Template Based Micro-Doppler Signature Classification

Graeme E. Smith*, Karl Woodbridge and Chris J. Baker
Department Of Electronic And Electrical Engineering
University College London, London, UK
Tel: +44 (0)20 7679 0153
Email: g.smith@ee.ucl.ac.uk

Abstract—The micro-Doppler signature of a target is a time varying frequency modulation imparted on the radar echo signal by moving components of the target. Battlefield radar output the baseband signal as audio and soldiers listening on headphones are able to identify the target from its micro-Doppler signature. Automation of this capability is desirable for improved reliability and reduction in classification time. For the first time dynamic time warping (DTW), a speech recognition technique, has been applied to the problem. Its performance has been compared with the common k-nearest neighbour (k-NN) classification method since both approaches utilise a template library.

Index Terms—Radar target recognition, Doppler radar, Pattern classification

I. INTRODUCTION

The benefits of using the micro-Doppler signature for target identification include: reduction in processing overhead compared with high range resolution (HRR) imaging systems where the signature is considered noise that must be removed [1]; improved classification time and accuracy compared to human operators [2], [3]; reduced dependence on expensive equipment as low range resolutions are acceptable; and the possibility of retrofitting to older systems.

To date attempts at automating micro-Doppler signature classification have relied on techniques that ignore its time varying nature. However, this discards information that may assist the classifier. Speech processing algorithms exploit the time variance in speech patterns to classify signals and identify words. Applying these algorithms to the micro-Doppler signature could allow its time variation to be exploited in a similar manner.

The micro-Doppler signature is a frequency modulation of the radar echo caused by moving components of a target with a cyclic nature that has been modelled for basic types of motion [4]. The micro-Doppler signature of a sinusoidally vibrating point scatterer is [5]

\[ f_D(t) = \frac{2\pi}{\lambda_c} [2v_r + 2D\omega \cos(\xi) \cos(\omega t + \psi)] \]  

(1)

where \( v_r \) is the radial velocity, \( D \) the vibration magnitude, \( \omega \) the angular frequency of vibration, \( \psi \) the starting phase of the vibration, \( \xi \) the angle between the plane of vibration and the radar line of site (LOS) and \( \lambda_c \) the wavelength of the carrier signal. If there is no vibration then (1) reduces to the standard expressions for Doppler shift [6]. The cyclic nature of this signature is clear despite the simplicity of the target—the \( \cos(\omega t + \psi) \) term provides a sinusoidal oscillation in frequency with a period \( T = \frac{2\pi}{\omega} \) that depends on the target, for higher frequency vibrations \( T \) is shorter. It is impractical to derive equations for the signatures of more complex targets but the cyclic nature may be seen by taking a short-time Fourier transform (STFT) of the received signal. Fig. 1 shows the spectrogram for the three targets used in this study. The cyclic nature of the signature is most clear in the personnel spectrogram—the swinging limbs of the person may be considered sinusoidally oscillating scatterers that will generate micro-Doppler frequencies in accordance with (1). For normalized time less than \( \approx 0.4 \) a sinusoid is visible in the spectrogram.

Several of the terms in (1) have a random nature. The initial phase of the vibration depends on when it is first detected and the angle to LOS depends on the radar-target geometry that differs for each encounter. Even the parameters to be used for classification are unclear. The vibration frequency may seem a good choice but it too depends on the specific encounter. If several vibrating scatterers were used to represent the limbs

\[ \lambda_c \]

\[ \xi \]

\[ \omega \]

\[ \psi \]

\[ T \]

\[ f_D(t) \]

\[ \frac{2\pi}{\lambda_c} \]

\[ 2v_r + 2D\omega \cos(\xi) \cos(\omega t + \psi) \]
of a person then the vibration frequencies would differ for walking and running. Despite the variations the target class is still personnel in both situations. Any classifier used will need to be robust to these difficulties.

II. METHODS OF CLASSIFICATION

Speech signals are a form of time series data that, broadly, are either classified by probabilistic models or template matching algorithms. In the former approach models such as linear-Gaussian state-space models and hidden Markov models [7] are used to assess the likelihood of a speech signal being generated by a particular word. The latter approach is simpler with unknown data frames being classified against a pre-classified library.

This study makes use of two template classifiers: \( k \)-nearest neighbours (\( k \)-NN), the most common template classifier; and dynamic time warping (DTW) a more advanced technique that allows for some variance between the test and reference signals. DTW has been chosen as it simple to implement and, as shall be discussed, is robust to some of the difficulties outlined in section I.

A. \( k \)-Nearest Neighbours

\( k \)-NN classification works through a simple system of majority voting. A test data frame—a series of regularly sampled feature vectors—is classified by finding the \( k \) reference frames nearest to it (the nearest neighbours) from the library [7]. The test frame is attributed to whichever class is most common among the nearest neighbours. The nearest neighbours are found by treating the test and reference frames as vectors and finding the Euclidean distance between them. The reference frames with the \( k \) smallest Euclidean distances are selected from the library.

The setting of \( k \) is critical to performance and forms a training stage in which various values are tried to see which gives the best performance. Typically hold one back cross validation [7] is used during training. \( k \) must be odd to prevent a tie in the majority vote and one greater than the number of classes to ensure each class has fair representation in the vote. Again to ensure fairness when voting there must be at least as many entries as \( k \) per class in the library.

It was not expected that \( k \)-NN would perform well when classifying micro-Doppler signatures. The problem of random starting phase, outlined in section I, means it is possible for the same target to generate test and reference frames with substantially different values at each sample. Additionally, because each frame is treated as a vector \( k \)-NN is limited to scalar values for feature vectors restricting the radar data used to incoherent intensity values with reduced information content.

B. Dynamic Time Warping

A comprehensive description of the DTW algorithm is given in [8]. Briefly, the algorithm may be considered a method for computing the normalized global distance between two data frames. To prevent the alignment problems encountered by \( k \)-NN the samples of each frame are expanded or compressed in time by DTW to align similar features in the frame prior to comparison [9]. It is even possible to compare frames of different length as can occur when the rate of data generation is variable.

If the reference frame contains \( M \) feature vectors \( \{r_m\} \) and the test frame \( N \), \( \{t_n\} \), then the normalized global distance to any point along the two frames is [8]

\[
D(r(m), t(n)) = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} d(r(m), t(n)) \tilde{W}(m,n)}{N(W)}
\]

where \( D(r(m), t(n)) \) is the normalized global distance to the point \( r(m), t(n) \); \( d(r(m), t(n)) \) is the local (Euclidean) distance between points in the reference, \( r(m) \), and test, \( t(n) \), frames; \( \tilde{W}(m,n) \) is a weighting for the arc of path to \( r(m), t(n) \); and \( N(W) \) is a normalization factor that is a functional of the weights. It is apparent that there are many paths to the point but the desired one is that which gives the lowest normalized global distance, \( \tilde{D} \), defined as \( D(r(M), t(N)) \) minimized over \( M, N, r(m) \) and \( t(n) \) [8].

Fig. 2 shows the optimal path between two simple frames found by applying DTW. Each point in the test and reference series is a scalar representing the amplitude of a signal. The intensity plot indicates the distance between the points of the two frames and green line is the optimal, or warping, path. It can be considered that the frames are projected onto the warping path before comparison.

DTW was expected to outperform \( k \)-NN because it is robust to the classification difficulties outlined in section I [9]. It can also operate on coherent complex samples with higher information content.

III. DATA ANALYSIS & EXPERIMENTAL METHOD

Data was available for three classes of target: tracked vehicles, wheeled vehicles and personnel. Although there was data for each class at various velocities and orientations a single test file for each was selected in which the velocities of the vehicles were comparable and all targets were moving radially towards the radar. Each data file contained 80,000 complex samples of the received signal. Fig. 1 shows the
spectrogram for each class—it can be seen that each class’s signature is distinct.

The duration of the available data is far longer than a typical radar dwell. To increase realism the data was divided into sections with durations comparable to the dwell time of scanning radar. Each section was then considered an individual data frame some of which were used to form the template library with the remainder used as the test set. The classification tests were repeated for frame durations of 8, 16, 32, 48 and 64 ms.

The selection of frames for the template library was based on the decorrelation time of the signal. An average value of decorrelation time was found by considering each frame in turn. One standard deviation was then subtracted from the average to give the time, \( t_{\text{Decorr}} \), used for selecting frames for the library. From the first frame every frame starting after an interval of \( t_{\text{Decorr}} \) was added to the template library. If this resulted in a library containing insufficient entries for the \( k \)-NN classifier \( t_{\text{Decorr}} \) was reduced by one frame duration until the library had sufficient entries. Unselected frames were used as test data. Multiple values of \( t_{\text{Decorr}} \) were found with decorrelation considered to be when correlation fell below 0.6, 0.7, 0.8 and 0.9.

The classification tests were performed for all permutations of frame duration and decorrelation time. Performance was measured as the probability of correct classification.

The study used three classes so the minimum value of \( k \) was 5 and the upper value was capped at 19, forcing each library to have at least 21 entries per class, see section II. The upper cap prevented the test set becoming too small when long frames were used. DTW did not require such large libraries but the same ones were used to allow comparison. While \( k \)-NN could only be applied to incoherent data DTW was applied to coherent and incoherent.

IV. DISCUSSION OF RESULTS

The variation of \( t_{\text{Decorr}} \) with decorrelation point is shown in Fig. 3. \( t_{\text{Decorr}} \) was almost identical for wheeled and tracked vehicles but substantially different for personnel. As was expected \( t_{\text{Decorr}} \) decreased as the decorrelation point rose for all classes. It should be noted that in many instances \( t_{\text{Decorr}} \) had to be reduced as described in section III.

Fig. 4 compares the average classification time, \( < t_{\text{class}} > \), of \( k \)-NN and DTW. For both methods \( < t_{\text{class}} > \) increases as the library’s decorrelation point rises. This was expected since libraries with a higher decorrelation point—the correlation at which the signal was considered decorrelated—contain more entries so more comparisons must be made per classification. Varying the frame duration has little effect on \( < t_{\text{class}} > \) for the \( k \)-NN classifier while for DTW there is large variation. For the lowest decorrelation point the longest frame takes around forty times longer than the shortest to classify. Most significantly the DTW algorithm is much slower than \( k \)-NN. For \( k \)-NN the maximum value of \( < t_{\text{class}} > \) is 0.013s while for incoherent DTW the minimum value is 3.3s and the maximum 212s and for coherent DTW the minimum and maximum are 3.5s and 224.4s respectively. DTW is slower because the underlying algorithm is more complex to implement and relies on a recursive routine to find the optimum warping path that searches a large search-space—a substantial overhead compared to just finding the Euclidean distance. As the frame length increases search-space increases also explaining the increase in \( < t_{\text{class}} > \) with frame duration.

The probability of correct classification averaged over the three classes, \( < P_{\text{class}} > \), is shown in Fig. 5. DTW is significantly more accurate than \( k \)-NN: \( < P_{\text{class}} > \) never falls below 0.8 for incoherent and coherent DTW and never rises above 0.6 for \( k \)-NN. There is little difference, around 0.01, in the maximum of \(< P_{\text{class}} >\) between the two types of DTW. A superior metric to \( < P_{\text{class}} > \) for understanding the
classifier performance is the confusion matrix since it allows visibility of the probability of correct classification for individual classes. Tables I–III show a representative confusion matrix for each classifier used in the study (the frame duration is 48 ms and the decorrelation point 0.9). The \textit{k}-NN classifier appears more likely to select the personnel class regardless true class of the input data; the DTW classifiers are move even in their selections. There is little between the two DTW techniques. The coherent classifier is a little better at identifying wheeled vehicles but slightly worse for tracked vehicles and personnel.

V. CONCLUSIONS & SUMMARY

This study has shown that speech recognition techniques have good potential for micro-Doppler signature classification. DTW has been shown to significantly out perform the benchmark \textit{k}-NN classifier for accuracy, but is much slower. Between coherent and incoherent DTW there is no significant difference in classification accuracy. Coherent classification takes a little longer to perform suggesting incoherent processing to be superior. This result was surprising and future research will investigate whether incoherent processing remains preferable when more target classes are introduced. When compared with the performance of human operators reported in [3] DTW provides superior accuracy of classification and \textit{k}-NN comparable accuracy.

The performance of the classifier depends on the decorrelation point of the library and the duration of the data frame. All classifiers are more accurate when library correlation is higher. The \textit{k}-NN classifier shows less improvement in classification accuracy with increased frame duration than DTW. The maximum frame duration considered was 64ms that is comparable with the dwell time of mechanically-scanned radar.

DTW is much slower than \textit{k}-NN. While \textit{k}-NN is able perform classification of long frame durations in a fraction of a second, but DTW takes several minutes. There is little difference between coherent and incoherent processing time. This is a substantial drawback to DTW classification since an aim of automation is improved classification speed. Bilik’s comparison of human and automated system [3] allowed the human operators 4s to perform classification. This is faster than DTW but much slower than \textit{k}-NN.

The DTW algorithm may be enhanced to improve its classification accuracy and speed. Investigating this will form the basis of future work. The other techniques highlighted in section I, which do not rely on template libraries and so are expected to execute much faster, will also be investigated.

ACKNOWLEDGEMENTS

The authors would like to thank Dr A. G. Stove and Thales UK Ltd at Crawley for supplying the test data used in this study.

REFERENCES
