

Independent component analysis of nanomechanical responses of cantilever arrays

Rick Archibald^{a,*}, Panos Datskos^a, Gerald Devault^a, Vincent Lamberti^b,
Nickolay Lavrik^{a,c}, Don Noid^a, Michael Sepaniak^c, Pampa Dutta^c

^a Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^b Y-12 National Security Complex, Oak Ridge, TN 37831, USA

^c University of Tennessee, Knoxville, TN 37996, USA

Received 21 July 2006; received in revised form 23 October 2006; accepted 2 November 2006

Available online 11 November 2006

Abstract

The ability to detect and identify chemical and biological elements in air or liquid environments is of far reaching importance. Performing this task using technology that minimally impacts the perceived environment is the ultimate goal. The development of functionalized cantilever arrays with nanomechanical sensing is an important step towards this goal. This report couples the feature extraction abilities of independent component analysis (ICA) and the classification techniques of neural networks to analyze the signals produced by microcantilever-array-based nanomechanical sensors. The unique capabilities of this analysis unleash the potential of this sensing technology to accurately identify chemical mixtures and concentrations. Furthermore, it is demonstrated that the knowledge of how the sensor array reacts to individual analytes in isolation is sufficient information to decode mixtures of analytes—a substantial benefit, significantly increasing the analytical utility of these sensing devices.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Independent component analysis; Functionalized cantilever arrays; Neural networks; Nanomechanical sensors

1. Introduction

Specially designed arrays of nanomechanical sensors have unique advantages for sensing chemical and biological agents in both gaseous and liquid environments. Demonstrated applications of such devices range from DNA analysis to selective detection of explosive compounds, to highly selective, antibody-mediated discrimination of enantiomers [1–5]. Selectivity in sensing is achieved by incorporating highly specific bioaffinity phases or using arrays of functionalized sensors exhibiting modest and distributed selectivity among the sensing elements [6,7]. At this stage, however, realization of the full potential of microcantilever sensing arrays depends largely on the availability of a suitable data analysis algorithm [8–10].

The general schematics of the nanomechanical chemical sensor are depicted in Fig. 1. At the heart of this device is an

array of beam-shaped transducers modified on one side with molecular recognition phases (MRPs). It also includes a “read” photon source, such as a laser or an LED, and an optical position sensitive detector (PSD), such as a CCD or a quadrant photodiode, which responds to spatial changes in the “read” beam. Judicious choice for each MRP in the array provides multiple isolated interaction surfaces for sensing the environment. When a particular chemical or biological agent binds to a transducer, the effective surface stresses of its modified and uncoated sides change unequally and the transducer begins to bend. Hence, each transducer is referred to as a microcantilever. The extent of bending depends upon the specific interactions between the microcantilever’s MRP and the analyte. Thus, the readout of a multi-MRP array is a complex multi-dimensional signal that out of necessity must be analyzed using advanced mathematics.

Given a properly designed array of cantilevers, the presentation of specific chemical or biological species at a particular concentration produces identifiable and reproducible dynamics. Specially designed MRPs are used for each cantilever

* Corresponding author. Tel.: +1 865 576 5761; fax: +1 865 574 0680.
E-mail address: archibaldrk@ornl.gov (R. Archibald).

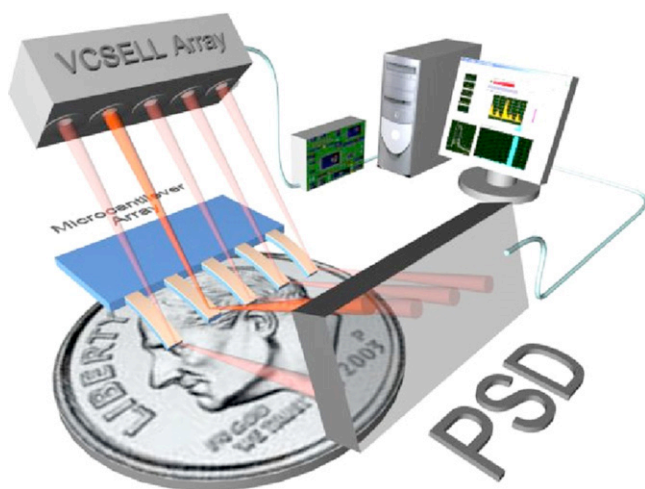


Fig. 1. Schematic illustration of a micro-electro-mechanical system (MEMS) sensor array with a readout based on sequential optical probing of the cantilevers with a vertical cavity surface emitting laser (VCSEL) array. The array used in this work is similar to this scheme except that a 12-cantilever array was used.

such that targeted environmental components (analytes) can be continually absorbed and desorbed. The dynamic equilibrium between adsorption and desorption varies the cantilever bending as a function of the concentration of the analytes and allows for system recovery. The bending action of each cantilever is therefore directly correlated to the presence of these targets, and the system as a whole produces a unique dynamic signature. Inherent noise is distributed across the system and the multi-dimensionality of the signal provides a natural mechanism to minimize the damaging effects of noise in analysis and classification.

The construction of this multi-cantilever sensor makes the readout signal specifically amenable to independent component analysis (ICA). The concepts surrounding ICA were first introduced in 1991 by Jutten et al. [11] and later formalized in 1994 by Comon [12]. Over the course of the last decade, ICA has proven to be a powerful technique in many different applications. It can provide a (linear) sparse coding representation of natural image data [13], and has also been adopted in bio-medical imaging, where it transposes EEG/MEG data into physiologically and functionally distinct sources [14,15]. Important to this study is the successful use of ICA in vision and auditory studies for feature extraction [16] and information processing [17]. It is a natural extension to adapt the ICA method to extract features and information for the current analysis of cantilever arrays that in essence act as an ‘electronic nose or tongue’.

This paper describes how ICA can be used as a data pre-processing method for neural network classification of nanomechanical sensor data. Appropriate coupling of these mathematical methods can accurately capture information contained within the array of sensors to identify chemical mixtures and concentrations. Specifically, how the sensor array interacts with individual analytes in isolation is sufficient information to decode mixtures of analytes. As a result, the usefulness of this device is greatly increased for complex applications.

2. Methods

2.1. Independent component analysis

A brief introduction of relevant properties of the ICA method, within the context of this study, is given here. We refer interested readers to Hyvärinen et al. for detailed information [18]. Suppose that $x(t)$ represents an N -dimensional measured time series vector of sensor signals. Assume that linear mixing of independent sources produces this measured data, or

$$x(t) = As(t), \quad (1)$$

for $s(t)$ the M -dimensional time series vector of independent sources and A the $N \times M$ mixing matrix. Further assume that $M \leq N$ and that A has full rank. ICA is a de-mixing procedure that, given only the measured data $x(t)$, recovers W , the mixing matrix, such that

$$W = SPA. \quad (2)$$

Here, S and P are scaling and permutation matrices, respectively. Amazingly, under these assumptions, both the mixing matrix and corresponding sources can be recovered based only on the knowledge of the measured data, to an arbitrary scaling and permutation. For systems where the measured data is the linear mixing of independent sources, ICA is a tool that can transform measurements into source information.

There are many variants of the ICA method and, in a general sense, they can be separated into two categories, based on how the mixing matrix is determined from the measured data. Instantaneous algorithms consider each data point in isolation, whereas summary algorithms generate summary statistics from a collection of data points. Generally, ICA methods that use summary algorithms have a greater tolerance to noise. For this reason, we use SOBI [19], an ICA method from the category of summary algorithms.

For this study, the feature extraction capability of ICA is used to pre-process readout data from cantilever arrays in order to concisely portray this information to the neural network for classification. Specifically, the multi-dimensional signal produced from the bending of the array of cantilevers is used as the measured data in the ICA method. Columns of the estimated mixing matrix are used as features, providing a vectorized input for neural network classification. Direct input of the entire signal, without pre-processing, will overwhelm the network and result in poor classification accuracy.

One fundamental issue that arises in this feature extraction procedure is the fact that ICA can determine the mixing matrix only to an arbitrary scaling and permutation. Therefore, it is necessary to impose a systematic procedure that induces consistent scaling and permutation. In this study, the scaling requirement will be that estimated sources are normalized and the permutation requirement will be that the mixing matrix columns are sorted in descending order by a distance norm. These requirements insure that consistent features are extracted.

The nature of the system of cantilevers lends itself to analysis from ICA. Through experiment we have found that for

this system, ICA has two significant properties. First, the dominant *features* depend strongly on the analyte–MRP interactions. Second, the dominant *sources* are similar across both chemical species and concentration. The explanations of these properties can be traced to the physics of the cantilevers. Since different MRP's will respond differently to a given analyte, these attributes will translate to the feature space. The estimation of similar sources stems from the fact that each cantilever to a certain degree undergoes similar types of characteristic bending, a direct result of the fundamental interfacial interactions between the transducer and the analyte. In other words, each source corresponds to distinctive bending dynamics that, in turn, can be ascribed to specific physical and chemical processes involved in the analyte–transducer interactions.

The vector structure of the extracted features can be exploited for efficient recovery of the identities and concentrations of the constituents of a gas sample from its microcantilever fingerprint. Each column of the mixing matrix represents an independent feature of the data, and a subset of the columns is sufficient for input to an artificial neural network. We have discovered that the same dominant independent sources are recovered regardless of concentration.

2.2. Neural networks

A premier method of data modeling and classification over the past two decades involves the use of neural networks. Since these methods have been extensively reviewed in the literature, a brief summary of our use is all that is needed in this report. Back-propagation networks were created with the MATLAB[®] Neural Network Toolbox. The typical architecture contained a single hidden layer and employed the hyperbolic tangent sigmoid transfer function for the input and hidden layers and the linear transfer function for the output layer. Training was accomplished through the Levenberg–Marquardt algorithm, a variant of Newton's method. Training was halted at a specified tolerance value.

2.3. Classification pseudo-algorithm

Consider a time period where M samples are taken from the readout signal of an array of N cantilevers. We notate this segment of signal as the matrix $x \in R^N \times R^M$. Classification of this portion of signal proceeds by the following steps:

- (I) Find the mixing matrix $W \in R^N \times R^N$ and corresponding sources $S \in R^N \times R^M$, as described in Eqs. (1) and (2) by using the ICA method on the signal segment.
- (II) Sort the mixing matrix such that (permutation requirement):

$$\sum_{i=1}^N |W(i, 1)| \geq \dots \geq \sum_{i=1}^N |W(i, N)|.$$

- (III) Normalize the mixing matrix such that sources satisfy (scaling requirement):

$$\sum_{i=1}^M S(1, i)^2 = \dots = \sum_{i=1}^M S(N, i)^2 = 1.$$

- (IV) Form the feature vector as

$$V(i) = W(i, \text{floor}(i/N) + 1), \quad i = 1, \dots, n \times N,$$

utilizing the $n \leq N$ most dominant features (the columns of the sorted and scaled mixing matrix).

- (V) Determine the identities and concentrations of analytes by applying the trained neural network to this feature vector.

3. Experimental

The experiments tested the analytes diisopropylmethylphosphonate (DIMP), ethanol, and propanol at a range of head space concentrations diluted in an inert gas (nitrogen or helium) and, subsequently, two room temperature gases, H₂ and CO₂, at volume percents of 10% or less in the diluent gas. Concentrations were adjusted by varying the flow rates of the analyte gases using programmable syringe pumps. These experiments were performed using arrays of silicon microcantilevers coated successively with nanostructured metal and MRPs [7]. The length, width, and thickness of each microcantilever were, respectively, 400 μm, 100 μm, and approximately 1 μm. Twelve different MRPs (macrocyclic compounds, GC phases, etc.) were thermally evaporated, one coating per cantilever, using the physical vapor deposition approach. This procedure resulted in a MRP thickness that was roughly 300 nm [7].

The microcantilever array was mounted in a stainless steel flow cell with a glass window and coupled to the optical readout. The total volume of the flow cell was approximately 150 μL. Inlet and outlet ports provided delivery and exhaust of the analyte mixtures. Beams of 12 VCSELs were focused onto the tips of the microcantilevers in the array. The beams reflected off the microcantilevers were captured and monitored by a single position sensitive detector (PSD, see Fig. 1). The deflection of a microcantilever resulted in a corresponding motion of the reflected beam across the PSD. Throughout the experiments, a constant flow of carrier gas was maintained through the flow cell at a rate of 4.0–6.0 mL min⁻¹. Analyte flow (and hence concentration) was controlled by the programmable syringe pumps. Each analyte was typically injected for a period of 30–60 s, followed by a 90–180 s period of a carrier gas flow only, thereby creating repetitive test periods.

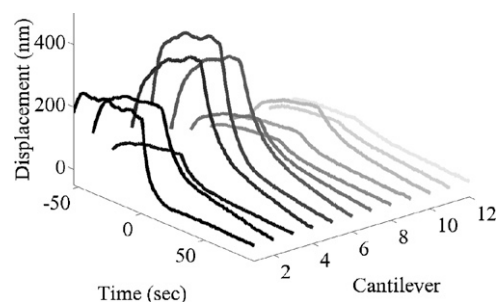


Fig. 2. The movement of all 12 cantilevers during one trial in which 1.2 mL min⁻¹ of ethanol is presented to the system. Time zero corresponds to end of the analyte injection and displacement is relative to end state.

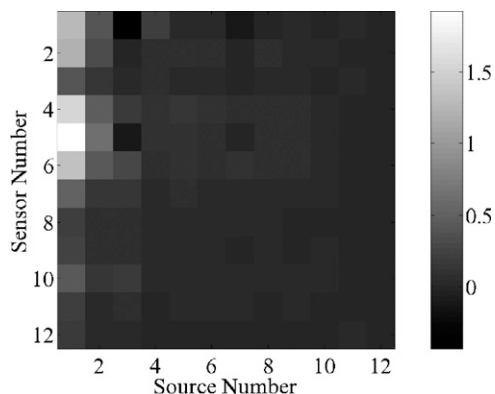


Fig. 3. Computed estimation of ICA mixing matrix for the cantilever array signals in Fig. 2.

4. Discussion and results

Depicted in Fig. 2 is the movement of all the MRP-modified cantilevers during one trial in which 1.2 mL min^{-1} of ethanol is presented to the system. Fig. 3 is the corresponding ICA decomposition of this signal. It can be seen that the magnitude of the columns of the mixing matrix drop rapidly. Similar results were observed for all the trials (50 for each chemical and each concentration). Therefore, only the first column is needed to represent the observed data and will be referred to as the feature vector.

Fig. 4 depicts the average of the first feature vector for ethanol, propanol, and DIMP at each of five concentrations. It can be seen that the feature vector provides a unique representation for each chemical and the magnitude of this feature is linearly related to the concentration. Neural network classification of these features supports these observations. Leave one out cross-

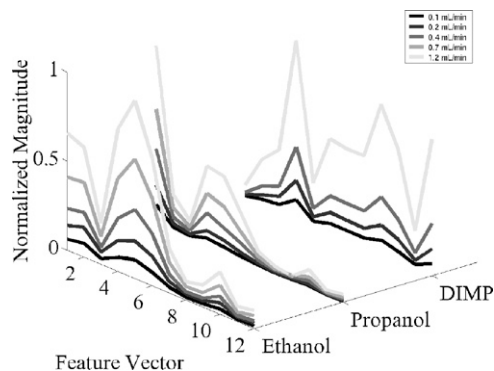


Fig. 4. Normalized averaged feature vector for each concentration of ethanol, propanol and DIMP. Concentrations were altered by varying the analyte headspace flow rate within a constant background flow of 6.0 mL min^{-1} of carrier gas.

validation was performed on all trials as a group with a training goal of 10^{-8} and a maximum number of training epochs of 2500. It was observed that the three different analytes could be accurately distinguished and their concentration predicted to an absolute average error of $5.1 \times 10^{-5} \text{ mL min}^{-1}$. The reliability of the extracted features can be explained by two important properties. First, the system of cantilevers response scales with the concentration rate of a given analyte, and the morphology of this response is different according to the chemical species. This property is a necessary condition for classification to be accurate and reliable. Second, the ICA estimates for dominant sources are similar across different chemicals and concentrations. Fig. 5 shows the average source estimation for ethanol, propanol, and DIMP. The first three sources averages are similar with correlation rates greater than $r=0.97$ among the signals. These sources

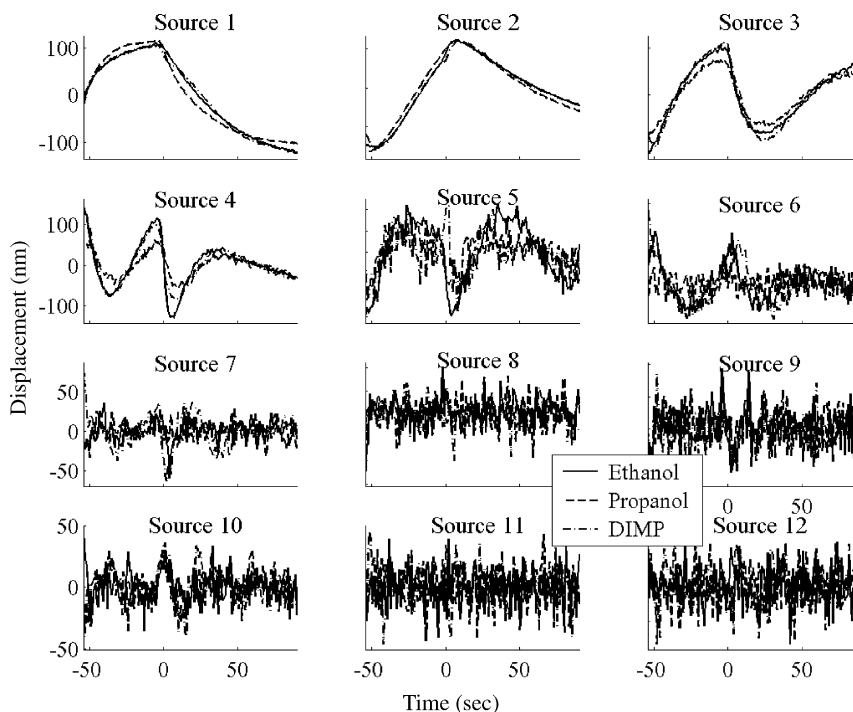


Fig. 5. Average of the 12 estimated sources for each molecule ordered in terms of signal strength contribution.

Table 1
Predicted vs. actual concentration in the mixture of CO₂ and H₂

Actual mixture of CO ₂ and H ₂	Predicted mixture of CO ₂ and H ₂ (mL min ⁻¹)
0.9:0.1	0.8947:0.0843
0.8:0.2	0.8548:0.1000
0.7:0.3	0.6279:0.3202
0.6:0.4	0.4946:0.4469
0.5:0.5	0.2603:0.5902
0.4:0.6	0.4369:0.5336
0.3:0.7	0.2603:0.5902
0.2:0.8	0.3017:0.7628

are heavily weighted in the mixing matrix and make up over 85% of the observed signals.

The combination of these two properties not only forces consistency among the features, allowing for precision in classification, but also implies linearity in the features for mixtures of chemicals. In other words, knowing only the features that correspond to chemical species presented to the system of cantilevers in isolation, the feature that results from any mixture of the chemical species can be determined directly through linear combinations of isolated chemical features. The result of the experimental validation of this phenomenon is recorded in Table 1. In this experiment, a constant flow of helium carrier gas (4.0 mL min⁻¹) was maintained. Two separate syringe pumps containing 50% H₂ and 50% CO₂ were mixed with the carrier gas such that the total analyte flow was 1.0 mL min⁻¹ with increasing H₂ (0.0–1.0 mL min⁻¹) and decreasing CO₂ (1.0–0.0 mL min⁻¹) flow rates.

Using only the feature vectors for CO₂ and H₂, where each is presented to the system in isolation, the neural network is trained on a computational synthetic mixture of these isolated feature vectors. Table 1 displays the accuracy of this neural network in predicting concentration rates from features generated from the action of the cantilevers for various real mixtures of CO₂ and H₂. Knowledge of how these analytes react to the system of cantilevers in isolation is sufficient to predict the concentration rates of mixtures to an absolute average error of 6.5×10^{-2} mL min⁻¹.

5. Summary

The detection and identification of chemicals through the use of functionalized nanomechanical sensor arrays is augmented by coupling this technology with the feature extraction abilities of ICA and the classification performance of neural networks. Further, based only on the knowledge of the features produced from individual chemical interactions with the sensor array, it is possible to train neural networks to identify the identities and concentrations of mixed chemicals, thereby greatly increasing the range of classification of this device.

Acknowledgments

R. Archibald would like to thank the Householder fellowship that is supported under the Mathematical, Information, and Computational Sciences Division; Office of Advanced Scientific Computing Research; U.S. Department of Energy (DE-AC05-00OR22725). We would also like to acknowledge support from the Defense Advanced Research Projects Agency. This work was partially supported by the Laboratory Director's Research and Development Program of ORNL. Additional support in part by U.S. Department of Energy, Basic Energy Sciences (DE-FG02-02ER15331) and the Environmental Protection Agency (EPA-83274001) under grants awarded to the University of Tennessee, Knoxville. Support was also provided by the Y-12 National Security Complex Plant Directed Research and Development program.

References

- [1] J. Fritz, M.K. Baller, H.P. Lang, H. Rothuizen, P. Vettiger, E. Meyer, H. Guntherodt, C. Gerber, J.K. Gimzewski, *Science* 288 (5464) (2000) 316–318.
- [2] K.M. Hanson, J.-F. Ji, G. Wu, R. Datar, R. Cote, A. Majumdar, T. Thundat, *Anal. Chem.* 73 (7) (2001) 1567–1571.
- [3] L.A. Pinnaduwaage, V. Boiadjev, J.E. Hawk, T. Thundat, *Appl. Phys. Lett.* 83 (7) (2003) 1471–1473.
- [4] L.A. Pinnaduwaage, D.L. Hedden, A. Gehl, V.I. Boiadjev, J.E. Hawk, R.H. Farahi, T. Thundat, E.J. Houser, S. Stepnowski, R.A. McGill, L. Deel, R.T. Lareau, *Rev. Sci. Instrum.* 75 (11) (2004) 4554–4557.
- [5] P. Dutta, C.A. Tipple, N.V. Lavrik, P.G. Datskos, H. Hofstetter, O. Hofstetter, M.J. Sepaniak, *Anal. Chem.* 75 (2003) 2342–2348.
- [6] N. Backmann, C. Zahnd, F. Huber, A. Bietsch, A. Pluckthun, H.P. Lang, J.-J. Guntherodt, M. Hegner, C. Gerber, *PNAS* 102 (41) (2005) 14587–14592.
- [7] L.R. Senesac, P. Dutta, P.G. Datskos, M.J. Sepaniak, *Anal. Chim. Acta* 558 (2006) 94–101.
- [8] H.-K. Hong, C.H. Kwon, S.-R. Kim, D.H. Yun, Y.K. Lee, K. Sung, *Sens. Actuators B: Chem.* 66 (1) (2000) 49–52.
- [9] M.K. Baller, H.P. Lang, J. Fritz, C. Gerber, J.K. Gimzewski, U. Drechsler, H. Rothuizen, M. Despont, P. Vettiger, F.M. Battiston, J.P. Ramseyer, P. Fornaro, E. Meyer, H.-J. Güntherodt, *Ultramicroscopy* 82 (2000) 1–9.
- [10] T.A. Dickinson, J. White, J.S. Kauer, D.R. Walt, *Nature* 382 (1996) 697–700.
- [11] C. Jutten, J. Hérault, P. Comon, E. Sorouchiary, *Signal Process.* 24 (1991) 1–29.
- [12] P. Comon, *Signal Process.* 36 (3) (1994) 287–314.
- [13] A.J. Bell, T.J. Sejnowski, *Vis. Res.* 37 (23) (1997) 3327–3338.
- [14] S. Makeig, S. Debener, J. Onton, A. Delorme, *Trends Cogn. Sci.* 8 (5) (2004) 204–210.
- [15] A.C. Tang, B.A. Pearlmutter, Invited Chapter in *Magnetic Source Imaging of the Human Brain*, Lawrence Erlbaum and Associated, 2003.
- [16] A.J. Bell, T.J. Sejnowski, *Network* 7 (1996) 261–266.
- [17] H.B. Barlow, in: C. Koch, J.L. Davis (Eds.), *Large-Scale Neuronal Theories of the Brain*, MIT Press, Cambridge, MA, 1994.
- [18] Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- [19] Belouchrani, K. Abed-Meraim, J.F. Cardoso, E. Moulines, *International Conference on Digital Signalling*, Nicosia, Cyprus, July, 1993, pp. 346–351.