Multi-Perspective and Biologically-Inspired NCTR

Hugh Griffiths / Alessio Balleri
Dept. Electronic and Electrical Engineering
University College London
Gower Street, London WC1E 6BT
UNITED KINGDOM

h.griffiths@ee.ucl.ac.uk / a.balleri@ee.ucl.ac.uk

ABSTRACT

Non-Cooperative Target Recognition by radar represents a substantial challenge to signal and image processing. In these lectures we consider two specific approaches to the problem. In the first we consider the additional information that may be exploited by observing a target from more than one perspective. It is shown that this can lead to a substantial increase in classification performance, though it is important that the additional perspectives are correctly chosen. In the second we consider sensor systems in nature which have benefitted from optimisation through millions of years of evolution. In particular, we show that echolocating bats use sophisticated waveforms and adaptive processing in the pursuit of their prey. Finally, we include some recent results on discrimination of human targets in terms of their acoustic micro-Doppler signatures.

1.0 INTRODUCTION

The ability to detect and locate targets on a 24 hour basis, over wide areas, regardless of weather conditions has long made radar a key sensor in many military and civilian radar applications. It is well recognised that the utility of the information supplied by a radar system would be hugely enhanced if targets could additionally be classified. This might be by type (e.g. to differentiate helicopters from civilian airliners from military jets) or even better, within type (e.g. a Chieftain tank from a T32 tank). It would be even more advantageous if this could be automated, at least partially. Huge quantities of imagery are produced by SAR and place unrealistic demands on a human image interpreter. Thus algorithmic aids are necessary if the data is ever to be fully exploited. However, this has proved far from straightforward and robust, and reliable classifiers able to cope with wide target and clutter sets do not exist. Most of the approaches to classification have concentrated on exploiting higher and higher spatial resolutions both in one [Ron97] and two [Nov97] dimensions. There has also been research examining the role of polarimetric information [Nov97] [Sad02]. However, there has been little examining the utility of angular diversity for improved classification performance [Shi05] [Xue02] [Run99] although it is clear that (i) there must be additional information in multiple perspectives and (ii) it is extremely easy for existing systems to acquire such data. In fact, to date, the variation of target signature with angle has been a source of performance degradation in single perspective classifications approaches. Alternatively, multi-perspective classification can be exploited by networks of co-operating radar systems [Bak03]. These have the added advantages of improved system sensitivity, tailored coverage and reduced vulnerability to electronic and physical attack.

There are no rigorous and universally accepted definitions for the terms ‘classification’ and ‘recognition’. In this lecture these terms are used interchangeably to address the ability of the radar system and subsequent processing to determine the class to which a target belongs (after defining a set of sub-populations such as a particular version of a tank). To perform Automatic Target Recognition (ATR) in a useful manner it is necessary to achieve a sufficiently high degree of confidence that a target has been correctly classified and hence it is extremely important to minimise the probability of misclassification. This is because in both military and civilian applications there are likely to be safety issues that demand high level of performance and a full understanding of its limits, variability and robustness.

ATR can be based either on 1-D signatures, i.e. features derived from a range profile of the target, or on 2-D imagery. 1-D signatures are often employed due to their simplicity in terms of implementation and signal processing, although they can lead to performance that is prone to low classification rates due to the likely inclusion of a clutter components which are scenario dependent. 2-D imagery is less likely to contain clutter (although there will still be a significant multipath or forward component) and hence would seem to guarantee a more detailed and authentic (but nevertheless still corrupted) representation of the target backscattering [Nov91]. This requires more sophisticated signal processing and, if the data is collected over an extended period of time (as in a frequency stepped radar system), corrections for irregular target movements. Such corrections can be far from straightforward especially because the detailed form of the motion corruption is commonly unknown.

When High Range Resolution Profiles (HRRPs) are used, slightly different target orientations due to rotational or translational motion of the target (assuming no range migration occurs) result in differences in relative range between scatterers which lead to interferences and bin fluctuations in the range profile. This is known as speckle. In addition to this, HRRPs collected at different aspect angle are likely to present differences due to shadowing; this occurs when one or more scatterers are occluded by other parts of the target (local shadowing) or by other targets (global shadowing) and therefore not illuminated by the incident waveform. As a consequence, a portion of the target is unavoidably masked and the signature further corrupted. These effects may result in a loss of information and may negatively affect classification performance .The employment of multiple perspectives is a way to overcome some of these effects and to exploit the additional information that is aspect dependent.

Here, the problem of radar target classification is examined for the case when more than one perspective or viewing angle of the target is available to the sensor. Using full-scale target signature measurements as the source data, it is shown how, multiple perspectives enhance classification performance. Classification performance is assessed both as a function of the number of perspectives and of the Signal-to-Noise Ratio (SNR). In particular we present three approaches to high range resolution profile multi-perspective classification to removes any possible bias that could be introduced by a single individual classifier.

We then look at the natural world to investigate what lessons may be learned from understanding how echolocating bats detect and select prey during foraging. A real feeding buzz (the sequence of pulses emitted by the bat) is analysed to show how waveform diversity is used to achieve a high degree of detection and classification performance. Details on how bats classify moving and static targets, such as insects or flowers of bat-pollinated plants are also given, with some recent results.

Although micro-Doppler target classification is the subject of another lecture we present some latest results on human recognition and motion classification by acoustic micro-Doppler signatures.

2.0 MULTI-ASPECT NCTR

In this section multi perspective target classification is exploited on a database consisting of HRRPs of real radar targets. Performance is assessed under the assumption that targets are already detected and tracked by the illuminating radar system and that the reflection signature to be recognised is a 1-D HRRP. The complete database used is collected at all angles (i.e. over 360°) with the radar system stationary and receiving the echoes from a target mounted on a rotating platform. During the period of collection and processing, a number of factors affect the range profiles, causing variability, although the orientation of the target and system parameters remains the same. These factors include measurement noise, together with rotational and translational range migration. Other possible causes that might affect radar measurements are clutter, jamming (signals transmitted in order to deliberately confuse the system attempting to detect it), and Electro-Magnetic Interference (EMI) caused, for example, by other communications or radar systems close to the radar. In this work we have attempted to minimise the effects of noise clutter, forward scatter and external interference. Although this is not wholly realistic, the purpose of the research is to determine whether or not multiple viewing perspectives improve classification performance, i.e. there is an in-built assumption that any improvement observed will also be apparent for targets in natural environments. However, due to the additional variability that will occur, this assumption will require careful testing and verification.

The time domain backscattering properties of radar targets can be used to form feature vectors for subsequent classification. These backscattering properties can take the form of either HRRPs or inverse synthetic aperture radar (ISAR) imagery. Both are well known methods of expressing the backscatter from a radar target [Weh95]. HRR data of full-scale real vehicle targets from the Thales airborne data acquisition system (ADAS) has been used as the signature set to examine M-P target classification. The data collection geometry is depicted in Figure 1. Two stationary trihedral targets are located in front and behind a turntable,

and provide a means for computation of the radar point spread function. For collection of the signatures for classification the two rotating corner reflectors are replaced by vehicles. In the figure two moving reflectors are located on opposite sides of the rotating turntable. During the 360° target rotation, HRR profiles are collected at an angular interval of 2 min of arc. High range resolution is achieved by transmitting a linear frequency-modulated (LFM chirp) pulsed waveform [Sul04]. On reception pulse compression is utilised to generate an HRRP. For example, for a pulse duration t_p, the chirp waveform is given by

$$t_p = \frac{1}{\sqrt{t_p}} rect\left(\frac{1}{t_p}\right) e^{jt(\omega_c + \pi\gamma t)}$$

where ω_c is the carrier frequency and γ the rate of frequency variation with time. The instantaneous frequency is

$$f(t) = \frac{1}{2\pi}\frac{d}{dt}\left(\omega_c t + \pi\gamma t^2\right) = \frac{\omega_c}{2\pi} + \gamma t$$

Therefore, the bandwidth is $B = \gamma t_p$. The pulse compression ratio (PCR) is equal to the time-bandwidth product Bt_p and represents the range compression factor (i.e. the ratio of the pulse length before time compression to its length after compression). The PCR is a gauge of the range resolution improvement given by pulse modulation. The ADAS data chirp parameters are shown in **Error! Reference source not found.**. These enable HRRPs with a range resolution of just over 30 cm after weighting. The format of the data is such that a series of pulses can be collected where the bandwidth of each pulse is stepped by half the bandwidth of any of the individual pulses. These can be combined to provide a very high range resolution signature with a resolution of 6 cm. However, here we limit investigations of M-P classification performance to the 30 cm range profiles only. Figure 2 shows an example of the form of the radar data. It depicts the intensity modulated time history of HRRPs from the two stationary and the two rotating trihedral reflectors for the geometry depicted in Figure 1. The two stationary trihedrals show a constant response at near and far range as expected. For the two rotating trihedral targets, when the line-of-sight is on the trihedral bisector, a peak of reflection occurs. This is consistent with the expected theoretical response [Kno85]. As the trihedral targets rotate, the backscattered field decreases progressively until a point is reached where there is a peak of specular reflection. This is a reflection from one of the sides making up the trihedral which is orthogonal to the illuminating radar system (i.e. it faces the radar beam and looks like a flat plate reflector). At increasing rotation angles the RCS of the target drops because the orientation of the trihedral is such that it tends to reflect incident radiation away from the radar. This angular dependency of the RCS of a known reflector, such as a trihedral, begins to illustrate how the backscattering properties of real targets may vary with the orientation of observation. For example, if a target has part of its structure that mimics a trihedral it will only show this feature over a limited angular range. Thus in an M-P environment, different angular samples of a target signature should improve the likelihood of observing a corner or corner-like reflector. Such shapes are common on many man-made structures and are often quite dominant features that may prove useful for classification. In Figure 3, the complete angular ambit of range profiles spanning 360° from a Land Rover vehicle rotating on the turntable is shown. This highlights a number of different scattering behaviours: the strong peaks from specular reflections (0°, 90°, 180°, . . .) appear over a very limited angle range and obscure nearby point like backscattering. Corner-like returns can be observed at a far range (6 m) for two range angular spans ([10°– 60°] and [130°–180°]). These returns correspond to the trihedral-like structures formed at the rear of the Land Rover. This is a vehicle without the rear soft top and has a metallic bench seat that makes a corner where it joins the rear bulkhead of the driver's cabin. At 8 m range there is a double bounce return corresponding to one of the corners. This type of effect increases the information that can be

extracted which would be otherwise impossible to reconstruct by a traditional single-perspective approach. Classification has been performed on a subpopulation consisting of a total of three ground vehicles (A, B, and C) exploiting multiple perspectives of HRRPs from each of the vehicles. For each target, there are 10 500 profiles spanning 360°. The data are divided into contiguous 10° sections, and 36 signatures randomly selected from each of the 10° sectors are given to the classifier to build the internal representation of the class. These signatures are then removed completely from the data set. The remaining profiles then provide the input to the classifiers.

2.1 Data pre-processing

After pulse compression and weighting, the signal is Fourier transformed to form the total range profile. Subsequently, the HRR profile is cut out of the illuminated area preserving 52 range bins, each with a sampled resolution of 30 cm, and an actual resolution of approximately 39 cm (because of the window applied). However, the data exhibit a low-level contribution from zero Doppler clutter (ZDC). ZDC is stationary or quasi-stationary clutter that contributes to the HRRPs. By observing a sequence of range profiles it becomes clear how the ZDC affects the classification process. Grass and trees show a relatively fast fluctuating term whose effect at DC can be observed over relatively small turntable rotation angles. This is in contrast to the more persistent response given by spatially more restricted scatterers such as static ground discontinuities and mounds of gravel. This also means that the ZDC can be quite different from one target measurement to another because as the vehicle approaches the turntable it modifies the local clutter structure and hence the ZDC component of the return in the signature. As a consequence, if ATR is implemented using the same data for both training and testing the classifier, the overall accuracy could be significantly upset by the ZDC contribution to backscattering, as the classifier would otherwise use the clutter as an input. Thus the ZDC has been estimated and subtracted. This has been performed using data derived in the frequency domain following the method outlined in [Sho98]. The mean value of ZDC over about 50° of turntable rotation is subtracted from those frequency signatures, and then the final HRRP is extracted. After ZDC subtraction, a mean filter over a neighbourhood of three echoes is applied to a sequence of range profiles. Since two consecutive HRRPs are spaced by only 2° of target rotation the non-coherent averaging partially averages out those phenomena that are independent of the geometry. In addition, the signal to noise ratio (SNR) is increased as demonstrated in [Sko80] by a factor $\sqrt{N}$ where N is the number of target echoes. Prior to non-coherent averaging, it is usually necessary to align the range profiles. However, here, because the input patterns for recognition are collected from the turntable measurements, no additional alignment is needed between training and test profiles. Finally, the signatures are normalised following [Zyw96] to make the target area in the HRRP more prominent. A target mask is generated to select the target only as part of the range profile by measuring the noise mean value μ and standard deviation σ. A threshold $\mu + 3\sigma$ is applied to eliminate the non-target zone. The target masks are used so that target-only data is presented to the classifier.

2.2 Feature extraction

In a typical pattern recognition problem it is often necessary to reduce the data dimension of the input of the classifier. This is mainly because of an intrinsic degree of redundancy within the data where there are regions where no information is contained. It also helps to reduce the very high quantities of data that otherwise have to be processed. Thus we may consider range profiles as providers of feature vectors that are to be separated. After focusing on those range resolution cells whose intensities depict the target backscattering, the number of elements representing the backscatter can be reduced with an information loss which is assumed negligible. Furthermore, dimensional reduction of this type also attempts to emphasise the differences between patterns and hence enhance classification performance. The PCA [Theo] is a statistical method that enables the data to be represented in a different vector basis such that it is possible to remove similarities (which therefore do not contribute to the classification process). After subtracting the mean $\bar{f}$ from each of the vectors of the training set *F* and producing a zero-mean set of data, the covariance matrix Cov can be formed

$$Cov(f) = \frac{1}{N}\sum_{n}(f_n - \bar{f})(f_n - \bar{f})^T$$

After calculating the Eigenvectors of $Cov(f)$, the *P*-most significant eigenvectors with the largest Eigenvalues are selected and form a new basis vector $V = (v_{1,}v_{2,},....,v_{P,})$. The test and training feature vectors can then be transformed as follows

$$f' = V^T(f_n - \bar{f})$$

The number of principal components PC = P is chosen as a function of the classification rate achieved. This usually becomes stable once the PCs necessary to fully describe the data have been selected. After testing the classifiers, their mean value correct classification rates (CCRs) are plotted against the number of principal components representing the feature vectors. For the data used here the dimension of the feature vector is 52. However, as can be seen in Figure 4, the probability of correct classification P_{cc} is almost unaltered until less than approximately 12 components. At this point they are no longer representative of the different target classes. Therefore, in order to depict the target with more robustness, the feature vectors used in the following sections are obtained using the first 21 principal components.

2.3 Multi-perspective classifiers

In order to remove any bias introduced by a single algorithm three classifiers have been implemented. These are: (1) a naive Bayesian classifier, (2) a nearest neighbour classifier and (3) a neural network classifier. All three are described in detail in [Ves07]. A traditional single-perspective classifier, after training the ATR algorithm with a set of templates, is tested with profiles collected from all orientations of the target. In an *N*-perspective scenario, the parameter that distinguishes the perspective topology is the vector $\Phi = \{\Delta\varphi_{i,j} : i, j = 1,..., N\}$ where the element $\Delta\varphi_{i,j} = \varphi_i - \varphi_j$ represents the angular displacement between the node i and j. For example in a two-perspective (2-P) scenario, the parameter that distinguishes the perspective node locations is their relative angular displacement $\Delta\varphi_{1,2} = \varphi_2 - \varphi_1$. Hence, after fixing $\Delta\varphi_{1,2}$, the 2-P classifier is tested with all possible pairs of HRR profiles displaced by that angle covering all the possible orientations of the target. Having a test set consisting of N profiles, the same number of pairs can be formed to test the 2-P classifier. As shown in Figure 5, to test the Multi Perspective (M-P) classifier, the sequence of all the possible pairs of HRR profiles displaced by $\Delta\varphi_{1,2}$ is given as an

input, and the accuracy for a 2-P classifier is subsequently computed. Thus to evaluate the 3-P recognition performance, after choosing the set $\Phi = \{\Delta\varphi_{1,2}, \Delta\varphi_{1,3}\}$, the classification accuracy is measured as the average correct classification over all the possible HRR profiles triples that could be formed from the M-P topology Φ. This procedure is extended for networks having four and five perspectives. The training set of representative vectors for each class is made up of 36 range profiles, taken approximately for every 10° of target rotation. The 'testing set' of each class consists of the remaining range profiles neglecting the templates. The angular displacement between nodes is not processed as information by the M-P classifier. Although this is only one of the possible approaches to M-P classification it is simple to implement. Furthermore, the computational burden is reduced to a minimum. Although the target has already been assumed to have been detected and tracked, the angular displacements F of the nodes are unknown to the network. Therefore, in case of a single sensor collecting multiple perspectives from a moving target, it is assumed that to combine the signatures, it is not necessary to estimate the motion parameters and the centre of rotation of the target. Here, we consider a network of radars consisting of N nodes, each single perspective j of the target is represented by the HRR profile X_j. Consequently, the set of signatures collected by the network is represented by $\{X_j : j = 1,.., N\}$ and is directly related to the vector of angular node displacements Φ. To test classification performance, the signatures $\{X_j : j = 1,.., N\}$ correspond to the HRRPs after applying PCA.

2.4 Multi-perspective classification performance

Classification performance is evaluated in a forced decision environment using confusion matrices (e.g. see [Tait]) where each column represents the instances in a predicted class, whereas the rows represent the output distribution for an actual class. In Table 2, the confusion matrices of the three classifiers are shown with respect to the number of perspectives involved in the decision processing. The single-perspective results represent the more usual single perspective classifier performance. It should be noted that even with a single perspective the classification performance is very good. This is a function of using turntable data together with a small number of sub-populations. However, it does allow the trends with a number of perspectives to be explored.

Table 2 clearly shows an improvement in classification performance, as the number of perspectives is increased. This trend is observed for all three classifiers and hence we can conclude that it is not a function of the classification process but because of the additional perspectives. Interestingly, the biggest increase in performance occurs when just one additional perspective is employed. A parameter to measure the goodness of a classifier is the CCR (i.e. the Probability of Correct Classification PCC on a finite test dataset). It is the mean value of the elements on the diagonal of the confusion matrix. In Table 3, the CCR of the three implemented classifiers are displayed using different numbers of perspectives. Again all three show the same trend of improvement in performance as the number of perspectives is increased.

Figure 6 shows this graphically and again highlights that the increase in classification performance is independent of the particular classifier approach used. Furthermore, it emphasises that the M-P improvements are not linear with the number of perspectives involved. On average, an 8.8% improvement in the correct classification rate, with respect to the single perspective classifier, is reached when two perspectives are used, whereas the benefit in relation to the 2-P classifier is 3.6% when three perspectives are used. It is 2.51% when going from three to four perspectives and finally 1.5% when going from four to five perspectives. This is a consequence of those signatures that relate to particular target orientations where specular reflections occur and the profile is thus dominated by only a very few scattering centres. These patterns are common to all the target classes and consequently are more easily misclassified. Nevertheless, for these target orientations, the decision confidence of the classifier is low and is therefore significantly aided by a second perspective. On the other hand, if both the first and second perspectives correspond to

those with reduced information signatures, a third perspective is less effective in improving performance. A third perspective has less influence to correct the weights of the others and thus the improvement from subsequent perspectives is progressively attenuated. This trend may be different if the perspective displacement of the nodes of the network is processed by the classifier as useful information. Nonetheless, the nonlinear relationship between CCR and number of radar perspectives suggests that only the employment of a small number of nodes in the network is necessary in order to achieve the best trade-off between system complexity and improvement in classification performance.

The benefits of using multiple perspectives to improve classification performance are now examined as a function of different SNR levels. The SNR of the data is measured as the distance in dB from the maximum magnitude in the HRR profile belonging to a particular target and the noise level after ZDC. From each target class '*i*', the noise level is measured in the non-target area before target normalisation. The target signal level is estimated as the mean value of the maximum target backscattering over the 360° range profile history and the $SNR_{class-i}$

$$SNR_{class-i} = \frac{1}{KN}\sum_{n=1}^{N}\sum_{k=1}^{K} 20\log\left(\frac{r_{n,k}}{\mu_n}\right)$$

where μ_n is the mean noise intensity for the n th, k -dimensional profile $RP_n = (r_{n,1}, r_{n,2}, ..., r_{n,K})$, and N is the number of profiles collected for the ith target class. The result of applying this procedure to the original data leads to an $SNR = (1/n_c)\sum_i SNR_{class-i} = 28.4dB$. The profiles are subsequently progressively corrupted with different levels of Gaussian noise applied to the I and Q channels, before matched filtering. Then, the SNR is measured as above, and classification attempted on PCs extracted from the 1D signatures. The original training dataset is preserved. As there was no bias introduced by any of the three classifiers only the FANNs classifier is used to examine performance as a function of SNR. In Figure 7, the CCR results for just the FANNs classifier are depicted against the number of perspectives used and the SNR levels. In all cases increasing the number of perspectives improves classification performance. In comparison to a single-perspective classifier, the CCRs remain stable above a SNR threshold of 17 dB and then a very rapid deterioration of the classification performance can be observed. This is because of the information loss caused by the increasing noise that corrupts the range profile. The relatively high SNRs at which this occurs can be partly explained by the fact that this is a combination of smaller magnitude and larger magnitude scatterers that provide the classification information. The smaller magnitude scatterers are more severely affected by the increasing noise and hence quickly have an impact on reducing performance. As the number of network nodes increases, the information is partially recovered with the help of other perspectives because the noise is not correlated and hence modifies the profiles differently. This result does highlight the importance of having adequate SNR such that small magnitude scatterers are preserved as an input to the classification process.

2.3 Summary

A novel radar target classification approach has been introduced. The method is based on the processing of real, full-scale target signatures collected at different perspectives in a network of monostatic radars. In a first single perspective stage, depending on the methodology adopted, the 1D imageries are processed in parallel, providing the following M-P stage with partial outputs eventually used for the final decision. The benefits of the M-P classifier implementation have been analysed and show a nonlinear but very clear CCR improvement with the number of perspectives. As the number of radar nodes increases, the M-P CCR stabilises, suggesting that most benefit accrues when using a small network of only two or three nodes and hence avoids any significant increase of system implementation complexity. For example, this enables a

simple approach to be adopted for a single monostatic radar that obtains multiple perspectives as it traverses an object or area of interest (provided sufficient angular diversity is obtained). The performance has also been investigated as a function of different noise levels. The incidence of multiple perspectives on correct classification performance improvement is observable for any value of the SNR range of values examined. In addition, the M-P benefits are more pronounced for low SNR levels that are for those SNR values that are typical of not controlled real-world experiments. Further research is required to investigate the effects of the separation of perspectives used to evaluate any relationship between angles of perspectives selected and the resulting classification performance. This is particularly important if single monostatic radar systems are to be used as it could minimize the time at which targets are illuminated.

3.0 BIOLOGICALLY-INSPIRED NCTR

In the natural world, echolocating mammals use waveform diversity as an inherent component of their normal behaviour. The constantly changing time and frequency structure, and location and direction of their transmitted signals represent a pro-active approach to interrogation of the surrounding environment. In addition, a multiplicity of processing streams collectively extracts information from received echoes to build up an accurate picture that is supplemented by long term 'experiential' memory. In this way bats, whales and dolphins are able to 'perceive' their environment, albeit in a very different form to humans, such that bats can autonomously navigate, feed, socialize and otherwise conduct their lives. Here we examine the use of diversity in natural sensing systems with a view to exploitation in synthetic sensors, with particular emphasis to target classification.

3.1 Waveform design

Echolocating mammals such as bats, whales and dolphins have been using waveform diversity for over 50 million years. In contrast, synthetic systems such as sonar and radar have been in existence for less than 100 years. Echolocating mammals vary parameters including the Pulse Repetition Frequency (PRF), power and frequency content of their transmitted waveforms. Recent developments in technology mean that it is now possible to replicate such diversity in radar and sonar systems – so that with modern digital technology high-precision wide-bandwidth waveforms can be generated and varied potentially on a pulse-to-pulse basis,. Thus echolocating mammals potentially offer valuable insights that may enable capabilities such as autonomous navigation and automatic target classification which have hitherto proved elusive. If we understand how bats exploit echolocation for autonomous navigation, collision avoidance and recognition, we can build this into synthetic systems. Potentially, this can provide a step function change in the utility of unmanned systems that can take advantage of the attributes of radar and sonar. Signal designs utilised by bats have been categorised by Jones and Teeling. The categorisation is based around signals emitted when bats are searching for prey: intra-specific (and indeed intra-individual) variation in call design can be substantial, and the scheme was introduced to illustrate patterns of convergent evolution. The following eight signal design categories may be summarised as follows:

- Most Old World fruit bats (*Pteropodidae*) do not use echolocation for orientation, and instead appear to rely largely on vision.
- Brief, broadband tongue clicks: the clicks are produced in pairs (one from each lip) by raising the tongue from the floor of the mouth. The clicks are of short duration (typically < 1 ms) and the mechanism of their production is quite different from that used in laryngeal echolocation by all other echolocating bats.
- Narrowband signals dominated by the fundamental harmonic: these calls have narrow bandwidths, and are relatively long (often > 5 ms).
- Narrowband multiharmonic signals: Each harmonic is narrowband, but several prominent harmonics

feature in the call. The dominant harmonic is usually not the fundamental.

- Short, broadband calls with a dominant fundamental harmonic. These calls are the typical 'chirps', or frequency-modulated (FM) calls that have dominated research on broadband echolocation.
- Short, broadband multi-harmonic signals.
- Long duration multi-harmonic broadband calls: most energy is in the second harmonic.
- Pure constant frequency (CF) signals are long duration signals whose dominant component has zero bandwidth and duration can be long (> 30 ms).

Call design varies within species and even within individual bats. In the next section we see that bat calls may change radically during insect capture. Prey in the air are searched for, detected and located by a series of echolocation pulses that ultimately ends in a 'terminal buzz'. During a terminal buzz, calls typically become shorter, are repeated at a higher rate, and have shorter intervals between them. In most bats, pulse-echo overlap is usually avoided during prey capture (because the calls become shorter as the prey is approached). The factors which determine the nature of the waveform are: *intensity, harmonic structure, frequency, bandwidth, duration, repetition time and duty cycle.*

3.2 Analysis of a feeding buzz

A real 'feeding buzz' is examined to highlight how a bat constantly exploits diversity in order to intercept and feed on prey. A typical situation is shown in Figure 8 in which the blue line traces the trajectory of a bat until it intercepts its prey (shown by the red dot). The black circles show the position at which a call (or waveform) was emitted.

A number of simple observations may be made:

1. The trajectory undergoes a significant and non-linear manoeuvre. This may be thought of as an example of spatial diversity where the bat is gathering improved information by taking multiple and different perspectives of the environment and target. The significance of spatial diversity is not clear but improved information will be gleaned by exploitation of multiple perspectives

2. Calls or waveforms are not transmitted at a constant PRF. This might be thought of as sampling or time diversity. Again the logic and physical explanation of this is unclear. It is known that the bat expends significant energy in making a call. Therefore it is likely that the number and frequency will be kept to a necessary minimum. Coherent processing of the type typically employed in radar will be different in bats.

3. There are two very distinctive PRF regimes, the first being relatively low and the second in the final intercept phase being much higher. Clearly, the need for acquiring information at a higher rate is necessary as the bat approaches the target. The precise nature of this information is again unclear but could be associated with the need for fine spatial positioning of the target to enable an effective intercept.

Figure 9 plots the feeding buzz sequence emitted by an *Epetesicus nilssoni* bat while first searching and then attacking a slow moving target. During data collection the bat was constantly changing its orientation, such that it viewed the target over a total angle range of approximately 270 degrees, and gradually getting closer and closer to the target.

It is evident that the buzz sequence can be divided into two main phases. The first phase, typically called the *searching phase*, corresponds to the initial part of the buzz in which it is believed the bat is looking for a possible target and performing classification. Ideally, this phase should include the task of target detection although, for this particular experiment, not enough information on the experiment is available to decide

whether the bat had already detected the target or not before the recording started. This phase is composed of echolocation calls characterised by amplitudes of about 0.3 V which are separated in time from each other by a PRI (Pulse Repetition Interval) of about 0.2 sec. The second phase, called the *terminal phase*, is the last part of the buzz and corresponds to the time when the bat attacks the prey that it has selected. Pulses belonging to this phase of the buzz are separated by a much lower PRI (about 1 msec) and present amplitudes up to 5-6 times lower than those recorded in the searching phase. It is believed this is because in the terminal phase the bat is closer to the target and therefore can operate with lower power transmissions in order to save precious energy that can instead be used to transmit with a higher PRF to keep tracking the target before the final attack.

Figure 10 shows the spectrogram, the normalised mean spectrum and the ambiguity function with range and Doppler cuts for the first pulse in the searching phase. The pulse is characterised by a time duration of about 9 msec and it is composed of three non-overlapping harmonics (in frequency) with peaks in power at about 30 kHz, 60 kHz and 90 kHz, respectively. Most of the energy is concentrated in the fundamental harmonic which can be fitted well to a hyperbolic function. The mean spectrum clearly shows the three non-overlapping harmonics which present a non-constant power/unit bandwidth that is typical of non-linearly modulated chirps. The range cut of the ambiguity function shows that the bat is operating with a range resolution of about 1.8 cm. The Doppler cut instead present a very narrow peak indicating that the bat is trying to acquire fine Doppler information on the target. This is in very good agreement with the previous literature showing that Doppler information is critical for the task of classification, and supports the hypothesis that target classification takes place in the searching phase.

Figure 11 gives the same plots for a pulse in the final phase that is composed of two overlapping (in frequency) harmonics. The fundamental harmonic is well fitted to a linear chirp that spans frequencies between 75 kHz and 25 kHz, while the second harmonic decreases from about 90 kHz to 55 kHz. The pulse length is reduced to 4 msec, more than 50% with respect to the pulses described above, in order to conserve energy and avoid eclipsing. At the same time the degree of hyperbolic curvature increases to extend the bandwidth and improve range resolution.

As expected, the range cut of the ambiguity function is now much narrower that the one associated with the pulse extracted from the searching phase, and shows that the bat now operates with a range resolution of about 8 mm. The Doppler cut instead has become much larger providing a poorer Doppler resolution. Overall, it is common to waveforms in the terminal phase to provide tolerance to any differential Doppler. This is likely to indicate that the bat at this stage has already gained the Doppler information for classification and is gathering the range information before the final attack. Unfortunately, there is no direct knowledge of how the bat was changing its position and orientation with respect to the target during this specific recording.

It is clear that diversity, even at this relatively coarse level of detail, is being utilized and utilized every time sensing takes place. Whilst we can speculate about the reasons why the bat operates in this way it is probably safe to say that it is done for good reason as is the habit of ‘Mother Nature’. Indeed, the success and rate at which bats are able to take insects on the wing is testament to the effectiveness of diversity exploitation. This seems a powerful indicator that by understanding how and why natural systems use diversity we can create a much more able capability in synthetic sensing systems than currently seems possible. The more demanding question that remains to be addressed is what are the appropriate adjustments that need to be made to maximise classification performance? This would be an interesting and challenging topic for future research.

3.3 Classification of insects

Classification in bat echolocation is defined as the use of patterns of information in echoes to categorize targets [Sch03]. Classification of insects by bats that emit continuous frequency (CF) signals has been at the centre of various research work.

Figure 12 shows an example of a CF waveform that is composed of three main harmonics at about 30 kHz, 60 kHz and 120 kHz that are initiated and terminated by frequency modulated sweeps. It has been argued that the portions at the beginning and end of the waveform function in target localization, while the long constant frequency component eases the detection and the classification of targets [Tian97]. Indeed, this type of waveform is not being used by any radar system and the fact that mammals perform amazingly by using it makes worth exploiting the impact that its use would have on radar target classification.

In flight most species of CF bats, such as horseshoe bats, lower their call frequency in relation to their flight speed in order to compensate for Doppler shifts induced by their own movement so that echoes always return at a best suited frequency for the hearing of the bats [Schu79]. Bats that emit CF signals can detect and classify fluttering insects from amplitude and frequency modulations of the echoes caused by the movement of the insect's wing [Ost84] [Emde86]. These modulations, called 'acoustical glints', potentially provide information about the wing beat rate and flight angle, and are very specific for each species of insect. Glints turn out to be useful not only for the detection of the prey but also for its identification. Amplitude modulations are the result of the fact that echo strength is strongest when the insect wings are perpendicular to the sound source and gets weaker as the insect wing moves away from the perpendicular position. Interestingly, the precise timing of the glint depends on the angular orientation of the insect [Sch83], then amplitude modulations give information about the target elevation that affect the timing of the glint production [Kob90].

Figure 13 shows an example of amplitude modulated echo from a flying moth *Autographa gamma* at three different angles with equidistant starting phase. In the figure the amplitude modulation is given with its corresponding wingbeat phase at each time. In the experiment the loudspeaker was placed at the same height as the insect. The plot shows that at 90 degrees the glint is produced at the top of the stroke, i.e. when the moth wings are perpendicular to the sound source. At 0 degree (frontal) and 180 degrees (rear), the glint occurs two phases after the upstroke suggesting that, in both cases, the same part of the wings produces the glint [Sch83].

The wing movement towards and away from the receiver induces Doppler shifts in the echoes, taking information of wing beat movement, that are important signatures for species identification since insect wing beat frequency scales with body size [Sot47]. Because the way in which a given insect species moves its wings is highly specific, the structure of glints varies across different species and even insects with the same wing beat may provide different spectral signature in the echo [Emde90].

Figure 14 shows the spectrogram of four echoes from four different insects: *Deilephila elpenor*, *Scotia exclamationes*, *Melolontha melolontha* and *Tipula oleracea*. These were flying with the same wing rate and were illuminated from 0, 90 and 180 degrees with respect to the sound source. The figure shows that spectrograms related to each species present different characteristics, although the fluttering frequency is the same. It also shows that for the same insect the property of the spectrogram are also dependent on the angle from which the sound source was coming, confirming that spectral cues provide information on angular position as well [Kob90] [Sch87]. Finally, the trace of the time oscillation confirms that the same considerations are valid in the time domain. Schnitzler showed that at 0 degrees (i.e. frontal view) spectral broadenings in echoes due to glints, typically fall below the carrier frequency (i.e. the micro-Doppler signature concentrated below the carrier frequency), while at 180 degrees they show negative and positive Doppler shift, and that the width of the broadening is related to the wing beat frequency of the insect [Sch87]. Roverud showed that bats that use shorter signals need greater differences in wing beat frequencies

than bats that emit longer signals, in order to discriminate between different fluttering targets [Rov91]. However, behavioural experiment tests show that neither amplitude nor frequency modulations alone are sufficient to achieve recognition of insects, and that differences in glints from series of echoes play an important role in target classification. This highlights the fact that whilst it is clear that bats are able to use echo location sources in a comprehensive way, it is not fully understood how they do this. For example, von der Emde and Schinitzler showed that greater horseshoe bats are even able to discriminate insects of a given species when presented with echoes from the insect illuminated at angles that they had not previously experienced [Emde90].

The way CF bats classify fluttering insects has a very close parallel with the attempt, in radar and sonar systems, to classify targets by using micro-Doppler information in the echo spectrogram. A big difference though is that bats are able to combine information that derives, at least, from both amplitude modulations and frequency modulations and manage to obtain remarkable performance in a very difficult task. Bats, in fact, have shown a peculiar ability to distinguish between insects that present the same wing beat rate and that have very similar dimension and obtain performance which currently is enormously better than what a radar system can do. It is evident that all this might open the way to a new branch of research that not only looks at the micro-Doppler signature of a moving target but also at how this movement modulates the amplitude of the echo and at how this information can be possibly combined and utilized to enhance radar target classification performance.

3.4 Nectar feeding bats and bat-pollinated plants

Although several works have looked at how bats recognize moving targets, the literature shows a lack of knowledge on how these mammals perform classification of static targets. Nectar feeding bats play an important role in the process of pollination of plants, which, with their flowers, represent a very interesting class of organisms for the study of target classification. Firstly, flowers are motionless and silent so that bats cannot rely on Doppler information or passive echolocation based on target sounds, and secondly their habitat is often a densely cluttered environment. Finding and approaching a flower is a gradual process that involves all the bats' senses. While bats can only use vision at dusk, they can rely on their sense of smell and on echolocation to search for flowers in darkness. Nectar feeding bats have a highly developed sense of smell and they are highly attracted by the scent of sulphur-compounds which are produced by many bat-pollinated plants [Helv00]. Although the sense of smell certainly plays an important role, and in particular for long range attraction, latest research results support the fact that nectar feeding bats largely rely on echolocation to plan their approach flight and to detect and select the flowers in the proximity of the plant. Indeed, it has been shown that bats are capable to detect and select objects by echolocation only.

The flowers of the bat-pollinated vine *Mucuna holtonii* plant release the greatest amount of nectar reward only on the first visit by the bat. By landing into the flower corolla the bat ignites the explosion of the flower's vexillum, which functions as a trigger for the pollen and the nectar release. Behavioral experiments have shown that bats can successfully select unvisited flowers within an inflorescence in darkness and this is a task that can only be done by echolocation. The scent of the flower, in fact, remains unchanged before and after the explosion. It has also been shown that modifying the echo-acoustic properties of these flowers, by inserting a small pad of cotton in cavity of the vexillum (i.e. without altering the scent and visual aspect of the flower) results in a drop of the success rate of the bat [Helv99] [Helv03a]. A few individuals of *Glossophaga soricina* bats were successfully trained to discriminate between an unscented hollow hemisphere and an unscented paraboloid, irrespective of size and in darkness, to show that bats have the ability to generalize object shapes by echolocation [Helv04]. In another experiments, bats were able to discriminate between unscented hollow hemispheres of different sizes in darkness [Simon06]. These experiments clearly suggest that echolocation alone plays a key role for the task of recognition.

Choosing the most suitable flowers to be visited within an inflorescence is a task for which echolocation certainly plays a fundamental role and, although classification of flowers in such environment is demanding, nectar-feeding bats succeed in their pollination task. Because it is in the interest of both the plant and the bat that pollination takes plays it is widely hypothesized that co-evolution has shaped bat-pollinated flowers in order to ease classification by bats [Helv03b]. Investigating if there are and which are the features of these flowers that aid the bat search for nectar is under investigation and it is key to understanding the strategy underpinning the task of classification of flowers by bats. This may well offer an interesting insight to radar and sonar engineers working on target classification and recognition as well. There are a number of similarities with the radar/sonar case. In the previous section we have seen that the types of waveforms deployed by bats are similar to linear and hyperbolic chirps commonly used by synthetic systems. Also the number of range cells covering a flower, i.e. the ratio between the flower size and the range resolution achievable by bats, assumes similar values to its counterpart in common radar/sonar applications.

3.5 Classification of flowers

Here, we analyse a set of data containing high range resolution profiles (HRRP) of two *Cobaea scandens* flowers taken at different stage of maturity and then manually modified. A photo of a C. *scandens* is given in Figure 15.

Classification performance of a *k*-NN classifier trained to distinguish between these flowers is assessed in order to investigate and quantify if the information contained in their echo acoustic signatures can be used to perform recognition with particular emphasis to differences in echoes due to the flower wilting process and to the manual removal of the flower petals.

3.5.1 Data collection

The data consists of high range resolution profiles (HRRP) of a few individuals flowers of *Cobaea scandens*. These were collected at the School of Biological Sciences of the University of Bristol in October 2008. The flowers to be ensonified were impaled on a thin pin that was placed at the centre of a turntable rotating with an angular resolution of 1°. A linear down-chirp spanning the frequencies between 50 kHz and 250 kHz, in order to cover the range of frequencies used by nectar feeding bats, was transmitted towards the flower with a custom-built loudspeaker [Strei03]. The echo, recorded with an ultrasound microphone (G.R.A.S. type 40 BF), was sampled at a rate of 500 kHz using a National Instruments PCIe-6251 card. During the experiments, both the loudspeaker and the microphone were placed into a custom built artificial bat head, capable of hosting one loudspeaker and two microphones in order to reproduce the spatial arrangement of the two ears and the mouth of the bat. The setup allowed measurements with a range resolution of about 1.5 mm. Three different horizontal images were recorded, each one representing HRRP of a *C. scandens* in three different cases; a flower in ideal condition for pollination, a desiccated flower, and a flower whose petals were removed by hand. The raw data was processed offline using Matlab.

3.5.2 Data pre-processing and results

The mean value was removed from each measurement $s_j(n)$, where j indicates the j^{th} angular perspective and n represents the discrete time index. The analytic signal $\hat{s}_j(n)$ of $s_j(n)$ defined as $\hat{s}_j(n) = \mathrm{Re}\{s_j(n)\}$ was computed with the *hilbert.m* Matlab function and matched filter with Matlab to the analytic signal of a copy of the transmitted chirp. The output of the matched filter $z_j(n)$ represents the HRRP corresponding to the j^{th} angular perspective. The background static clutter was removed by removing the mean value from each range bin, i.e. by removing the zero Doppler component in the frequency-range plot.

For each angular perspective the target HRRP $\mathbf{x}_j$ was formed with the element of $\mathbf{z} = \begin{pmatrix} z(1) & z(2) & .. & z(n) \end{pmatrix}^T$ corresponding to the range bins from 18 cm to 32 cm, i.e. those range bins that contained the echo from the targets. Similarly a noise profile $\mathbf{n}_j$ was formed with the elements of $\mathbf{z}$ corresponding to the range bins that contained noise only.

The images showing the magnitude of the HRRPs for each angular perspective are given in Figure 16, Figure 17 and Figure 18. These are related to the angular perspectives ranging between -90 and +90 degrees, where the 0 degree line corresponds to the case when the flower was facing the microphone and the loudspeaker. The images were normalised to the same value in order to allow a fair comparison.

The magnitude of the HRRPs of the individual of *C. scandens* that was ready for pollination, given in Figure 16, exhibits strong reflections at a distance of between 21 cm and 23 cm from the microphone that are originating from the petals of the corolla, and weaker reflections between 20 cm and 21 cm are originated from the anthers protruding from the corolla. An overall weaker scattering originates from the bell-shaped inner part of the flower's corolla between 25 cm and 27 cm. This is the part of the corolla that contains nectar and from which the pistil grows. Weaker scattering due to the echo generated by the ring of sepals that cover the external back side of the flower is also visible from about 28.5 cm to 29.5 cm.

It is interesting to observe how the petals can scatter a considerable amount of power over a wide range of angles that goes between -60 degrees and +60 degrees. Exploiting how the shape of the corolla can change as the flower wilts, i.e. stops producing nectar hence losing attractiveness to bats, could lead to understanding which features are responsible for recognition. The HRRPs of the desiccated *C. scandens* that had started to wilt due to an overnight frost are displayed in Figure 17. The time plot shows that the scattering due to the petals between 21 cm and 23 cm is less complex with respect to that in Figure 16. A wide scattering that goes between –60 degrees and 0 degrees is still present at about 22 cm and probably due to one petal that was still in good condition at the time of the measurement. However, most of the scattering from the petals is no longer visible, leading to a marked loss of information. This is because the petals tend to fold back during the wilting process and the mirror effect that is generating when they are in the upright position diminishes.

The loss in turgor might also affect reflective properties of the plant tissue. The scattering from inside the corolla does not show any loss of complexity but it is weaker than before with respect to the maximum value of the image. To investigate what the image of the flower would be without petals at the corolla, the front ends of the petals from the same individual were removed manually using a pair of scissors. Figure 18 shows the HRRPs image obtained from this modified flower. The scattering from the petals disappears over all angles. A fraction of power is still scattered at 0 degrees probably originated by some residuals of petals that might have not been removed properly or simply by the line delimiting the aperture of the corolla. The fact that the image is very similar to the one obtained from the wilting flower suggests that the scattering form the petals tends to disappear when the flower is no longer suitable for pollination. Scattering from the inside of the flower is now more visible as compared to Figure 17 but no relevant change in the complexity of the echo is to be reported.

One cue available to bats is the overall power of the echo. The mean power of the noise was estimated from the noise profiles $\mathbf{n}_j$ and similarly the mean power of the range bin containing target plus noise was estimated from all perspectives $\mathbf{x}_j$. The results show differences in reflected power of more than 3 dB between the good flower and the other two that are within the sensitivity shown by bats and therefore could be used by bats to perform discrimination.

Differences in the echo acoustic structure were quantified by means of a *k*-NN classifier trained with HRRPs extracted from each image with an angular step of 3° to 5°. The remaining HRRPs for each image were used to test the classifier. For each HRRP the Principal Components were extracted with the Principal Component Analysis algorithm as described in the previous Chapter. The maximum value of the cross-covariance was used as a measure to identify the k nearest neighbours to each test. The profile was assigned to the class owning the highest number of neighbours between the k selected. When the *k*-NN classifier was set to make a decision with $k = 3$ nearest neighbours 77% correct classification was achieved. This was obtained with an angular separation between consecutive training profiles equal to 3° degrees corresponding to a total number of 60 training profiles and 120 test profiles. Dimensionality reduction was performed through the PCA algorithms by selecting the principal 20 features for each profile. Classification performance improved significantly up to values over 90% correct classification when the classifier was trained with profiles separated by 2° and similarly performance dropped to 68% when the decisions were made on $k = 5$ neighbours whilst keeping the angular separation between the training profiles equal to 3 degrees. Results suggest that high-level classification performance can be achieved and highlight how important it is to account for the cross-correlation properties between different angular perspectives in each target.

Overall results show that high level classification performance can be achieved. They support the hypothesis that the echo acoustic signatures of bat-pollinated flowers contain information that changes together with the stage of maturity of the flower and that can be used to achieve high level classification performance. In particular scattering from the petals and the way this changes as a function of the state of the flower could be a way of letting bats know which flower is most suitable for pollination, i.e. holds a nectar reward, and hence play an important role for recognition and selection of the flowers. The results of the power reflectivity analysis show that differences in the backscattered power can also be used to aid recognition as these are in the detectable range shown by bats.

4.0 ACOUSTIC MICRO-DOPPLER

The echo produced by a moving target that is illuminated by a radar or ensonified by a sonar system contains frequency modulations caused by the time varying delay occurring between the target and the sensor. The main bulk translation of the target towards or away from the sensor induces a frequency or Doppler shift of the echo as a result of the well known Doppler effect [Gil65]. Additional movements of small parts of the target contribute with frequency modulations around the main Doppler shift and these are commonly called micro-Doppler modulations. Micro-Doppler modulations contain a signature of the target that can be used for target recognition.

Classification of targets by micro-Doppler signatures has been widely investigated for the Radio Frequency (RF) regime and in particular for applications related to radar systems. V.C. Chen *et al* have modelled the radar micro-Doppler phenomenon and simulated micro-Doppler signatures for various targets, such as rotating cylinders, vibrating scatterers and personnel targets [Che03] [Che06] [Che00]. The authors also show that a time-frequency analysis of the radar return can be used to extract micro-Doppler signatures from the received signal, leading to additional information on the target that can be used for classification and recognition.

An easy way to perceive target micro-Doppler modulations is to listen to an appropriate audible version of the radar return. Indeed, this method has been used to perform classification of targets but it has the disadvantage of requiring the presence of a human operators. For these reasons, the most recent challenge that has to be addressed is to develop micro-Doppler classification methods that allow Automatic Target Recognition (ATR). Various studies have looked at classifying targets automatically through the analysis of their micro-Doppler signatures. For example, classification by micro-Doppler signatures of a tracked vehicle, a wheeled vehicle and a personnel target was performed in [Smi06a] [Smi062b] [Smi08]. Here, classification performance was analysed via a K-Nearest Neighbourhood classifier and a Naive Bayesian classifier by

testing the main features of the micro-Doppler signatures that were extracted with both the Principal Component Analysis and the dynamic time warping algorithms.

The use of micro-Doppler signatures for biometric purposes and, more specifically, to identify humans, has been at the centre of a number of works in the last few years. In particular, there have been attempts to relate specific components of micro-Doppler gait signatures to parts of the body for identification purposes [Tah09]. In [Tiv10] a technique introduced for visual pattern recognition is applied to human gait signatures to distinguish between different human motions. The exploitation of micro-Doppler information for classification has been recently extended to the field of radar imaging and in particular to ISAR [Vig09] [Gha08][Che09], and there has also been an attempt to exploit signatures of moving personnel targets through multi-static radar measurements [Sam10].

Although target classification by micro-Doppler signatures has been widely exploited in the RF regime, there has been rather little research done on the same topic in the acoustic regime. An acoustic Doppler sensor operating at 40 kHz was developed in 2007 to characterise a person's gait with the final goal to perform human recognition [Kal07]. The system was deployed to measure Doppler signatures of men and women walking towards and away from the sensor which were used to assess classification performance of a Bayesian classifier.

An additional active acoustic sensing system, still operating at 40 kHz, was developed at Johns Hopkins University [Zha07] [Zha07b]. The system was used to gather micro-Doppler signatures of humans and four-legged animals that were used to perform target classification [Zha08]. It is desirable for an active surveillance system to operate indiscreetly, ideally be undetectable at all times and have minimal impact on the environmental surroundings including on any animals other than humans. This requires an acoustic radar to operate at frequencies outside the hearing sensitivity of, at least, most common non-human targets. Although humans can hear up to about 20 kHz, common pets' hearing system is sensitive up to much higher frequencies; dogs for example can hear up to about 40 kHz and cats can hear up to 60 kHz. For these reasons, it may be an operational requirement that an acoustic radar operates at frequencies above 40 kHz.

The use of higher frequencies leads to an additional advantage; as the Doppler shift is directly proportional to the carrier frequency, given a certain frequency resolution, at higher frequencies small movements induce a wider frequency shift and therefore a clearer contribution to the micro-Doppler signature. On the other hand, the increased acoustic attenuation at higher frequencies lowers the operational range of the system. Here, a fully coherent acoustic radar operating at 80 kHz is described together with a set of experimental data, containing micro-Doppler signatures of personnel targets performing various actions. The data was collected at University College London between 2010 and 2011. Classification performance of a K-Nearest Neighbourhood classifier and a Naive Bayesian classifier trained to distinguish between different human motions and different personnel targets are then presented.

4.1 Description of the acoustic radar

The acoustic radar developed at UCL is composed of a signal generator, a loudspeaker, a microphone, two pre-amplifiers that amplify the signals before transmission and after recording and a Data AcQuisition (DAQ) card that digitizes the recorded signal. The waveforms are generated by a National Instrument NI PXI-6733 card capable of generating up to 500 MSamples/sec (32 bit resolution) per channel. Each card can transmit on 8 channels and therefore 8 different waveforms could potentially be transmitted simultaneously.

The generated waveform is given as in input to the preamplifier (Ultra Sound Advice S55A) capable of generating a maximum output of 140 V peak to peak. The loudspeaker provides a monitor output which is 1/100 of the actual output of the amplifier. This allows monitoring of the output and of the gain of the amplifier. The signal is transmitted by a loudspeaker (Ultra Sound Advice) which can nominally operate between 20 kHz and over 200 kHz. It can generate an output level greater than 105 dB S.P.L. (Sound Pressure Level) between 20 kHz and 50 kHz (measured at 0.25 m) and greater than 85 dB S.P.L. up to 150 kHz. The round-shaped active area of the loudspeaker has got a 50 mm diameter.

In the receive section, echoes are captured by a microphone capsule (Ultra Sound Advice UM3 capsule). The sensitive element of the capsule is a 1.8 cm diameter disk made of a very thin aluminised polyester film. Nominally, its sensitivity is better than -57 dB from 20 kHz to 120 kHz and better than -70 dB at 180 kHz. The maximum gain obtainable by the amplifier is nominally x220 at 100 kHz. The amplified received signal is digitised by a DAQ card (NI PXI-6133) which is capable of receiving 8 channels simultaneously with a sampling frequency up to 500 kHz per channel. The digitised data is stored in a file and the signal processing is performed off line with Matlab. The data containing micro-Doppler signatures were gathered at UCL between 2010 and 2011. The microphone and the loudspeaker were arranged in a pseudo mono-static configuration next to each other as shown in Figure 19.

4.2 Description of the experiments

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

- walking
- running
- walking while carrying an object in one hand
- walking while carrying an object with both hands
- walking with a heavy backpack on the shoulders
- shadow boxing
- picking up an object form the floor
- oscillating with the torso while standing on the spot

The physical characteristic of the targets are given in Figure 20.

When walking or running the targets were moving on a Pro Fitness manual treadmill and were ensonified from the back to avoid any masking effects. This was deployed to remove the main Doppler shift contribution from the Doppler signatures, and to keep the Signal-to-Noise ratio of the received signal as constant as possible during the recording. Also the treadmill removed any movement constraint on the targets and therefore allowed longer recordings.

4.3 Data pre-processing

A recording of the background clutter was removed from the raw data for each target recording. The analytic signal of the difference signal was mixed with the complex conjugate of the analytic signal of the transmitted tone to obtain the baseband signal containing the micro-Doppler signatures. The analytic signals were obtained by applying the Matlab *Hilbert.m* function to the original real signals. The mixer output was filtered with an ideal low pass filter cutting off all the frequencies over 2.5 kHz. It was experimentally observed, in fact, that the micro-Doppler signatures were contained within the frequencies below 2.5 kHz. The resulting signal was down-sampled of a factor 50 leading to a final sampling frequency of 10 kHz. This was done to reduce the load of the data without altering the information contained in the frequencies below 2.5 kHz. The Short Time Fourier transform was applied to all the recordings to visualise the recorded micro-Doppler signatures. This was calculated by using 30 msec long 50% overlapping windows weighted with a Hanning function and by computing their respective Fourier transform as described in Section II. The measured micro-Doppler signatures for one of the two targets taking all the action listed above are given in Figure 21 to Figure 22.

4.4 Classification Performance Results

The baseband data obtained after pre-processing containing the micro-Doppler signatures were used to assess and compare the performance of a Bayesian classifier and a K-Nearest Neighborhood classifier. In particular, our main focus was to distinguish between different personnel targets undertaking the same action or between different actions undertaken by the same target. The time sequences corresponding to each class were divided in 50% overlapping windows of duration T_w. The analysis was repeated for three different window durations; $T_w = 30$ msec, $T_w = 100$ msec, $T_w = 150$ msec. This resulted in a total number N_w of windows for each class. The content of each window was organised in column vectors $\mathbf{x}_j^i$ where j indicates the j^{th} window belonging to the i^{th} class. The first N_{Train} windows for each class, corresponding to the first 1.5 sec of the recorded signal, were used to train the classifiers, so that the information used for training remained the same in all three cases.

The Principal Component Analysis (PCA), Cepstrum and MEL Cepstrum algorithms were used to extract N_f features from each window to reduce the dimensionality of the data and reduce the computational load of the classifiers [Duda] [Theo].

Table 1 shows the classification performance that was obtained from the comparison between the walking gait of *Target A* and *Target C* for $T_w = 30$ ms. This is divided in a sub-table per classifier, each one containing the three confusion matrices related to the three feature extraction algorithms that were used. This allows a straightforward comparison of the results obtained with different feature extraction algorithms given a specific classifier. In each confusion matrix, the percentage number is the ratio between the number of assignments and the total number of windows under test, given the same class. The averaged Probability of Correct Classification (P_{cc}) is given below each confusion matrix. This was calculated under the assumption each class had the same *a priori* probability. The parameter k of the *k*-NN classifier was set equal to 3. Results show that high level classification performance can be obtained. In particular, for this case, the combination PCA plus Naive Bayesian classifier leads to a rate of correct classification over 90%. The lowest rate, instead, is given by the PCA followed by the *k*-NN classifier presenting a high number of wrong decisions when Target A is given. Overall, in all other cases, correct classification performance is above 80%. Figure 23 shows the magnitude of the first two Mel-Cepstrum features of each class, in the 2D plane, giving a visual idea of the degree of separability between the two classes [Theo].

The confusion matrices obtained from the comparison of the walking gaits of all three targets, for $T_w = 30$ ms, is given in Table 2 The parameter k of the *k*-NN classifier was set to 5, and all cases that could not univocally be assigned to any of the three classes were treated as *unknowns* and counted to average performance. Classification performance drops for both the classifiers and all feature extraction algorithms, all presenting a high rate of wrong decisions between *Target B* and *Target C*. The highest rate of correct classification is achieved by the PCA algorithm followed by the Naive Bayesian classifier and this is equal to 64%. The drop in performance may be due to differences in the gait of *Target B* and *Target C* that cannot be resolved by the feature extraction algorithms. This appears clear by looking at Figure 24, where the magnitude of the first two Mel-Cepstrum features of each class is plotted versus each other on the 2D plane. A substantial overlapping between the features related to *Target B* and those related to *Target C* can be observed. Classification performance improves when the duration of the windows increases.

Table 3 reports the results for the same case but with the window duration set to $T_w = 150$ msec. The probability of correct classification on the average improves of 20% up to a correct rate of 71% given by the combination of the PCA algorithm followed by the Naive Bayesian classifier. The plot of the magnitude of the first two Mel-Cepstrum features, given in Figure 25, shows a higher degree of separability between the three classes that clearly leads to the improvement in performance. The same behaviour is observed for all feature extraction algorithms.

Table 4 reports the results obtained when the classifiers were trained to distinguish between the same personnel target, *Target B*, undertaking two different actions for $T_w = 30$ ms. In particular results are related to the comparison between Walking Normally (Class *B-WN*}) and Walking with a BackPack (Class *B-WBP*). The parameter k of the *k*-NN classifier was set to 3. Results show that it is possible to achieve correct classification performance over 80%. The highest level of performance is achieved by the PCA algorithms and the Mel-Cepstrum algorithm for the Naive Bayesian classifier. Figure 26 shows the magnitude of the first two Mel-Cepstrum features for each class in the 2D plane.

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

The impact that using longer time durations has on classification performance is currently under investigation. Longer window durations allow a better characterisation of the micro-Doppler signature, for example by looking at how the Cepstral coefficients vary as a function of time, and allow extraction of additional features, such as the signature period and maximum bandwidth, that otherwise would not be available. Clearly there is the disadvantage of a longer time gaps between decisions, which can be an issue for some applications. Preliminary results obtained on the same dataset show that there is a consistent improvement in classification performance when the signatures are tested in blocks of 2.5 sec.

Table 5 show classification performance of a maximum likelihood classifier that uses a Gaussian mixture distribution to model the data and that was trained to distinguish between a set of 5 different actions performed by various targets. Results show that a rate of correct classification higher than 90%. A further analysis is required to corroborate these results.

4.5 Summary

An acoustic radar operating at 80 kHz was developed for collection of micro Doppler signatures of moving targets, doubling the frequency used in previous works. The acoustic radar was deployed in a set of experimental trials in which micro-Doppler signatures of various personnel targets undertaking a number of action were collected in a highly indoor cluttered environment. These were used to test K-NN and Naive Bayesian classifiers trained to distinguish between walking gaits of different targets or between different actions undertaken by the same targets. Results show that the acoustic radar can be successfully deployed at short ranges to collect micro-Doppler signatures of moving targets. Classification performance results show that the information contained in these acoustic signatures can be used to perform identification and recognition of personnel targets.

Data has been gathered and analysed for a range of personnel target motions and good levels of classification have been achieved. Although the acoustic radar can only survey short ranges due to the high attenuation of sound waves in air, it can be successfully deployed in indoor environments, such as airports, to monitor, for example, the flow of passengers through doors or security checks with the potential to identify suspicious behavior. In addition to this, the acoustic radar offers an easy and inexpensive way to collect micro-Doppler signatures of a wide range of targets, often difficult to obtain, that can be used to analyse human walking gaits or other type of behaviour in a number of applications. Future work will look at understanding how additional features such as geometrical features (i.e. the signature period or the signatures maximum Doppler shift) or how the history of the current features (i.e. the time variation of the Cepstrum or MEL-Cepstrum coefficients) can be deployed to enhance classification performance. Extending the work presented in this paper to the underwater regime is also under consideration.

5.0 REFERENCES

[1] [Bak03] C.J. Baker and A.L. Hume, “Netted Radar Sensing”, *IEEE AES Magazine*, vol. 18, pp. 3–6, 2003.

[2] [Che03] V. Chen, F. Li, S.-S. Ho and H. Wechsler, “Analysis of micro-doppler signatures,” *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–6, 2003.

[3] [Che06] V. Chen, F. Li, S.-S. Ho and H. Wechsler, “Micro-doppler effect in radar: phenomenon, model, and simulation study”, *IEEE Trans. Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, Jan. 2006

[4] [Che00] V. Chen, “Analysis of radar micro-doppler with time-frequency transform,” in *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing*, 2000.

[5] [Che09] V. Chen, W. Miceli and B. Himed, “Micro-doppler analysis in ISAR - review and perspectives,” *IEEE International Radar Conference 2009*, pp. 1 –6, 2009.

[6] [Duda] R. Duda, P. Hart and D. Stork, *Pattern Classification*, John Wiley and Sons, 2001.

[7] [Emde86] G. von der Emde and H.-U. Schnitzler, “Fluttering target detection in hipposiderid bats," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 159, pp. 765-772, 1986.

[8] [Emde90] G. von der Emde and H.-U. Schnitzler, “Classification of insects by echolocating greater horseshoe bats," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 167, pp. 423-430, 1990.

[9] [Gha08] A. Ghaleb, L. Vignaud and J. Nicolas, "Micro-doppler analysis of wheels and pedestrians in ISAR imaging", *IET Signal Processing*, vol.2, no.3, pp.301 –311, September 2008.

[10] [Gil65] T. P. Gill, *The Doppler Effect, an Introduction to the Theory of the Effect*, Logos Press, Limited, 1965. March 16, 2011.

[11] [Helv99] D. von Helversen and O. von Helversen, "Acoustic guide in bat-pollinated flower," *Nature*, vol. 398, pp. 759–760, 1999.

[12] [Helv00] O. v. Helversen, L. Winker and H. Bestmann, "Sulphur-containing "perfumes" attract flower-visiting bats," *Journal of Comparative Physiology A*, vol. 186, pp. 143–153, 2000.

[13] [Helv03a] D. v. Helversen and O. v. Helversen, "Object recognition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 189, pp. 327–336, 2003.

[14] [Helv03b] D. v. Helversen, M. W. Holderied and O. v. Helversen, "Echoes of bat-pollinated bell-shaped flowers: conspicuous for nectar-feeding bats?" *Journal Experimental Biology*, vol. 206, no. 6, pp. 1025–1034, 2003.

[15] [Helv04] D. v. Helversen, "Object classification by echolocation in nectar feeding bats: size-independent generalization of shape," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 190, pp. 515–521, 2004.

[16] [Kal07] K. Kalgaonkar and B. Raj, "Acoustic Doppler sonar for gait recognition," *IEEE Conference on Advanced Video and Signal Based Surveillance*, 2007. AVSS, pp. 27 –32, 2007.

[17] [Kno85] E.F. Knott, J.F. Shaeffer and M.T. Tuley, *Radar Cross Section*, Artech House, 1985.

[18] [Kob90] R. Kober and H.-U. Schnitzler, "Information in sonar echoes of uttering insects available for echolocating bats," *The Journal of the Acoustical Society of America*, vol. 87, no. 2, pp. 882-896, 1990.

[19] [Nov91] L.M. Novak, "A comparison of 1-D and 2-D algorithms for radar target classification", *IEEE International Conference on Systems Engineering*, pp. 6–12, August 1991.

[20] [Nov97] L.M. Novak, S.D. Halversen, G. Owirka and M. Hiett, "Effects of polarization and resolution on SAR ATR", *IEEE Transaction. Aerospace and Electronic Systems*, vol. 33, pp. 102–116, 1997.

[21] [Ost84] J. Ostwald, "Tonotopical organization and pure tone response characteristics of single units in the auditory cortex of the Greater Horseshoe Bat," Journal of Comparative Physiology A: Neuroethology, Sensory, *Neural, and Behavioral Physiology*, vol. 155, pp. 821-834, 1984.

[22] [Ron97] H. Rong and Z. Zhaoda, "Researches on radar target classification based on high resolution range profiles", *IEEE Aerospace and Electronics Conference*, July 1997, vol. 2, pp. 951–955.

[23] [Rov91] R. C. Roverud, V. Nitsche and G. Neuweiler, "Discrimination of wing-beat motion by bats, correlated with echolocation sound pattern," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 168, pp. 259-263, 1991.

[24] [Run99] P.R. Runkle, P.K. Bharadwaj, L. Couchman and L. Carin, "Hidden Markov models for multiaspect target classification", *IEEE Trans. Signal Processing*, vol. 47, pp. 2035–2040, 1999.

[25] [Sad02] F. Sadjadi, "Improved target classification using optimum polarimetric SAR signatures", *IEEE Trans. Aerospace and Electronic Systems*, vol. 38, pp. 38–49, 2002.

[26] [Sam10] P. Sammartino and J. Fortuny-Guash, "Space and frequency diversity for moving personnel spectrogram estimation," *IEEE International Conference 2010*, Washington DC.

[27] [Sch83] H.-U. Schnitzler, D. Menne, R. Kober and K. Heblich, "The acoustical image of fluttering insects in echolocating bats," *Neuroethology and Behavioral Physiology*, pp. 235-250, 1983.

[28] [Sch87] H.-U. Schnitzler, *Echoes of Fluttering Insects: Information for Echolocating* Bats, Cambridge University Press, pp. 226-243, 1987.

[29] [Sch03] H.-U. Schnitzler, C. F. Moss and A. Denzinger, "From spatial orientation to food acquisition in echolocating bats," *Trends in Ecology and Evolution*, vol. 18, no. 8, pp. 386{394, 2003

[30] [Schu79] G. Schuller and G. D. Pollak, Disproportionate frequency representation in the inferior colliculus of doppler compensating greater horseshoe bats. Evidence for an acoustic fovea," *Journal of Comparative Physiology*, vol. 132, pp. 47-54, 1979.

[31] [Shi05] J. Shihao, L. Xuejun and L. Carin, "Adaptive multiaspect target classification and detection with hidden Markov models", *IEEE Sens. J.*, vol. 5, pp. 1035–1042, 2005.

[32] [Sho98] Showman, G.A., Richards, M.A. and Sangston, K.J., "Comparison of two algorithms for correcting zero-Doppler clutter in turntable ISAR imagery", *Conference on Signals, Systems & Computers*, 1998, vol. I, pp. 411–415.

[33] [Simon06] R. Simon, M. W. Holderied and O. von Helversen, "Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat," *Journal of Experimental Biology*, vol. 209, pp. 3599–3609, 2006

[34] [Sko80] M.I. Skolnik, *Introduction to Radar Systems*, McGraw-Hill, 1980.

[35] [Smi06a] G.E. Smith, K. Woodbridge and C.J. Baker, "Template based micro-doppler signature classification," IET Seminar on High Resolution Imaging and Target Classification, 2006, pp. 127–144.

[36] [Smi06b] G. Smith, K. Woodbridge and C. Baker, "Micro-doppler signature classification," *CIE International Radar Conference 2006*, Shanghai, pp. 1 –4.

[37] [Smi08] G. Smith, K. Woodbridge and C. Baker, "Naive Bayesian radar micro-doppler recognition," *International Radar Conference 2008*, pp. 111–116.

[38] [Sot47] O. Sotavalta, "The flight tone (wing stroke frequency) of insects," *Acta Entomologica Fennica*, vol. 4, pp. 5-117, 1947.

[39] [Strei03] A. Streicher, R. Muller, H. Peremans and R. Lerch, "Broadband ultrasonic transducer for a artificial bat head," *2003 IEEE Symposium on Ultrasonics*, vol. 2, pp.1364-1367, 2003.

[40] [Sul04] R.J. Sullivan, *Radar Foundations for Imaging and Advanced Concepts* , SciTech Publishing, 2004.

[41] [Tait] P. Tait, *An Introduction to Radar Target Recognition*, IEE, 2005, ISBN: 0863415016.

[42] [Tah09] D. Tahmoush and J. Silvious, "Radar micro-doppler for long range front-view gait recognition," *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, 2009, pp. 1 –6.

[43] [Theo] S. Theodoridis and K. Koutroumbas, *Pattern Recognition* - Fourth Edition. Academic Press, 2009

[44] [Tian97] B. Tian and H.-U. Schnitzler, "Echolocation signals of the greater horse-shoe bat (rhinolophus ferrumequinum) in transfer flight and during landing," *The Journal of the Acoustical Society of America*, vol. 101, no. 4, pp. 2347{2364, 1997

[45] [Tiv1010] F. Tivive, A. Bouzerdoum, and M. Amin, "Automatic human motion classification from doppler spectrograms," in 2nd International Workshop on Cognitive Information Processing (CIP), pp. 237 –242.

[46] [Ves07] M. Vespe, C.J. Baker and H.D. Griffiths "Radar target classification using multiple perspectives," *IET Radar, Sonar & Navigation*, vol.1, no.4, pp.300-307, Aug. 2007.

[47] [Vig09] L. Vignaud, A. Ghaleb, J. Le Kernec and J.-M. Nicolas, "Radar high resolution range & micro-doppler analysis of human motions," *International Radar Conference 2009*, pp. 1 –6.

[48] [Weh95] Wehner, D.R.: *High Resolution Radar* (Artech House, USA, 1995).

[49] [Xue02] L. Xuejun, P. Runkle and L. Carin, "Identification of ground targets from sequential high-range-resolution radar signatures", *IEEE Trans. Aerospace and Electronic Systems*, vol. 38, pp. 1230–1242, 2002.

[50] [Zha07] Z. Zhang, P. O. Pouliquen, A. Waxman and A. G. Andreou, "Acoustic micro-Doppler radar for human gait imaging," *Journal of the Acoustical Society of America Express Letters*, vol. 121, no. 3, pp. 110–113, 2007.

[51] [Zha07b] Z. Zhang, P. Pouliquen, A. Waxman and A. Andreou, "Acoustic micro-doppler gait signatures of humans and animals," *41st Annual Conference on Information Sciences and Systems*, 2007, pp. 627 –630

[52] [Zha08] Z. Zhang and A. Andreou, "Human identification experiments using acoustic micro-doppler signatures," in Argentine School of Micro-Nanoelectronics, Technology and Applications, 2008. pp. 81 –86

[53] [Zyw96] Zyweck, A. and Bogner, R.E., "Radar target classification of commercial aircraft", *IEEE Trans. Aerospace and Electronic Systems*, vol. 32, pp. 598–606, 1996

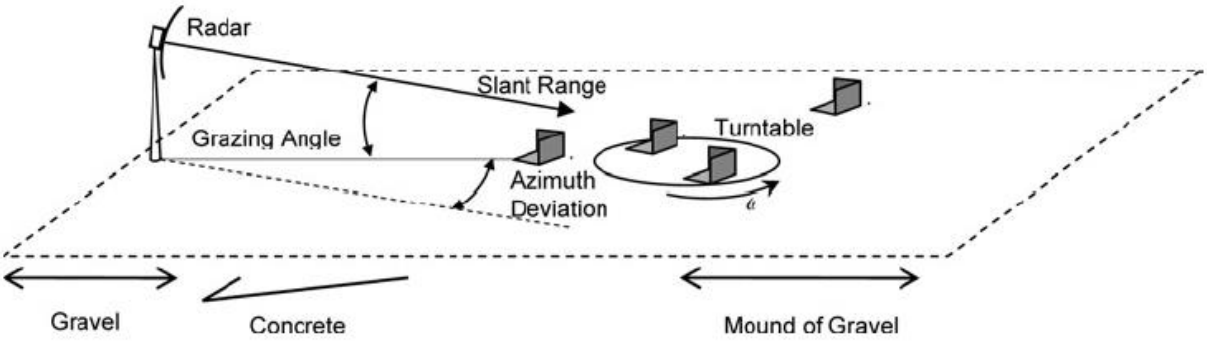

Figure 1: ISAR geometry: two stationary corner reflectors are in front and behind the turntable, while two rotating reflectors are placed on the turntable.

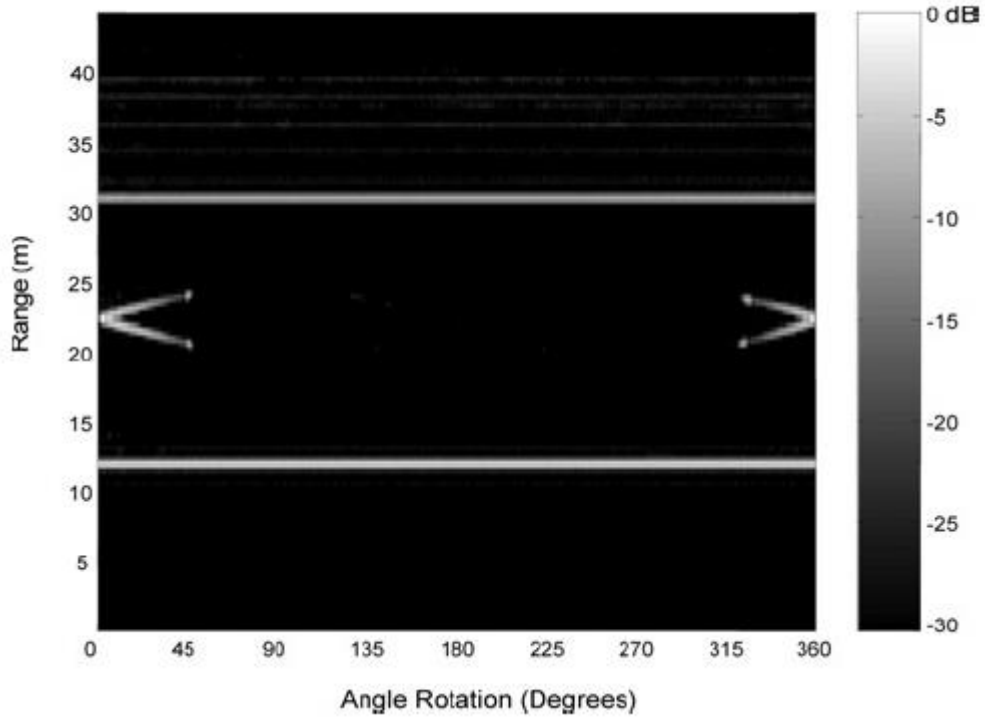

Figure 2: History of HRRPs (30 cm range resolution) from four corner reflectors, two rotating and two stationary.

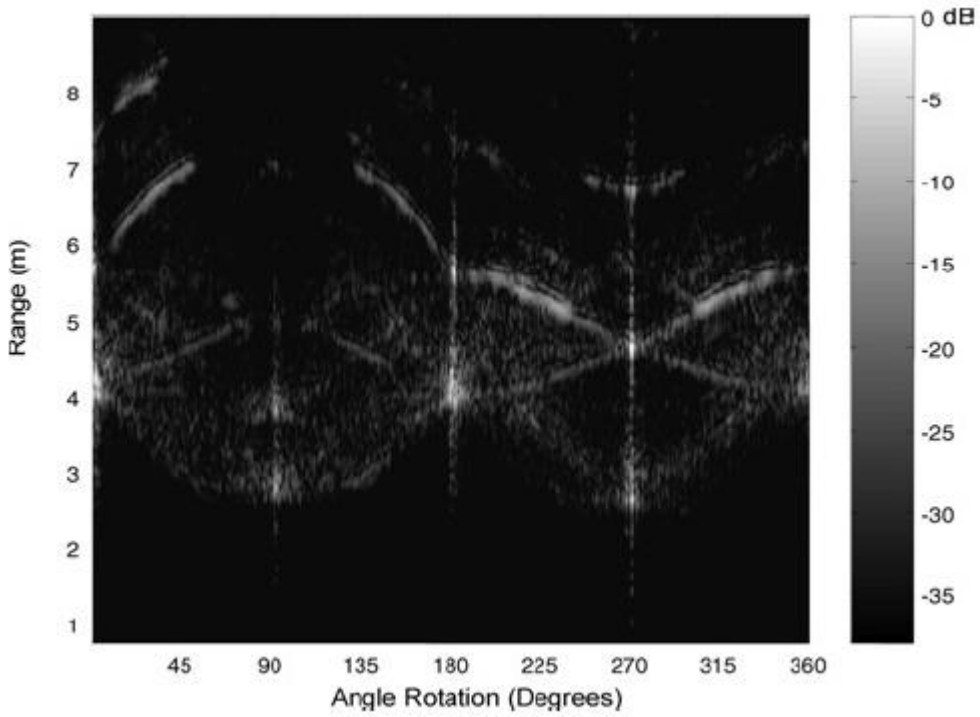

Figure 3: History of HRRPs (8 cm range resolution) from a series of X-band stepped frequency chirps illuminating a ground vehicle as it rotates over 360 degrees with the turntable.

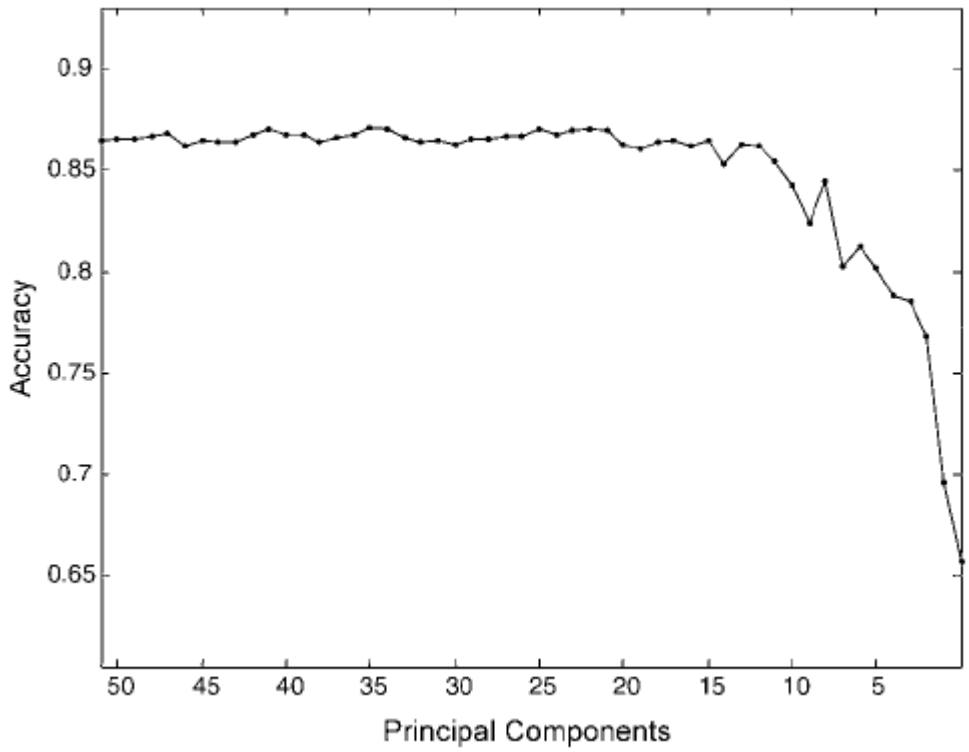

Figure 4: Classification accuracy versus number of Principal Components obtained with the

PCA algorithm.

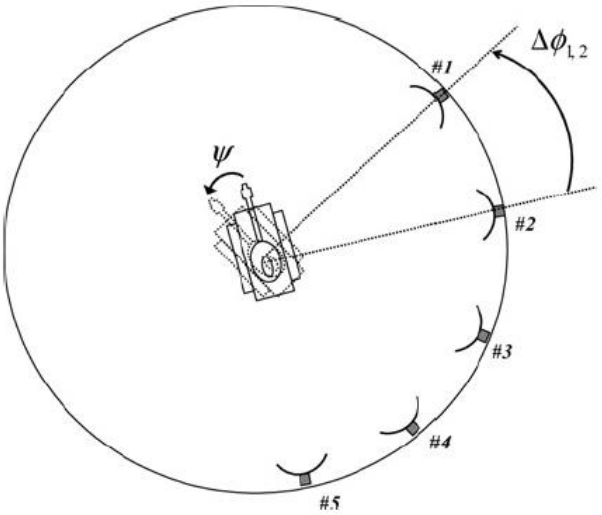

Figure 5: Multi-perspective environment approximation.

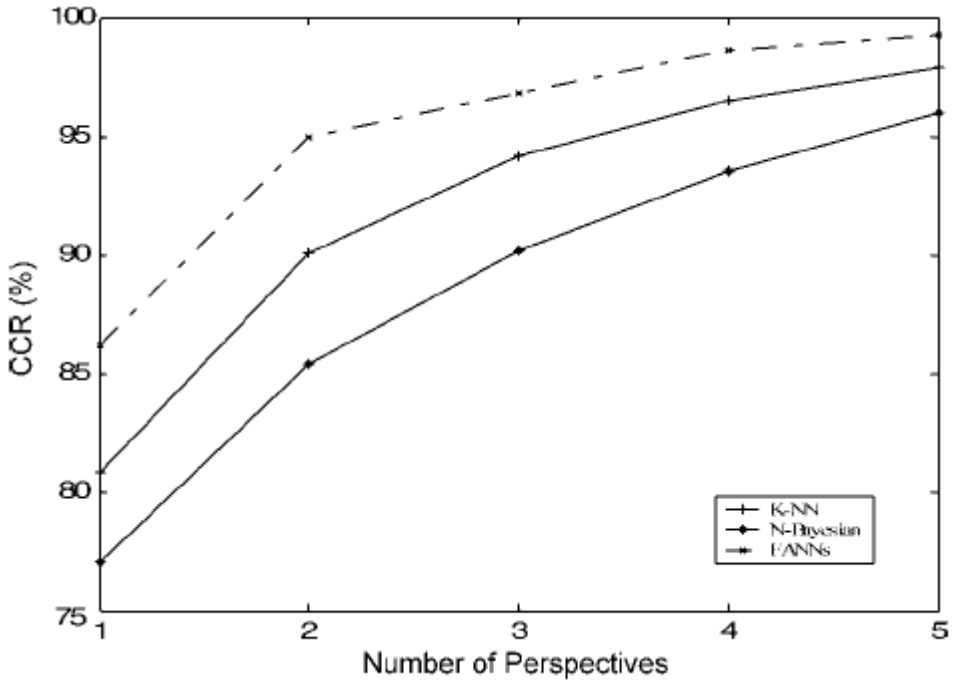

Figure 6: Correct classification rates of three classifiers using different numbers of perspectives.

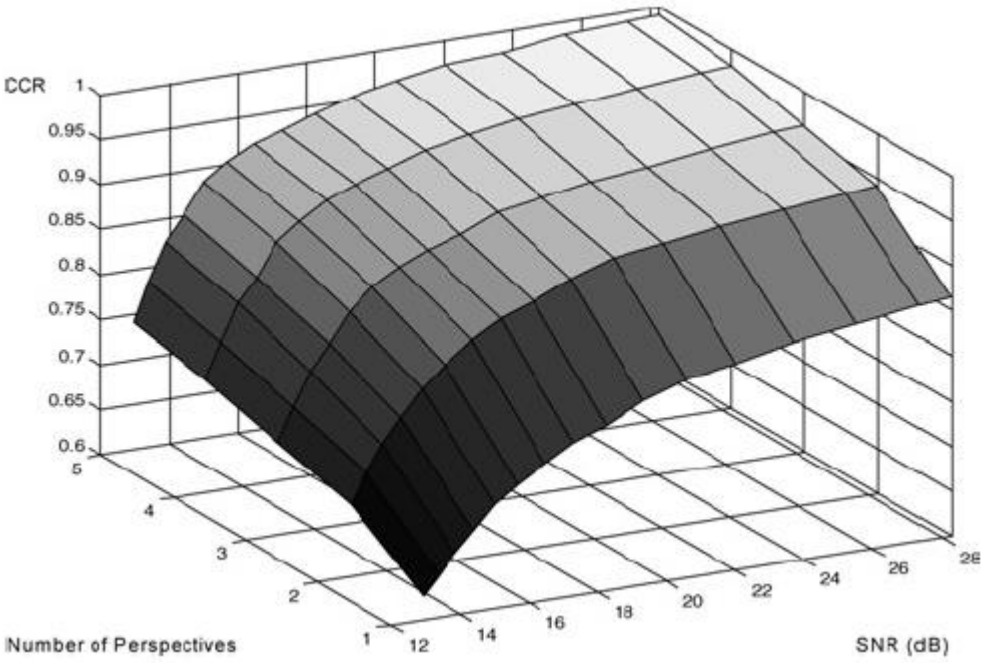

Figure 7: Correct classification rates for an M-P FANNs classifier with different SNR test sets.

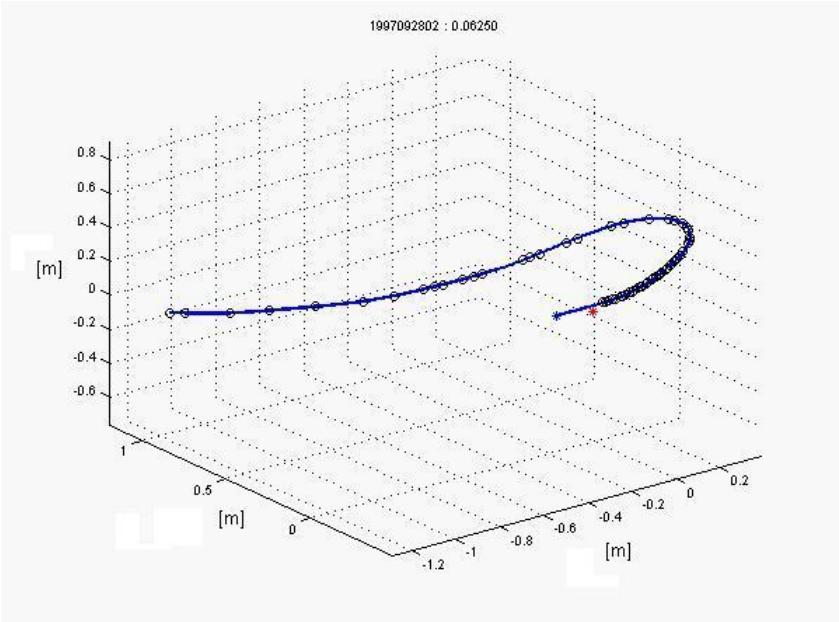

Figure 8: Trajectory of a bat (blue) with respect to a static insect (red) in a similar experiment performed at the University of Maryland (taken from http://www.bsos.umd.edu/psyc/batlab/).

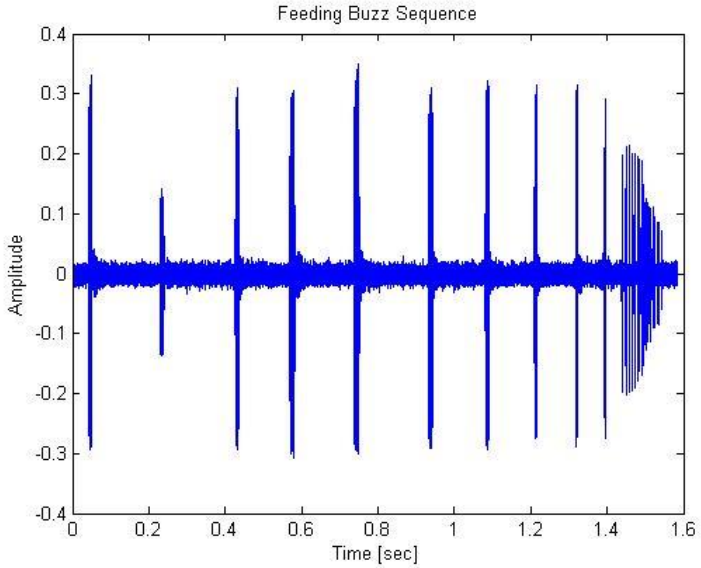

Figure 9: Feeding buzz of an *Epetesicus nilssoni* bat.

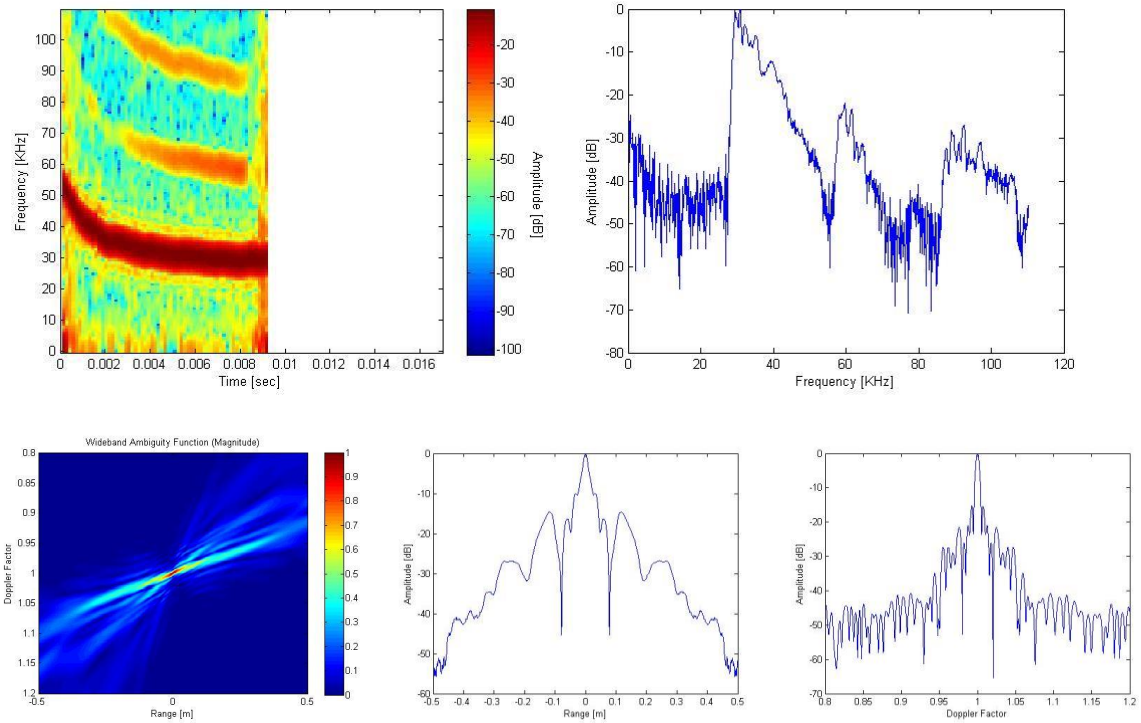

Figure 10: Spectrogram, mean spectrum and ambiguity function with relative range and Doppler cuts for the first pulse in the searching phase.

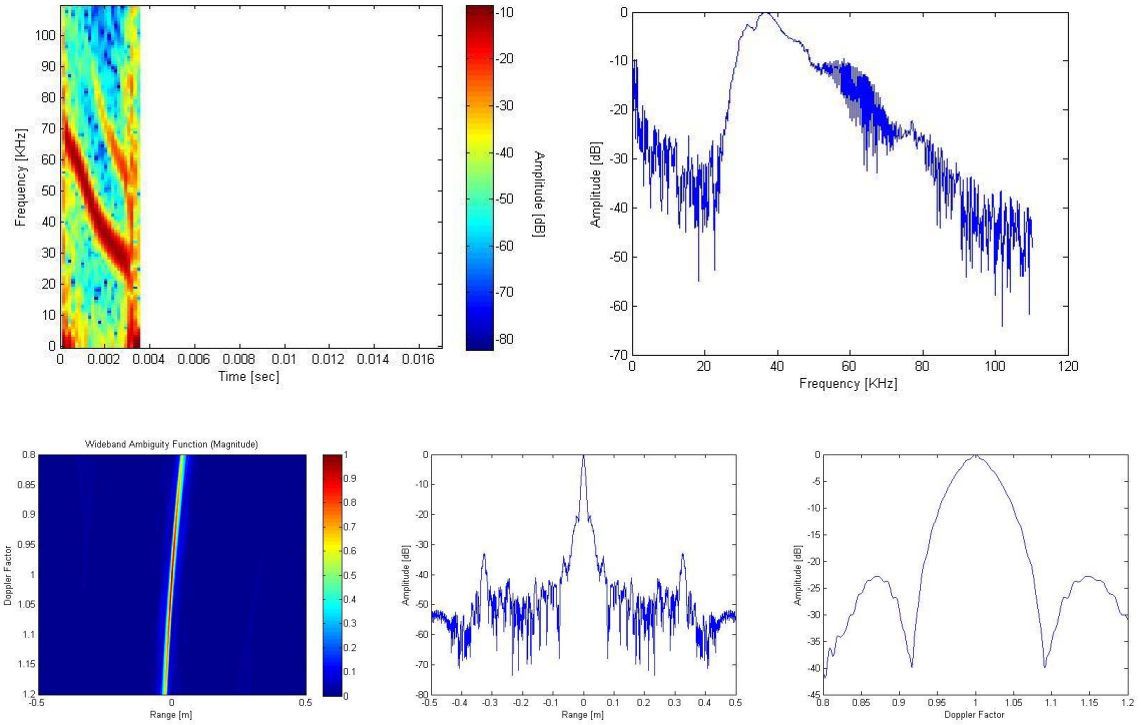

Figure 11: Spectrogram, mean spectrum and ambiguity function with relative range and Doppler cuts for a pulse in the terminal phase.

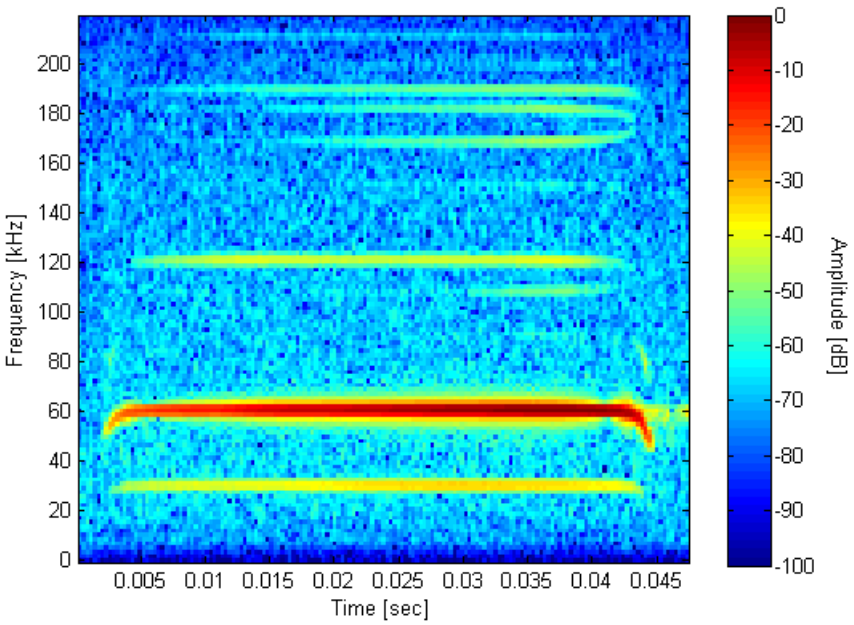

Figure 12: Example of a multi-harmonic CF signal with initial and final sweep.

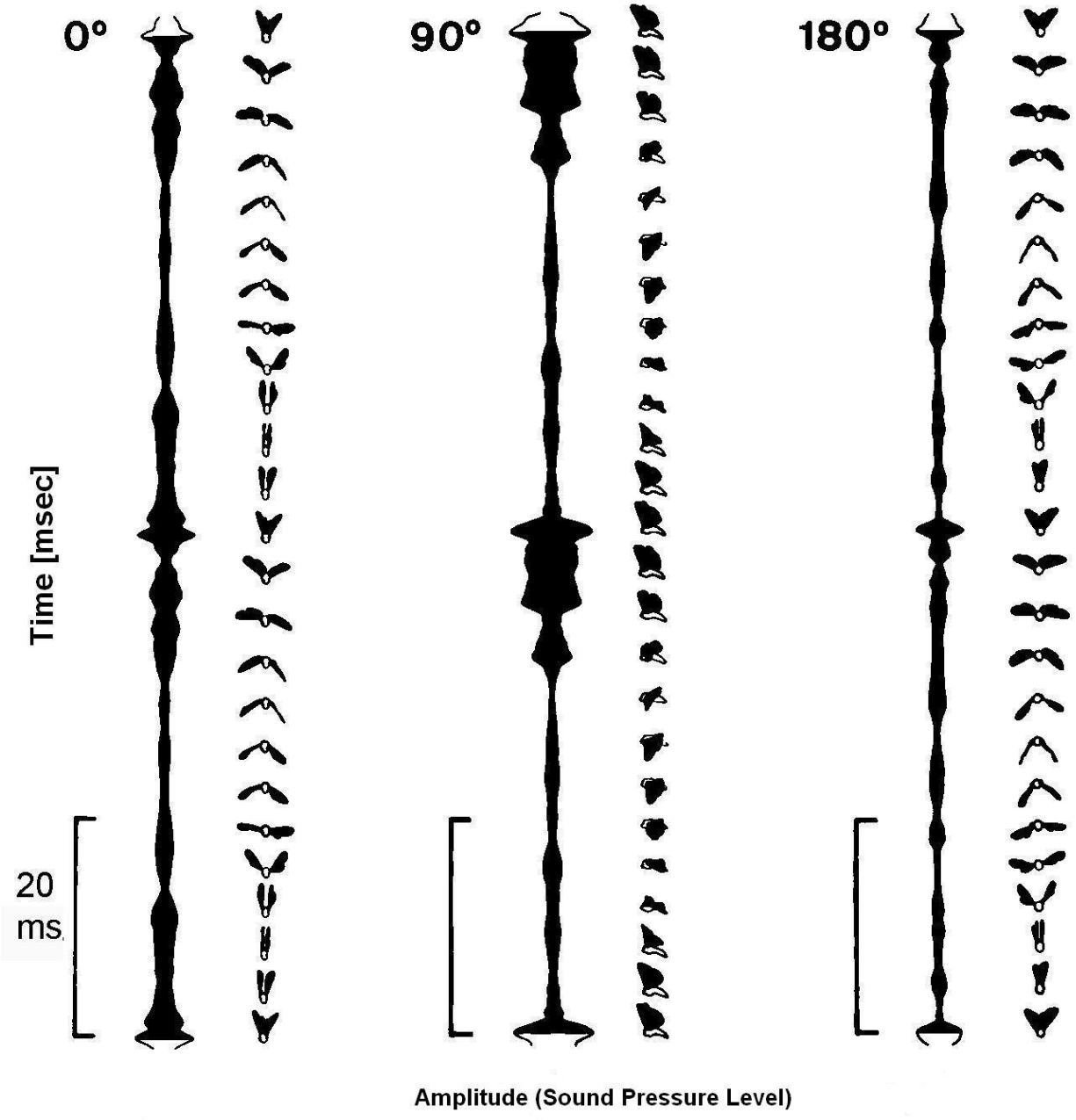

Figure 13: Amplitude-modulated echoes from a flying Autographa gamma at three different angles with equidistant phases of 2 wing beat cycles and their corresponding wingbeat phases. At 90 degrees the glint is produced at the top of the stroke, i.e. when the moth wings are perpendicular to the sound source. At 0 degree (frontal) and 180 degrees (rear), the glint occurs two phases after the upstroke suggesting that, in both cases, the same part of the wings produces the glint. Taken from [Sch83].

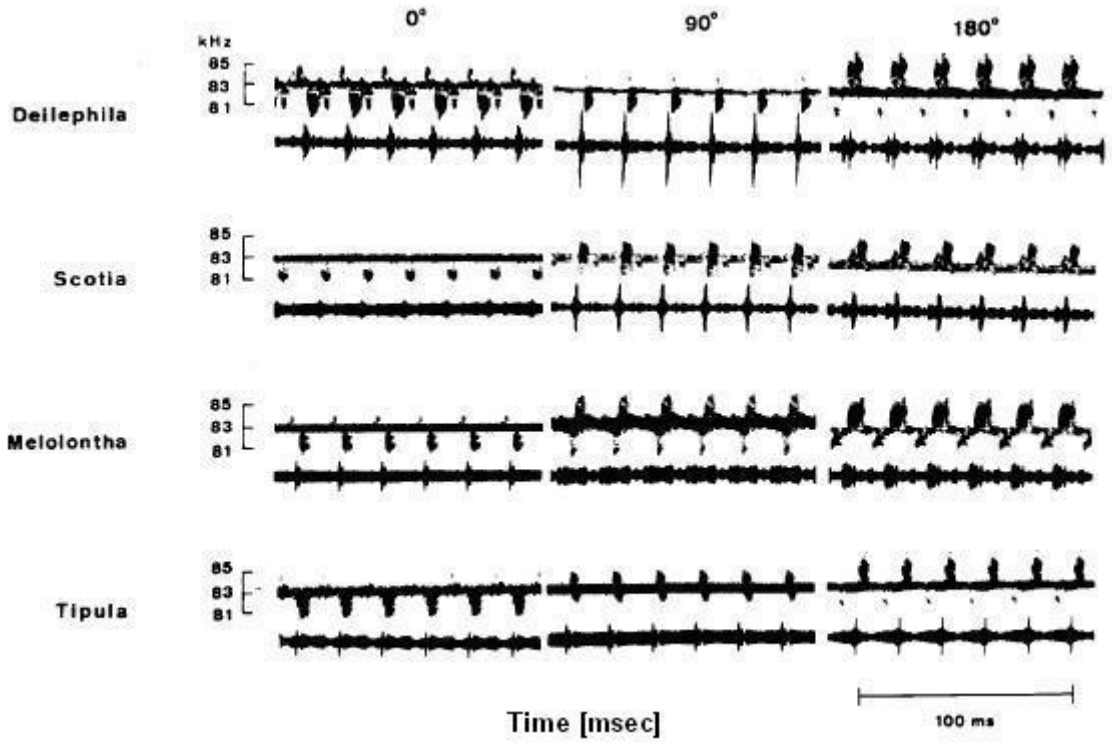

Figure 14: Echoes from four insect species all fluttering at 50 Hz but with different spectral patterns in echoes from their wing beats. For each insect, the upper trace represent the spectrogram of the echo and the lower trace the time oscillation. Taken from [Emde90].

Figure 15: Photo of a desiccated *Cobaea scandens* flower with protruding anthers, a bell-shaped corolla of largely merged petals whose unmerged petal ends fold back, and a ring of partly merged sepals at its base.

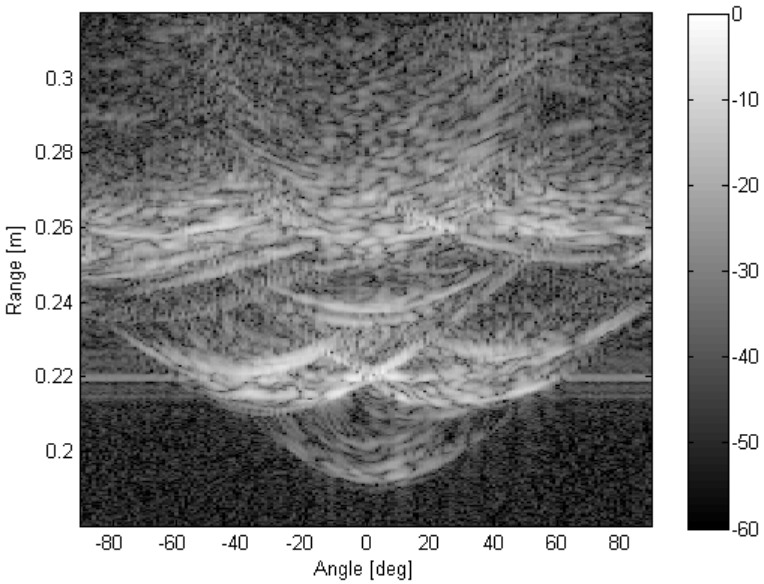

Figure 16: Magnitude of HRRPs over 180 degrees for a *C. scandens* ready for pollination. 0 degrees is to the front of the flower. HRRPs direction is horizontal, i.e. from left to right with respect to the flower's natural orientation. Scaled in relative [dB].

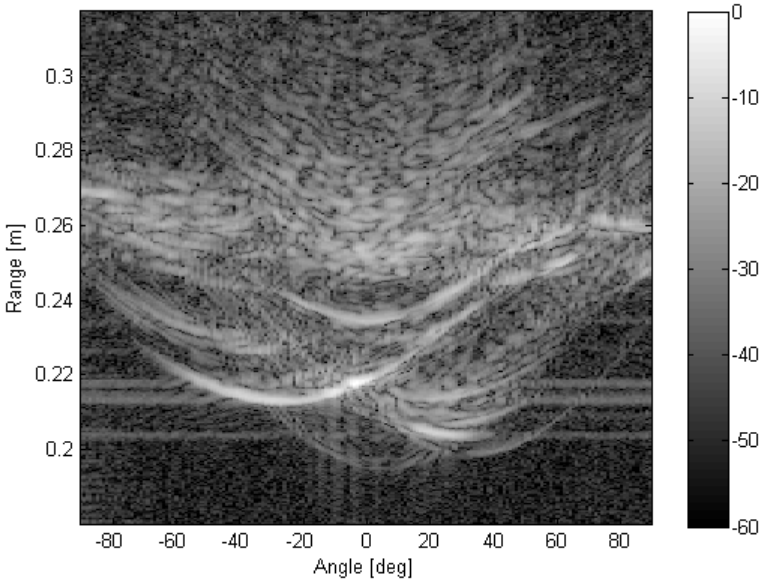

Figure 17: Magnitude of HRRPs over 180 degrees for a desiccated *C. scandens*.

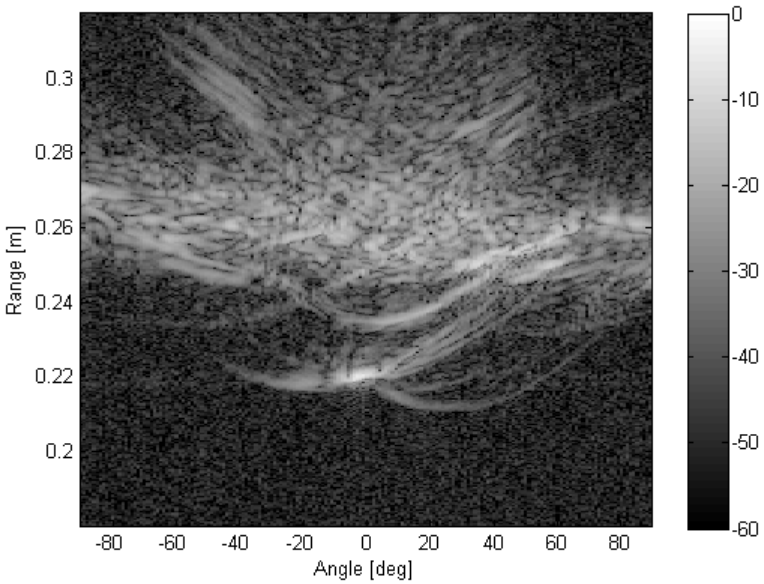

Figure 18: Magnitude of HRRPs over 180 degrees for a modified *C. scandens*.

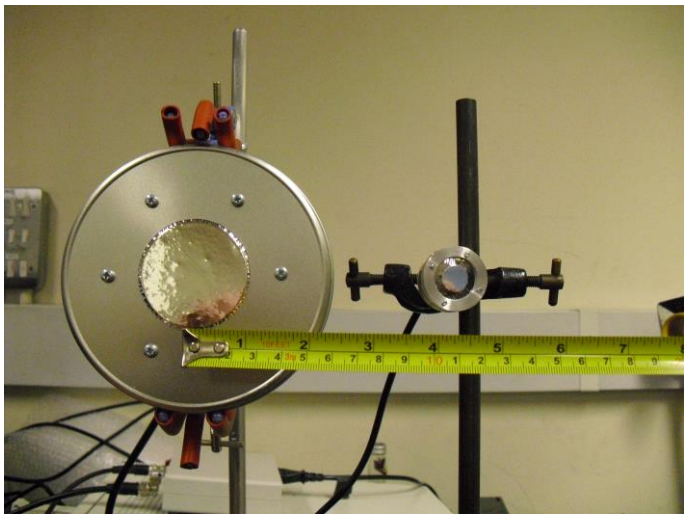

Figure 19: Frontal and lateral view of the space arrangement of the microphone and the loudspeaker during the experiment.

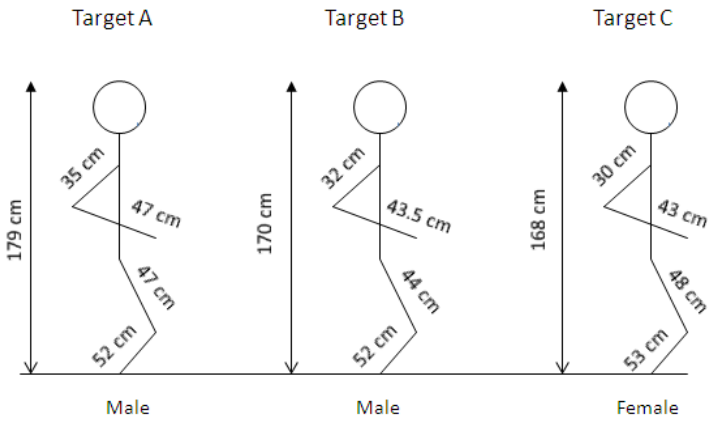

Figure 20: Physical characteristics of the three personnel targets, a female and two males, used for the experiments.

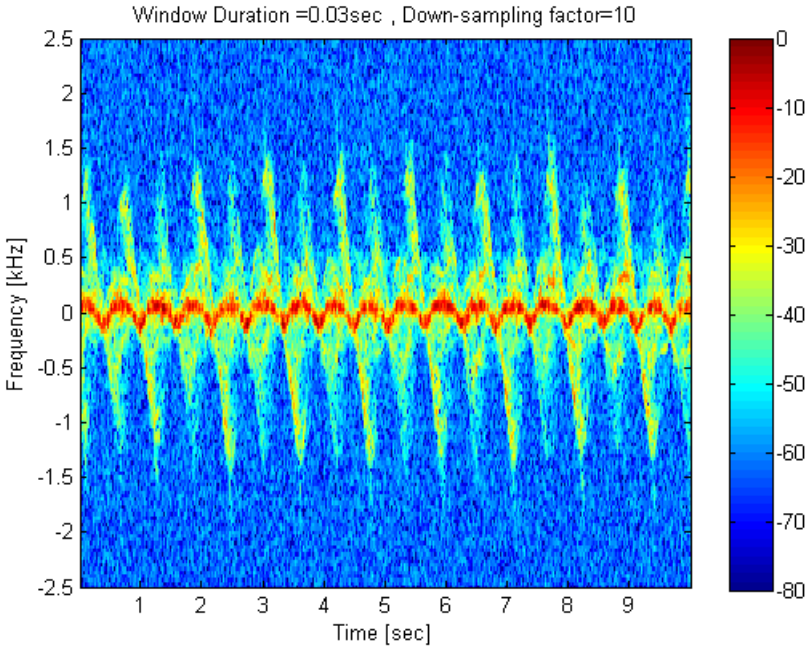

Figure 21: Micro-Doppler signature of *Target A* walking on a treadmill, facing the acoustic radar at a distance of 2m.

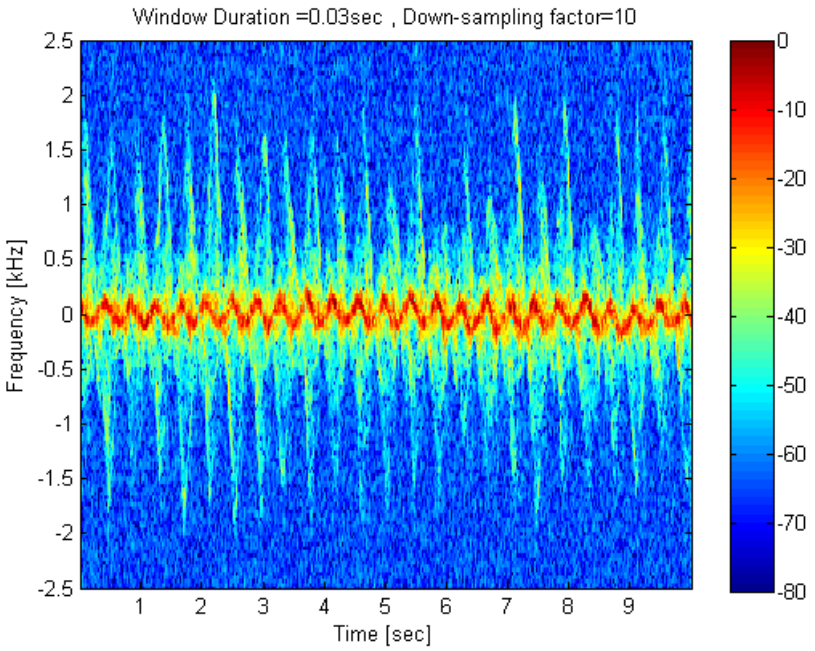

Figure 22: Micro-Doppler signature of *Target A* running on a treadmill.

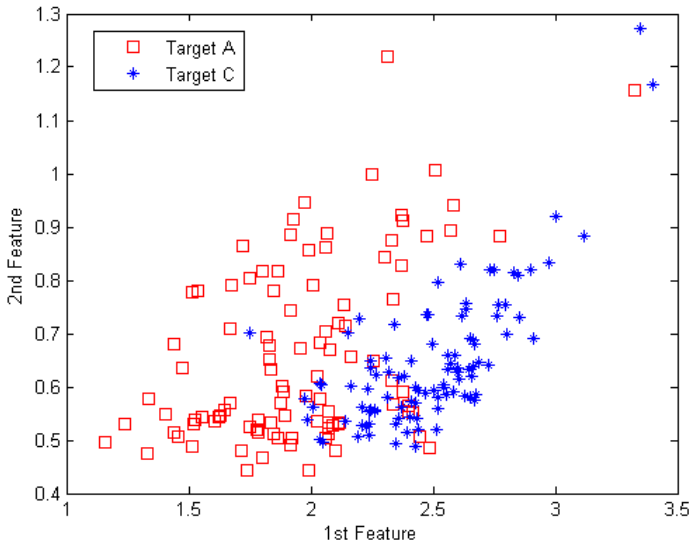

Figure 23: Plot of the first two Mel-Cepstrum features in the 2D plane for *Target A* (red) and *Target C* (blue).

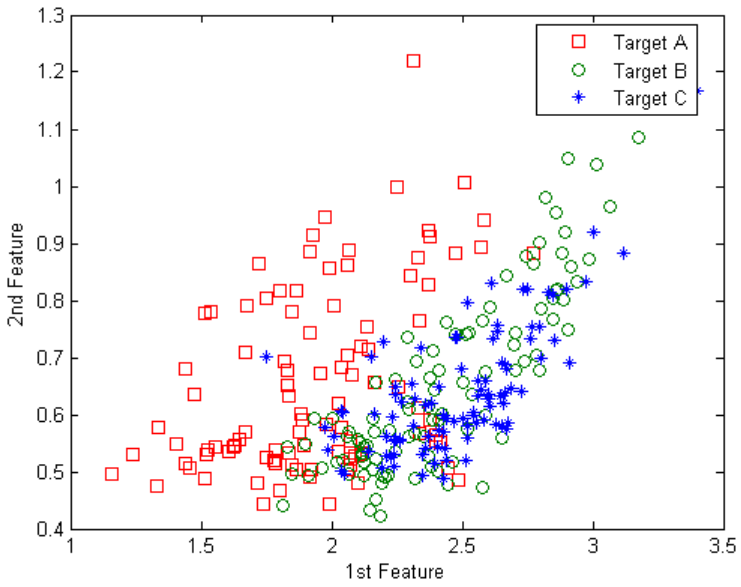

Figure 24: Plot of the first two Mel-Cepstrum features in the 2D plane for *Target A* (red), *Target B* (green) and *Target C* (blue).

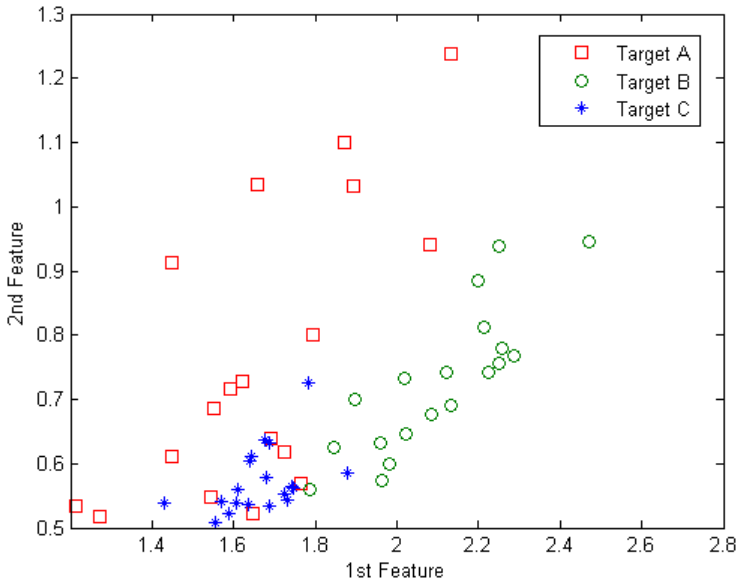

Figure 25: Plot of the first two Mel-Cepstrum features in the 2D plane for *Target A* (red), *Target B* (green) and *Target C* (blue).

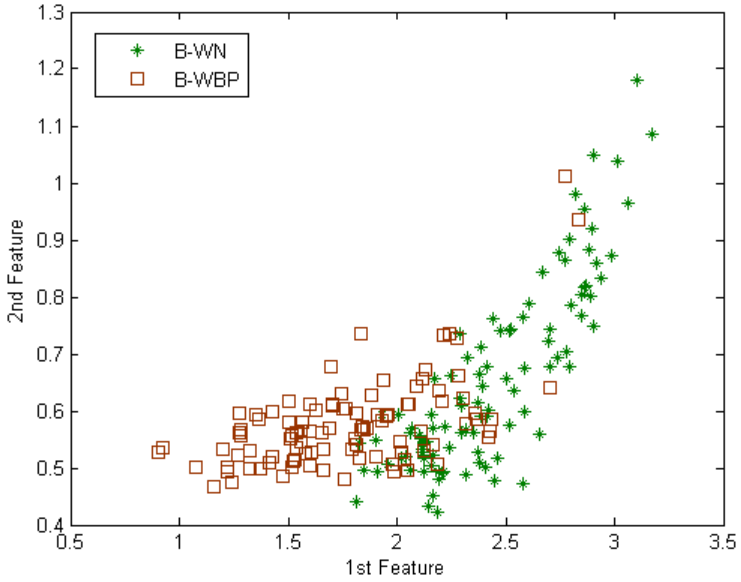

Figure 26: Plot of the first two Mel-Cepstrum features in the 2D plane for *Target B* walking normally (Class *B-WN*, green) and *Target B* walking with a heavy backpack on the shoulders (Class *B-WBP*, brown).

Table 1: ADAS single chirp parameters.

pulse length	T	341 ns
pulse repetition interval	PRI	2.048 ms
centre frequency	f_c	9.25 GHz
chirp rate	γ	1.46×10^{15}
pulse compression ratio	PCR	170.5
transmitted bandwidth	B_{tx}	500 MHz
sampling frequency	f_s	500 MHz
range resolution	Δr	30 cm

Table 2: Multi-perspective confusion matrices for three classifiers (K-NN, Naive Bayesian and FANNs) using different numbers of perspectives.

Classifier		*K*-NN, %			Bayesian, %			FANNs, %		
persp.:	class	*A*	*B*	*C*	*A*	*B*	*C*	*A*	*B*	*C*
1	*A*	79.37	1.72	18.91	78.39	6.90	14.71	89.73	1.4	8.87
	B	2.92	85.82	11.26	2.10	67.99	29.91	2.58	82.06	15.36
	C	1.92	20.85	77.23	4.86	10.34	84.80	1.63	11.49	86.88
2	*A*	91.61	0.48	7.91	90.38	1.74	7.88	96.77	0.15	3.08
	B	1.11	94.92	3.97	2.04	77.82	20.12	0.39	94.58	5.03
	C	0.25	15.97	83.78	3.83	8.20	87.97	0.45	5.97	93.58
3	*A*	92.26	0.07	7.67	95.27	0.71	4.02	98.12	0.05	1.83
	B	0	97.88	2.12	1.84	84.52	13.64	0.19	96.41	3.40
	C	0.04	7.59	92.37	2.58	6.68	90.74	0	4.03	95.97
4	*A*	96.88	0.01	3.11	98.24	0.19	1.57	99.44	0	0.56
	B	0	99.65	0.35	1.55	89.86	8.59	0.05	98.66	1.29
	C	0	7.05	92.95	1.78	5.66	92.56	0	2.14	97.86
5	*A*	97.34	0	2.66	98.93	0.09	0.98	99.75	0	0.25
	B	0	99.82	0.18	0.94	95.18	3.88	0	99.36	0.64
	C	0	3.5	96.5	1.09	4.99	93.92	0	1.20	98.80

Table 3: Multi-perspective correct classification rates (CCR) for the three classifiers using a progressive number of perspectives.

	K-NN, %	Bayesian, %	FANNs, %
1	80.81	77.06	86.22
2	90.10	85.39	94.97
3	94.17	90.17	96.83
4	96.49	93.55	98.65
5	97.88	96.01	99.30

Table 4: Total number of training windows N_{Train} and test windows N_{Test} used to assess the performance of the classifier for each window duration T_w.

	PCA		Cepstrum		Mel Cep.	
Naïve Bayesian	*Target A*	*Target C*	*Target A*	*Target C*	*Target A*	*Target C*
Target A	97%	3%	91.7%	8.3%	85%	15%
Target C	11.9%	88.1%	27.6%	72.4%	24%	76%
	$P_{cc} = 92.57\%$		$P_{cc} = 82.03\%$		$P_{cc} = 80.52\%$	

	PCA		Cepstrum		Mel Cep.	
K-NN	*Target A*	*Target C*	*Target A*	*Target C*	*Target A*	*Target C*
Target A	47.2%	52.8%	89.8%	10.2%	80.3%	19.7%
Target C	1%	99%	18.8%	81.2%	12.3%	87.7%
	$P_{cc} = 73.1\%$		$P_{cc} = 85.52\%$		$P_{cc} = 84\%$	

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

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

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

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

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

Table 6: Confusion matrices reporting the performance of the Naive Bayesian classifier and the K-NN classifier $k = 5$ testing the walking gait of *Target A*, *Target B* and *Target C* for all feature extraction algorithms ($T_w = 30$ ms).

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

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

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

Table 7: Confusion matrices reporting the performance of the Naive Bayesian classifier and the K-NN classifier ($k = 5$) testing the walking gait of *Target A*, *Target B* and *Target C* for all feature extraction algorithms ($T_w = 150$ ms).

	PCA		Cepstrum		Mel Cep.	
Naïve Bayesian	*B-WN*	*B-WBP*	*B-WN*	*B-WBP*	*B-WN*	*B-WBP*
B-WN	87.5%	12.5%	60%	40%	70%	30%
B-WBP	23.4%	76.6%	9.6%	90.4%	14.6%	85.4%
	$P_{cc} = 82.05\%$		$P_{cc} = 75.21\%$		$P_{cc} = 77.7\%$	

	PCA		Cepstrum		Mel Cep.	
K-NN	*B-WN*	*B-WBP*	*B-WN*	*B-WBP*	*B-WN*	*B-WBP*
B-WN	96.5%	3.5%	57.8%	42.2%	77.2%	22.8%
B-WBP	46.5%	53.5%	7%	93%	19%	81%
	$P_{cc} = 75\%$		$P_{cc} = 75.38\%$		$P_{cc} = 79.12\%$	

Table 8: Classification performance of a maximum likelihood classifier using Gaussian mixture models and trained to distinguish between various personnel targets; walking normally (Class 1), walking with an object in one hand (Class 2), walking with an object in two hands (Class 3), running (Class 4) and walking with a backsack (Class 5).

	Class 1	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 1	96.9%	0	0	1%	2.1%
Class 2	3.2%	88.9%	4.8%	0%	3.2%
Class 3	0%	3.2%	93.6%	0%	3.2%
Class 4	1.7%	0%	0%	98.3%	0%
Class 5	0%	0%	0%	0%	100%