

Real-time odor discrimination using a bioelectronic sensor array based on the insect electroantennogram

A J Myrick¹, K-C Park², J R Hetling¹ and T C Baker²

¹ Department of Bioengineering, University of Illinois at Chicago, SEO 232, MC 063, 851 South Morgan Street, Chicago, IL 60607-7052, USA

² Department of Entomology, Chemical Ecology Laboratory, Pennsylvania State University, University Park, PA 16802, USA

Received 30 September 2007

Accepted for publication 7 October 2008

Published 7 November 2008

Online at stacks.iop.org/BB/3/046006

Abstract

Current trends in artificial nose research are strongly influenced by knowledge of biological olfactory systems. Insects have evolved over millions of years to detect and maneuver toward a food source or mate, or away from predators. The insect olfactory system is able to identify volatiles on a time scale that matches their ability to maneuver. Here, biological olfactory sense organs, insect antennae, have been exploited in a hybrid-device biosensor, demonstrating the ability to identify individual strands of odor in a plume passing over the sensor on a sub-second time scale. A portable system was designed to utilize the electrophysiological responses recorded from a sensor array composed of male or female antennae from four or eight different species of insects (a multi-channel electroantennogram, EAG). A computational analysis strategy that allows discrimination between odors in real time is described in detail. Following a training period, both semi-parametric and k -nearest neighbor (k -NN) classifiers with the ability to discard ambiguous responses are applied toward the classification of up to eight odors. EAG responses to individual strands in an odor plume are classified or discarded as ambiguous with a delay (sensor response to classification report) on the order of 1 s. The dependence of classification error rate on several parameters is described. Finally, the performance of the approach is compared to that of a minimal conditional risk classifier.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Detection and identification of airborne volatile chemical compounds (i.e. odors) have many potential applications in military, industrial, clinical and research areas. To this end, development of the artificial nose approach, whose concept was first published in 1982 [1], has been accelerating for the past decade [2]. The artificial nose can be distinguished from other chemical detectors (such as pH or NO electrodes) by the promise of detecting a number of different compounds with the same device. Artificial noses can complement more established analytical methods such as gas chromatography, mass spectrometry, infrared spectroscopy and ion mobility spectroscopy [3]. Artificial

noses, mimicking their biological counterparts, generally consist of an array of sensors that demonstrate specificity when information from all sensors is combined. Such arrays have been constructed using several sensor types reviewed in [4, 5]. A valuable tool for the detection of volatile compounds, known as the electroantennogram (EAG), has been suggested for the purpose of detecting compounds with a single insect antenna [6]. Sensors that change resistivity in response to odorants include conducting polymers, metal oxides, and carbon black/polymer. Devices that detect minute changes in mass due to adsorption of various compounds include surface acoustic wave and quartz crystal devices. Other sensors include coated optical fiber sensor arrays and porous gate MOSFETs. The response patterns from an array of

sensors must be interpreted by a pattern recognition scheme; this has been accomplished using, amongst other techniques, statistical, computational neural networks approaches and non-parametric classifiers such as the k -nearest neighbor technique (k -NN) [7], which was pioneered in 1951 by Fix and Hodges [8], and later analyzed in detail by Cover and Hart [9].

There are two important limitations of current artificial nose technologies. First, the long response times (tens of seconds to minutes) of most approaches limit them to steady-state measurements, where steady-state may take impractically long times to reach under transient field conditions, even when response is accelerated by techniques such as increasing the flow rate across the surface of polymer based resistor sensors [10]. The EAG response to a single short duration filament of odor can be as short as 15–75 ms, and can carry ~ 20 bits/s of information (species *Spodoptera exigua* and *Cadra cautella*) [11]. Some unique applications of fast odor sensors include estimating the location of an odor source at a close range, or detecting odor plumes at a high speed in aerial searches. Individual strands will in general be highly concentrated (rather than being diffused and mixed with other odorants) but short-lived. The ability to rapidly detect discrete packets of odor in the air stream also supports the ability to detect an odor of interest when it is surrounded by interfering sources. If the sensor detected the mean odor concentration over tens of seconds, two discrete sources, A and B, would likely be perceived as a unique source, C, consisting of a mixture of A and B. When scalar concentrations in the odor plume remain unmixed, the current system would detect a series of A and B strands instead of the mixture. A second limitation of most artificial nose technologies is low sensitivity [5, 12].

Properties of the biological olfactory system include short response time and high sensitivity. The principal arrangement of the biological olfactory system is quite well conserved across phyla, from insects to mammals. Olfactory receptor neurons (ORNs) exhibit a response when airborne molecules bind to metabotropic membrane receptors and activate G-protein-coupled cascades, providing amplification and eventually leading to membrane potential changes and characteristic trains of action potentials [13–16]. In insects [17], information obtained by ~ 40 types of ORNs, numbering in the thousands to hundreds of thousands, is organized and transmitted to the antennal lobe [18, 19]. In moths, the odorant sensitive dendrites of ORNs reside in the antennae, encapsulated in structures called sensilla [20], where volatile compounds enter the sensillar lymph through nanoscale pores in the cuticle. Sensillar lymph contains high concentrations of odorant binding proteins, which ferry odorant molecules to the receptors on ORN dendrites [18]. The subsequent mechanism of molecular recognition is complex and has not been precisely described. However, structure-activity studies (chain-elongation, double-bond position, functionality) performed on noctuid moth olfactory neurons *in vivo* have been particularly enlightening in understanding that ligand–receptor interactions can behave according to conformational energy and electron distribution

models and not merely to space-filling [21–23]. The result of the sensitive transduction and the antennal lobe processing is a system that exhibits a remarkably high sensitivity with broad-band detection and discrimination.

Insect antennae are highly sensitive to odors of survival interest but also to compounds such as explosives and controlled drugs (unpublished data). The EAG, the biopotential developed between two points on an insect antenna, is the result of the massed response of the ORNs to an odor stimulus; several groups have shown the potential use of insect antennae and the EAG in a hybrid-device biosensor [24–26]. However, each of these studies made use of a single antenna, which cannot provide discrimination between odors. For the past several years, a hybrid system for odor detection based on an array of insect antennae has been under development [27–29]. The proof of concept for a system that can potentially discriminate between different odors was demonstrated when averaged electroantennogram responses to controlled puffs from living insects of five different species showed species-specific EAG response spectra to 20 volatile compounds tested. The EAG response profiles were then reconstructed for each compound across the five insect species. Most of the compounds could be distinguished by visually comparing the response spectra. A four-antenna array was then implemented to discriminate among three odorants based on the relative EAG amplitudes evoked when the probe was placed in a wind tunnel or in an outdoor field [27]. Stable EAG responses could be simultaneously recorded from four different insect antennae, and different volatile compounds could be distinguished in real time by visually comparing relative EAG responses recorded from a combination of differently tuned insect antennae. The first computational odor discrimination on individual strands of odor was reported using pheromonal components [28], and subsequently preliminary results involving both pheromonal and non-pheromonal odors [29] using four EAG channels were presented. Additions to the system include an increase from four to eight channels, the addition of a semi-parametric classifier used for detection, and an offline minimal conditional risk classifier. Using this system, we have shown the ability to detect and classify eight non-pheromonal odors.

Here we describe hardware and software suitable for field use that can computationally discriminate between *single strands* of both pheromonal and non-pheromonal odors based on the EAG response recorded simultaneously from antennae of four or eight different antennae in odor plumes created in a windtunnel. The pattern recognition system as applied to an electronic nose (EN) has application-specific requirements such as the need to reject outliers (unknown odors), to reject patterns which are ambiguous to reduce the false alarm rate and to assign a degree of certainty to each classification operation [30].

Several problems are inherent in the detection and classification of strands in a natural plume. Odor plumes contain background odors, and EAG waveforms also contain noise, giving rise to mixture probability densities. We cannot measure true positives, where true positives, in the limit, approach a transduction event involving one molecule

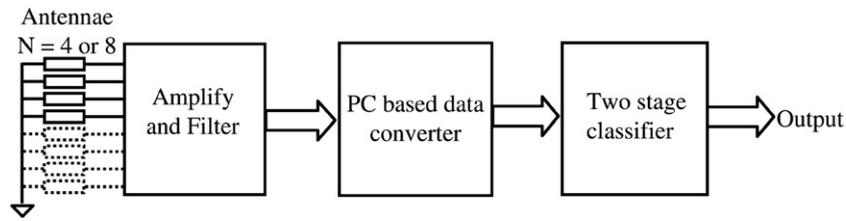


Figure 1. Block diagram of the EAG based biosensor system depicting major operations.

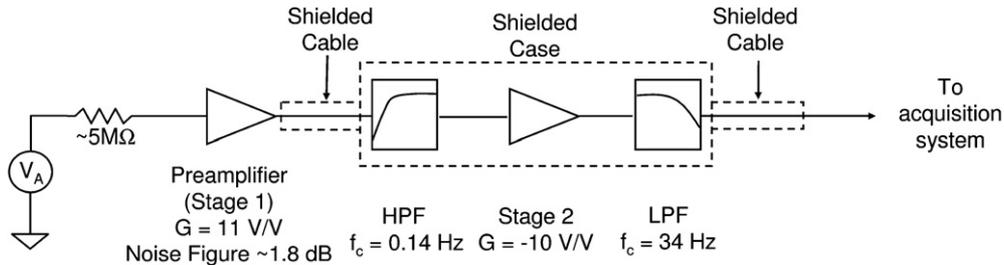


Figure 2. Block diagram of one channel of the EAG amplifier. Four channels are implemented in parallel to record from four different antennae simultaneously. Eight channels are achieved by utilizing two amplifiers. Amplified and filtered signals are passed to a 16 channel acquisition system and then to a classifier implemented on a portable PC. G, gain; f_c , filter corner frequency; HPF, high-pass filter; LPF, low pass filter; V_A , antenna voltage.

during some time interval (noise–odorant mixture). The general problem is identified in the literature as learning with an imperfect teacher or supervisor [31, 32], and mixture densities in particular can be addressed using the expectation-maximization (EM) technique. Therefore, error rate cannot be estimated through error counting of the classifier, but it can be estimated with varying degrees of accuracy through modeling. Regardless, it must be decided which positives are worth classifying and which are not, which is correctly approached using a minimal conditional risk classifier. An analogous problem without imperfect teaching is encountered in the design of active radar systems and likely radar warning receivers, which must passively receive and classify emitters. Solutions to the automatic target recognition problem in radar imaging usually divide the procedure into several steps to reduce computation, namely detection, discrimination and classification [33], and this approach is also employed here, with certain differences. Further, in a natural plume, the multiple antennae may not necessarily be identically and simultaneously exposed, which adds further uncertainty to the measurements. However, this does not prevent use of faster sensors for identification of odors over longer periods of time, or even the use of an odor delivery system optimized for the response of the sensors.

Semi-parametric and k -NN classifiers comprise a simple way to implement a real-time electronic nose that demonstrates an actual realization of the idea proposed in [27]. In this study, we compare our results in an ad hoc fashion to a more optimal conditional risk minimizing classifier offline.

2. Methods

Figure 1 depicts a block diagram of the complete biosensor. Hardware and software design are described separately below.

2.1. Hardware

Below, a brief description of the hardware built and employed to measure multiple simultaneous EAGs is given. Spectral characterization of the amplifier noise, EAG baseline (no odors) and EAG signal in a plume are summarized.

2.1.1. Design. The system described here acquires up to 16 EAG signals simultaneously for real-time processing and classification on a laptop PC. A block diagram for one channel of the EAG acquisition system is shown in figure 2. Each four-channel preamplifier connects via a shielded cable to a second stage amplifier. The second stage amplifier is connected via another cable to an analog to digital converter (ADC) and a laptop computer.

The preamplifier is small in size, exhibits sufficiently low noise and virtually non-existent input bias current as a result of amplifier selection (the quad CMOS input operational amplifier LMC6084) and board design which includes guard rings surrounding each input. The small size minimizes disruption of the air stream passing by the sensor and allows for future expansion to more EAG channels. A concern related to the preamplifier is the noise it adds to the input signal. Johnson noise, unavoidable noise that is present across any source of real finite impedance (the antenna in this case), is created by thermal motion of charge carriers. The noise figure, expressed in decibels, is a measure of how much noise the amplifier adds to the signal over thermal noise present at the input:

$$NF = 10 \log_{10} \left(\frac{S_i/N_i}{S_o/N_o} \right), \quad (1)$$

where S_i/N_i is the signal to thermal noise ratio at the input of the amplifier. The calculated noise figure of each amplifier from 0.1 to 25 Hz is 1.8 dB when a nominal antenna resistance of 5 MΩ is attached to the LMC6084. The finished preamplifier,

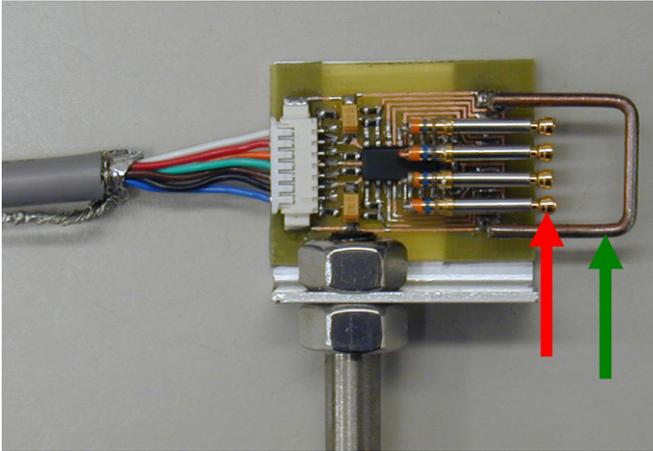


Figure 3. Photograph of a completed EAG preamplifier, including the four electrodes used to contact the four excised insect antennae. Each antenna extends from one recording electrode (red (left) arrow) to the common ground electrode (green arrow). Recording electrodes are length-adjustable to accommodate the various lengths of antennae from different species.

including the electrodes that contact the antennae, is shown in figure 3.

The second stage utilizes a general purpose JFET quad opamp (LF147) incorporating dc-block and anti-aliasing filters ($f_c = 0.14$ and 34 Hz, respectively), and is housed in a shielded box (National Instruments model SCB-68). The analog to digital converter (National Instruments model 6036E) interfaces to a Laptop PC (Dell Inspiron 8200) running Labview 6.1. The gain of the second stage (-10 V/V) was chosen to render the noise, both in-band and out-of-band, added by the acquisition card insignificant.

2.1.2. Hardware performance: EAG signal and acquisition system noise spectra. Some characterization of the acquisition system and the EAG signal in our wind tunnel was performed by measuring the spectral content of several recordings. Recordings consisted of four different 132 second (100 Sa s^{-1}) segments: 1. Amplifier input shorted to ground. 2. Amplifier input connected to ground through a 4.7 M Ω resistor, 3. Amplifier input configured to measure the EAG of a male *Helicoverpa zea* moth under ‘background’ (no odors introduced into wind tunnel—see methods for details) and the same EAG signal measured while responding to a plume of thujone under the conditions outlined for experiments B and C. Welch periodograms were obtained using 1024 point FFTs and Blackman windows with 608 averages in 20 point increments. Computer generated Gaussian Johnson noise at 298 K and 4.7 M Ω is shown for reference. Figure 4, which depicts 1024 point periodograms, shows that the EAG background noise power is approximately 30 dB above the amplifier flicker noise level, and has a similar $1/f$ frequency dependency, perhaps due to the high dc bias (~ 200 mV) present across the antenna. The $1/f$ spectrum is also reminiscent of that present across tin-oxide sensors, which is used to glean more information about the odorant present [34]. We propose no mechanism for its generation, although it is often associated with dc.

2.2. Software

Both real-time and post-acquisition programs have been developed that utilize the same lower level functions for data processing and analysis. In the real-time program, data are oversampled at $200/N$ kSa s^{-1} where N is the number of channels, after which a 20th order infinite impulse response (IIR) 25 Hz Butterworth lowpass filter (LPF) is applied prior to

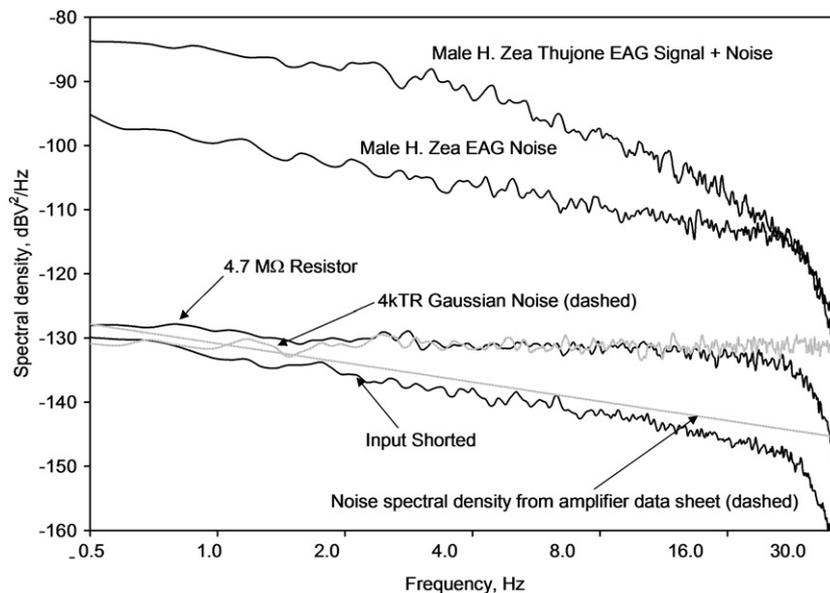


Figure 4. Measured power spectra of various signals. Roll-off at 25 Hz is due to an anti-aliasing filter. Spectra are labeled.

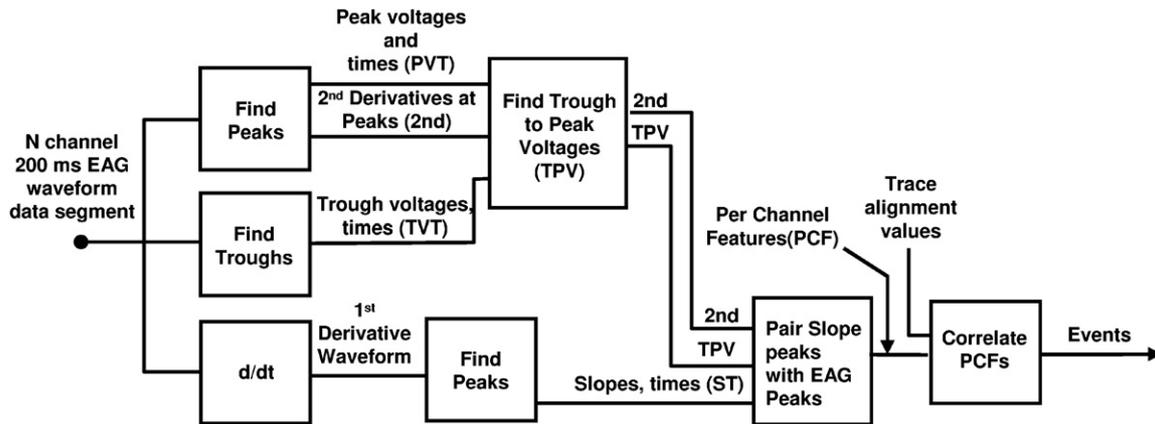


Figure 5. Block diagram of the event identification and feature extraction processes, implemented in Labview. Three features are measured for each event on each of the channels (trough-to-peak amplitude, TPV; maximum slope of the leading edge of the response, S; second derivative at the response peak, 2nd). EAG response events to discrete odor packets are identified by locating coincident peaks in the EAG biopotential waveforms recorded on the N channels. The largest TPV in each event is compared to a very low amplitude threshold of 10 V, and super-threshold events are passed to the classifier.

decimation to 100 Sa s^{-1} . The decimated data are recorded to a disk in a text format for a post-acquisition analysis. During the analysis (real-time or post-acquisition) a 1.7–11.8 Hz (-3 dB) finite impulse response symmetric linear phase bandpass filter is used to remove $1/f$ type noise in the EAG signal.

Feature extraction, event identification and event classification are carried out in real time via software implemented in Labview 6.1 and some C++ code for more time intensive portions of the program. The time from EAG depolarization to a classification report (which may be identified by an optional audible tone) is variable, but has a typical value of approximately 1 s. The delay is due to the non-causal filtering, properties of the waveform (the length of the EAG depolarization) and the delay due to processing of 200 ms data segments. Classification of events, once identified, is virtually instantaneous. Figure 5 depicts the algorithm for identifying EAG events, measuring several features of each event and grouping time-correlated events occurring on the N antennae. Each time-correlated event corresponds to a ‘strand’ of odor passing over the array of antennae.

2.2.1. Feature extraction. As a result of the high time resolution of the EAG, odorants in a natural unmixed plume can be detected as short depolarizations when filaments of odor pass over the antenna. Illustrative raw EAG recordings and responses to individual odor plume strands are shown in figure 6. It can be seen that male *H. zea* antennae responded strongly to Z11–16:Ald strands, a compound that is present in female *H. zea* pheromone. Male *T. ni* antennae responded as well, but less strongly. Conversely, male *T. ni* antennae responded to Z8–12:Ac strands more strongly than male *H. zea*. Females also responded to pheromone, but responses were weaker (not shown). Some modeling results of the moth EAG transduction (that is, the relation between concentration and measured EAG) in a wind tunnel plume was reported by Justice *et al* [11], and was found to be well approximated

(although not perfectly, see figure 7 for non-linearities) by a simple linear single pole lowpass filter. Their work shows the nature of the wind tunnel plume stimulus as measured by a photoionization detector (PID) simultaneously with EAG recordings. Further work on controlling the odor stimulus was performed by French and Meissner [35], confirming the linear system model. Since evidence shows the antennal sensor may be modeled with a linear system, it is probable that with an adequate signal and noise model, optimal estimates of the stimulus could be made. It is also possible that optimal controlled stimuli could be delivered to the antennae. As a first step, however, we identify and classify ‘depolarizations’ found in the waveform, which are likely to be due to high concentration packets of unmixed odor. Here we make the simplifying assumption that these depolarizations are independent even though generally they are not. For instance, under natural conditions, as a plume sweeps over the sensor, several depolarizations are encountered, so it is likely that depolarizations due to separate strands in close proximity in time are due to the same odor.

Depolarizations on each individual channel are detected by finding peaks of second degree least squares polynomial fits five points in length to the recorded waveform. Here peak refers to the peak of the negative-going depolarization. This function is performed by Labview’s proprietary ‘peak finder’ function, which is able to operate on discontinuous waveform segments. The best peak fits are determined by the function, but are specified to have amplitudes greater than $1 \mu\text{V}$. Using a similar method, troughs are also identified. The nearest trough preceding each peak is found, and following this, the amplitude difference between the peak to nearest preceding trough is retained as a measure of the magnitude of each response (trough-to-peak voltage, TPV), along with the time the peak occurred. Second derivatives are found from the second degree polynomial fits to the depolarizations. To obtain slopes, the waveforms are approximately differentiated by subtracting adjacent points. A peak search on the derivative waveform is then used to identify the nearest preceding peak

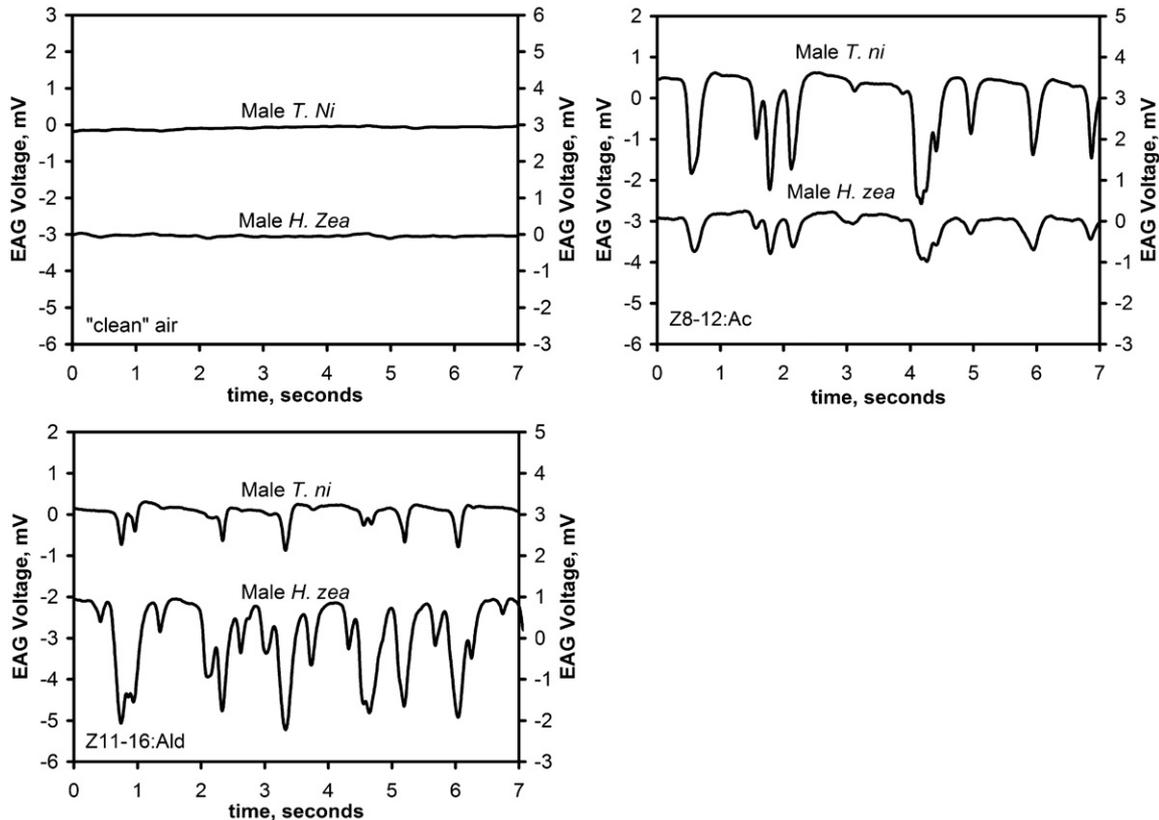


Figure 6. Electroantennogram (EAG) responses recorded simultaneously in two species. Upper trace in each panel from *Trichoplusia ni*, lower trace from *Helicoverpa zea*. Top-left panel, responses to background air. Top-right panel, responses to Z8-12:Ac, similar to Z7-12:Ac, a major component in *T. ni* pheromone. Bottom-left, responses to Z11-16:Ald, a major component in *H. zea* pheromone. The vertical scale for *T. ni* is indicated on the left of each graph, the vertical scale for *H. zea* is indicated on the right.

in the differentiated waveform to each TPV. The value of the polynomial fit at the peak of the differentiated waveform is used as the slope value. All three features, the associated channel and the depolarization times (peak times) are stored for each depolarization. The peaks associated with the features may span adjacent (or any number of) 200 ms data segments, which have been handled by the software.

2.2.2. Channel alignment. As a result of the high time resolution of the EAG, differences in signal time of arrival at the antenna can be significant, up to 40 ms here (visible in figures 8 and 9). Therefore, a simple channel alignment procedure is implemented after training recordings are collected (when the user presses the ‘classify’ button). This is accomplished by loading the (100 Sa s^{-1}) training segments from the disk, resampling at a $10\times$ rate and cross-correlating each channel with a reference channel (channel 1) to obtain time offsets, assuming an appreciable activity is present on all channels. The offsets are added to the time values obtained by the peak detector during the event generation (see figure 5).

Resampling and cross-correlation are performed in the frequency domain by calculating the discrete Fourier transform (DFT) sequences for each channel to be (circularly) cross-correlated first. The complex conjugate of the reference

channel is computed so that when the spectra of channels are multiplied, the circular cross-correlation is obtained. Subsequently, the DFT spectra products are split near the middle, so that the bins corresponding to or below $f_s/2$ comprise the lower half. Between halves, the DFT is padded with a number of zeros that is nine times the length of the DFT. Following this, the inverse DFT is computed and the location of the maximum cross-correlation result relative to $t = 0$ is stored.

2.2.3. Combining features from multiple channels—event generation. In the event generator, times of larger depolarizations serve as points of reference. Given that multiple depolarizations occur on the same channel near (in time) a reference, it is assumed the largest is more likely to be the ‘correct’ choice. The event generator maintains a list of depolarizations and outputs a list of events whenever a depolarization occurs that is out of range (>50 ms) of all previous depolarizations, or the function is called (processing 200 ms at a time) and no depolarizations have occurred. The out of range condition is encountered on average six times a second, so events are generally received one or two at a time. Events are assembled from depolarizations in the following manner. Starting with the largest stored depolarization, a search width of ± 50 ms is used to find peaks occurring

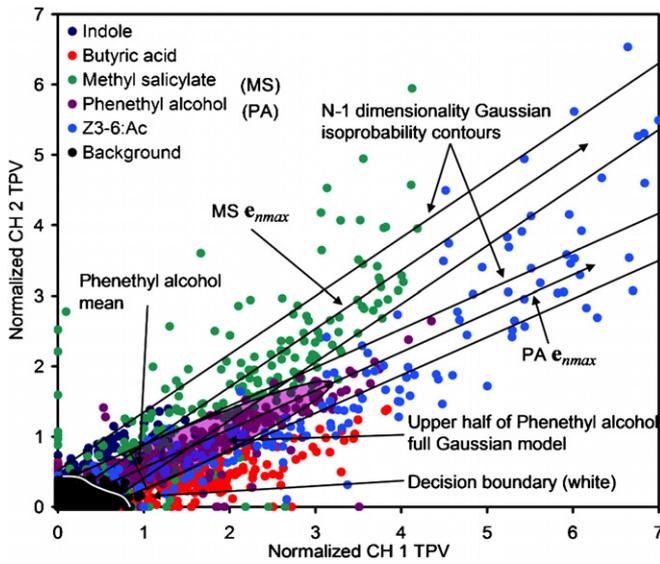


Figure 7. Illustration of training data from experiment B for two channels showing model used to remove background in stage 1 classifier. (Lines drawn are purely illustrative) CH 1: male *Anticarsia gemmatilis*. CH 2: male *Trichoplusia ni*. Illustrative eigenvectors with largest eigenvalues are labeled for both phenethyl alcohol (PA) and methyl salicylate (MS). Isoprobability contours are drawn for the reduced dimensionality Gaussian model (PA and MS) and the full Gaussian model (PE only). Data used to estimate the covariance matrix are taken above the hyperplane (line in this case) above the odor mean. Black data points show density near origin due to background (i.e. true negatives). Background is modeled as the product of independent channel densities, each estimated with Gaussian kernels. Possible location of decision boundary is drawn.

near the large peak. The largest peak from each channel is used to form an event vector. These peaks are marked as

used. Depolarizations are processed in decreasing order of size until the event threshold ($10 \mu\text{V}$, a very small value) is reached. Following construction of the events, events containing marked depolarizations are removed. Thus when events share depolarizations, the event with the larger size is kept, and the smaller is removed, as a result of the order of depolarization marking. Although TPVs are used to construct the events, the amplitude along with the associated features, slope and second derivative, are assembled into a $N \times 3$ feature vector for each event.

2.2.4. Background odors, noise and ambiguous responses. Under most conditions, background odors and especially noise on the EAG baseline will be present. The EAG waveforms acquired in this experiment have dc offsets of approximately 200 mV, which is very large when compared to the EAG signal ($\sim 1 \text{ mV}$) obtained from changes in odor concentration. The dc offset is not reliably constant in the absence of odors, and is thus filtered out. A source of background noise consists of summed activity of neurons with background firing rates. Interestingly, action potential trains from individual neurons may be discerned in wideband EAG recordings. Some of these neurons may have activity levels that can be decreased or increased, depending on the stimulus. Also, neural firings associated with mechanoreceptors due to turbulence are present. Some background odors tend to be diffuse and well mixed and therefore cause small variations of a low frequency in the baseline. Other times, background odors result in sharper depolarizations. Noise on the baseline is Gaussian in nature and uncorrelated across channels. See figure 8(A).

During the presentation of an odor, an event may be the result of background (true negatives) or background + odor

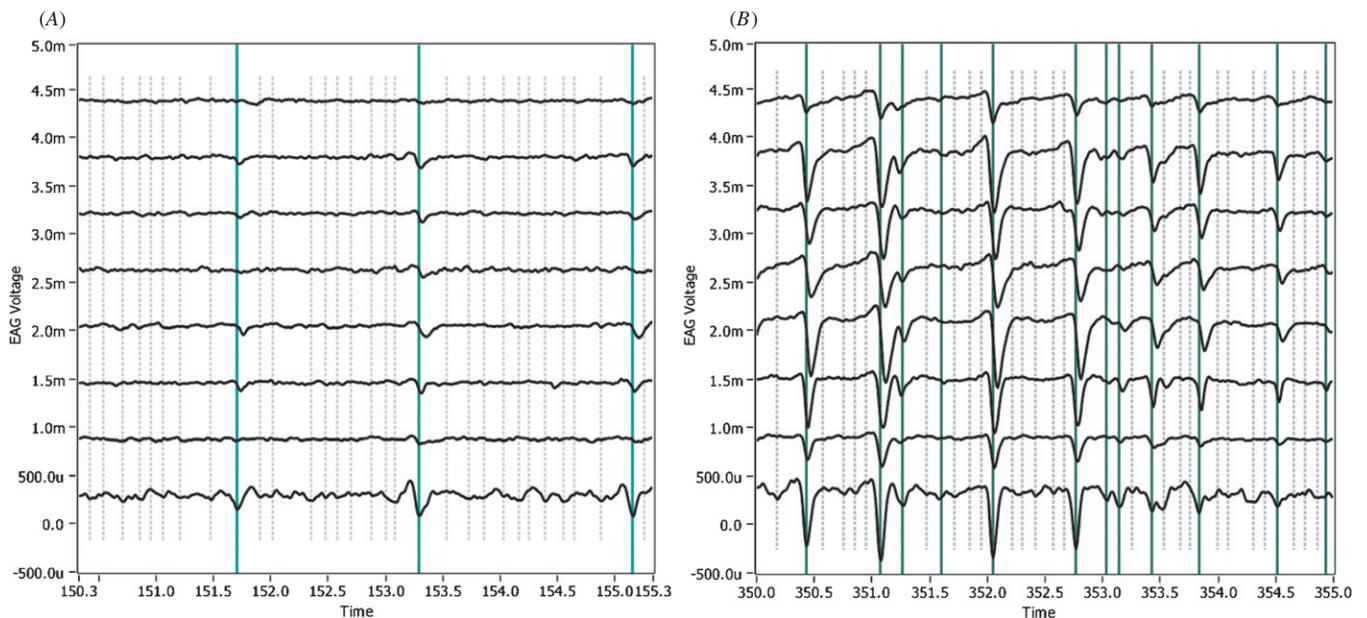


Figure 8. EAG recordings excised from experiment C illustrate stage 1 detection. Longer, green lines indicate events passed on to stage 2. Shorter, gray dotted lines indicate events removed by the detection procedure. (A) Background segment. (B) Citronellal segment.

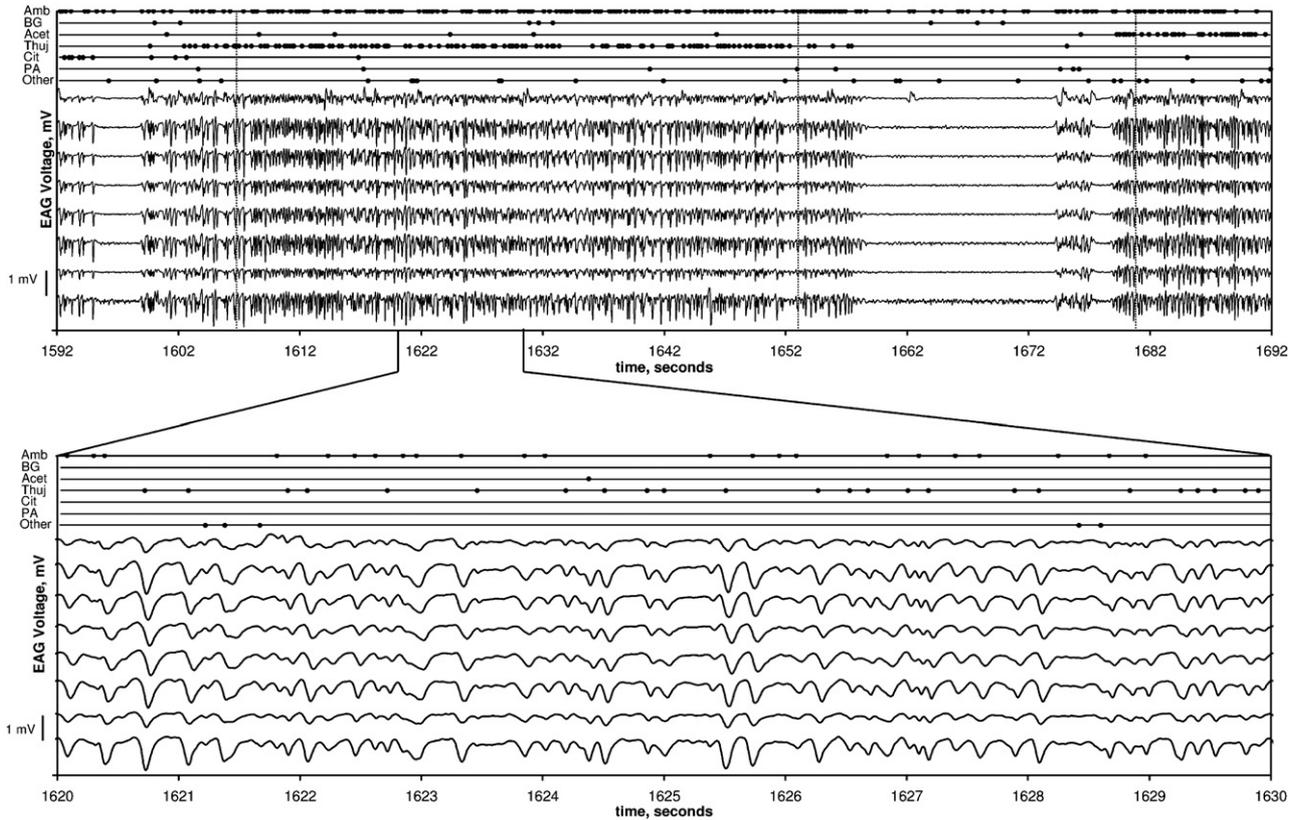


Figure 9. Eight-channel real-time EAG classification results while the antennal array was exposed to odor plume strands of thujone during experiment C. Channels are numbered 1 through 8 from bottom to top. Filled circles represent stage 2 classified EAG depolarizations. The vertical location of the circles is used to designate the classification. Labels for the classifications are given on the left-hand side where Amb, BG, Acet, Thuj, Cit, PA and Other refer to ambiguous, background, acetophenol, thujone, citronellal, phenethyl alcohol and other odor, respectively. Starting at ~1680 s, the antennal array was exposed to a plume of acetophenol and ~1595 s is the time at which exposure to a plume of citronellal ceased. The ambiguity reject parameter was set to 0.6. Vertical dotted lines indicate test segment time boundaries.

(1 molecule or more) (true positives). Measuring these is not possible (i.e. imperfect teaching). Therefore the error rate *cannot* be estimated through error counting. However, it is possible to measure background alone to estimate its density, and to estimate the odor probability densities. If the estimates are correct (known perfectly), decision boundaries can be created that minimize the error rate or risk. The true error rate or risk can then be estimated from the densities and decision boundaries. This is not a task we have chosen to undertake.

Variance in the odorant training data is also contributed to by the method of natural exposure (rather than artificial puffs) to the flow field, which does not result in uniform exposure of all antennae to the dispersing scalar odorant concentrations. Even given perfect teaching, in practical situations where individual strands are to be identified, a substantial fraction of true positives are not worth the risk to classify. For this reason, we have utilized a variation of Chow’s [36] ambiguous reject. Using Chow’s definition of error rate (error rate amongst samples not rejected) makes the error rate dependent on the rate of rejection. Thus any reported error rate with an ambiguous reject class should specify the thresholds used for rejection.

The process of detection, or classifying events as positives and negatives, is a step we have used to efficiently reduce the amount of data to be processed both in the data stream

and training sets. As a result of imperfect teaching, the performance of the detection stage cannot be estimated through error counting, except to count false positives. However, when rejecting a substantial fraction of true positives due to ambiguity (see equation (7)), most false negatives that would otherwise increase the error rate in the absence of an ambiguous reject class will be rejected. Further, any false positives will tend to either increase the error rate estimate, or be rejected. Therefore, the creation of decision boundaries designed to remove true negatives (detection) does not increase the reported error rate in a significant way when the ambiguous reject rate is significant. In this way, the conditions are similar to the conditions utilized in an automatic target recognition problem in radar, where false negatives are kept low and false positives are acceptable before the discrimination stage. However, in this case, it is also desirable to reduce the false positive rate using a semi-parametric classifier, since *k*-NN is less consistent in its behavior.

In our real-time implementation, background activity is removed in three ways. Low frequencies present on the EAG due to well mixed background odors are removed by the digital bandpass filter. Events due to noise and background odors are removed by two stages of classification, which could be referred to as detection (first) and discrimination (second). Events occurring during background conditions which remain

after detection are either false positives or due to background odors, which lead to more correlated depolarizations not modeled by the semi-parametric classifier. These background depolarizations are the target of the stage 2 background class. The same reasoning regarding imperfect teaching that applies to the detection stage also applies to the discrimination stage.

In summary, because background odorants and test odorants are present in the training and test plumes, the error rate result may be distorted by the inclusion of background odorants when they are incorrectly classified as an odorant (false positives) and vice versa (false negatives). This is less likely to occur as the ambiguous reject rate increases. As a result, we refer to the empirical error rate in the classification stage.

2.2.5. Classifier.

2.2.5.1. Training and normalization. The system is trained by collecting data while exposing the antennae to known odor source plumes. These data may be represented by the set $\{(X_i, C_i); i = 1, \dots, N\}$, where X_i are d -dimensional training data points associated with class C . Let the $L + 2$ odor classes be represented by $\Omega = \{c_0, c_1, \dots, c_L, c_b\}$, where c_b is the background reject class and c_0 is the ambiguity reject class, which will be explained further. The present system may be trained to any number of known odor sources. The maximum number of odor sources that can be distinguished reliably is a function of many variables [2]. Training events for each odor are first decimated to the desired number of training points. Before classification, the training set feature type values are normalized to have a standard deviation of 1 over the entire training set [37]. That is, the standard deviation of each feature type (rather than each feature) is calculated after data from every channel are combined into events. The normalization constant applied to each feature type (TPV, slope or second derivative) is applied to incoming events on all channels, so that data associated with background activity and noise are comparable across channels. Note that since the normalization constants apply to feature type, signal strength information is retained in the events. Training is accomplished in two stages, first to obtain the stage 1 model, and second after discarding events in each training segment that fall into the background class of the stage 1 classifier. This second training set is used to classify events in the stage 2 classifier. Two sets of normalization constants are used to scale incoming real-time events, corresponding to each classification task.

2.2.5.2. Stage 1—detection. A suboptimal detection procedure is used to remove small events that resemble the background class in the stage 1 classifier. Representative results of the stage 1 detection process are visible in figure 8. This method uses TPVs only for classification. Noise and background are modeled assuming channel independence, so that the density is the product of the TPV density on each channel. Each channel's density is estimated using Gaussian kernels and stored in lookup tables 100 points in length amongst the range encountered in the data, and assumed zero outside the range. The kernel width is chosen using

Silverman's rule [38]. Linear interpolation is used to obtain the density given any TPV value. Since the noise and background are Gaussian, it is also possible to model the density; however this was not done, since a better feature extraction method is likely to exist.

The probability densities of *each* odor are modeled using a multi-dimensional Gaussian density. The distribution is derived from an N -dimensional model, where variation along the covariance matrix eigenvector with the largest eigenvalue is ignored, so that the dimensionality of the density is reduced from N to $N - 1$. The component of any feature vector that points along this eigenvector is considered to be a measure of signal strength. The multi-dimensional Gaussian density may be expressed by the following equation [37]:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right), \quad (3)$$

where \mathbf{x} is a vector of d elements containing the feature values (i.e. $d =$ the N measurements from multiple antennae), Σ is the sample covariance matrix for a given odor and $\boldsymbol{\mu}$ is the vector containing the feature means. The exponent of the MDG may be written as a summation in terms of the eigenvectors and eigenvalues of the covariance matrix. We write it this way to explicitly remove dependence of the probability density in the direction of the eigenvector with the largest eigenvalue. Let the index of the largest eigenvector and corresponding eigenvalue be n_{\max} . The reduced dimensionality MDG may be written as

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2} \sum_{n=1, n \neq n_{\max}}^d \frac{(\mathbf{e}_n \bullet (\mathbf{x} - \boldsymbol{\mu}))^2}{\lambda_n}\right), \quad (4)$$

where \mathbf{e}_n and λ_n are the n th eigenvector and eigenvalue of the sample covariance matrix. Although changes along the largest eigenvector/eigenvalue combination are ignored, the normalizing constant of the density function remains unchanged. This is an effort to extrapolate the density of each odor to the origin, where the marginal density along the n_{\max} th eigenvector remains approximately constant, but is inversely proportional to the standard deviation along the n_{\max} th eigenvector. This point is not crucial, since similar results are achieved using a multitude of Gaussian models. In addition, the rate of background classification was found to be relatively insensitive to changes in the background class prior probability (not shown). The sample covariance matrix is calculated from the points in the first training data set that reside above the hyperplane perpendicular to the mean vector. Class membership of a new vector is determined using the Bayesian inference assuming each odorant has an equal prior probability. Any event classified as background is then discarded. Figure 7, which uses normalized TPV values from experiment B, is included to help visualize the above equations as applied to our data.

2.2.5.3. Stage 2— k -NN classification. The nonparametric k -NN algorithm was used to classify events in the stage 2 classifier using a class-dependent weighted voting scheme.

The following equation is used to calculate the posterior probability of membership in each class, c_j :

$$\hat{f}(c_j|x) = \frac{\frac{1}{n_j} \sum_{i=1}^N 1[D(x, X_i) \leq D_k]1[C_i = c_j]}{\sum_{l=1}^{L+1} \frac{1}{n_l} \sum_{i=1}^N 1[D(x, X_i) \leq D_k]1[C_i = c_l]}, \quad (5)$$

where $1(F)$ is the indicator function, which is 1 if F is true and 0 if F is false. D_k is the Euclidean distance to the k th nearest neighbor in the training set to the test point x . n_j and n_l refer to the number of training points obtained for a particular odor. These compensate for unequal training set sizes so that each class is assigned an equal prior probability. All analyses in this work utilize $k = 10$ nearest neighbors. Albeit not convergent to the true densities, this estimate may be more successful in calculating the relative contribution of each prior density in our data than a volumetric k -NN classifier based upon prior densities calculated using the Loftsgaarden and Quesenberry [39] estimate (either an Euclidean or a class-dependent Mahalanobis distance metric not shown). This becomes apparent when applying a threshold equation (11), but not when simply selecting the class with the highest posterior probability. Volumetric k -NN was least successful for experiment C, which involves higher dimensionality, $d = 24$. Other simple techniques for adjusting prior probabilities include re-sampling [40] and weighing the neighbor distances [41].

An optimal method for rejecting ambiguous events was described by Chow [36], which assigned a minimum threshold on the posterior probability estimate. Soon after, the technique was applied to nearest neighbor classifiers [42]. Baum and Pursley [43] defined a classifier with a loss matrix similar to that of ours (equivalent to Chow's ambiguous reject) to receive or erase symbols in communication systems. In this implementation, one parameter is used to control the tendency to reject by varying separate thresholds for each odor. The parameter is meant to control the rate of rejection. This is accomplished by calculating the probability estimate $\hat{f}(c_j|x)$ equation (5) for each training point belonging to its own class using the leave-one-out technique. These are then sorted. A user-selected value between 0 and 1, the reject rate parameter, then determines the probability within the sorted list to serve as the threshold for class membership, which is done for each class individually. For example, a value of 0.5 would use the median value in the sorted list for each odor as the threshold for classification of a new event; a probability below the median value would result in classification of that event as ambiguous. In effect, the user is using an estimate of the reject rate from the training data to exclude a certain fraction of data classified in any class.

Let j enumerate odors 1 to L , and l enumerate odors 1 to $L + 1$. t_j is the threshold assigned to odor j which is determined using the leave-one-out technique described above. Events for which

$$\arg \max_l \hat{f}(c_l|x) = j \quad \text{and} \quad \hat{f}(c_l|x) \geq t_l \quad (6)$$

are declared as class c_j . In the event of a tie, a pseudo-random uniform distribution number generator utilizing a repeatable seed (available with Labview) is used to choose the class

amongst those that share the same posterior probability. Below threshold events for which

$$\arg \max_l \hat{f}(c_l|x) = j \quad \text{and} \quad \hat{f}(c_l|x) < t_l \quad (7)$$

are declared ambiguous, or c_0 . Finally, events for which

$$\arg \max_l \hat{f}(c_l|x) = L + 1 \quad (8)$$

are unconditionally classified as background, c_b .

Results are summarized in real time in a histogram that displays the number of events classified as each of the target odors during any arbitrary time interval. Classification results are also communicated audibly by tones assigned to each target odor.

Certainty estimates are commonly used in electronic nose applications. Each classification is accompanied by a weight, the posterior probability associated with the decision, which is used to control the volume of the audio signal associated with each classification.

2.2.6. Empirical error rate calculation. The error rate of a test session during which odor j is presented for the purposes of evaluating the classifier performance is calculated by error counting:

$$E_j = 1 - \frac{\sum_{m=1}^M 1[\arg \max_l \hat{f}(c_l|x_m) = j]1[\hat{f}(c_j|x_m) \geq t_j]}{\sum_{n=1}^L \sum_{m=1}^M 1[\arg \max_l \hat{f}(c_l|x_m) = n]1[\hat{f}(c_n|x_m) \geq t_n]}, \quad (9)$$

where m enumerates the above threshold events detected during the test session. Since removal of background events is part of detection and discrimination, the error rate is considered amongst the odorants and is termed the empirical error rate. Note that as discussed in section 2.2.4, the measured error rate may be underestimated for lower reject rates.

2.2.7. Minimum conditional risk classifier for performance reference. As a measure of performance, the results of the real-time classifier were compared to a minimum conditional risk classifier utilizing the k -NN density estimator (equation (5)) with the loss matrix in table 1 through error counting, even though the density estimates do not meet the same criteria necessary for estimating the Bayes risk [44].

Classified events may have utility u or loss, ls . A correct reception has value, an incorrect reception is associated with a loss and a miss has no associated value or loss. When events are classified as background, no value is associated; it is the same as having no measurement system at all. The well-known

Table 1. Loss matrix.

Actual	Decision			
	Odor 1	Odor 2 . . .	Odor L	Background
Odor 1	$-u$	ls	ls	0
Odor 2 . . .	ls	$-u$	ls	0
Odor L	ls	ls	$-u$	0
Background	ls	ls	ls	0

conditional risk minimizing classifier [45, 46] chooses class c_l where

$$l = \arg \min_q \sum_{p=1}^{L+1} \lambda_{pq} \hat{f}(c_p|x), \quad (10)$$

where λ_{pq} is a loss matrix element involving the classification of an event of type p as q . When simplified, the decision rule becomes:

$$\text{if } \max_l \hat{f}(c_l|x) > \frac{ls}{ls+u} \text{ assign class } \arg \max_l \hat{f}(c_l|x), \\ \text{else assign class } L+1(\text{background}), \quad (11)$$

which is seen to combine ambiguous events as defined by Chow and some background events into the background class. Note that if the background posterior probability is high, an event is more likely to be placed into the background class. To evaluate performance, the average risk of each approach was estimated, given values of u and ls using the k -fold cross-validation technique ($k = 10$) on the training data. Let c_m denote the decision made by either the real-time classifier or the minimal risk classifier for the m th event. The empirical risk associated with class l is evaluated through error counting.

$$R_l = \frac{1}{M} \sum_{n=1}^{L+1} \sum_{m=1}^M 1_{[c_m = n]} \lambda'_{ln}, \quad (12)$$

where m enumerates all events encountered during the test session when odor l was presented. The prime on the loss matrix element λ'_{ln} is used to indicate that the risk estimate may employ a different loss matrix than that for which the classifier was designed equation (10). The average empirical risk is calculated as

$$R = \frac{1}{L+1} \sum_{l=1}^{L+1} R_l. \quad (13)$$

Multiple 10-fold cross-validation runs on the training data were performed for each technique (real-time and minimum conditional risk) offline. The real-time ambiguity reject parameter was varied from 0.05 to 1.0 in increments of 0.05, while the minimal risk classifier was run so that $ls/(ls+u)$ varied from 0.05 to 0.95 in increments of 0.05. Runs for both methods used the maximum number of available training points. After storing the confusion matrix results for all the runs, performance comparisons were made in a separate program offline using an additional Labview code.

To perform ad hoc comparisons, the ambiguity reject parameter was chosen for each loss-utility threshold ($ls/(ls+u)$) entered into the error counting loss matrix λ' for which the counted risk was minimized. The minimal risk estimated through error counting when considering the minimal conditional risk classifier results does not generally correspond to the loss-utility threshold employed in the classifier. To avoid giving the real-time method a biased advantage, the loss-utility threshold that was entered in the classifier loss matrix, λ , was chosen so as to minimize the risk measured through error counting given each loss-utility threshold employed in the error counting loss matrix, λ . Results are given in table 2. The best error-counted risk of both methods for all three experiments is depicted in figure 12 for the purpose of comparison of the two methods.

2.3. Recording the EAG

The electroantennogram was recorded simultaneously from excised antennae obtained from insects of different species and sex, yielding a multi-channel response to selected volatile compounds. Three configurations were employed to generate data used to demonstrate the system, denoted by A, B and C. Each antenna was fixed to the preamplifier; an electrical contact was made with a conductive gel (detailed methods for EAG recording can be found in [27]).

2.3.1. *Experiment A.* Antennae used in configuration A were from male insects of the following species:

1. *Platynota idaeusalis* (tufted apple budmoth)
2. *Helioverpa zea* (corn earworm)
3. *Ostrinia nubilalis* (European cornborer) and
4. *Cydia pomonella* (codling moth).

In experiment A, the antennal array was positioned 0.3 m downwind from the odor source in a wind tunnel; flow rate was 1.5 m s^{-1} . The wind tunnel used for this experiment consisted of a box-fan blowing through a furnace filter into a $2 \text{ ft} \times 2 \text{ ft} \times 10 \text{ ft}$ rectangular clear plastic duct (1 ft \simeq 30 cm). Air was removed from the other end of the duct by a laboratory hood. The antennae were oriented horizontally and perpendicular to the direction of flow, so that the leading sensor was channel 1. The odor source consisted of $100 \mu\text{g}$ of a chosen compound in solvent (hexane) applied to a piece of filter paper ($\sim 1 \text{ cm} \times 4 \text{ cm}$) and placed in the tunnel after the solvent had evaporated. The compounds used in configuration A [(Z)-11-hexadecenal (Z11-16:Ald), (Z)-11-tetradecenyl acetate (Z11-14:Ac), (E, E)-8,10-dodecadien-1-ol (E8,E10-12:OH), (E)-11-tetradecen-1-ol (E11-14:OH)] are major components in the pheromones of each of the insects used in this study.

2.3.2. *Experiments B and C.* Configuration B utilized antennae from the following:

1. Male *Anticarsia gemmatalis* (velvetbean caterpillar moth)
2. Male *Trichoplusia ni*
3. Male *Heliothis virescens* (tobacco budworm)
4. Male *Helicoverpa zea*.

Configuration C, employing eight antennae, was realized using the following species.

1. Male *Anticarsia gemmatalis* (velvetbean caterpillar moth)
2. Male *Trichoplusia ni*
3. Male *Heliothis virescens* (tobacco budworm)
4. Male *Heliothis subflexa*
5. Female *Trichoplusia ni*
6. Female *Anticarsia gemmatalis*
7. Female *Heliothis virescens*
8. Female *Heliothis subflexa*.

For experiments B and C, the antennal array was placed 1.5 m from the odor source with a wind speed of 0.5 m s^{-1} , verified with an ultrasonic anemometer. This wind tunnel has a width of 1.2 m, a length of 2.8 m and a height of 1 m at its peak (an arched shape based on [47]). In configuration B,

the antennae were also oriented horizontally and perpendicular to the direction of flow so that the leading sensor was channel 1. The additional antennae in configuration C were mounted on a second amplifier, opposed and ~ 1 mm below the antennae corresponding to channels 1–4, so that channels 1 and 8 were the leading sensors. Relative humidity was $\sim 70\%$ and the temperature was ~ 65 °F (~ 18 °C) as measured using a digital hygrometer. Humidity was obtained by spray atomizing deionized water into the wind tunnel air intake, dropping the temperature slightly from room temperature. High humidity is necessary when using excised antennae to extend their useful lifetime. Compounds used in configurations B and C included indole, butyric acid, citronellal, thujone, acetophenol, methyl salicylate phenethyl alcohol and (*Z*)-3-hexenyl acetate (Z3–6 Ac). The indole source was prepared by placing 50 μl of 10 $\mu\text{g } \mu\text{l}^{-1}$ hexane solution onto a piece of filter paper (~ 1 cm \times 4 cm). The remaining compounds were used to impregnate the filter paper directly, using 25 μl of each compound.

2.3.3. Experimental procedure. The plume was aligned with the detector prior to starting experiments using a Z11–16:Ald filter paper source while monitoring the EAG traces. Following plume alignment, EAG recordings were started. Odor sources were prepared immediately prior to being placed into the wind tunnel, after which the real-time program was instructed to train at least 40 s. Some additional time was usually allowed to pass before deactivating the train button, and replacing the odor source. Following training the odorants, with the exception of background, were cycled through the wind tunnel in the same order as trained,³ for approximately 45–60 s each. Classification was started a few seconds after placing the odor source in the tunnel and stopped prior to removing the source. The order of presentation and duration of presentations is indicated in tables 3, 4 and 5. Average time between the end of training an odor and the beginning of test of that odor in experiments A, B and C was 392, 401 and 492 s, respectively. Experiments A and B were cycled twice, and experiment C was cycled three times.

2.3.4. Non-ideal experimental conditions. Because *Anticarsia gemmatilis* adults were available over a short period of time while experimental protocol was being established, some non-ideal experimental conditions were encountered. In configuration C, a third presentation of the odors was performed as a result of a plume shift (i.e. the odor plume missed the sensor array) that occurred during the training session. The plume shift was due to non-steady-state humidity and temperature caused by the humidification method. The training and test sessions analyzed in configuration C are the second two odor presentation cycles. As a result, this experiment does not reflect the real-time outcome of the software. A similar occurrence in configuration B led the investigator to retrain the first odor, indole, during the training cycle (after the third training odor, citronellal). Also in configuration B, six odors were presented ('out of a hat'); however the number of odors was reduced to five

(citronellal was removed) since two of the odors were difficult to distinguish.

2.3.5. Pheromonal compounds. For readers unfamiliar with the pheromonal compounds listed, we have included several references that characterize the pheromonal mixtures present in female glands of the listed species: *Platynota idaeusalis* [48], *Helicoverpa zea* [49], *Ostrinia nubilalis* [50], *Cydia pomonella* [51] and *Trichoplusia ni* [52]. In summary, E11–14:OH and E11–14:Ac are the two main components of *Platynota idaeusalis* pheromone; however the antennae of this species will also respond to Z11–14:Ac, whereas Z11–16:Ald is the major component of the *Helicoverpa zea* pheromone. The pheromone of *Cydia pomonella* is composed of mainly E8,E10–12:OH, and the pheromone of both strains of *Ostrinia nubilalis* contains Z11:14:Ac and Z11:14:OH. Finally, the major component of *Trichoplusia ni* pheromone is Z7–12:Ac, an analog of which is Z8–12:Ac.

3. Results

The classification results presented include several 10-fold cross-validation runs on training data intended to evaluate the performance (error rate, equation (9) and risk, equation (13)) of the system and tabular results of real-time runs. Real-time runs are performed on separate training and test segments. A principal component analysis is used to visualize some of the training data and classified data. Cross-validation is used to show the effect of training points and the ambiguity reject parameter on empirical error rates. Further, these results are evaluated using an ad hoc method with respect to the minimum conditional risk classifier with the loss matrix defined in table 1. Real-time confusion matrices for each experiment are included as well as a figure superimposing real-time classifier results on top of EAG recordings.

3.1. Real-time results—EAG waveforms

A representative non-pheromonal eight-channel recording including real-time classification results from experiment C is shown in figure 9.

Examination of figure 9 reveals the nature of the recordings after digital filtering and the performance of the event detection. Events shown do not include those removed by the stage 1 classifier. See figure 8 for representative events removed by the stage 1 detector. Classification results are the same real-time results presented in table 5. Also, it should be noted that the wrong odorant (phenethyl alcohol; PA) was briefly placed into the wind tunnel during this test from ~ 1674 to 1677 s. Waveforms visible in figure 9 are due to odorants that were placed into the wind tunnel in the following order: citronellal, thujone, (phenethyl alcohol; PA) and acetophenone. Inactive times between are representative of background activity where events are for the most part either not present due to stage 1 removal, or marked as ambiguous. The frequency of their appearance, though lower than that during active times, indicates that the background noise is not stationary, since they were not removed by the

³ Except for experiment B, see section 2.3.4.

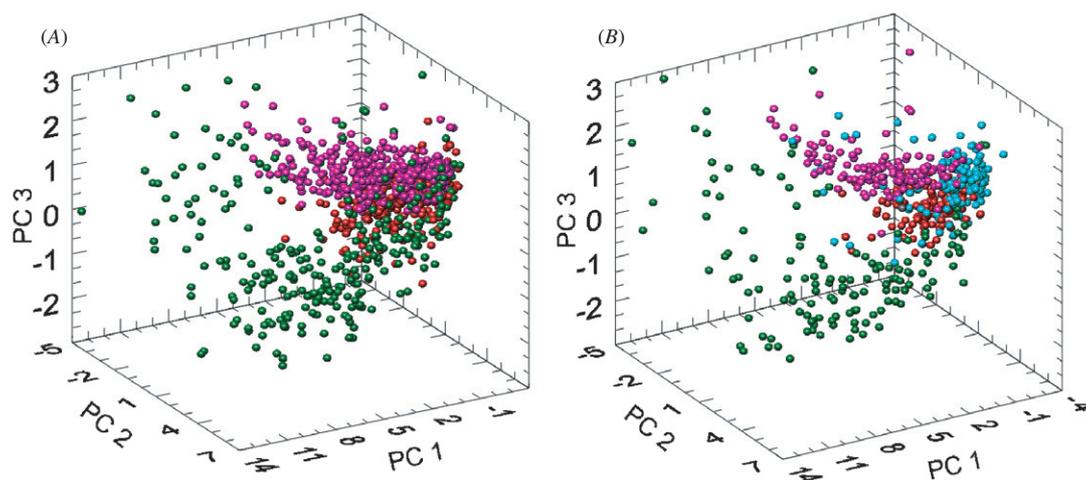


Figure 10. Experiment C (eight channels, 24 feature) PCA plots of single plume strands of butyric acid (red), methyl salicylate (purple) and Z3-6:Ac (green) data where component directions are computed from the second-stage training set of 8 odors plus background. (A) Training data. (B) (Real-time) classified data amongst the three test segments plotted along the same principal components. Ambiguous events are plotted in cyan.

stage 1 classifier. Also note the waveforms are time-shifted due to the physical location of the antennae; these shifts were compensated for in the waveform alignment procedure. Some erratic activity visible on channel 8 is likely due to the leakage current between the power supply and the amplifier input over a dirty soldermask layer. Such activity in the past has been removed by rinsing the amplifier in ethyl alcohol. Note the testing times do not include events before the plumes had stabilized, and hand odors and/or mechanical disturbances due to disruption of the air flow were no longer present. Because of the digital filtering, the low frequency disturbances are not visible on the EAG waveforms shown. Classification of thujone does seem to be affected, especially at the beginning. However, classification performance of citronellal and acetophenone appear to be unaffected. It is possible that training could take place under these background conditions to see the effect on performance. It may prove beneficial to investigate other methods (other than high-pass filtering) for resolution of stimulation due to superimposed odors arising from different types of plume sources. In this regard, determining how much of each odor is present, rather than classifying, may be beneficial.

3.2. Training and classification results

PCA plots of both training and classified real-time data are presented in figure 10 (ambiguous reject parameter = 0.6, see section 2.2.5.3). Three fairly well separated odors (butyric acid, methyl salicylate and Z3-6:Ac) were chosen so one can see the groupings. Many of the odors overlap along the first three principal components. The first principal component (77% of the total variance) accounts mainly for concentration changes, while the remaining two account for an additional 10% of the variance, which is approximately 40% of the variance not in the first principal component.

3.3. Parameter dependent error rate—*k*-fold performance

In an effort to evaluate the performance, the empirical error rate of the classifier equation (9) was evaluated as a function of the ambiguity reject parameter and number of training points using 10-fold cross-validation on the training data. Results are shown in figure 11. Panels A and B are results from experiment A, panels C and D are results from experiment B, and panels E and F are results from experiment C. Panels A, C and E show the results of varying the number of training points with an ambiguity reject parameter of 0.6.

The maximum number of training points tested was chosen so that the number of training points available to the non-background odor with the least number of training points was not exceeded (see tables 3–5). A ‘sufficient’ number of training points appears to be about 150 (21 s) for experiment A, 150 (21 s) for experiment B and 300 (43 s) for experiment C, though these observations are subjective. Items in parentheses indicate the training time in seconds when an event rate of seven events per second is assumed. Panels B, D and F plot the rejection rate versus the empirical error rate as a function of the value of the ambiguity reject parameter using the maximum number of training points per odor for experiments A, B and C, respectively. The empirical error rate of the classifier generally decreases as the ambiguity reject parameter decreases and the reject rate increases. It should be noted that the error rate may be underestimated to an unknown degree when the ambiguity reject parameter is low (see section 2.2.4). Although not shown directly, the rejection rate is consistently underestimated by the rejection rate parameter.

3.4. Performance in comparison to minimal conditional risk classifier

Figure 12 shows the ad hoc empirical risk results equation (13) for both the two-stage approach and the minimum conditional risk method, which use the parameters outlined in table 2.

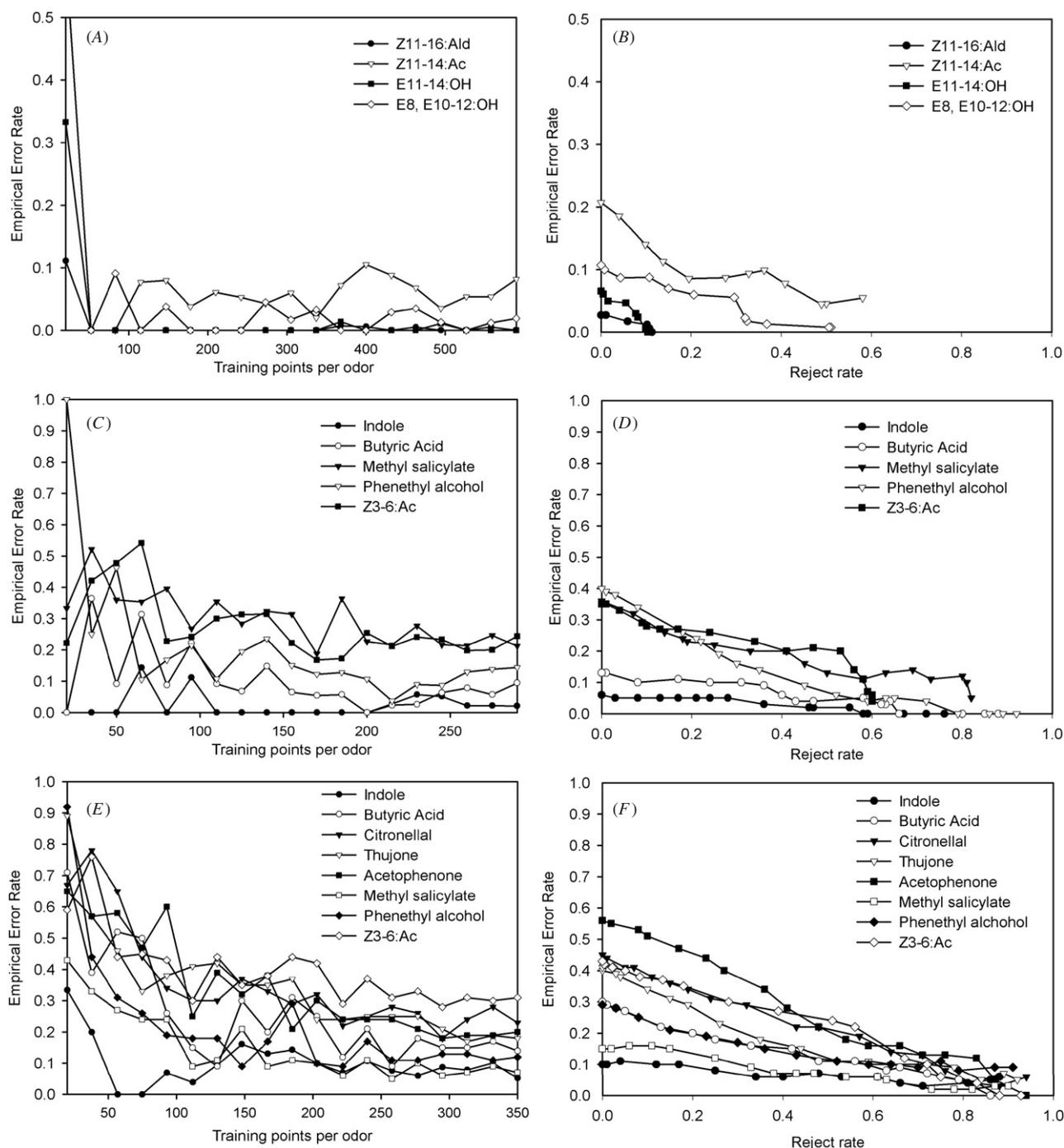


Figure 11. Empirical error rate evaluated as a function of stage 1 training points per odor and the ambiguity reject parameter. Analyses are performed using 10-fold cross-validation on the training data. Panels (B), (D) and (F) plot the empirical error rate versus the stage 2 rejection rate for the maximum number of training points (see tables 3, 4 and 5) as the ambiguity reject parameter is varied from 0.00 to 0.95 in increments of 0.05.

The risk was estimated given the optimal ambiguity reject parameter as ls varied, while u remained constant at unity. The empirical risk method clearly outperforms the two-stage approach for experiment A; however, for experiments B and C, both methods are fairly equivalent over a wide range of loss per error values. Note the minimum risk method outperforms the real-time method for experiments B and C when background and ambiguous events are not removed at lower loss/error values; however when the loss-utility threshold increases

above ~ 0.5 , the performance difference is negligible. When odors are more discriminable, such as in experiment A, the two-stage approach may not be optimal.

The real-time results are obtained using an ambiguity reject parameter of 0.6. For comparison to the minimal conditional risk method, the loss-utility ratio that approximately corresponds to the chosen ambiguity reject parameter of 0.6 was calculated. This represents the loss-utility ratio entered into the loss matrix that performs best when

Table 2. Comparative loss-utility ratios and ambiguous reject parameters

Error counting loss-utility ratio	Experiment A		Experiment B		Experiment C	
	Best loss-utility ratio in classifier	Best ambiguous reject	Best loss-utility ratio in classifier	Best ambiguous reject	Best loss-utility ratio in classifier	Best ambiguous reject
0.05	0.05	0.10	0.05	0.10	0.05	0.10
0.11	0.05	0.10	0.05	0.10	0.25	0.10
0.18	0.18	0.10	0.25	0.10	0.25	0.10
0.25	0.25	0.10	0.33	0.10	0.43	0.10
0.33	0.43	0.10	0.54	0.10	0.43	0.15
0.43	0.67	0.10	0.67	0.10	0.67	0.35
0.54	0.67	0.15	1.00	0.25	0.82	0.40
0.67	0.82	0.20	1.00	0.30	1.00	0.40
0.82	1.22	0.20	1.22	0.30	1.00	0.45
1.00	1.86	0.20	1.22	0.45	1.00	0.55
1.22	1.86	0.20	1.50	0.45	1.00	0.55
1.50	2.33	0.20	1.50	0.55	1.50	0.60
1.86	4.00	0.20	1.86	0.65	1.50	0.60
2.33	5.67	0.20	3.00	0.65	1.86	0.65
3.00	5.67	0.40	3.00	0.65	2.33	0.70
4.00	5.67	0.50	4.00	0.70	3.00	0.75
5.67	5.67	0.50	5.67	0.85	3.00	0.80
9.00	9.00	0.60	19.00	0.95	9.00	0.85
19.00	19.00	0.60	19.00	0.95	19.00	0.95

Table 3. Experiment A real-time results

Actual	Decision							Removed by stage 1	Total	Empirical error rate	Test time (s)	Stage 2 TPs	Stage 1 TPs
	Z11-16:Ald	Z8-12:Ac	E8,E10-12:OH	E11-14:OH	BG	Ambiguous							
Z11-16:Ald	116	0	0	0	21	6	207	350	0.00	51	273	622	
Z8-12:Ac	0	120	0	3	17	71	404	615	0.02	80	194	591	
E8,E10-12:OH	0	0	190	0	10	26	501	727	0.00	102	225	931	
E11-14:OH	0	0	0	121	13	60	572	766	0.00	86	266	787	
BG*	3	0	3	0	19	4	1693	1722	N/A	205	29	1722	

Table 4. Experiment B real-time results

Actual	Decision							Removed by stage 1	Total	Empirical error rate	Test time (s)	Stage 2 TPs	Stage 1 TPs
	Indole	Butyric acid	Methyl salicylate	Phenethyl alcohol	Z3-6:Ac	BG	Ambiguous						
Indole	58	1	2	1	0	22	117	280	481	0.06	56	152	404
Butyric acid	0	96	0	0	0	17	66	199	378	0.00	47	184	394
Methyl salicylate	13	1	105	7	0	5	93	94	318	0.17	43	224	296
Phenethyl alcohol	9	2	9	66	0	10	141	224	461	0.23	58	186	337
Z3-6:Ac	2	1	20	8	181	16	125	158	511	0.15	70	186	292
BG*	1	0	0	0	0	26	2	704	733	N/A	87	29	733

empirical risk using the identical loss-utility ratio is compared to the risk obtained with an ambiguous reject parameter of 0.6. The following results were obtained: Configuration A: 19.0 (either optimal or optimum may be >19.0). Configuration B: 2.33. Configuration C: 1.86. Although not shown here, the class-dependent thresholds (versus optimal ambiguous reject) have the effect of reducing the range of empirical error rates amongst the classes given a particular reject parameter. Further, it may be easier to manually configure the ambiguity reject parameter to effect an improvement on an odor, independent of its statistics.

3.5. Real-time results—tabulated

Final real-time results for experiments A, B and C are summarized in tables 3, 4 and 5, respectively (ambiguous reject parameter = 0.6, see section 2.2.5.3). Each table includes a confusion matrix and additional information including the empirical error rate, the duration of the test session and the number of training points in stages 1 and 2. Since the experiments were not initially designed to show the ‘false positive’ rate, (background events classified as an odor), no background test segment is available other than the training

Table 5. Experiment C real-time results.

Actual	Decision										Removed by stage 1	Total	Empirical error rate	Test time (s)	Stage 2 TPs	Stage 1 TPs
	Indole	Butyric acid	Citronellal	Thujone	Acetophenol	Methyl salicylate	Phenethyl alcohol	Z3- 6:Ac	Background	Ambiguous						
Indole	66	0	0	0	0	0	7	0	34	69	164	340	0.10	47	201	361
Butyric acid	13	79	0	0	0	0	3	0	40	108	127	370	0.17	52	248	385
Citronellal	1	2	97	4	20	3	3	1	4	158	118	411	0.26	61	308	386
Thujone	4	2	1	111	5	2	3	1	3	122	75	329	0.14	62	270	352
Acetophenol	0	1	3	2	119	16	5	1	0	166	64	377	0.19	53	299	360
Methyl salicylate	0	0	1	0	15	105	1	0	2	105	96	325	0.14	48	314	415
Phenethyl alcohol	9	2	0	0	1	0	98	0	11	123	86	330	0.11	49	281	409
Z3-6:Ac	0	3	1	5	10	0	15	114	5	133	75	361	0.23	51	294	357
BG*	2	0	0	0	0	0	0	0	6	6	530	544	N/A	79	14	544

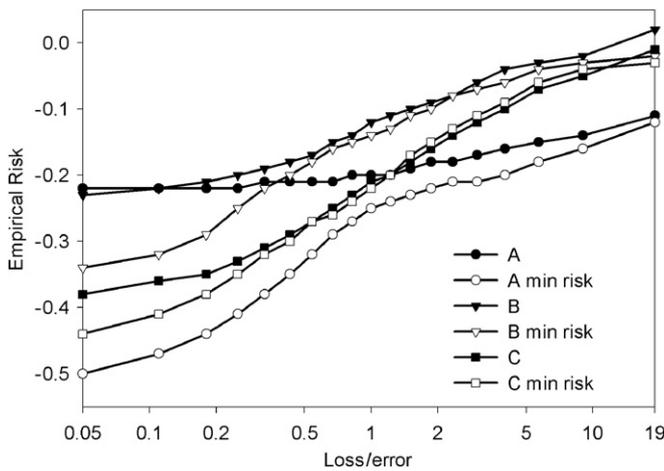


Figure 12. Empirical risk equation (13) of the two-stage method for experiments A, B and C as ls is varied and u remains at $u = 1$ (see tables 1 and 2).

segment. Performance of the classifier using the resubstitution technique on the background training segment is compiled in this row. It can be noted that the rate of background events is different depending on the odor. This can be attributed in some cases to the sensitivity of the antennae to the particular odor. For instance, indole generally results in small EAG depolarizations, many of which may be superimposed on background noise, and consequently placed in the ambiguous or background classes.

Lower error rates are observed for real-time tests when compared to the cross-validation runs. The average empirical error rate for an ambiguity reject parameter of 0.6 is 0.032 and 0.006 for experiment A, 0.15 and 0.12 for experiment B, 0.25 and 0.17 for experiment C in the cross-validation results and real-time results, respectively. This may be due to the increased number of training points.

4. Discussion

A real-time bioelectronic nose utilizing insect antennae as the odor sensor has been implemented using k -NN and semi-parametric methods for classification. The results presented here demonstrate the ability to discriminate odors by analyzing an electrophysiological signal derived from an array of differentially tuned insect olfactory organs. ORNs respond quickly to the arrival of an odor which contributes to the rapid response time of biological olfactory systems. The EAG impulse response is complete in approximately 15–75 ms, allowing sub-second odor classifications to be made using the present biosensor system. This can be contrasted with artificial noses, which typically detect the mean odorant concentration over tens of seconds or minutes.

The k -NN classifier has been modified to reject odors as ambiguous using an ambiguity reject parameter entered by the user. Further, the effects of several parameters on classifier performance have been evaluated, including number of training points per odor and the ambiguity reject parameter. Using data sets recorded under varying conditions (species,

odorants), classifier performance was consistently high after ambiguous responses were removed. The overall error rate was low ($\lesssim 0.2$) when compared to that of chance. Training time was found to be approximately 20 s per odor for the four-channel experiments and 40 s per odor for the eight-channel experiments. The consistent performance of the sensor/classifier predicts good performance using a variety of insect species and odorants. The very low error rate obtained when using the experiment A data suggests that the species and odorants can be optimally chosen to maximize system performance, and an error rate near 0% can be obtained under ideal circumstances. An ad hoc comparison of the two-stage approach was made to a more ideal conditional risk minimization classifier. It was found that generally the conditional risk classifier was better than the two-stage approach, under many circumstances the performances were equivalent. Under the real-time conditions, it was found that the rate of rejection chosen resulted in performance equivalent to penalizing errors at approximately twice the rate of rewarding correct classifications for the experiments involving non-pheromonal odors. This value was undetermined for experiment A, where pheromonal odors were easier to classify.

System performance was quantified using the empirical error rate equation (9) and empirical risk equation (13). These accuracy measures are distinct from a direct measure of classifier accuracy. A measure of accuracy would require absolute knowledge of the odorant in contact with the sensor array for each response event recorded. This difficulty is known in the literature as the imperfect teacher problem. A limiting assumption of the accuracy measure is that the response events recorded during presentation of a given odor are due either to that odor or to ‘background’ odorants.

It is evident that in some cases, additional diffuse background activity on the EAG waveforms does not interfere with classification performance, but in other cases does. It is possible that the high time resolution could be utilized to provide an even better resolution of superimposed odors.

An important limitation of the current system is the longevity of the excised antennae, which typically provide reliable responses for 30–90 min, after which new antennae need to be mounted and the system retrained. One possible means of overcoming this limitation is to record from the antennae *in situ*, while the intact insect is restrained. Stable EAG recordings can be made for days using this approach (unpublished data). A second approach would be to use a partially dissected preparation, or voltage sensitive dyes and to record directly from the glomeruli in the antennal lobe of a single insect. Direct access to the different glomeruli would be the basis for a sensor with discrimination approaching that of the insect; recordings from 21 individual glomeruli of honeybees has recently been used to discriminate between several odors using principal component analysis [53]. This approach is not currently practical for field applications due to the significant difficulty in applying optical recording techniques.

Acknowledgments

This research was initially funded by the Controlled Biological Systems Program of Defense Advanced Research Projects Agency (DARPA). It was subsequently funded by the Office of Naval Research (ONR) and the Defense Threat Reduction Agency (DTRA), through grants to TCB at Iowa State University (DARPA) and at Penn State University (ONR Counter-IED Program; DTRA). This research was also supported by a Keystone Alliance grant from the State of Pennsylvania, through Penn State University. The authors would like to thank Bryan Banks, Penn State University, for rearing the test insects and for his after-hours assistance in setting up the wind tunnel for humidification. We would also like to thank Emily Kuhns, Penn State University, for insects used in this study.

References

- [1] Persaud K and Dodd G 1982 Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose *Nature* **299** 352–5
- [2] Pearce T C, Schiffman S S, Nagle H T and Gardner J W 2002 *Handbook of Machine Olfaction* (Weinheim: Wiley)
- [3] Murray G M and Southard G E 2002 Sensors for chemical weapons detection *IEEE Instrum. Meas. Mag.* **5** 12–21
- [4] Dickinson T A, White J, Kauer J S and Walt D R 1998 Current trends in ‘artificial-nose’ technology *Trends Biotechnol.* **16** 250–8
- [5] Arshak K, Moore E, Lyons G M, Harris J and Clifford S 2004 A review of gas sensors employed in electronic nose applications *Sensor Rev.* **24** 181–98
- [6] Ziegler C, Gopel W, Hammerle H, Hatt H, Jung G, Laxhuber L, Schmidt H L, Schutz S, Vogtle F and Zell A 1998 Bioelectronic noses: a status report. Part II *Biosensors Bioelectron.* **13** 539–71
- [7] Arshak K, Lyons G M, Cunniffe C, Harris J and Clifford S 2003 A review of digital data acquisition hardware and software for a portable electronic nose *Sensor Rev.* **23** 332–44
- [8] Fix E and Hodges J L 1951 Discriminatory analysis: nonparametric discrimination: consistency properties *USAF School of Aviation Medicine, Project 21-49-004, Report Number 4* 261–79
- [9] Cover T M and Hart P E 1967 Nearest neighbor pattern classification *IEEE Trans. Inf. Theor.* **IT-13** 21–7
- [10] Tan S-L, Covington J A and Gardner J W 2006 Velocity-optimized diffusion for ultra-fast polymer-based resistive gas sensors *IEE Proc.: Sci. Meas. Technol.* **153** 94–100
- [11] Justice K A, Carde R T and French A S 2005 Dynamic properties of antennal responses to pheromone in two moth species *J. Neurophysiol.* **93** 2233–9
- [12] Grate J W 2000 Acoustic wave microsensor arrays for vapor sensing *Chem. Rev.* **100** 2627–48
- [13] Schild D 1988 Principles of odor coding and a neural network for odor discrimination *Biophys. J.* **54** 1001–11
- [14] Torre V, Ashmore J F, Lamb T D and Menini A 1995 Transduction and adaptation in sensory receptor cells *J. Neurosci.* **15** 7757–68
- [15] Getchell T V, Margolis F L and Getchell M L 1984 Perireceptor and receptor events in vertebrate olfaction *Prog. Neurobiol.* **23** 317–45
- [16] Anderson P A V and Ache B W 1985 Voltage- and current-clamp recordings of the receptor potential in olfactory receptor cells in situ *Brain Res.* **338** 273–80
- [17] Greiner B, Gadenne C and Anton S 2004 Three-dimensional antennal lobe atlas of the male moth, *Agrotis ipsilon*: a tool to study structure-function correlation *J. Comp. Neurol.* **475** 202–10
- [18] Tegoni M, Campanacci V and Cambillau C 2004 Structural aspects of sexual attraction and chemical communication in insects *Trends Biochem. Sci.* **29** 257–64
- [19] Martin E C, Jefferis G S X E, Komiyama T, Zhu H and Luo L 2002 Representation of the glomerular olfactory map in the drosophila brain *Cell* **102** 243–55
- [20] Shields V D C and Hildebrand J G 2001 Recent advances in insect olfaction, specifically regarding the morphology and sensory physiology of antennal sensilla of the femal sphinx moth *Manduca sexta* *Microsc. Res. Tech.* **53** 307–29
- [21] Liljefors T, Thelin B and van der Pers J N C 1984 Structure-activity relationships between stimulus molecule and response of a pheromone receptor cell in turnip moth, *Agrotis segetum* *J. Chem. Ecol.* **10** 1661–75
- [22] Liljefors T, Thelin B, van der Pers J N C and Löfstedt C 1985 Chain-elongated analogues of a pheromone component of the turnip moth, *Agrotis segetum*. A structure-activity study using molecular mechanics *J. Chem. Soc., Perkin Trans. 2* **12** 1957–62
- [23] Liljefors T, Engtsson M and Hansson B S 1987 Effects of double-bond configuration on interaction between a moth sex pheromone component and its receptor *J. Chem. Ecol.* **13** 2023–40
- [24] Baker T C and Haynes K F 1989 Field and laboratory electroantennographic measurements of pheromone plume structure correlated with oriental fruit moth behavior *Physiol. Entomol.* **14** 1–12
- [25] Karg G and Sauer A E 1994 Spatial distribution of pheromone in vineyards treated for mating disruption of the grape vine moth *Lobesia botrana* measured with electroantennograms *J. Chem. Ecol.* **21** 1299–314
- [26] van der Pers J N C and Minks A K 1998 A portable electroantennogram sensor for routine measurements of pheromone concentrations in greenhouses *Ent. Exp. Appl.* **87** 209–15
- [27] Park K C, Ochieng S A, Zhu J and Baker T C 2002 Odor discrimination using insect electroantennogram responses from an insect antennal array *Chem. Senses* **27** 343–52
- [28] Hetling J R, Myrick A J, Park K C and Baker T C 2003 Odor Discrimination Using a Hybrid-Device Olfactory Biosensor *Proc. 1st Int. IEEE EMBS Conf. on Neural Engineering, 2003*
- [29] Myrick A J, Baker T C, Park K-C and Hetling J R 2005 Bioelectronic Artificial Nose Using Four-Channel Moth Antenna Biopotential Recordings *Proc. 2nd Int. IEEE EMBS Conf. on Neural Engineering, 2005*
- [30] Shaffer R E, Rose-Pehrsson S L and McGill R A 1999 A comparison study of chemical sensor array pattern recognition algorithms *Anal. Chim. Acta* **384** 305–17
- [31] Shanmugam K and Breipohl A M 1971 An error correcting procedure for learning with an imperfect teacher *IEEE Trans. Syst. Man Cybern.* **SMC-1** 223–9
- [32] Krishnan T 2001 Imperfect supervision in statistical pattern recognition *Pattern Recognition from Classical to Modern Approaches* ed S P Pal and A Pal (River Edge, NJ: World Scientific) pp 25–65
- [33] Principe J C, Radisavljevic A, Fisher I John, Hiatt M and Novak L M 1998 Target prescreening based on a quadratic gamma discriminator *IEEE Trans. Aerosp. Electron. Syst.* **34** 706–15
- [34] Solis J L, Seeton G E, Li Y and Kish L B 2005 Fluctuation-enhanced multiple-gas sensing by commercial taguchi sensors *IEEE Sensors J.* **5** 1338–45
- [35] French A S and Meisner S 2007 A new method for wide frequency range dynamic olfactory stimulation and characterization *Chem. Senses* **32** 681–8

- [36] Chow C 1970 On optimum recognition error and reject tradeoff *IEEE Trans. Inf. Theor.* **16** 41–6
- [37] Gose E, Johnsonbaugh R and Jost S 1996 *Pattern Recognition and Image Analysis* (Upper Saddle River, NJ: Prentice Hall)
- [38] Silverman B W 1986 *Density Estimation* (London: Chapman and Hall)
- [39] Loftsgaarden D O and Quesenberry C P 1965 A nonparametric estimate of a multivariate density function *Ann. Math. Stat.* **36** 1049–51
- [40] Karagiannopoulos M G, Anyfantis D S, Kotsiantis S B and Pintelas P E 2007 Local cost sensitive learning for handling imbalanced data sets *Mediterranean Conf. Control & Automation* pp 1–6
- [41] Brown T A and Koplowitz J 1979 The weighted nearest neighbor rule for class dependent sample sizes *IEEE Trans. Inf. Theor.* **IT-25** 617–9
- [42] Dasarathy B V 1991 *Nearest Neighbor (NN) Norms: Nn Pattern Classification Techniques* (Los Alamitos, CA: IEEE Computer Society)
- [43] Baum C W and Pursley M B 1992 Bayesian methods for erasure insertion in frequency-hop communication systems with partial-band interference *IEEE Trans. Commun.* **40** 1231–8
- [44] Fralick S C and Scott R W 1971 Nonparametric Bayes-Risk Estimation *IEEE Trans. Inf. Theor.* **IT-17** 440–4
- [45] Theodoridis S and Koutroumbas K 1999 *Pattern Recognition* (San Diego, CA: Academic)
- [46] Chen Z and Fu K-S 1977 Nonparametric Bayes risk estimation for pattern classification *IEEE Trans. Syst. Man Cybern. SMC-7* 651–6
- [47] Miller J R and Roelofs W L 1978 Sustained-flight tunnel for measuring insect responses to wind-borne sex pheromones *J. Chem. Ecol.* **4** 178–98
- [48] Hill A, Carde R, Comeau A, Bode W and Roelofs W 1974 Sex pheromones of the tufted apple bud moth (*Platynota idaeusalis*) *Environ. Entomol.* **3** 249–52
- [49] Pope M M, Gaston L K and Baker T C 1984 Composition, quantification, and periodicity of sex pheromone volatiles from individual *Heliothis zea* females *J. Insect Physiol.* **30** 943–5
- [50] Roelofs W L, Du J W, Tang X H, Robbins P S and Eckenrode C J 1985 Three European corn borer populations in New York based on sex pheromones and voltinism *J. Chem. Ecol.* **11** 829–36
- [51] Roelofs W L, Comeau A, Hill A and Milicevic G 1971 Sex attractant of the codling moth: characterization with electroantennogram technique *Science* **174** 297–9
- [52] Berger R S 1966 Isolation, identification, and synthesis of the sex attractant of the cabbage looper, *Trichoplusia ni* *Ann. Entomol. Soc. Am.* **59** 767
- [53] Galan R F, Sachse S, Galizia C G and Herz A V M 2004 Odor-driven attractor dynamics in the antennal lobe allow for simple and rapid olfactory pattern classification *Neural Comput.* **16** 999–1012