Netted Sensors-based Vehicle Acoustic Classification at Tier 1 Nodes

Garry M. Jacyna[a], Carol T. Christou[a], Bryan George[a] and Burhan F. Necioglu[a]

[a]The MITRE Corporation, 7525 Colshire Dr., McLean, Virginia

ABSTRACT

The MITRE Corporation has embarked on a three-year internally-funded research program in netted sensors with applications to border monitoring, situational awareness in support of combat identification, and urban warfare. The first-year effort emphasized a border monitoring application for dismounted personnel and vehicle surveillance. This paper will focus primarily on the Tier 1 acoustic-based vehicle classification component. In a hierarchical network topology, distributed clusters of less capable nodes (Tier 1) at the lowest layer are in communication with more capable super nodes (Tier 2) at the next highest layer. Specifically, coarse information is aggregated at the lowest layer and communicated to the next layer in the network where more refined processing and information extraction can take place. We determined that fairly sophisticated classification processing can be performed at a Tier 1 node. This results in sizeable energy savings since data transfer rates are reduced across and between network layers. In general, acoustic-based vehicle classification is a difficult signal processing problem independent of the considerable hardware challenges. The acoustic waveforms are highly non-stationary and lack discernible harmonic structure, particularly for commercial vehicles that are designed to be intentionally quiet. In addition, the waveforms are rather complicated functions of the vehicle speed, engine RPM rate, and sensor-to-vehicle aspect angle. We discuss the development and implementation of a robust linear-weighted classifier on a Mica2 Crossbow mote using feature extraction algorithms specifically developed by MITRE for mote-based processing applications. These include a block floating point Fast Fourier Transform (FFT) algorithm and an 8-band proportional bandwidth filter bank. Results of in-field testing are compared and contrasted with theoretically-derived performance bounds.

Keywords: acoustics, feature extraction, linear discrimination, distributed classification, Mica2 Crossbow motes

1. INTRODUCTION

Netted Sensing (NS) has the potential to radically change how we address difficult sensing problems. Our concept of sensing has evolved over time from very early attempts to network homogeneous sensors together for specific missions to interoperable sensor netting approaches where heterogeneous sensors provide complementary views of the environment using a mixture of centralized and distributed node control, processing, and decision making. Netted Sensing reflects yet a further evolution along this path – low-cost pervasive monitoring and surveillance of the environment using highly distributed but expendable networked sensors. What sets netted sensing apart from previous approaches is the potential to support simultaneous missions using large and dynamically changing numbers of diverse heterogenous sensors for both remote and proximal sensing. There is no centralized control, processing, or decision making – information is distributed and fused within the network to provide overlapping views of the environment across space, time, and phenomenology.

Netted Sensing reflects a convergence of several different disciplines, including signal and information processing, computer architectures, software and hardware design, communications and networking, and information management. As a result, the problems associated with the design and implementation of netted sensing systems are enormous. Over two years ago, MITRE established a company-wide working group tasked with developing a netted sensors systems perspective. We solicited the views of our government sponsors and derived a list of technical issues relevant to a range of mission requirements. In addition, we conducted a survey of industry, academic, and government research and development (R&D) activities and products to establish what was state-of-the-art as opposed to state-of-the-practice. From this information, we identified technology shortfalls consistent with government mission requirements. Based on these identified needs, we recommended a multi-year R&D program in netted sensing that was subsequently funded in FY04.

Further author information: (Send correspondence to G. M. J.)
G. M. J. : E-mail: gjacyna@mitre.org, Telephone: 1 703 883 6972

MITRE's multi-year netted sensing program is driven by a set of sponsor challenge problems that provide a focal point for our research, experimentation, and development activities. Central to the success of this effort is the development of a test bed infrastructure that supports the validation and verification (V&V) of in-house algorithms while eventually supporting more extensive government V&V activities. With the inception of the program in FY04, we identified three challenge problems in the areas of border monitoring, situational awareness in support of combat identification, and urban warfare. Border monitoring is basically a perimeter surveillance problem. Situational awareness and urban warfare are, respectively, area and volume surveillance problems. Since the level of technical complexity grows as a function of spatial dimension, we logically chose to address the border monitoring problem in FY04. This effort led to the development of a number of core technology components that are being leveraged to address the area and volume surveillance problems in FY05 and FY06, respectively.

In FY04, MITRE demonstrated an approach to border monitoring using a combination of Tier 1 Mica2 Crossbow motes[1] and Tier 2 Sensoria nodes[2] to detect and classify dismounted personnel and commercial vehicles. This paper will focus on one aspect of this problem – a netted sensor-based solution for vehicle classification using a distributed collection of Tier 1 nodes. The concept of operations (CONOPS) is fairly simple in scope – we use a field of Tier 1 nodes equipped with an acoustic sensor to detect activity in an area, determine coarse vehicle type (car, bus, or truck), and estimate vehicle speed. This low data-rate information is communicated to a set of Tier 2 nodes, each containing a web-enabled camera. For a pre-specified coarse vehicle type, a picture of the vehicle is taken and relayed to a laptop computer through a standard 802.11b wireless communications link where a map of the area is displayed along with vehicle detection locations, estimated speeds, as well as pictures of select vehicles. Details regarding other aspects of the vehicle surveillance demonstration by B. Flanagan[3] and B. Necioglu[4] as well as information regarding the dismounted personnel demonstration[5] can be found elsewhere in these Proceedings.

Reliable vehicle classification is based on the use of multiple sensing modalities such as acoustic, video/EO, infrared, and magnetometric. In our FY04 vehicle surveillance demonstration, we restricted our purview to acoustic-based sensing, partly to limit overall system complexity and partly to determine how much useful information could be gleaned from acoustic vehicle signatures. A vehicle is a rather complicated noise mechanism. There are several sources of sound generation radiating from distinct areas around the vehicle such as the engine compartment, powertrain, manifold, exhaust, and tires. Analysis is complicated by the fact that the the aggregate frequency spectra and power levels are functions of the vehicle speed and vehicle-to-sensor aspect angle. One of the more interesting outputs of our study has been the observation that commercial vehicles lack exploitable narrowband harmonic structure, a characteristic that has been richly exploited for classifying military-style vehicles. There are two probable reasons – the harmonic signatures are overwhelmed by other broadband sources of sound generation, in particular, tire whine noise for vehicle speeds above 30 MPH and the fact that the commercial car industry has been successful in quieting annoying sources of narrowband noise radiating from the engine compartment and powertrain. We have developed a high-fidelity acoustic vehicle simulator to help us better understand the complex interplay between these noise mechanisms and how they can be exploited to refine our classification approach. A detailed description of the simulator by C. Christou can be found elsewhere in these Proceedings.[6]

This paper is organized as follows. Section 2 outlines the development of the feature extraction and classification components for the Mica2 Crossbow mote. Section 3 examines theoretically-derived bounds on classification performance as a function of vehicle type. Section 4 examines classifier performance using field data collected around the MITRE campus in the August/September 2004 time frame and compares these results to performance bounds derived in Section 3. Section 5 briefly summarizes the major conclusions of this work.

2. FEATURE EXTRACTION AND CLASSIFICATION

Fairly sophisticated vehicle classification can be performed at a Tier 1 (Mica2 Crossbow) node using a sparse collection of acoustic features and a low-complexity linear-weighted classifier. Figure 1 summarizes the classification processing flow within the Mica2 mote. Both the acoustic sensor and frontend data acquisition software/hardware are supplied from Crossbow – this includes the analog amplification/calibration circuitry and the A/D converter*. The signals are sampled at a 3750 Hz Nyquist rate and partitioned into 125 msec data-frames that are output every 250 msec. This allows us to balance our processing needs against onboard memory constraints. The feature extraction and classification operations are

*Only 8 of the available 10 bits provided by the A/D is used to conserve memory. We judiciously select which 8 bits to keep (bit selection) to preserve as much of the dynamic range as possible.

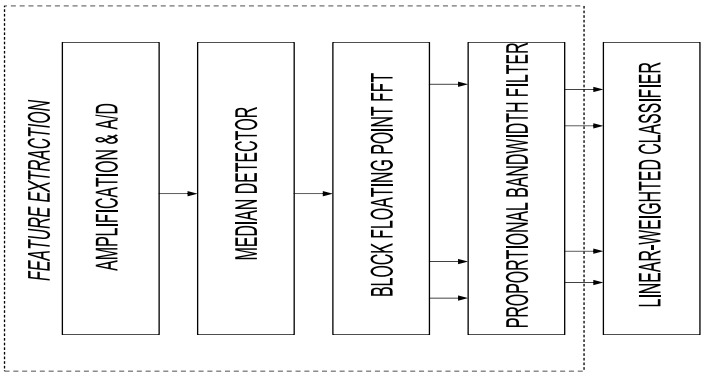

Figure 1. Flow chart of classification processing in the Mica2 Crossbow mote detailing the block floating-point FFT, eight-band proportional bandwidth filter, and linear-weighted classifier operations.

fairly processing intensive and are initiated only if a vehicle is detected. MITRE has developed a robust median-based detection algorithm that is both computationally efficient and provides reasonable detection sensitivity at low false alarm rates. Details regarding other aspects of this detector can be found elsewhere in these Proceedings.[3] In addition, MITRE has designed, implemented, and tested a 256 pt. block-floating point Fast Fourier Transform (FFT) algorithm that performs in near-real time on a Mica2 mote. The FFT outputs are used to construct the acoustic-based features.

Let $X(f_j)$ denote the FFT output corresponding to the j^{th} frequency bin. These outputs are magnitude-squared and summed to produce a set of proportional bandwidth-scaled powers of the form: $v_i = \sum_{j \in S_i} |X(f_j)|^2$, where S_i is the set of frequency bin indices included in the i^{th} filter band. Our feature vector is defined by the sequence $v_1, v_2, \ldots, v_I$. Specifically, we implement an eight-band proportional bandwidth filter ($I = 8$) with the following band characteristics: 60-100 Hz, 100-150 Hz, 150-225 Hz, 225-350 Hz, 350-650 Hz, 650-850 Hz, 850-1300 Hz, and 1300-1875 Hz. Figure 2 shows two plots of the proportional bandwidth filter output as a function of time for a representative car and truck class based on a small subset of collected field data. Both plots have been image-normalized, so that comparisons based on absolute power differences between the car and truck are not possible. Most of the radiated energy from the car is confined to mid-range frequencies and short time intervals. In contrast, most of the radiated truck energy is spread across mid- and high-range frequencies and longer time intervals. We developed the proportional bandwidth filter to accentuate these spectral differences. The observations in Fig. 2 are generally true for the majority of car and truck data examined. However, there are exceptions – a car can sound like a truck if the manifold, muffler, or powertrain is damaged.

Each of the eight feature components v_i; $i = 1, 2, \ldots, 8$ are input to the linear-weighted classifier. Under zero-mean Gaussian signal assumptions, it is possible to show that v_i is χ^2 distributed with $2N_i$ degrees of freedom, where N_i is the number frequency bins in the i^{th} filter band. In addition, v_i and v_j are statistically independent whenever $i \neq j$. It is shown in Appendix A that the optimal *a posteriori* classifier under Gaussian signal assumptions is a linear-weighted classifier of the form:

$$\ell = \sum_{i=1}^{8} \gamma_i v_i + \beta \,, \tag{1}$$

where $\gamma_i = 1/\sigma_{2i}^2 - 1/\sigma_{1i}^2$, $\beta = \sum_{i=1}^{8} N_i \log(\sigma_{2i}^2/\sigma_{1i}^2)$, and σ_{1i}^2 and σ_{2i}^2 are the average powers (variances) for the i^{th} filter band under hypotheses H_1 and H_2, respectively. The classifier is trained by estimating the parameters γ_i and β from the following expressions developed in Appendix A: $\hat{\gamma}_i = N_i(1/\hat{P}_{2i} - 1/\hat{P}_{1i})$ and $\hat{\beta} = \sum_{i=1}^{8} N_i \log(\hat{P}_{2i}/\hat{P}_{1i})$, where $\hat{P}_{1i}$ and $\hat{P}_{2i}$ are the *estimated* powers derived from the training data for the i^{th} filter band under hypotheses H_1 and H_2, respectively.

As discussed later in Sect. 3, the classifier is quite sensitive to absolute power differences between vehicle classes. Thus, for example, if the training data reflects the fact that trucks are always louder than cars, then a vehicle will always be classified as a truck if it is loud, independent of its spectral shape. This would create a problem for certain cars that are actually louder than trucks due to a number of factors such as poor maintenance or design. A more logical approach

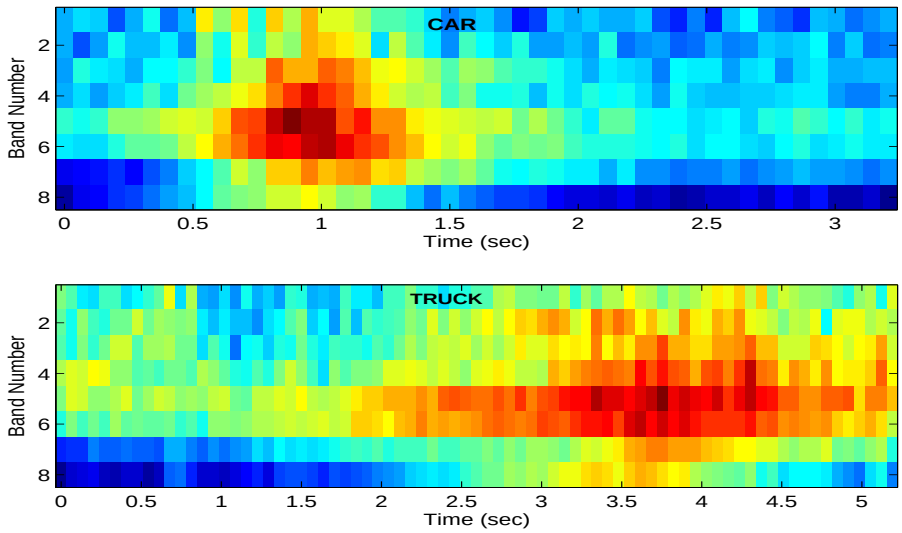

Figure 2. Log-power of the output of an eight-band proportional bandwidth filter as a function of time for a representative car and truck class.

would base classification decisions on spectral shape. This requires a simple modification to the above expressions. We simply redefine $\hat{P}_{ki}$ as $\hat{P}_{ki}/\sum_{i=1}^{8}\hat{P}_{ki}$, for $i = 1, 2, \ldots, 8$ and $k = 1, 2$, i.e., the power is normalized to unity under each hypothesis.

3. THEORETICAL CLASSIFICATION PERFORMANCE

In this section, we examine two methods for estimating the performance of the linear-weighted classifier. The first method provides an exact characterization of classifier performance when the acoustic signals are assumed to be zero-mean Gaussian random processes. The second method establishes an upper bound on the classification error probability useful in determining the effects of spectral shape and variations in the absolute power level.

3.1. Exact Performance Characterization

This analysis is based on the Fourier Transform relationship between a probability density function and its characteristic function. It is shown in Appendix A that the density function for the instantaneous powers at the output of an I-band proportional bandwidth filter is a product of χ^2 distributions with $2N_i$ degrees of freedom, where the characteristic function can be written in closed-form. Using similar arguments, the characteristic function for the classifier sufficient statistic ℓ can also be written in closed-form. A classification error occurs whenever $\ell < 0$ under the H_1 hypothesis or $\ell > 0$ under the H_2 hypothesis. The probabilities associated with these cases are: Prob($\ell < 0|H_1$) and Prob($\ell > 0|H_2$). Using characteristic functions, it is shown in Appendix B that:

$$\begin{aligned} \text{Prob}(\ell < 0|H_1) &= \int_{-\infty-j\epsilon}^{\infty-j\epsilon} [C_\ell(\omega|H_1)/j\omega]\, d\omega/2\pi \\ \text{Prob}(\ell > 0|H_2) &= -\int_{-\infty+j\epsilon}^{\infty+j\epsilon} [C_\ell(\omega|H_2)/j\omega]\, d\omega/2\pi\,, \end{aligned} \tag{2}$$

where the characteristic functions $C_\ell(\omega|H_1)$ and $C_\ell(\omega|H_2)$ are of the form:

$$\begin{aligned} C_\ell(\omega|H_1) &= \exp(-j\omega\beta)\prod_{i=1}^{I}(1/(1+j\omega(\psi_i-1))^{N_i} \\ C_\ell(\omega|H_2) &= \exp(-j\omega\beta)\prod_{i=1}^{I}(1/(1+j\omega(1-1/\psi_i))^{N_i}\,, \end{aligned} \tag{3}$$

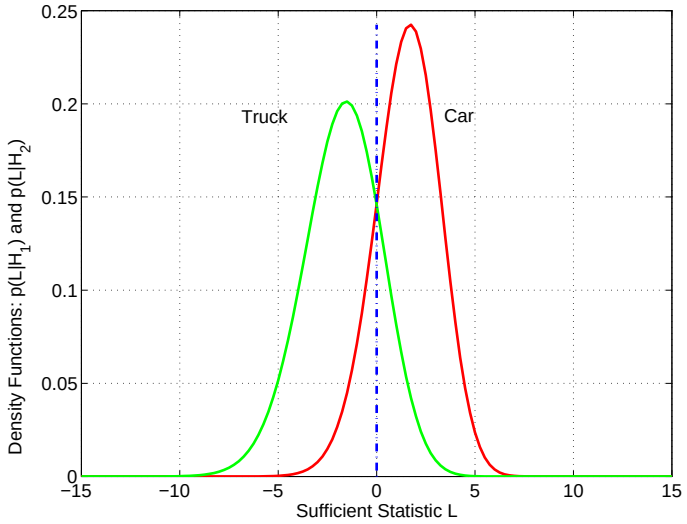

Figure 3. Theoretically-derived density functions $p(\ell|H_k)$ for representative car (H_1) and truck (H_2) classes using the estimated filter band power levels from collected field data.

where $\psi_i = \sigma_{1i}^2/\sigma_{2i}^2$, $\beta = -\sum_{i=1}^{I} N_i \log(\psi_i)$, ϵ is a small positive number, and j is the imaginary number. The classification error probability $\text{Prob}_\text{C}(\text{error})$ is determined from a weighted sum of the individual error probabilities: $\text{Prob}_\text{C}(\text{error}) = P_1\text{Prob}(\ell < 0|H_1) + P_2\text{Prob}(\ell > 0|H_2)$, where P_1 and P_2 are the *a priori* probabilities associated with hypotheses H_1 and H_2, respectively. These equations can be implemented efficiently using standard numerical integration packages.

The characteristic functions in Eq. (3) can also be used to determine the density functions $p(\ell|H_k)$ through the inverse Fourier Transform. Figure 3 is a representative plot of these density functions for typical car (H_1) and truck (H_2) classes. The power levels in each filter band were estimated from collected field data to compute ψ_i for an 8-band proportional bandwidth filter (I = 8). Note that the dotted line in the figure is the decision threshold for the hypothesis test, i.e., if $\ell > 0$, a car is declared; otherwise, a truck is declared. Since $p(\ell|H_1) \neq 0$ for $\ell < 0$ and $p(\ell|H_2) \neq 0$ for $\ell > 0$, the classification error probability is non-zero. In fact, based on numerical results using Eqs. (2) and (3), the error probability is approximately 12%. A more detailed set of comparisons with collected field data is deferred to Sect. 4.

3.2. Chernoff Upper Bound

Although Eqs. (2) and (3) result in a refined characterization of classification performance, these expressions are not amenable to simple analysis. However, the Chernoff bounds normally lead to closed-form analytical expressions for the classification error probabilities. A Chernoff bound is derived from the moment generating functions associated with $p(\ell|H_1)$ and $p(\ell|H_2)$. Since both densities are products of χ^2 distributions, it is shown in Appendix B that the error probability $\text{Prob}_\text{C}(\text{error})$ under equi-probable hypotheses is upper bounded by the following expression:

$$\text{Prob}_\text{C}(\text{error}) \leq \exp(\mu(s))/2\,, \tag{4}$$

where $\mu(s) = -(1-s)\beta - \sum_{i=1}^{I} N_i \log(s + (1-s)\psi_i)$, s is chosen such that $d\mu(s)/ds = 0$ for $0 \leq s \leq 1$, and ψ_i and β are defined above. We now examine classification performance as a function of constant power offset (e.g., one vehicle much louder than another vehicle) and variations in spectral shape.

3.2.1. Constant power offset

Suppose that the ratios of powers (ψ_i) at the output of a I-band proportional bandwidth filter are constant and that vehicle 1 (H_1) is louder than vehicle 2 (H_2) so that $\psi_i = \eta > 1$. Using Eq. (4), we can write: $\mu(s) = (1-s)N\log(\eta) - N\log(s + (1-s)\eta)$, where $N = \sum_{i=1}^{I} N_i$. Let s^* denote the value of s for which $\mu(s)$ is minimized. In Appendix B, it is shown that: $\mu(s^*) = N + N\log(\eta)/(1-\eta) - N\log(-(1-\eta)/\log(\eta))$. Some simple analysis also shows that $\mu(s^*) \leq 0$ and $d\mu(s^*)/d\eta < 0$ for $\eta > 1$. When $0 \leq \eta < 1$, it follows that $d\mu(s^*)/d\eta > 0$. Since $\text{Prob}_\text{C}(\text{error})$ is a monotonically

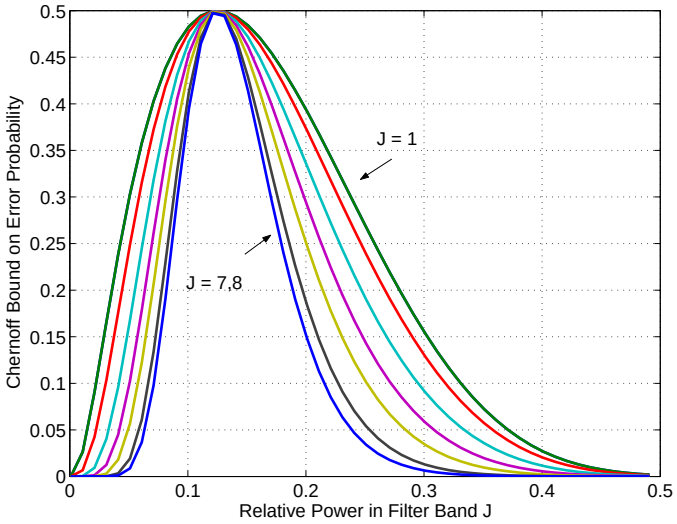

Figure 4. Chernoff bounds on classification error probabilities as a function of relative power level η_j and filter band selection where the overall power across bands is a constant.

increasing function of $\mu(s^*)$, we conclude that the classification error probability decreases as vehicle 1 becomes louder than vehicle 2. In fact, the converse is also true – if vehicle 2 becomes louder than vehicle 1, then the error probability also decreases. This follows from the monotonically decreasing dependence of $\mu(s^*)$ on η when $0 \leq \eta < 1$.

The implication of this analysis is that the linear-weighted classifier is extremely sensitive to variations in the *absolute power* levels between vehicle types. Unfortunately, absolute power level is *not* a reliable classification measure – some cars are actually louder than trucks, particularly if damage has been done to the manifold or exhaust system. We found that spectral shape is a more reliable indicator of coarse vehicle type under a constant power assumption.

3.2.2. Spectral shape under a constant power assumption

In order to keep the analysis tractable, we vary the power level in only one of I filter bands such that the overall power across bands is a constant. In particular, let $\psi_j = I\eta_j$ and $\psi_i = I\eta$, where $\eta_j + (I-1)\eta = 1$ so that $\eta = (1-\eta_j)/(I-1)$. Using Eq. (4), we can derive a closed-form expression for $\mu(s)$ in terms of η_j, η, N_i, and s. The details are outlined in Appendix B (Eq. (20)), where it is shown that $\text{Prob}_\text{C}(\text{error}) \leq \exp(\phi)/2$ for $\phi = \min_s(\mu(s))$.

Figure 4 is a plot of the Chernoff bounds for the classification error probabilities as a function of the relative power level in the j^{th} filter band for $j = 1, 2, \ldots, 8$ under a constant overall power assumption. As described in Sect. 2, the proportional bandwidth filter is designed such that $N_i > N_j$ for $i > j$, i.e., the bandwidth increases as the band-centered frequencies increase. We first note that the error probabilities reach a maximum at $\eta_j = 1/8$, independent of the particular filter band. This corresponds to the case $\psi_i = \psi_j = 1$ where both vehicle classes have the same spectral shape, i.e., the outcome is comparable to a coin toss (50%). We also note that the error probabilities increase as η_j increases over the range $0 \leq \eta_j < 1/8$ and subsequently decrease as η_j increases for $\eta_j > 1/8$. This is simply stating the obvious – the larger the difference in relative spectral shape between the two vehicle classes the lower the corresponding classification error.

The envelope of error probabilities leads to some interesting conclusions. For the lower filter bands ($j = 1, 2$), power variations have less of an effect on the error probabilities than corresponding variations at the upper bands ($j = 7, 8$). Keeping in mind that $N_i > N_j$ for $i > j$, it is understandable that the higher filter bands would accentuate relative power differences whereas the lower bands would be less sensitive to such differences. The proportional bandwidth filter was designed specifically to exploit these effects. Cars, trucks, and SUVs pretty much sound alike at low frequencies (ignoring diesel trucks that radiate appreciable energy at low frequencies). It is only at the higher frequencies that spectrally distinguishing features emerge. These higher frequencies are emphasized in the filter while allowing for some capability at the lower frequencies to discriminate diesel trucks from everything else.

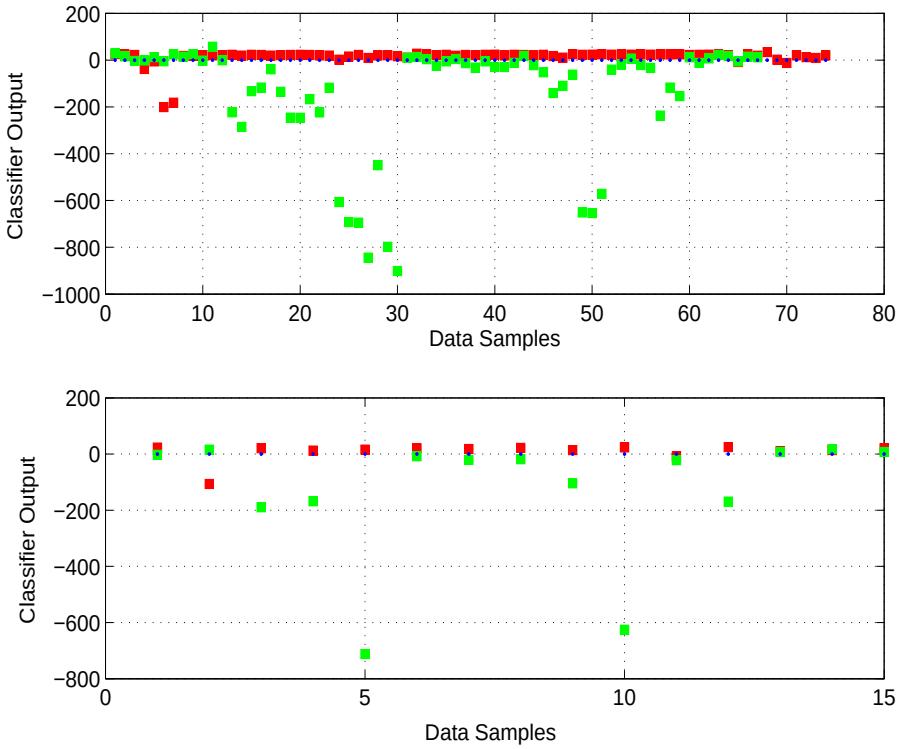

Figure 5. Linear-weighted un-averaged (top figure) and averaged (bottom figure) classifier threshold crossings for a representative sampling of cars (red) and trucks (green) based on collected field data.

4. CLASSIFICATION PERFORMANCE USING COLLECTED FIELD DATA

The Mica2 mote and acoustic sensor were enclosed in a foam-lined plastic container with holes drilled in the bottom to protect it from the elements. A fabric liner was inserted in the bottom to act as a windscreen. The container was fitted with four legs and stood roughly two inches off the ground. Since we were interested in collecting data for testing purposes, each mote was connected to an interface board using a 51-position ribbon cable through a slit cut in the side of the container. Each interface board was connected to a Linux notebook computer via a serial port. Data collections were performed in August/September 2004. During post processing, the time-stamped audio segments were manually categorized into vehicle types using accompanying video footage. There were more than 100 minutes of total data collected comprising 3311 car and 415 truck audio segments.

A small but representative sampling of the linear-weighted classifier results are shown in Figure 5 for a combination of car and truck classes. The upper plot depicts the classifier threshold crossings for distinct (un-averaged) 125 msec data segments, where the threshold is the blue dotted line extending across the zero ordinate. Note that the data has *not* been normalized to unity power so relative power level differences between vehicle classes are important. The cars are denoted by red dots and trucks by green dots. Cars (trucks) above (below) the threshold are classified correctly. Otherwise, misclassifications occur. The number of false threshold crossings are accumulated for each vehicle type and normalized by the total number of vehicles to produce an estimate of the error probabilities Prob($\ell < 0|$car) and Prob($\ell > 0|$truck). The *a priori* probabilities P_1 and P_2 for each vehicle type are determined from the known number of car and truck instantiations, where the error probability is subsequently computed from: P_1Prob($\ell < 0|$car) + P_2Prob($\ell > 0|$truck). The estimated error probability was determined to be roughly 13%, which is in good agreement with the theoretical prediction of 12% in Sect. 3. When classification is based on spectral shape (unity power) alone, the estimated error probability increases to roughly 23%, a result validated by our theoretical performance predictions.

The second plot in Figure 5 depicts the threshold crossings for a variant of the linear-weighted classifier where the classifier output statistic ℓ is averaged over blocks comprising five 125 msec data segments each. The estimated error probabilities are roughly the same. For Mica2 implementation, this approach is more desirable than the previous method

from a communications perspective. Only one classification decision needs to be communicated to a Tier 2 node for each vehicle. The previous approach would have required multiple classification decisions and multiple communications to a Tier 2 node for each vehicle.

A more extensive examination of vehicle classification performance for Gaussian mixture and linear-weighted classifiers are addressed elsewhere in these Proceedings.[4]

5. SUMMARY & CONCLUSIONS

This paper has outlined a procedure for classifying various coarse vehicle types on a Mica2 Crossbow mote using only an onboard acoustic sensor. MITRE developed several novel algorithms that allowed this processing to be performed in near real time given the considerable hardware and memory challenges. The classification logic included a median-based detector, a 256 pt. block floating-point FFT algorithm, an 8-band proportional bandwidth filter, and a linear-weighted classifier. This paper has specifically discussed the design, implementation, and testing of the proportional bandwidth filter and classifier. We outlined the development of the classification algorithm from first principles and developed two performance models using both a numerically-exact characteristic function formulation and a Chernoff bound. We also examined how classification performance changes as a function of the absolute power differences between vehicle types and the spectral shape. Comparisons relative to collected field data showed good agreement with these theoretical predictions.

We are extending this work in FY05 to address vehicle classification and tracking using a combination of multi-modal sensors – acoustic, magnetometric, seismic, and passive IR. To date, we have developed a low-complexity tracking algorithm on a Mica2 mote that uses closest-point-of-approach (CPA) times along with the maximum powers near CPA to develop piecewise-linear track solutions for vehicles passing through a field of motes. We are in the process of extending this approach to include feature-based tracking for multi-target applications. This work is in support of our FY05 challenge problem focus on situational awareness.

APPENDIX A. THE LINEAR-WEIGHTED CLASSIFIER

This appendix outlines the development and training of the Crossbow mote-based linear-weighted classifier. The first subsection derives the optimal *a posteriori* classifier under Gaussian signal assumptions for an I-band proportional bandwidth filter. The second subsection discusses how the classifier was trained.

A.1. Classifier Development

The output of an I-band proportional bandwidth filter can be written as $\mathbf{v}_k = [v_{k1}, v_{k2}, \ldots, v_{kI}]$, where the component $v_{ki} = \sum_{j \in S_i} |X_k(f_j)|^2$, $X_k(f_j)$ is the Fourier Transform at frequency f_j for the H_k hypothesis, and S_i is the set of frequency bin indices that are included in the i^{th} filter band. The signals are assumed to be zero-mean Gaussian processes. This implies that $X_k(f_j)$ is a zero-mean complex Gaussian random variable. For sufficiently large time-bandwidth products, $X_k(f_j)$ and $X_k(f_{j'})$ are uncorrelated and independent for $j \neq j'$. Writing $X_k(f_j) = X_k^i(f_j) + jX_k^q(f_j)$, where X_k^i and X_k^q are the in-phase (real) and quadrature (imaginary) components of $X_k(f_j)$, respectively, it is easy to show that X_k^i and X_k^q are both uncorrelated and independent.

Let $E(|X_k(f_j)|^2) = \sigma_{ki}^2$, where $E(\cdot)$ is the expectation operator and σ_{ki}^2 is the (constant) variance associated with the complex Fourier Transform for all frequencies f_j such that $j \in S_i$. Both the in-phase and quadrature components have identical variances of the form $\sigma_{ki}^2/2$. We conclude that v_{ki} is a χ^2 random variable with $2N_i$ degrees of freedom, where N_i are the number of elements (cardinality) in S_i. The corresponding density function can be written as[7]:

$$p(v_i|H_k) = \frac{v_i^{N_i-1}\exp(-v_i/\sigma_{ki}^2)}{\sigma_{ki}^{2N_i}(N_i-1)!};\ v_i \geq 0\,. \tag{5}$$

We now show that v_i and v_j $(i \neq j)$ are independent random variables by using a characteristic function argument. The multi-dimensional characteristic function can be written as:

$$M(\alpha_1, \alpha_2, \ldots, \alpha_I) = \int\int \cdots \int \exp\left(j\sum_{i=1}^{I}\alpha_i v_i\right) p(X(f_1), X(f_2), \ldots, X(f_N))\, d\mathbf{X}\,, \tag{6}$$

where N denotes the total number of frequency bins, $p(\cdot)$ is the corresponding joint density function, and the limits of integration implicitly extend from $-\infty$ to ∞. For notational simplicity, we have suppressed the subscript k on the Fourier Transforms $X(f_j)$ and v_i. Now, since the Fourier Transforms at different frequencies are uncorrelated under a large time-bandwidth product assumption, both transforms are also statistically independent since the corresponding density functions are complex Gaussian. This implies that the joint density can be factored into a product of marginal densities of the form:

$$\prod_{l=1}^{N} p(X(f_l)) = \left(\prod_{j \in S_1} p(X(f_j)) \right) \cdots \left(\prod_{j \in S_I} p(X(f_j)) \right) . \tag{7}$$

Substituting this result into Eq. (6) and simplifying, it is easily shown that $M(\alpha_1, \alpha_2, \ldots, \alpha_I) = \prod_{i=1}^{I} M(\alpha_i)$. Since the multi-dimensional characteristic function can be factored into a product of individual characteristic functions, the random variables $v_1, v_2, \ldots, v_I$ are statistically independent and we can write $p(v_1, v_2, \ldots, v_I) = \prod_{i=1}^{I} p(v_i)$. Using the definition of $\mathbf{v}$ and Eq. (5), the joint density $p(\mathbf{v}|H_k)$ can be expressed as:

$$p(\mathbf{v}|H_k) = \prod_{i=1}^{I} \left(\frac{v_i^{N_i - 1} \exp(-v_i/\sigma_{ki}^2)}{\sigma_{ki}^{2N_i} (N_i - 1)!} \right) . \tag{8}$$

For a maximum *a posteriori* classifier,[8] we choose hypothesis H_1 if $P_1 p(\mathbf{v}|H_1) > P_2 P(\mathbf{v}|H_2)$ and hypothesis H_2 otherwise, where P_1 and P_2 are the *a priori* probabilities associated with hypotheses H_1 and H_2, respectively. Without loss of generality, we set $P_1 = P_2 = 1/2$. Since the logarithm is a monotonically increasing function of its argument, we can simplify the above expression with the help of Eq. (8) so that the resulting classifier can be written in the following more manageable form:

$$\begin{aligned} &\text{Choose } H_1 \text{ if:} \quad \textstyle\sum_{i=1}^{I} \gamma_i v_i + \beta > 0 \\ &\text{Choose } H_2 \text{ if:} \quad \textstyle\sum_{i=1}^{I} \gamma_i v_i + \beta < 0 \,, \end{aligned} \tag{9}$$

where $\gamma_i = 1/\sigma_{2i}^2 - 1/\sigma_{1i}^2$ and the bias term is $\beta = \sum_{i=1}^{I} N_i \log(\sigma_{2i}^2/\sigma_{1i}^2)$.

A.2. Training as an Estimation Problem

Training entails estimating the parameters γ_i and β associated with Eq. (9). Since these parameters are functions of σ_{ki}^2 for $k = 1, 2$, the problem becomes one of estimating the individual variances associated with both hypotheses. Assume M independent training sets of the form $\mathbf{v}_k^{(m)}$; $m = 1, 2, \ldots, M$ under hypothesis H_k, where $\mathbf{v}_k$ was defined above. This implies that:

$$p(\mathbf{v}^{(m)}|\sigma_{k1}^2, \sigma_{k2}^2, \ldots, \sigma_{kI}^2; H_k) = \prod_{m=1}^{M} p(\mathbf{v}^{(m)}|H_k) \,, \tag{10}$$

where $p(\mathbf{v}^{(m)}|H_k)$ is the joint χ^2 density given by Eq. (8) for the m^{th} training set. We have explicitly included the dependence on the unknown variances. We can simplify the above expression by examining the logarithm of the density and noting that the components $v_1^{(m)}, v_2^{(m)}, \ldots, v_I^{(m)}$ are statistically independent as a result of the analysis above.

The maximum likelihood estimate of σ_{ki}^2 follows necessarily from the condition:

$$\partial \log(p(v_i^{(1)}, v_i^{(2)}, \ldots, v_i^{(M)}|\sigma_{ki}^2; H_k)/\partial \sigma_{ki}^2 = 0 \,, \tag{11}$$

where a little algebraic manipulation results in the following estimate for σ_{ki}^2: $\hat{\sigma}_{ki}^2 = \hat{P}_{ki}/N_i$, where $\hat{P}_{ki} = \sum_{m=1}^{M} v_{ki}^{(m)}/M$ is just the estimated average power in the i^{th} filter band under the H_k hypothesis. Using the expressions for γ_i and β above, we form the following estimates used for training the classifier:

$$\begin{aligned} \hat{\gamma}_i &= N_i(1/\hat{P}_{2i} - 1/\hat{P}_{1i}) \\ \hat{\beta} &= \sum_{i=1}^{I} N_i \log(\hat{P}_{2i}/\hat{P}_{1i}) \,. \end{aligned} \tag{12}$$

APPENDIX B. THEORETICAL PERFORMANCE CHARACTERIZATION

This appendix develops the theoretical performance bounds for the linear-weighted classifier. The first subsection outlines a numerical procedure that leads to an exact performance characterization under the stated assumptions. The second subsection examines another procedure based on the Chernoff bound. This bound is appropriate for gaining insight into classification performance when the average power within each filter band is varied.

B.1. Numerically-based Performance Characterization

The approach is based on the use of the characteristic function. Let ℓ be defined as the classifier sufficient statistic corresponding to Eq. (9). Since v_i is χ^2 distributed with $2N_i$ degrees of freedom, the characteristic function for v_i under hypothesis H_k can be written as[8]:

$$C_{v_i}(\omega|H_k) = \left(\frac{1}{1+j\omega\sigma_{ki}^2}\right)^{N_i} . \tag{13}$$

Including the contributions from γ_i and the bias term β, it is simple to show that γ_i enters into Eq. (13) as a scale term on the variance σ_{ki}^2 while the bias term transforms into a constant multiplier of the form $\exp(-j\omega\beta)$. Therefore, the characteristic function for the sufficient statistic ℓ is:

$$C_\ell(\omega|H_k) = \exp(-j\omega\beta)\prod_{i=1}^{I}\left(\frac{1}{1+j\omega\gamma_i\sigma_{ki}^2}\right)^{N_i} . \tag{14}$$

The density function $p(\ell|H_k)$ can be computed from the inverse Fourier Transform of the characteristic function. Equation (14) can be further simplified by noting that $\gamma_i\sigma_{1i}^2 = \psi_i - 1$ and $\gamma_i\sigma_{2i}^2 = 1 - 1/\psi_i$, where $\psi_i = \sigma_{1i}^2/\sigma_{2i}^2$. In addition, the bias term can be rewritten as $\beta = -\sum_{i=1}^{I} N_i \log(\psi_i)$.

In order to determine the classification error, we need to compute the following probabilities: Prob$(\ell < 0|H_1)$ and Prob$(\ell > 0|H_2)$ which can be determined from the density $p(\ell|H_k)$. However, these probabilities can be determined directly from the characteristic function in Eq. (14) by noting, for example, that:

$$\text{Prob}(\ell < 0|H_1) = \int_{-\infty}^{0}\int_{-\infty}^{\infty} \exp(j\omega\ell) C_\ell(\omega|H_1)\, d\omega\, d\ell/2\pi\,. \tag{15}$$

If we examine the integral $\Delta = \int_{-\infty}^{0}\exp(j\omega\ell)\,d\ell$ in Eq. (15), we note that Δ does not exist in the strict sense as long as ω is confined to the real axis. However, if we perturb ω off the real axis in the proper manner, then Δ is bounded. Since $\ell < 0$, we redefine ω as $\omega - j\epsilon$, where $\epsilon > 0$ and small[†]. In this case $\Delta = -j/(\omega - j\epsilon)$. For Prob$(\ell > 0|H_2)$, it is easy to show that we have to redefine ω as $\omega + j\epsilon$, in which case $\Delta = j/(\omega + j\epsilon)$. Therefore, the probability expression in Eq. (15) and the corresponding probability for $\ell > 0$ under H_2 can be rewritten in a more numerically convenient form as:

$$\begin{aligned}
\text{Prob}(\ell < 0|H_1) &= \int_{-\infty-j\epsilon}^{\infty-j\epsilon} [C_\ell(\omega|H_1)/j\omega]\, d\omega/2\pi \\
\text{Prob}(\ell > 0|H_2) &= -\int_{-\infty+j\epsilon}^{\infty+j\epsilon} [C_\ell(\omega|H_2)/j\omega]\, d\omega/2\pi\,.
\end{aligned} \tag{16}$$

Both integrals can be numerically approximated using standard integration packages provided that the integration step size $\Delta\omega$ is chosen smaller than ϵ. The classification error probability can now be written as: $\text{Prob}_\text{C}(\text{error}) = \text{Prob}(\ell < 0|H_1)/2 + \text{Prob}(\ell > 0|H_2)/2$ for equi-probable events.

[†]More formally, the perturbed path for ω is referred to as the Bromwich path and is closely related to the concept of the Cauchy Principal Value.[9]

B.2. Chernoff Bound

The Chernoff bound provides a closed-form upper bound on classification performance. It is based on knowledge of the moment-generating functions for the densities $p(\mathbf{v}|H_k)$ defined by Eq. (8).[10] For our particular problem, the bound can be written as $\text{Prob}_\text{C}(\text{error}) \leq \exp(\mu(s))/2$, where:

$$\mu(s) = \log\left(\int\int\ldots\int p^s(\mathbf{v}|H_1)p^{1-s}(\mathbf{v}|H_2)\,d\mathbf{v}\right), \tag{17}$$

and s is chosen such that $d\mu(s)/ds = 0$, provided $0 \leq s \leq 1$. The limits of integration in Eq. (17) extend from $-\infty$ to ∞. Substituting Eq. (8) into Eq. (17) and simplifying, we find:

$$\mu(s) = -(1-s)\beta - \sum_{i=1}^{I} N_i \log(s + (1-s)\psi_i), \tag{18}$$

where ψ_i is defined in the previous subsection. There are two cases we now wish to examine.

B.2.1. Constant power offset

Here we are interested in examining the effect of a power offset on classification performance. This can occur if one vehicle (truck) is much louder than another vehicle (car). In general, we would normalize the powers so that classification decisions are based on the *shape* of the frequency spectrum rather than on *absolute power* levels.

Let $\psi_i = \eta$; $i = 1, 2, \ldots, I$ so that the ratio of powers under the H_1 and H_2 hypotheses, respectively, is constant within each filter band. Substituting ψ_i into Eq. (18) and simplifying, we find: $\mu(s) = (1-s)N\log(\eta) - N\log(s + (1-s)\eta)$, where $N = \sum_{i=1}^{I} N_i$. A minimum of $\mu(s)$ necessarily occurs when $d\mu(s)/ds = 0$, for which $s^* = -\eta/(1-\eta) - 1/\log(\eta)$. This immediately implies that $\text{Prob}_\text{C}(\text{error}) \leq \exp(\mu(s^*))/2$, where

$$\mu(s^*) = N + N\log(\eta)/(1-\eta) - N\log(-(1-\eta)/\log(\eta)). \tag{19}$$

From this equation, it is not difficult to show that $\mu(s^*) \leq 0$ and $d\mu(s^*)/d\eta > 0$ for $0 \leq \eta < 1$ and $d\mu(s^*)/d\eta < 0$ for $\eta > 1$. Since the upper bound on the classification error probability in Eq. (19) is a monotonically increasing function of $\mu(s^*)$, we conclude that the error probability increases as η increases within the range $0 \leq \eta < 1$ and decreases as η increases for $\eta > 1$. Therefore, as vehicle 1 becomes louder or quieter than vehicle 2, the classification error probability decreases.

B.2.2. Variable spectrum under a constant power assumption

In this analysis, we examine the effect of spectral shape on classification performance under a constant power assumption. In order to keep the analysis tractable, it is assumed that the power level in only one of I filter bands can be modified. The power levels in the remaining $I-1$ bands are adjusted to maintain an overall constant power level.

Let $\psi_j = I\eta_j$ and $\psi_i = I\eta$ for all $i \neq j$. We normalize the overall power such that $\eta_j + (I-1)\eta = 1$, in which case $\eta = (1-\eta_j)/(I-1)$. It follows necessarily that $\psi_i = I(1-\eta_j)/(I-1)$ for all $i \neq j$. Note that other normalizations are also possible. Substituting these expressions into Eq. (19) and simplifying, we find that:

$$\begin{aligned}\mu(s) \quad = \quad & (1-s)(N-N_j)\log(\psi_i) + (1-s)N_j\log(\psi_j) \\ & -(N-N_j)\log(s+(1-s)\psi_i) - N_j\log(s+(1-s)\psi_j),\end{aligned} \tag{20}$$

where N_j are the number of frequency bins in the j^{th} filter band and $N = \sum_{i=1}^{I} N_i$. The upper bound on the classification error probability can now be expressed as: $\text{Prob}_\text{C}(\text{error}) \leq \exp(\phi)/2$, where $\phi = \min_s \mu(s)$. In Sect. 3, this upper bound is plotted as a function of η_j for $j = 1, 2, \ldots, 8$.

ACKNOWLEDGMENTS

This work was funded under the auspices of MITRE's Technology Program. The authors would like to thank Mr. Marcus Glenn and Dr. Anton Haug for their careful review of this paper.

REFERENCES

1. Crossbow Technology, "Smarter sensors in silicon," *URL: http://www.xbow.com*, 2004.
2. Sensoria Corporation, "Energy aware netted embedded systems for tactical unattended ground sensors," *URL: http://www.sensoria.com/pdf/SPIEPaperApril2003.pdf*, 2003.
3. B. Flanagan, "Robust distributed detection using low-power sensors," in *SPIE Defense & Security Symposium*, J. McConnell, ed., *Unattended Ground Sensor Technologies and Applications VII*, 2005.
4. B. F. Necioglu, "Vehicle acoustic classification in netted sensor systems using gaussian mixture models," in *SPIE Defense & Security Symposium*, J. McConnell, ed., *Signal Processing, Sensor Fusion, and Target Recognition XIV*, 2005.
5. M. Otero, "Application of a continuous wave radar for human gait recognition," in *SPIE Defense & Security Symposium*, J. McConnell, ed., *Signal Processing, Sensor Fusion, and Target Recognition XIV*, 2005.
6. C. T. Christou, "Simulation of vehicle acoustics in support of netted sensor research and development," in *SPIE Defense & Security Symposium*, J. McConnell, ed., *Unattended Ground Sensor Technologies and Applications VII*, 2005.
7. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1965.
8. H. L. VanTrees, *Detection, Estimation, and Modulation Theory: Part I*, John Wiley and Sons, New York, 1968.
9. A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.
10. H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Annals Math. Stat.* **23**, pp. 493–507, 1962.