, *23*, 116–130.

Dabrowski, J. J., & de Villiers, J. P. (2015b). A unified model for context-based behavioural modelling and classification. *Expert Systems with Applications*, *42*, 6738 – 6757.

De Witt, J., Alahmadi, M., & Alzamil, A. (2012). Design and use of a mobile, X-band, high range resolution, radar research facility, .

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)*, *39*, 1–38.

Deng, L., & O'Shaughnessy, D. (2003). *Speech processing: A Dynamic and Optimization-Oriented Approach*. CRC Press.

Duffy, R. (2016). War, by conservation. *Geoforum*, *69*, 238–248.

Fairchild, D. P., & Narayanan, R. M. (2014). Classification of human motions using empirical mode decomposition of human micro-Doppler signatures. *IET Radar, Sonar Navigation*, *8*, 425–434. doi:10.1049/iet-rsn.2013.0165.

Fioranelli, F., Ritchie, M., & Griffiths, H. (2015). Classification of unarmed/armed personnel using the netRAD multistatic radar for micro-Doppler and singular value decomposition features. *IEEE Geoscience and Remote Sensing Letters*, *12*, 1933–1937. doi:10.1109/LGRS.2015.2439393.

Fioranelli, F., Ritchie, M., Gurbuz, S. Z., & Griffiths, H. (2017). Feature diversity for optimized human micro-Doppler classification using multistatic radar. *IEEE Transactions on Aerospace and Electronic Systems*, *PP*, 1–1. doi:10.1109/TAES.2017.2651678.

Hershey, J. R., & Olsen, P. A. (2007). Approximating the Kullback Leibler divergence between Gaussian mixture models. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (pp. IV–317). IEEE volume 4.

Jordan, T. S. (2016). Using convolutional neural networks for human activity classification on micro-Doppler radar spectrograms. In *SPIE Defense+ Security* (pp. 982509–982509). International Society for Optics and Photonics.

Kim, Y., & Ling, H. (2008). Human activity classification based on micro-Doppler signatures using an artificial neural network. In *2008 IEEE Antennas and Propagation Society International Symposium* (pp. 1–4). doi:10.1109/APS.2008.4619933.

Kim, Y., & Ling, H. (2009). Human activity classification based on micro-Doppler signatures using a support vector machine. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 1328–1337. doi:10.1109/TGRS.2009.2012849.

Kim, Y., & Moon, T. (2016). Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, *13*, 8–12. doi:10.1109/LGRS.2015.2491329.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*, 159–178.

Li, J., Phung, S. L., Tivive, F. H. C., & Bouzerdoum, A. (2012). Automatic classication of human motions using Doppler radar. *The 2012 International Joint Conference on Neural Networks*, (pp. 1–6).

Ljung, L. (1999). *System Identification*. Wiley Online Library.

Mashala, P. (2013). Developing farmers feel the scourge of stock theft. *Farmers Weekly*, . [Online; posted 10-May-2013].

McLoughlin, I. V. (2016). *Speech and Audio Processing: A MATLAB-Based Approach*. Cambridge University Press.

Molchanov, P., Astola, J., Egiazarian, K., & Totsky, A. (2012). Classification of ground moving radar targets by using joint time-frequency analysis. In *Radar Conference (RADAR), 2012 IEEE* (pp. 0366–0371). IEEE.

Molchanov, P. O., Astola, J. T., Egiazarian, K. O., & Totsky, A. V. (2011). Target classification by using pattern features extracted from bispectrum-based radar Doppler signatures. In *Radar Symposium (IRS), 2011 Proceedings International* (pp. 791–796). IEEE.

Nanzer, J. A., & Rogers, R. L. (2009). Bayesian classification of humans and vehicles using micro-Doppler signals from a scanning-beam radar. *IEEE Microwave and Wireless Components Letters*, *19*, 338–340. doi:10.1109/LMWC.2009.2017620.

Nebabin, V. (1995). *Methods and Techniques of Radar Recognition*. Artech House radar library. Artech House Inc. URL: `https://books.google.co.za/books?id=NBNTAAAAMAAJ` (pp. 106-125).

Padar, M. O., Ertan, A. E., & ĝatay Candan, Ç. (2016). Classification of human motion using radar micro-Doppler signatures with hidden Markov models. In *Radar Conference (RadarConf), 2016 IEEE* (pp. 1–6). IEEE.

Qi, F., Li, Z., Liang, F., Lv, H., An, Q., & Wang, J. (2016). A novel time-frequency analysis method based on HHT for finer-grained human activity using SFCW radar. In *Progress in Electromagnetic Research Symposium (PIERS)* (pp. 2536–2539). IEEE.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*, 257–286.

Scheme, E., & Englehart, K. (2015). A comparison of classification based confidence metrics for use in the design of myoelectric control systems. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 7278–7283). doi:10.1109/EMBC.2015.7320072.

Van Dorp, P., & Groen, F. (2008). Feature-based human motion parameter estimation with radar. *IET Radar, Sonar & Navigation*, *2*, 135–145.

van Eeden, W., de Villiers, J., Nel, W., Kloke, K., & Blasch, E. (2015). A comparative cepstral based analysis of simulated and measured S-band and X-band radar Doppler spectra of human motion. In *Radar Conference, 2015 IEEE* (pp. 283–288). IEEE.