Active Concentration-Independent Chemical Identification With a Tunable Infrared Sensor

Jin Huang, Rakesh Gosangi, and Ricardo Gutierrez-Osuna, *Senior Member, IEEE*

***Abstract*— This paper presents an active-sensing framework for concentration-independent identification of volatile chemicals using a tunable infrared interferometer. The framework operates in real time to generate a sequence of absorption lines that can best discriminate among a given set of chemicals. The active-sensing algorithm was previously developed to optimize temperature programs for metal-oxide chemosensors. Here, we adapt it to tune a nondispersive infrared spectroscope on the basis of a Fabry–Pérot interferometer (FPI). We also extend this framework to allow the identification of chemical samples irrespective of their concentrations. Therefore, we use nonnegative matrix factorization to create concentration-independent absorption profiles of different chemicals, and then employ linear least squares to fit sensor observations to the response profiles. We tested the framework on a simulated classification problem with 27 chemicals and compared against a passive sensing approach; the active-sensing consistently outperformed the passive sensing in terms of classification performance for various sensing budgets and at various levels of sensor noise. We also validated the approach experimentally using a commercial FPI sensor and a database of eight household chemicals. Our results show that the method can predict the sample identity irrespective of concentration.**

***Index Terms*— Active sensing, concentration normalization, Fabry–Perot interferometer, tunable sensors.**

I. Introduction

OVER the past sixty years [1], infrared (IR) spectroscopy has been extensively used for chemical analysis. One of the most popular IR spectroscopic techniques is Fourier transform infrared spectroscopy (FTIR). FTIR provides high-resolution spectra (down to 0.001 cm^{-1}) covering a wide spectral range ($200 - 10000$ cm^{-1} at 2 cm^{-1} resolution) [2]. However, FTIR is mostly limited to laboratory settings due to its cost, size, and power requirements. Recently, many affordable and portable alternatives to FTIR have become commercially available but these spectrometers come at the expense of lower spectral resolution and range. These instruments find use in application domains with simpler chemical identification problems (e.g. industrial monitoring, hazardous gas detection, indoor air quality monitoring). Along these lines, in previous work [3] we developed an IR spectroscope based on linear variable filter (LVF). The LVF consisted of a wedge-shaped interference filter, which provides a bank of transmission wavelengths from the thinnest end (short wavelengths) to the thickest (long wavelengths), atop an array of pyroelectric detectors. An alternative to LVF-based spectrometers is the Fabry-Perot interferometer (FPI), a device consisting of two partially reflective parallel mirrors whose distance can be tuned to filter out different wavelengths. Since FPIs consist of a single pixel, they are more cost-effective than LVFs, which require a detector array. FPIs can also be more power-efficient if a focusing lens is used to increase the signal to noise ratio (SNR). Finally, unlike LVFs, which acquire the entire spectrum at once, FPIs are tunable: they can measure absorption at a specific wavelength of interest.

In many chemical identification problems, the entire spectrum is not required in order to discriminate the various species. This idea has inspired various approaches for wavelength selection, where the goal is to find the most informative wavelengths for each discrimination problem. Wavelength selection can speed up the sensing process and also increase classification accuracy since irrelevant information is ignored. This idea has been around for more than two decades [4]–[8]. However, these algorithms work offline – a subset of wavelengths is selected using training data and the sensor is operated at these fixed frequencies, and generally delays any processing of sensor data until all wavelengths have been acquired.

In contrast with wavelength selection, the approach presented here focuses on active-sensing strategies to select wavelengths sequentially on-the-fly, based on sensor responses obtained thus far. This allows the sensor to adapt its sensing program in response to different chemical stimuli and their concentrations, as well as to environmental influences. Our approach leverages prior work on active temperature programming for metal-oxide (MOX) sensors [9], and models active sensing as a probabilistic state estimation process [10]. In this article,[1] we apply the active-sensing algorithm to Fabry-Perot interferometry. More importantly, we extend the approach to allow chemical identification at multiple concentrations. The approach consists of generating concentration-independent absorption profiles for each chemical target through nonnegative matrix factorization (NMF) [11], and fitting incoming

Manuscript received November 6, 2011; revised July 5, 2012; accepted July 18, 2012. Date of publication August 7, 2012; date of current version October 4, 2012. This work was supported in part by the National Science Foundation under Grant 1002028. The associate editor coordinating the review of this paper and approving it for publication was Prof. Julian W. Gardner.

The authors are with the Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: tonmey@cse.tamu.edu; rakesh@cse.tamu.edu; rgutier@cse.tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2012.2212186

[1] An earlier version of this paper was presented at the International Symposium on Olfaction and Electronic Nose (ISOEN 2011) as a two-page extended abstract [12].

sensor data to those profiles through linear least squares (LLS) [13]. We evaluate the concentration-independent active sensing algorithm on a database of IR absorption spectra from 27 chemicals, as well as experimentally on an 8-chemical discrimination problem using an FPI prototype.

II. Related Work

A handful of research groups have applied the idea of active sensing to machine olfaction. Priebe et al. [14] developed a statistical pattern recognition method based on the concept of Integrated Sensing and Processing (ISP). Given a feature vector, the method builds a decision tree that partitions feature space hierarchically; nodes close to the root of the tree select features based on their ability to provide good clustering of examples regardless of class labels, whereas nodes at the leaves select features based on their ability to discriminate examples from different classes. The authors evaluated the model on an experimental dataset from an array of 19 optical sensors exposed to trichloroethylene (a carcinogenic industrial solvent) in complex backgrounds. Their results show that the ISP method can reduce misclassification rates by 50%, while requiring only 20% of all the sensors to make any individual classification.

In previous work [9], we used Partially Observable Markov Decision Processes (POMPDs) for actively temperature programming with MOX gas sensors. We modeled the dynamic sensor response to a sequence of temperature pulses using an input-output hidden Markov model (IOHMM): the IOHMM predicted the sensor response to a sequence of temperature pulses. Once sensor models were learned, temperature programs were optimized as a POMDP. The method was validated experimentally on commercial MOX sensors.

More recently, Dinakarababu et al. [15] proposed an adaptive spectroscopic architecture called adaptive feature specific spectrometer (AFSS). Unlike a traditional IR spectrometer, AFSS has a tunable spectral filter (a digital micro-mirror device), which aids in multiplexing certain spectral bands and directing them onto a photo-detector. The system thus measures the projection of the incoming spectral density onto a set of basis vectors, rather than measuring the spectral density directly. The basis vectors are changed over time based on the information obtained from previous measurements. For a more thorough review on adaptive chemo-sensing systems, please refer to [16].

III. Methods

A. Problem Statement

Given a gas sample of unknown concentration but known to belong to one of chemical classes $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ and a tunable IR spectrometer with l spectral lines $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_l\}$, consider the problem of finding a sequence of actions $(a_1, a_2, \ldots, a_T)$ that minimizes the cost of discriminating gas samples. For this purpose, each action a_i has an associated cost: tuning the spectrometer to wavelength λ_r incurs a cost $C_s(\lambda_r) = c_r$ (e.g., power consumption), and classifying the gas sample based on available information carries a misclassification cost $C_c(u, v) = c_{uv}$ when a sample from class is incorrectly assigned to class ω_v.

We model this problem as that of probabilistic state estimation where each of the n classes is represented as a state, and maintain a probabilistic distribution b_T (where T denotes time) that represents our belief that the sample belongs to each class

$$b_T : \Omega \rightarrow [0, 1]; \sum_{\omega_i \epsilon \Omega} b_T(\omega_i) = 1. \quad (1)$$

Given an initial belief distribution $b_o(\omega_i)$, a sequence of actions $a_1, a_2, \ldots, a_T$, and the corresponding observations $o_1, o_2, \ldots, o_T$, the current belief b_T is defined as

$$b_T(\omega_i) = p(\omega_i | o_1, o_2, \ldots, o_T, a_1, a_2, \ldots, a_T). \quad (2)$$

In [9], we used a similar approach to actively modulate the operating temperatures of MOX sensors. However, that formulation did not account for the concentration of the gas sample; the framework assumed that each chemical was presented at a fixed concentration. Here, we extend the framework to identify samples at various concentrations.

Our approach can be broadly divided into two stages: sensor modeling and active sensing. During the *sensor modeling stage*, we first create concentration-independent absorption profiles for each chemical; these profiles remove the linear concentration effects while preserving chemical identity information. The resulting concentration-normalized response is then modeled with a Gaussian mixture model. During the *active sensing stage*, we first determine the optimal operating wavelength for the sensor at each time step; for this, we use a utility function that measures the difference between the sensing cost at each wavelength and the corresponding expected reduction in Bayes risk. Then we acquire absorption at the chosen wavelength, remove linear concentration effects and update the belief distribution accordingly.

B. Sensor Modeling

To a first-order approximation, the relation between light absorption and its travelling medium follows the Beer-Lambert law [17]. The absorption A of an IR beam transmitted through a spectrometer with a gas chamber of length l filled with a chemical of absorption coefficient ϵ at concentration can be estimated as $A = \epsilon l c$. Thus, absorption is linearly dependent on analyte concentration. We take advantage of this relation to remove the linear influence of concentration on absorption spectra, thus creating concentration-independent absorption profiles. We start by collecting absorption spectra for each chemical (ω_i) at m different concentrations and organizing the data as a matrix X_i of size $m \times l$, where l is the number of discrete wavelengths in the FPI. We then employ NMF to factorize X_i into a product of two matrices W_i and H_i such that error function $||W_i H_i - X_i||^2$ is minimized

$$X_i = W_i H_i + U_i \quad (3)$$

where W_i is a column matrix of size $m \times 1$, H_i is a row matrix of size $1 \times l$, and U_i is a matrix of size $m \times l$. W_i represents the concentration of the m absorption spectra, H_i can be interpreted as the concentration-independent absorption

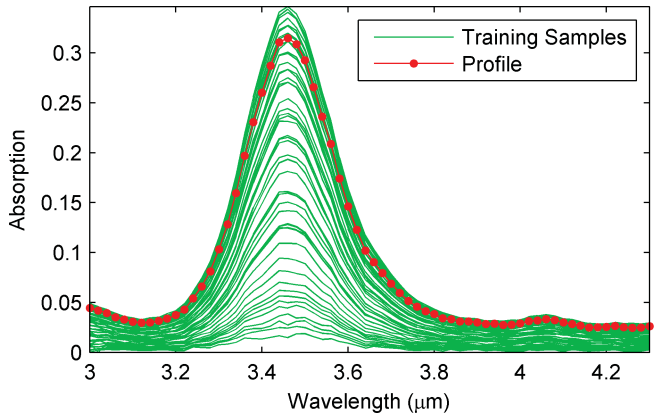


Fig. 1. Absorption spectra for acetone at 50 different concentrations (solid lines) and the estimated profile (dotted line).

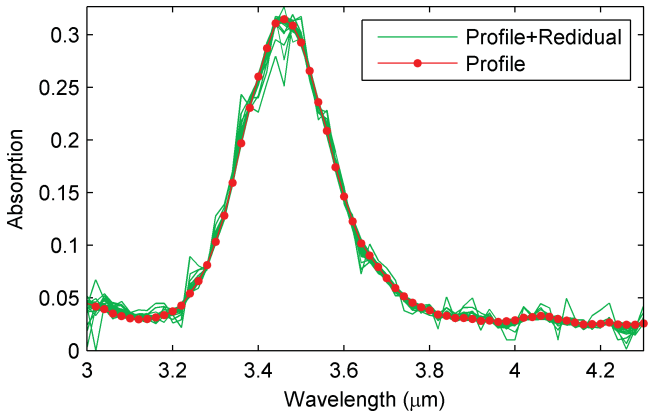


Fig. 2. Dotted line: absorption profile of acetone. Solid green lines: sum of the residual matrix and the absorption profile.

profile of chemical ω_i, and U_i is a residual matrix that captures sensor noise. Given a concentration ω_r, we can approximate the absorption spectra of chemical ω_i as $w_r H_i$. We chose to use NMF for this factorization process because NMF enforces a constraint that all the elements of W_i and H_i be non-negative; this constraint is necessary because concentrations W_i and absorption spectra H_i are strictly non-negative. This process is repeated n times to obtain absorption profiles for each chemical: $H_1, H_2, \ldots, H_n$. Fig. 1 illustrates the NMF process for the absorption spectra of acetone at 50 different concentrations ranging from 0% (air) to 100% (pure chemical).[2]

For each chemical, we then model the concentration-independent spectra (profile plus residual) with a Gaussian mixture model (GMM). These models will be used during the sensing stage to predict sensor responses. For each chemical ω_i, we first create matrix $\hat{X}_i$ as

$$\bar{X}_i = U_i + I_{m\times 1} H_i \tag{4}$$

where $I_{m\times 1}$ is an identity matrix of size $m \times 1$. Thus is the sum of sensor noise and concentration-independent absorption spectra for chemical ω_i. Fig. 2 shows an example of this matrix for the acetone dataset in Fig. 1.

Using a GMM, the response to chemical ω_i at wavelength λ_j can be expressed as follows:

$$P(\bar{x}|\omega_i, \lambda_j) = \sum_{k=1}^{M_{i,j}} \alpha_{i,j,k} N(o|\mu_{i,j,k}, \sigma_{i,j,k}) \tag{5}$$

where $M_{i,j}$ is the number of Gaussian components. $\alpha_{i,j,k}$, $\mu_{i,j,k}$, and $\sigma_{i,j,k}$ are the mixing coefficient, mean, and standard deviation of each Gaussian component, respectively. These models are trained on $\hat{X}_i$, i.e., the j^{th} column of matrix $\hat{X}_i$ is used to learn the mixture model for chemical ω_i at wavelength λ_j. Model parameters are estimated using Expectation Maximization [18].

[2]Concentrations are specified as percentage dilutions of a saturated headspace, e.g., a 10% concentration corresponds to a mixture of 90% air and 10% of the analyte vapors obtained from a saturated headspace.

C. Active Sensing

The first step in the active sensing stage is to select the 'best' wavelength to which the FPI should be tuned. For this, we use a greedy approach that selects a wavelength that maximizes the following utility function:

$$U(b_T, \lambda_i) = \Delta R - c_i = (R_C(b_T) - R_E(b_T, \lambda_i)) - c_i \tag{6}$$

where $U(b_T, \lambda_i)$ is the utility of wavelength λ_i, b_T is the current belief distribution, ΔR is expected reduction in Bayes risk, and c_i is the sensing cost at wavelength λ_i. The expected reduction in Bayes risk (ΔR) is defined as the difference between the current Bayes risk $R_c(b_T)$ and the expected risk $R_S(b_T, \lambda_i)$ upon tuning the sensor to λ_i. The current Bayes risk is estimated as

$$R_C(b_T) = \min_{\omega_u} \sum_{\omega_v \in \Omega} c_{uv} b_T(\omega_v) \tag{7}$$

which reflects the expected risk of classifying the sample. The expected Bayes risk $R_S(b_T, \lambda_i)$ of tuning the sensor to at the next time step is computed as

$$R_S(b_T, \lambda_i) = \sum_{\forall o} \min_u \left(\sum_{\omega_v \in \Omega} c_{uv}\, p(o|\omega_v, \lambda_i)\, b_T(\omega_v) \right) \tag{8}$$

$R_s(b_T, \lambda_i)$ averages the minimum Bayes risk over all observations that may result from λ_i. If the utility of all l wavelengths is negative, we halt the sensing process and classify the sample based on eq. (7). For a continuous observation space, eq. (8) becomes an intractable integral; instead, we discretize the absorption space into a finite set of values (see Appendix).

The second step in the active-sensing process is to tune the spectrometer to the optimal wavelength λ_T and obtain the corresponding observation o_T. To remove linear concentration effects, we then fit the observation sequence $\vec{o} = o_1, o_2, \ldots, o_T$ (for wavelength sequence $\vec{\lambda} = \lambda_1, \lambda_2, \ldots, \lambda_T$) to the profile of each chemical using linear least squares. Namely, given concentration-independent profile H_i for chemical ω_i, we find coefficient k_i that minimizes the sum-squared error $\sum_{j=1}^{T} \left(k_i \cdot H_i(\lambda_j) - o_j\right)^2$; this results in fitted observations $\vec{F}_i = \vec{o}/k_i$. The process is repeated for each chemical, leading to fitted observations $\vec{F}_1, \vec{F}_2, \ldots, \vec{F}_n$ that

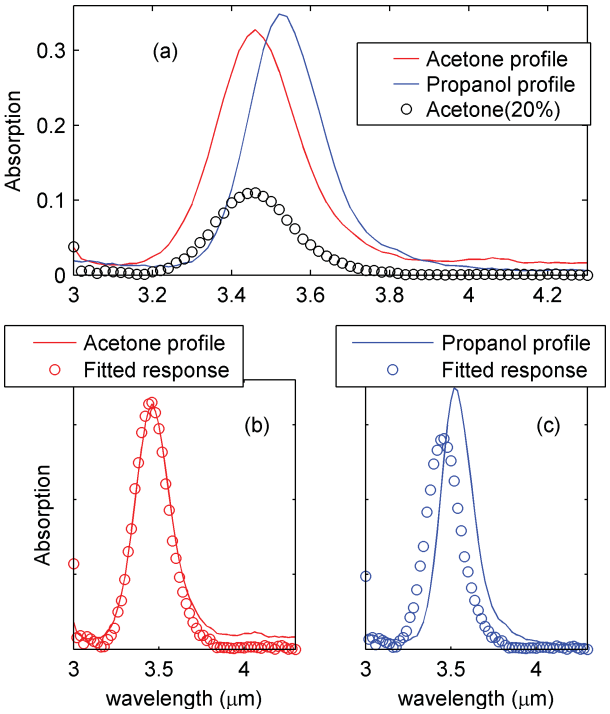


Fig. 3. Fitting sensor responses to concentration-independent profiles H_i. (a) Solid lines: absorption profiles for acetone and propanol. Circles: sensor observations in the presence of 20% acetone. Observation fitted to (b) acetone profile and (c) propanol profile.

are independent of concentration effects. Fig. 3 illustrates the entire process with an example. We first created absorption profiles for acetone and propanol (as described in section III.B) using data collected at 50 different concentrations ranging from 0% to 100%. Then, the sensor was exposed to acetone at a concentration of 20%, and we obtained responses at 66 wavelengths. The responses were then fitted to the concentration-independent profile of each chemical. As shown in the figure, the observations fit better to the acetone profile (mean square error:3.38×10^{-4}) than to propanol (4.4×10^{-3}).

The last step in the sensing process is to update the belief distribution using the fitted observations. Since the normalization step has to be performed on the entire observation sequence $(o_1, o_2, \ldots, o_T)$, we also recalculate the belief from time $t = 0$ using the fitted observation $\vec{F}_i$. The process is summarized in Table I.

In the above pseudo-code, η acts a normalization constant that ensures the belief distribution sums to 1. The value $p(\vec{F}_i(t)|\vec{\lambda}(t), \omega_i)$ is obtained using the probabilistic sensor models as described in section III.B.

IV. Simulation

We first tested the active-sensing framework on a large classification problem using simulated data; this allowed us to thoroughly compare the active-sensing approach against a passive feature selection strategy. To simulate the response of the FPI sensor to different chemicals, we used data from the NIST Chemistry WebBook [19], which provides high resolution FTIR spectra in the range 3–21 μm for over 40 000 chemicals in gas phase. We identified 27 chemicals

TABLE I
Pseudo Code for Belief Update

Belief_update $(\vec{F}_1, \vec{F}_2, \ldots, \vec{F}_n, \vec{\lambda})$
Initialize belief: $b_o(\omega_i) = \frac{1}{n} \forall i$
for $t = 1$ to $t = T$
  $\eta = 0$
  for $i = 1$ to $i = n$
  $b_t(\omega_i) = p(\vec{F}_i(t)|\vec{\lambda}(t), \omega_i) b_{t-1}(\omega_i)$
  $\eta = \eta + b_t(\omega_i)$
  end
  $b_t(\omega_i) = \frac{b_t(\omega_i)}{\eta}$
end

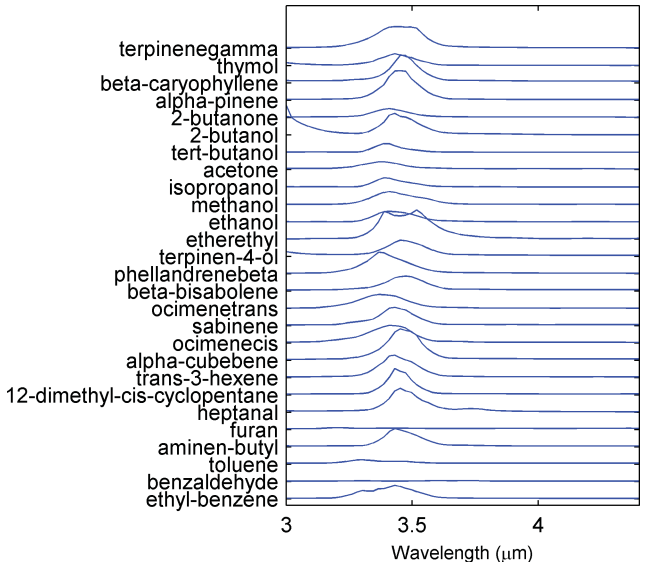


Fig. 4. Simulated absorption spectra for 27 chemicals. Spectra are plotted with an offset along the y-axis for visualization purposes.

that had at least one absorption peak in the operating range of our FPI (3-4.3 μm). To simulate the spectral resolution of the FPI, we downsampled the FTIR absorption spectra to 66 values, the number of unique wavelength tunings in the FPI. Simulated spectra are shown in Fig. 4. Using these spectra, we generated 30 absorption spectra for each chemical by adding Gaussian noise of variance 0.05 at each wavelength. The resulting dataset (30×27 spectra) was then used to train the sensor models, one for each chemical, as described in section III.B. We evenly discretized the observation space at each wavelength into 200 steps (see Appendix).

A. Results

First, we present a test case to illustrate the active-sensing process. In this case, the sensor was exposed to trans-3-hexene, and we assumed a 1-0 loss function for the classification costs $c_{uv} = I\,(u \neq v)$, and uniform sensing costs $c_i = 0.02$ for all wavelengths. The algorithm required 15 sensing actions before it classified the sample. Fig. 5(a) shows the average absorption spectrum of trans-3-hexene and the 15 operating wavelengths selected by the method, whereas Fig. 5(b) shows the belief distribution as a function of time. At time $t = 0$, all chemicals are equally likely. As observations are obtained, the

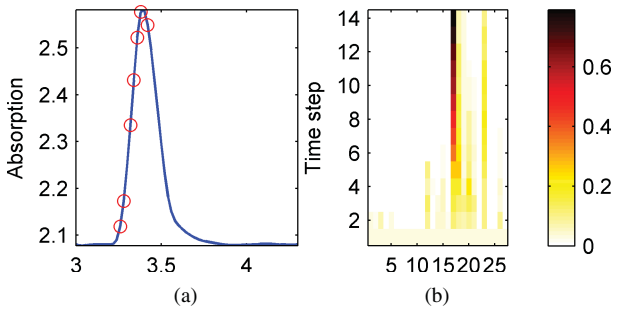


Fig. 5. (a) Average absorption spectrum of trans-3-hexene (solid line) and wavelengths chosen by the active sensing algorithm (circles); the following wavelengths were chosen: 3.4, 3.36, 3.3, 3.34, 3.28, 3.38, 3.38, 3.44, 3.38, 3.44, 3.44, 3.38, 3.38, 3.44, and 3.34. (b) Belief distribution as a function of time.

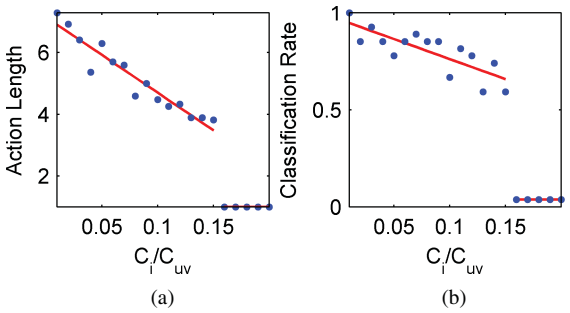


Fig. 6. (a) Average number of observations used. (b) Classification rate obtained by the active-sensing framework as a function of the ratio of sensing cost to misclassification cost (C_i/C_{uv}).

belief for trans-3-hexene, sabinene, and butyl-aminen increase. However, from time $t = 8$ the probability of trans-3-hexene starts dominating. At $t = 15$, the sensing process is halted since sufficient evidence is available where the utility based on eq.(6) for any further sensing becomes negative, and the sample is classified as trans-3-hexene.

Next, we tested the active sensing algorithm for various settings of the misclassification and sensing costs. Adjusting these costs allows us to balance the total cost of sensing against potential cost of misclassification. Without loss of generality, we used a 1-0 loss function for the classification costs $c_{uv} = I(u \neq v)$ and varied the sensing costs c_i from 0 to 0.2 in increments of 0.02. At each cost setting, we tested the algorithm 30 times on each of the 27 chemicals resulting in $27 \times 30 = 510$ test cases. Results are summarized in Fig. 6. As shown in the figure, the classification rate degrades with increasing sensing costs. Also, the average number of sensing actions used reduced with increasing sensing costs. Hence, the method balances the classification performance with sensing cost. We also observed that at $c_i \geqslant 0.16$ the cost of taking a sensing action is higher than the expected reduction in risk. Therefore, the algorithm halts the sensing process immediately, resulting in a classification performance of 3.7% which corresponds to chance level performance for a problem with 27 classes.

B. Comparison With Passive Sensing

We also compared the active sensing method against a feature subset selection strategy. Feature selection is a passive process where the optimal subset is obtained off-line using training data. In contrast, active sensing selects features on-line. We used sequential forward selection (SFS) coupled with a wrapper objective function [16] to obtain 'optimal' feature subsets of different cardinalities (p); the wrapper was based on a naïve Bayes classifier. To ensure a fair comparison between the two methods, we modified the stopping criterion of the active-sensing algorithm such that the algorithm stopped as soon as it acquired p observations.

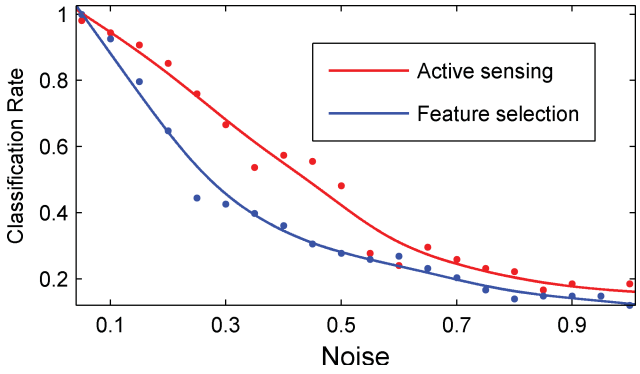


Fig. 7. Classification performance of the two methods as a function of the variance of the additive Gaussian noise.

First, we conducted an experiment to compare the performance of the two methods with increasing levels of measurement noise. For this purpose, we generated training data from FTIR spectra (see section IV), added Gaussian noise of variance 0.05 and then trained sensor models. To test the active sensing and feature selection methods, we generated 20 test sets, each containing 270 spectra (10 times per chemical), by changing the variance in the noise from 0.05 to 1 in steps of 0.05. Fig. 7 compares the classification performance of the two methods for $p = 15$ features. As shown, both strategies obtain nearly perfect classification performance at low noise levels. However, as noise levels increase active sensing consistently outperforms SFS. This is because active sensing selects features at measurement time in a way that adapts to noise levels, whereas SFS uses a pre-specified sequence that was computed off-line under more forgiving noise conditions.

To further emphasize the advantages of active sensing over feature subset selection, we conducted a second experiment that compares the classification performance of the two methods with increasing values of p, the number of features or wavelengths. As before, we generated training data with an additive Gaussian noise of variance 0.05. Then, we evaluated both methods on test data with additive noise of variance 0.4. For each value of p, we ran each method 270 times (10 times per chemical). Results are summarized in Fig. 8. As before, active sensing consistently outperforms SFS at all values of p.

V. Experimental

A. Experimental Setup

We also evaluated the active sensing framework on experimental data using an FPI device (LFP-3041L-337; Infratec, Inc). This device operates in the range of $(3-4.3\ \mu m)$ and has

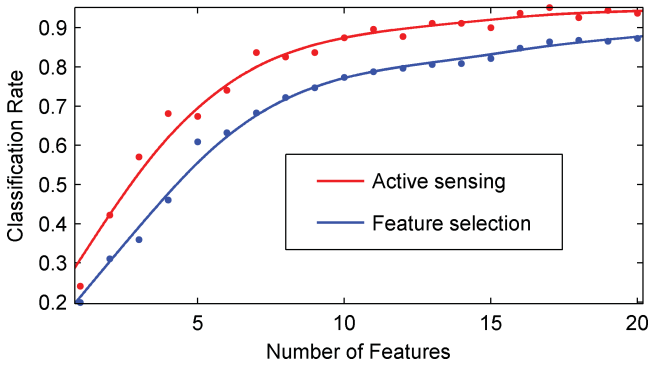


Fig. 8. Classification performance of the two methods as a function of the number of observations used.

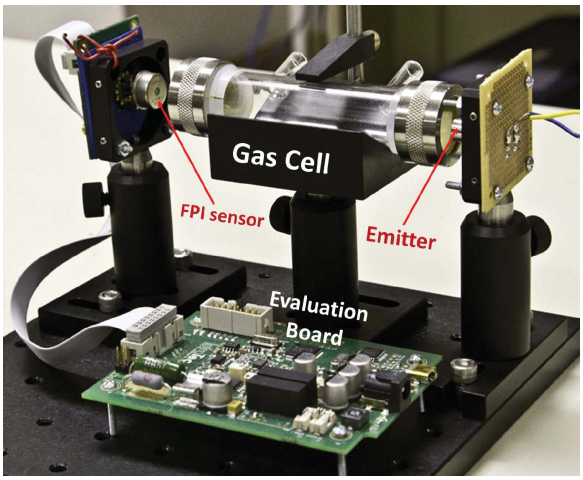


Fig. 9. Experimental prototype of the Fabry–Pérot spectrometer.

a resolving power $\lambda/\Delta\lambda$ of 60. We used a broadband infrared pulsable source (INTX 20-1000-R; Intex, Inc.) operated at a 10 Hz modulation frequency and 60% duty cycle. We mounted a 10cm gas cell (66001-10A; Specac, Inc.) with ZnSe window (602L08; Specac, Inc.) between the sensor and the IR source using an opto-mechanics fixture (Thorlabs, Inc.). This ensured a precise alignment of source, gas cell, and FPI sensor. The sensor was controlled using Matlab through a USB based evaluation board provided by the vendor. Chemicals were delivered to the system from 30 ml glass vials using negative pressure with a pump connected downstream from the sample cell. Fig. 9 shows the configuration of the device.

We tested the active-sensing algorithm and FPI prototype on a discrimination problem with eight chemicals; see Table II. We operated the FPI sensor at 66 different wavelengths ranging from 3 μm to 4.3 μm in steps of 0.02 μm. The sensor response was sampled at a rate of 1 KHz. Since the sensor response is modulated by the emitter, we can minimize interferences (e.g., external infrared sources, electronic noise) by extracting the power only at the modulation frequency (10 Hz) using Goertzel's algorithm [20]. We estimated transmittance as the ratio of the sensor response (power at 10 Hz) to the sample and to a reference gas (air), and converted transmittance Tr into absorption as $A = \log_{10}\frac{100}{Tr}$. The experiments were conducted in a laboratory environment at a temperature of 22.2 °C and

TABLE II
LIST OF CHEMICALS AND THEIR MAJOR COMPONENTS

Index	Chemical	Components
1	Air	
2	Brush cleaner	Raffinates, acetone, methanol
3	Lacquer thinner	Toluene, methanol, hexane, light aliphatic naphtha
4	Denatured alcohol	Ethyl alcohol, methanol
5	Acetone	Acetone
6	Xylene	Xylene (mixed isomers), ethylbenzene
7	Isopropyl alcohol	Isopropyl alcohol
8	Propanol	Propanol

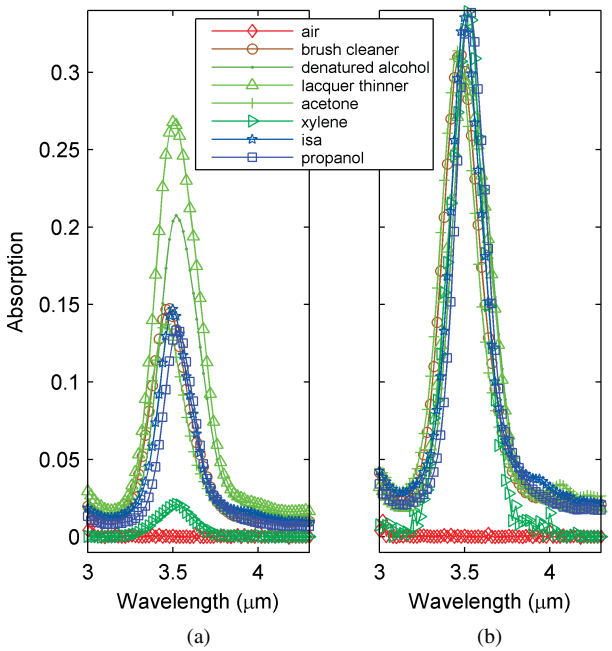


Fig. 10. (a) Average absorption spectra of all the chemicals including air. (b) Concentration-independent absorption profiles of all the chemicals.

standard atmospheric pressure of 760 mmHg. Before each experiment, we acquired the sensor's response to air and used these values as reference to estimate the absorption spectra. This helped us minimize the effects caused by changes in temperature, pressure, or humidity between experiments.

B. Experimental Data

We collected 50 absorption spectra for each chemical by varying the concentration from 0% to 100% in steps of 2%. Fig. 10(a) shows the average absorption spectra for the eight chemicals, obtained by averaging the 50 absorption spectra. We then applied NMF to obtain the concentration-independent absorption profiles, shown in Fig. 10(b). We generated training data for each chemical using the NMF profiles and residual matrices, as described in section III.B, from which GMMs were trained. We experimented with various numbers of Gaussian components per GMM, but GMMs with a single component proved sufficient since the NMF residual noise was fairly Gaussian-distributed.

We tested the framework 20 times for each chemical, resulting in 160 test cases. For each chemical, the concentration

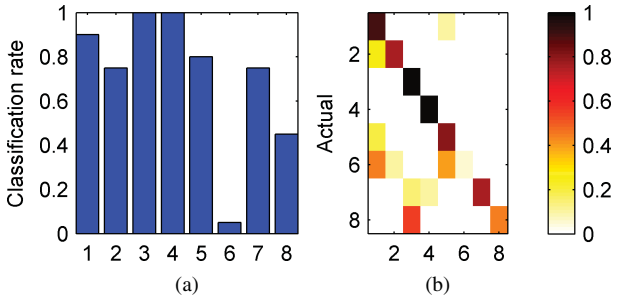


Fig. 11. (a) Classification performance (true positives) for different chemicals. (b) Corresponding confusion matrix.

of the test sample was randomly varied in the range of 20% to 80%. Chemicals were introduced in a randomized order to avoid any systematic errors or drift, and the gas cell was flushed with air for 2 minutes between exposures to remove residuals from the previous sample. When introducing a sample into the gas cell, we monitored the sensor response continuously until it stabilized; this ensured that the sample concentration had reached equilibrium. We used the sensor's average response (over 20 repetitions) to air as the reference.

C. Results

Based on the simulation results obtained in section IV.A, we chose to set the ratio of sensing cost and misclassification cost as 0.02 to promote high classification performance. Classification results and the corresponding confusion matrix are shown in Fig. 11(a) and (b), respectively. Denatured alcohol and lacquer thinner are correctly classified 100% of the times, followed by air, brush cleaner, acetone, and isopropyl alcohol, which are classified accurately on more than 75% of the test cases. In contrast, propanol is classified 50% of the times as lacquer thinner; the two chemicals have highest absorption strength (peaks) at same wavelength. Xylene is most often misclassified as air because of its low absorption strength in the sensor's spectral range, which is comparable to that of sensor noise variance. Also, the absorption profile of xylene obtained using NMF is significantly noisy compared to other chemicals (see Fig. 10(b)), especially in the range $3-3.3\ \mu m$ and $3.7-4.3\ \mu m$. We also observed that brush cleaner and acetone are often misclassified as air at low concentrations.

To test the robustness of the concentration-normalization method, we tested the framework on 10 samples of acetone at each of 10 concentrations, ranging from 100% (pure sample) down to 10% in steps of 10%, for a total of 100 test cases. Acetone concentration was controlled using a gas diluter [21]. Results are summarized in Fig. 12. The active sensing method accurately identified all samples in the concentration range 20-100%, and only failed at a 10% concentration, in which case all samples were classified as air. Fig. 12(b) shows the average number of actions used at different concentrations. The average number of observations used increased as concentration decreased, from 5 observations at a 100% concentration up to 9.3 observations at a 20% concentration. This result is consistent with the fact that the SNR decreases with concentration, and shows how the active sensing method can adapt the number of measurements required in order to obtain sufficient evidence for classification.

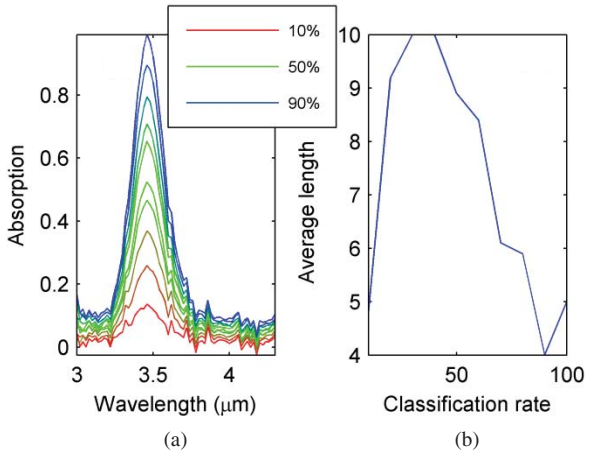


Fig. 12. (a) Average absorption spectra of acetone. (b) Average number of observations at the ten concentrations.

VI. Discussion and Future Work

We have presented an approach to actively select absorption wavelengths for a tunable IR interferometer in the context of concentration-independent discrimination of chemical samples. Our approach first creates concentration-independent absorption profiles for each target chemical using non-negative matrix factorization (NMF). The resulting normalized responses are then modeled using Gaussian mixture models. We formulated sensing as a decision-theoretic process, where we sequentially select wavelengths that are expected to provide the best reduction in Bayes risk. We validated the proposed method on both simulated and experimental data. Results on simulated data show that the active sensing can outperform passive sensing in terms of classification rates for various sensing budgets and at various levels of sensor noise. Using a Fabry-Perot interferometer, we further validated experimentally that the active sensing method can identify chemical samples independently of their concentration.

Our experimental results also show an interesting anomaly: samples at 10% concentration only triggered 4.8 observations on average and were all misclassified as air. We believe this result reflects the limitations of sensor's sensitivity. At low concentrations, the light beam interacts with fewer analyte molecules, which results in very weak absorption spectra; as shown in Fig. 1, the peak absorption values become comparable to sensor noise. As a result, the first few observations obtained drives up the belief associated with air, which tricks the algorithm into bringing the sensing process to an early termination. This limitation can be addressed, and the overall sensitivity of the system improved, by either: a) increasing the length of the optical path, which would increase the number of analyte molecules encountered by the light beam and thus increase the signal strength; or b) incorporating a pre-concentrator (PCT) into the gas delivery system, which would increase the concentration of the gas samples by 1-2 orders of magnitude.

To create concentration-independent absorption profiles, our approach uses NMF, a technique that it is only applicable to non-negative matrices. However, at low concentrations (or if an analyte has low absorption strength at certain wavelengths) the sensor response can become negative. Whenever this occurs, our current implementation replaces negative values with zeros before applying NMF. This affects the factorization process, since the negative noisy responses are replaced by zeros but their positive counterparts are not. This effect can be seen in xylene's absorption profile of Fig. 10(b). which is jagged as compared to the other chemicals. This is more evident in the wavelength ranges of $3-3.2\ \mu m$, and $4-4.3\ \mu m$, where many negative values have been replaced. An alternative solution to this problem would be to apply a denoising technique to the spectra prior to performing NMF.

At the time of this writing, we are extending the active sensing approach for multicomponent analysis by using sensor models that capture absorption spectra of gas mixtures. A PCT module might also be useful in this case to provide a broad separation of chemical species, e.g., by actively controlling the PCT temperature could be actively controlled to release different chemicals at different times. We expect the problems with interferents (e.g., background gases, water vapor), could be addressed with similar extensions.

APPENDIX

Eq. (8) is only applicable in discrete observation domains; in continuous spaces it becomes an intractable integration. To address this problem, we discretize the continuous observation space into a finite number of discrete values. For each wavelength λ_i, we uniformly discretize the corresponding observation space into a sorted set of d discrete values $\{\bar{o}_{i,1}, \bar{o}_{i,2}, \ldots, \bar{o}_{i,d}\}$. The posterior probability of the k^{th} discrete observation for chemical ω_j is calculated as

$$p(\bar{o}_{i,k}|\lambda_i, \omega_j) = \int_{\frac{\bar{o}_{i,k}+\bar{o}_{i,k-1}}{2}}^{\frac{\bar{o}_{i,k}+\bar{o}_{i,k+1}}{2}} \sum_{g=1}^{M} N(\bar{o}_{i,k}|\mu_{i,j,g}, \sigma_{i,j,g}). \quad (9)$$

The number of discrete observations d influences the accuracy and computational complexity of eq. (8). Therefore, after some experiments, we choose the d to be 200 to approximate the continuous observation space and still allow real-time operation.

REFERENCES

[1] K. Wilbur, "Near-infrared spectroscopy: I. Spectral identification and analytical applications," *Spectrochim. Acta*, vol. 6, no. 4, pp. 257–287, 1954.

[2] *MIR 8025™ FT-IR*. [Online]. Available: http://www.newport.com/Introduction-to-FT-IR-Spectroscopy/405840/1033/content.aspx, (Retrieved 2011).

[3] F. G. Nogueira, D. Felps, and R. Gutierrez-Osuna, "Development of an infrared absorption spectroscope based on linear variable filters," *IEEE Sensors J.*, vol. 7, no. 8, pp. 1183–1190, Aug. 2007.

[4] W. Wu and D. L. Massart, "Artificial neural networks in classification of NIR spectral data: Selection of the input," *Chemometr. Intell. Lab. Syst.*, vol. 35, no. 1, pp. 127–135, 1996.

[5] W. Wu, S. C. Rutan, A. Baldovin, and D.-L. Massart, "Feature selection using the Kalman filter for classification of multivariate data," *Anal. Chim. Acta*, vol. 335, nos. 1–2, pp. 11–22, 1996.

[6] W. Wu, B. Walczak, D. L. Massart, K. A. Prebbleb, and I. R. Last, "Spectral transformation and wavelength selection in near-infrared spectra classification," *Anal. Chim. Acta*, vol. 315, no. 3, pp. 243–255, 1995.

[7] Y. L. Mallet, D. H. Coomans, and O. de Vel, "Spectral data, modern classification methods for," in *Encyclopedia of Analytical Chemistry*. New York: Wiley, 2006.

[8] R. Leardi and A. L. González, "Genetic algorithms applied to feature selection in PLS regression: How and when to use them," *Chemomet. Intell. Lab. Syst.*, vol. 41, no. 2, pp. 195–207, 1998.

[9] R. Gosangi and R. Gutierrez-Osuna, "Active temperature programming for metal-oxide chemoresistors," *IEEE Sensors J.*, vol. 10, no. 6, pp. 1075–1082, Jun. 2010.

[10] S. Ji and L. Carin, "Cost-sensitive feature acquisition and classification," *Pattern Recogn.*, vol. 40, no. 5, pp. 1474–1485, 2007.

[11] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[12] J. Huang, R. Gosangi, and R. Gutierrez-Osuna, "Active sensing with Fabry–Perot infrared interferometers," in *Proc. AIP Conf.*, 2011, pp. 31–32.

[13] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Philadelphia, PA: SIAM, 1995.

[14] C. E. Priebe, D. J. Marchette, and D. M. Healy, "Integrated sensing and processing decision trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 699–708, Jun. 2004.

[15] D. R. G. D. V. Dinakarababu and M. E. Gehm, "Adaptive feature specific spectroscopy for rapid chemical identification," *Opt. Express*, vol. 19, no. 5, pp. 4595–4610, 2011.

[16] R. Gutierrez-Osuna, "Pattern analysis for machine olfaction: A review," *IEEE Sensors J.*, vol. 2, no. 3, pp. 189–202, Jun. 2002.

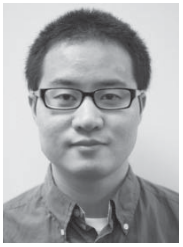
[17] A. Beer, "Bestimmung der absorption des rothen Lichts in farbigen Flüssigkeiten," *Ann. Phy. Chem.*, vol. 86, no. 5, pp. 78–88, 1852.

[18] A. Dempster, N. Laird, and D. Rdin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[19] S. E. Stein, "Infrared spectra," in *NIST Chemistry WebBook*, P. J. Linstrom and W. G. Mallard, Eds. Gaithersburg, MD: Nat. Inst. Standards Technol., (Retrieved Aug. 2011) [Online]. Available: http://webook.nist.gov

[20] G. Goertzel, "An algorithm for the evaluation of finite trigonomentric series," *Amer. Math.*, vol. 65, no. 8, pp. 3–35, 1958.

[21] *Model 1010 Precision Gas Diluter*. [Online]. Available: http://www.customsensorsolutions.com/M1010specs.html, (Retrieved 2007).



Jin Huang received the B.E. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2009. He is currently pursuing the Ph.D. degree at Texas A&M University, College Station.

His current research interests include pattern recognition, active sensing, and chemical sensors.

Rakesh Gosangi received the B.Tech. degree in computer science and engineering from the Indian Institute of Technology Madras, Chennai, India, in 2007. He is currently pursuing the Ph.D. degree at Texas A&M University, College Station.

His current research interests include pattern recognition, probabilistic models, active sensing, and chemical sensors.

Ricardo Gutierrez-Osuna (SM'00) received the B.S. degree in electrical engineering from the Polytechnic University of Madrid, Madrid, Spain, in 1992, and the M.S. and Ph.D. degrees in computer engineering from North Carolina State University, Raleigh, in 1995 and 1998, respectively.

He was a Faculty Member with Wright State University, Fairborn, OH, from 1998 to 2002. He is currently a Professor of computer engineering with Texas A&M University, College Station. His current research interests include pattern recognition, neuromorphic computation, chemical sensor arrays, and audio-visual speech processing.