Anal. Chem. **2003,** *75,* 4161−4167

Information Coding in Artificial Olfaction Multisensor Arrays

Keith J. Albert† and David R. Walt*

The Max Tishler Laboratory for Organic Chemistry, Department of Chemistry, Tufts University, Medford, Massachusetts 02155

High-density sensor arrays were prepared with microbead vapor sensors to explore and compare the information coded in sensor response profiles following odor stimulus. The coded information in the sensor−odor response profiles, which is used for odor discrimination purposes, was extracted from the microsensor arrays via two different approaches. In the first approach, the responses from individual microsensors were separated (decoded array) and independently processed. In the second approach, response profiles from all microsensors within the entire array, i.e., the sensor ensemble, were combined to create one response per odor stimulus (nondecoded array). Although the amount of response data is markedly reduced in the second approach, the system shows comparable odor discrimination rates for the two signal extraction methods. The ensemble approach streamlines system resources without decreasing system performance. These signal compression approaches may simulate or parallel information coding in the mammalian olfactory system.

Over the last two decades, there has been an increase in the development of odor detection systems that utilize cross-reactive sensor array responses in conjunction with pattern recognition techniques for odor interpretation, discrimination, and quantification.[1−10] These systems contain multiple cross-reactive vapor-sensitive elements in an array format.[10] Upon exposure to vapors, the cross-reactive sensors generate response patterns that are characteristic of the vapors. The response patterns are used to train a pattern recognition program to recognize and identify the vapors upon subsequent exposure. The combination of cross-reactive sensors with pattern recognition schemes is exemplified by the mammalian olfactory system; hence, such systems have been termed artificial, or electronic, noses.

Our laboratory has previously reported an electronic nose array platform[11,15] incorporating hundreds to thousands of sensor elements.[13−18] These arrays employ randomly distributed vapor-sensitive microbead sensors with the solvatochromic dye Nile Red adsorbed[10,15−21] to their surfaces. Upon vapor exposure, the sensors exhibit fluorescence changes that are characteristic of the vapor. There are advantages to employing high-density randomized arrays[11] such as their simple fabrication, signal-to-noise enhancements through signal averaging,[13,15] and the ability to incorporate new sensors as they are developed without increasing costs or system resources.[18] The burden of fabricating identical sensing elements is also alleviated because billions of microbead sensors are fabricated at once, i.e., 1 mL of a bead suspension contains $\sim 6 \times 10^9$ microspheres with virtually identical response characteristics.[11,13−15] Because the microsphere sensors within an array assembly are randomized, the identity of individual sensors in the array must be determined before the array can be used for an odor detection task; i.e., the sensors in the array must be *decoded* before the individual sensor−odor response profiles can be extracted. Multiple decoding strategies have been employed to optically identify the individual sensor types within high-density arrays before the array is used for subsequent detection tasks.[11,14,15,18,22] Depending on the complexity of the array and the number of sensor elements within it, decoding may take a long

* To whom correspondence should be addressed. Fax: 617.627.3443. E-mail: david.walt@tufts.edu.

† Presently at Binax, Inc., 217 Read St., Portland, ME 04103.

(1) Persaud, K.; Dodd, G. *Nature* **1982**, *299*, 352−355.
(2) Carey, W. P.; Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1987**, *57*, 1529−1534.
(3) Gardner, J. W.; Bartlett, P. N. *Electronic Noses: Principles and Applications*; Oxford University Press: Oxford, U.K., 1999.
(4) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298−2312.
(5) Stetter, J. R.; Jurs, P. C.; Rose, S. L. *Anal. Chem.* **1986**, *58*, 860−866.
(6) Shaffer, R. E.; Rose-Pehrsson, S. L.; McGill, R. A. *Anal. Chim. Acta* **1999**, *384*, 305−317.
(7) Jurs, P. C.; Bakken, G. A.; McClelland, H. E. *Chem. Rev.* **2000**, *100*, 2649−2678.
(8) Hines, E. L.; Llobet, E.; Gardner, J. W. *IEE Proc.-Circuits Devices Syst.* **1999**, *146*, 297−310.
(9) Gardner, J. W.; Hines, E. L.; Tang, H. C. *Sens. Actuators, B* **1992**, *9*, 9−15.
(10) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R. *Chem. Rev.* **2000**, *100*, 2595−2626.
(11) Michael, K. L.; Taylor, L. C.; Shultz, S. L.; Walt, D. R. *Anal. Chem.* **1998**, *70*, 1242−1248.
(12) Pantano, P.; Walt, D. R. *Chem. Mater.* **1996**, *8*, 2832−2835.
(13) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2000**, *72*, 1947−1955.
(14) Albert, K. J.; Gill, D. S.; Walt, D. R.; Pearce, T. C. *Anal. Bioanal. Chem.* **2002**, *373*, 792−802.
(15) Dickinson, T. A.; Michael, K. M.; Kauer, J. S.; Walt, D. R. *Anal. Chem.* **1999**, *71*, 2192−2198.
(16) Albert, K. J.; Gill, D. S.; Pearce, T. C.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 2501−2508.
(17) Bakken, G. A.; Kauffman, G. W.; Jurs, P. C.; Albert, K. J.; Stitzel, S. E. *Sens. Actuators, B* **2001**, *79*, 1−10.
(18) Stitzel, S. E.; Cowen, L.; Albert, K. J.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 5266−5271.
(19) Albert, K. J.; Myrick, M. L.; Brown, S. B.; James, D. L.; Milanovich, F.; Walt, D. R. *Environ. Sci. Technol.* **2001**, *35*, 3193−3200.
(20) Walt, D. R.; Dickinson, T.; White, J.; Kauer, J.; Johnson, S.; Engelhardt, H.; Sutter, J.; Jurs, P. *Biosens. Bioelectron.* **1998**, *13*, 695−699.
(21) Stitzel, S. E.; Albert, K. J.; Ignatov, S. G.; Walt, D. R. *Proc. SPIE-Int. Opt. Eng.* **2001**, *4575* (Chemical and Biological Early Warning Monitoring for Water, Food, and Ground).

10.1021/ac0264776 CCC: $25.00
Published on Web 07/04/2003

time. In this paper, we report a method to streamline, or even avoid entirely, the requirement for sensor decoding.

This paper focuses on two strategies used to extract sensor−odor response data from high-density arrays and reduce the resources required for the process. For most electronic nose arrays, sensor−odor data extraction is normally a trivial task because they consist of less than a dozen different sensor types and replicate sensors are not present. In addition, some electronic noses collect signal responses only at a specified time point after vapor exposure or wait until the system is at equilibrium. Thus, each sensor in the array provides a single data point. Although many other detection systems can collect transient sensor−odor response data, few, if any, other systems can simultaneously collect and analyze hundreds to thousands of sensor elements. Our optically based system collects temporal responses of megapixel fluorescent images during the entire vapor exposure sequence. The time response profiles from each sensor in the array are collected and stored separately. This collection of extracted sensor−odor response profiles is used to create a *fingerprint* for each odor.[3,7] These data extraction processes work well when the array contains a small number of sensor elements, but as the number of sensors increases, the amount of extracted data can become overwhelming. One goal of electronic nose systems is to discriminate/detect odors in a high-speed and accurate fashion, but the data extraction processes can limit both speed and accuracy of system performance. The use of high-density sensor arrays containing thousands of sensing elements can be a challenging prospect for data extraction processes due to the vast amounts of data produced for each odor exposure. Using our microsensor array platform and custom designed instrumentation, we explore different signal convergence themes to further understand the information coded in the sensor response. These signal convergence themes include the following: (1) a decoded approach, where each sensor type within the array is individually analyzed, and (2) an ensemble approach, where response profiles from *all* sensors are combined and collectively analyzed. The ensemble approach removes the sensor decoding processes; i.e., we do not identify each sensor within the high-density array. This study, which employs two data extraction methods with high-density arrays, is akin to a recent publication from Reich et al. in which they found that individual neurons in the primary visual cortex (V1) contain more response, or spike, information compared to average responses from neuronal clusters.[31] Simply stated, Reich et al. reported that response information was higher when neurons were independently monitored as the summing/averaging process removes neuron-to-neuron differences. By monitoring individual neurons, however, many more (and potentially costly) resources devoted to data collection are required. In a similar fashion, we explore information coded in the individual sensor responses in our optical arrays compared to the information coded in responses produced from a sensor ensemble.

EXPERIMENTAL SECTION

Materials. Acetone, 1-pentanol, 3-pentanol, 1-butanol, benzene, *p*-xylene, toluene, and 22 mm × 30 mm glass coverslips were purchased from Fisher Scientific (Fairlawn, NJ). The 90 ppb 1,3-dinitrobenzene aqueous sample was obtained from the Defense Advanced Research Projects Agency. Nile Red dye, 1,3-dinitrobenzene, and remaining solvents (HPLC grade or better) were purchased from Aldrich (Milwaukee, WI). All chemicals were used as received. The carrier gas used in the experiments is ultra zero grade air from Northeast Gas, Inc. (Salem, NH). The different types of porous silica microsphere packing materials were removed from HPLC columns from Phenomenex (Torrance, CA), Alltech (Deerfield, IL), or Keystone Scientific, Inc. (Bellefonte, PA) and employed as substrates to which the fluorescent dye was physisorbed.

Sensor and Array Fabrication. All sensors and coverslip arrays were prepared as described previously.[13] Briefly, each stock of porous silica microspheres was rinsed with ethanol and allowed to dry in room air overnight before being stained with a solution of 0.5 mg/mL Nile Red in toluene. Approximately 50 mg of microsphere material was placed on a vacuum filtration system, rinsed with toluene, and then excess Nile Red solution was passed over the beads. Because many of the microbead stocks were purchased from different manufacturers, each microbead stock has different physical/material characteristics. For instance, two microbead stocks with the same surface functionality (i.e., amino or hydroxyl groups) may differentially adsorb Nile Red dye or interact differently upon odor exposure because they were manufactured by different companies and have different characteristics (pore size, amount of surface coating, etc.). Once prepared, sensors were stored in individually labeled and capped 4-mL glass vials in the dark at room temperature until use. All vapor tests employed an array of sensors electrostatically held to the surface of a glass coverslip. All sensor stocks were fabricated many months prior to the testing events (see Supporting Information). The coverslip arrays were prepared by depositing a small amount of microsensors onto a glass coverslip and smearing the sensors with a latex glove. All microsensors stocks and array compositions are shown in the Supporting Information. Coverslip arrays were prepared within one or two days prior to the testing events.

Imaging and Vapor Delivery System. The fabricated sensor arrays were positioned onto a custom-built imaging and detection system as described previously.[23] The system employs an inverted Olympus fluorescence microscope incorporating a 75-W Xe excitation source with detection by a 640 × 480 pixel SensiCam high-performance CCD camera (Cooke Corp., Auburn Hills, MI). In all testing events, Zeiss microscope objectives were employed and the CCD chip was binned 2 × 2. All analytes were individually placed in sealed flasks for vapor delivery by a vacuum-controlled sparging apparatus.[23] For "air" (blank) delivery, an empty sealed flask was employed in order to purge the carrier (air) gas onto the sensors. It should be noted that vapor delivery was based on dilution of the vapor stream and not by employing calibrated vapor generators. Prior to delivery of analyte vapor onto a sensor, a diluent air stream was used to lower vapor concentrations below saturation. Analyte vapors were not preconcentrated prior to sensor array exposure. The odorants were selected because they represented simple and complex discrimination tasks and were used because they were available in our laboratory at the time of testing. The bacteria types employed in this study were made available from the Biology Department and were used to represent

(22) Steemers, F. J.; Ferguson, J. A.; Walt, D. R. *Nat. Biotechnol.* **2000**, *18*, 91−94.
(23) White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1996**, *68*, 2191−2202.

a difficult discrimination task. The types and concentrations of odorant components in each bacterium's headspace were not relevant to this study's focus.

Three-Bead (03-Bead) Randomized Array Used To Discriminate Bacteria Headspace Odors from Medium (Control). The randomized array was monitored before, during, and after a 2.38-s exposure to each of the four odor conditions (water, medium, *Escherichia coli*, and *Klebsiella pneumoniae*). Four replicate exposures were recorded for each odor for a total of 16 observations (4 odors × 4 replicates). CCD exposure was 100 ms/frame, and each image sequence consisted of 60 image frames (10 before, 20 after, and 30 after odor exposure) for a total duration of 7.13 s. Excitation and emission filters were 530 (band-pass 30) nm and 630 (band-pass 20) nm, respectively. Flow rates were set to 200 mL/min, and the headspace odor was not diluted with carrier air. Between each observation, a 30-s purge time was allotted to clear the vapor delivery system of possible carryover. The response profiles for the randomized array were extracted via the two methods discussed above: (1) a decoded approach, where the three sensor types were independently addressed, and (2) the ensemble, nondecoded approach, where responses from the three sensor types were combined.

Six-Bead (06-Bead) Randomized Array Used To Discriminate Live and Dead Bacteria Headspace Odors as well as Medium (Control). The randomly distributed sensors were monitored before, during, and after exposure to the 10 bacteria odor conditions (× 5 replicates each) and medium (10 replicates). Each movie sequence consisted of 50 image frames (5 images before, 20 during, and 25 after a 2.41-s odor exposure) for a total of 6.04 s. The excitation filter was set to 560 (band-passs 40) nm, and a CRI tunable emission filter (Cambridge Research & Instrumentation, Inc., Cambridge, MA) was positioned to 600 nm. CCD exposure was 100 ms/frame, and saturated odor flow rates were 200 mL/min. The sensor arrays were decoded with a previously reported approach.[14,21] Bacteria cultures were autoclaved to produce the dead bacteria samples. The medium (or control) was also autoclaved prior to sample delivery for 5 of the 10 samples. As with the 03-bead array discussed above, the extracted response profiles for the six-bead array were analyzed with the (1) decoded approach and the (2) ensemble approach.

Experiments I and II. Array Fabrication and Testing Protocols for 20 Different Odor Conditions. Although experiments I and II both employed the same 18 prepared sensor stocks, the order in which the sensors were combined was completely different. Using Microsoft Excel, 18 prepared sensor stocks were assigned a random number through a random number generator. The random numbers were then placed in sequential order, which signified the order in which sensors were randomly combined to fabricate multiple single-sensor (homogeneous) and randomized arrays. As defined by Supporting Information, each coverslip array consisted of 1−18 randomly distributed microbead sensor types. A total of nine high-density arrays were fabricated: three homogeneous (01-bead) arrays and six multiplexed randomized arrays (03-bead, 06-bead, 09-bead, 12-bead, 15-bead, and 18-bead array). The three sensor types used for the 01-bead arrays were combined to fabricate the 03-bead randomized array. Experiments I and II were performed more than three months apart (7 June 2001 and 18 September 2001). Each acquired movie sequence consisted of 45 images with 5 images captured before analyte delivery, 15 images captured during delivery, and 25 captured post odor delivery. The total duration of each image sequence was 5.4 s with a 1.8-s odor exposure period. The CCD exposure time was 100 ms, and odors were delivered at 50% saturation with a total flow rate of 200 mL/min. The excitation and emission filters were 530 (band-pass 30) nm and 630 (band-pass 20) nm. Each coverslip in these two experiments was exposed to 100 odor observations (5 repeats × 20 odors) and the arrays were not decoded.

Signal Preprocessing and Random Array Decoding. The preprocessing method employed on the extracted raw data is referred to as normalization. Briefly, normalizing the raw data is a simple calculation to make all sensor response profiles equally weighted. To normalize data, the entire response profile was divided by the first time point and 1 was subsequently subtracted from all time points. The first starting point therefore started at zero, and the profile's absolute maximum was scaled to 1 (or −1 for an absolute minimum). For the randomized arrays, odor response data, emission scan responses, or both were collected to positionally register the sensors within the array as previously discussed.[14,15,21]

Image Data Collection, Data Extraction via Region-of-Interest (ROI) Templates, and Data Analysis. After raw response profiles were extracted from the image sequences, and the data were normalized. The preprocessed data were subjected to singular value decomposition in MATLAB software. This technique was used as the standard technique to reduce the dimensionality of the data set so that data could be visualized directly in principle components (PC) space. The data were then used to assess the classification accuracy via a MATLAB analysis code written by Dr. Gregory Bakken (Penn State University).

RESULTS AND DISCUSSION

The vapor-sensitive microsensor arrays in this work were prepared by attaching Nile Red to the surfaces or interiors of chemically diverse 3- and 5-μm bead stocks.[13,16] Upon vapor exposure, Nile Red's spectral properties change and these changes are unique for each microsensor type. In the work described here, each high-density array was prepared containing between 1 and 18 different microsensor stocks. Because each microsphere type gives a unique and characteristic response to a particular vapor, it is relatively easy to identify the position of each sensor type when multiple sensor types are distributed within a randomized array.[14,15] We refer to this process as decoding the array. The unique fluorescence response profiles occurring before, during, and after odor stimulus for each sensor−odor combination are recorded with a custom-designed instrument,[23] which consists of a fluorescence microscope, CCD detector, 75-W xenon arc lamp excitation source, excitation and emission filter wheels, sparging vapor delivery system, and imaging software. The CCD camera is used to record three-dimensional video (time, sensor position, fluorescence intensity) for each odor stimulus event, and image processing software is then used to extract the sensor−odor response data from the movies using ROI templates. Briefly, these ROI templates allow us to independently extract response data for every microsensor within the array. With multiple sensor types and many replicates of each type in the array, there are enormous amounts of sensor−odor response data that potentially could be used for odor discrimination purposes.

As with conventional data extraction for electronic noses, each sensor−odor response is analyzed separately. For our system, we extract response data with one ROI per sensor and the number of extracted response profiles depends on the numbers of sensor types, sensor replicates, odor exposures, and exposure replicates.[13−18] Because there may be thousands of sensors in the array, the extracted data files can be very large and the data extraction process can take a long time. For instance, the total number of response profiles consisting of sensors (s) and odor exposures (o) is $(sr_1 + sr_2 + sr_3 + ...+ sr_n) \times (or_1 + or_2 + or_3 + ...+ or_n)$, where sr_1 refers to the number of replicates of sensor type 1, sr_2 refers to the number of replicates of sensor type 2, or_1 refers to the number of replicates of odor 1, etc. In a recent study, high-density multibead arrays were used for simple and complex odor discrimination[16] and the study consisted of 2683 total sensors and 65 total odor exposures. In all, 174 355 sensor−odor response profiles were recorded and each response profile consisted of 45 time points, creating a total of $\sim 71.848 \times 10^6$ data points, which had to be extracted from the image sequences. Even when response profiles for replicate sensors were combined and averaged to reduce data dimensions and enhance signal-to-noise levels, there were still 17 550 data points (6 sensor types × 1 averaged response profile × 65 odor exposures × 45 time points), which had to be processed further for subsequent odor discrimination. It is important to recognize that the above study employed multiple data reduction and analysis techniques, i.e., common preprocessing metrics, discriminant function analysis, and multivariate analysis of variance.[16] These techniques were not the focus of this study, and they were not used. Moreover, the present study did not focus on the lower levels of analyte detection or sensitivity measurements, as they have both been reported and discussed in more detail elsewhere by our group.[13,15] The focus of this paper, is to (blindly) compare results from two different data extraction techniques, which both employed the same fluorescent CCD-based image movies and the same unsupervised pattern recognition software. Table 1 shows the relative time required to extract *x* ROI response profiles from an acquired image sequence with 45 images. This table also shows the computer memory required to save each data file for the extraction process.

As seen in the table, the relative amount of time and memory required to extract sensor response profiles and save the data files for only 2000 sensor elements, i.e., 2000 ROIs, *per odor observation*, is 91 s and 1.11 MB, respectively. These values increase dramatically with a larger number of odor observations. For example, 2000 sensors exposed to 100 odor observations requires 2.53 h of data extraction time and 111 MB of memory.

In an effort to streamline the data extraction process, we explored methods to combine the sensor−odor responses for subpopulations of sensor elements and for all sensors contained in a high-density sensor array via different ROI templates. ROI templates can be used either to extract every individual microsensor's response (as with the example above) or to simultaneously extract data from a sensor ensemble (thousands of sensors). We compared the array's ability to discriminate odors when the individual sensor responses were individually extracted (decoded array) to an array where the ensemble sensor responses were collected (nondecoded array) (Figure 1). We compare the ability of the sensors to discriminate between various odors when the signals from sensors were *combined* and processed to the ability of the sensors to discriminate odors when data from the individual sensors were kept *separated* and processed. It is important to note that this study used unsupervised pattern recognition, i.e., principal component analysis, and the sensors and odors were not preselected.

Table 1. Relative Time and Memory Required To Extract Raw Response Profiles from Image Sequences by Varying ROI Numbers[a]

total no. of ROIs	data extraction time (s)	raw data file memory (*K*)
1	3	4
8	4	8
100	14	60
500	57	292
1000	65	584
2000	91	1110

[a] Data are presented for *each* odor observation and do not include time required to open image sequences or draw ROI templates.

For the first odor recognition task, we exposed the sensor array to the headspace odors of various bacterial cultures in sealed Erlenmeyer flasks. Three different sensor types (referred to as a 03-bead array) were employed to fabricate the randomized array containing thousands of replicates of each microsphere type. The 03-bead array was subjected to a 2.4-s odor stimulus from two bacterial headspace odors and two controls (medium, water). Response data for 16 odor exposures (4 odors × 4 replicates) were extracted via both the decoded and nondecoded array approach. In the first approach, response profiles from each individual microsphere in the array were independently extracted; i.e., the sensors in the random array were positionally identified/decoded, and response profiles for all replicate sensors were averaged. The three independent response profiles were then concatenated for each odor observation. These data were compared to a reduced response profile for the nondecoded array, where *all* sensor responses were combined into one response profile, to produce only one-third as many data points per observation (3 profiles vs 1 profile). Response data for both signal convergence methods for a different 03-bead array are shown in Figure 1. In this experiment, the three individual microspheres gave 100% odor discrimination while the combined ensemble signal from >2000 microspheres gave a 94% discrimination rate (Table 2, top). It should be noted that even though the signals from replicate sensors were combined, the data extraction process employed to obtain the individual responses from each sensor in the array is time-consuming and memory intensive.

In a similar but more complex experiment, a six-bead array was used to discriminate 11 odors (5 *live* bacteria, 5 *dead* bacteria, and medium). Each 2.4-s odor stimulus was repeated five times, except the control (medium), which had 10 replicates (60 total odor observations). Table 2 lists the different bacterial species from which the headspace odors were purged onto the six-bead array. The separate sensor signals from the decoded six-bead array produced an 85% classification rate as compared to 80% for the combined sensors signals (Table 2, bottom). The reduced performance is expected for the ensemble, but the performance for the ensemble response is still quite good and compares well with the separately processed sensor responses. See the Support-

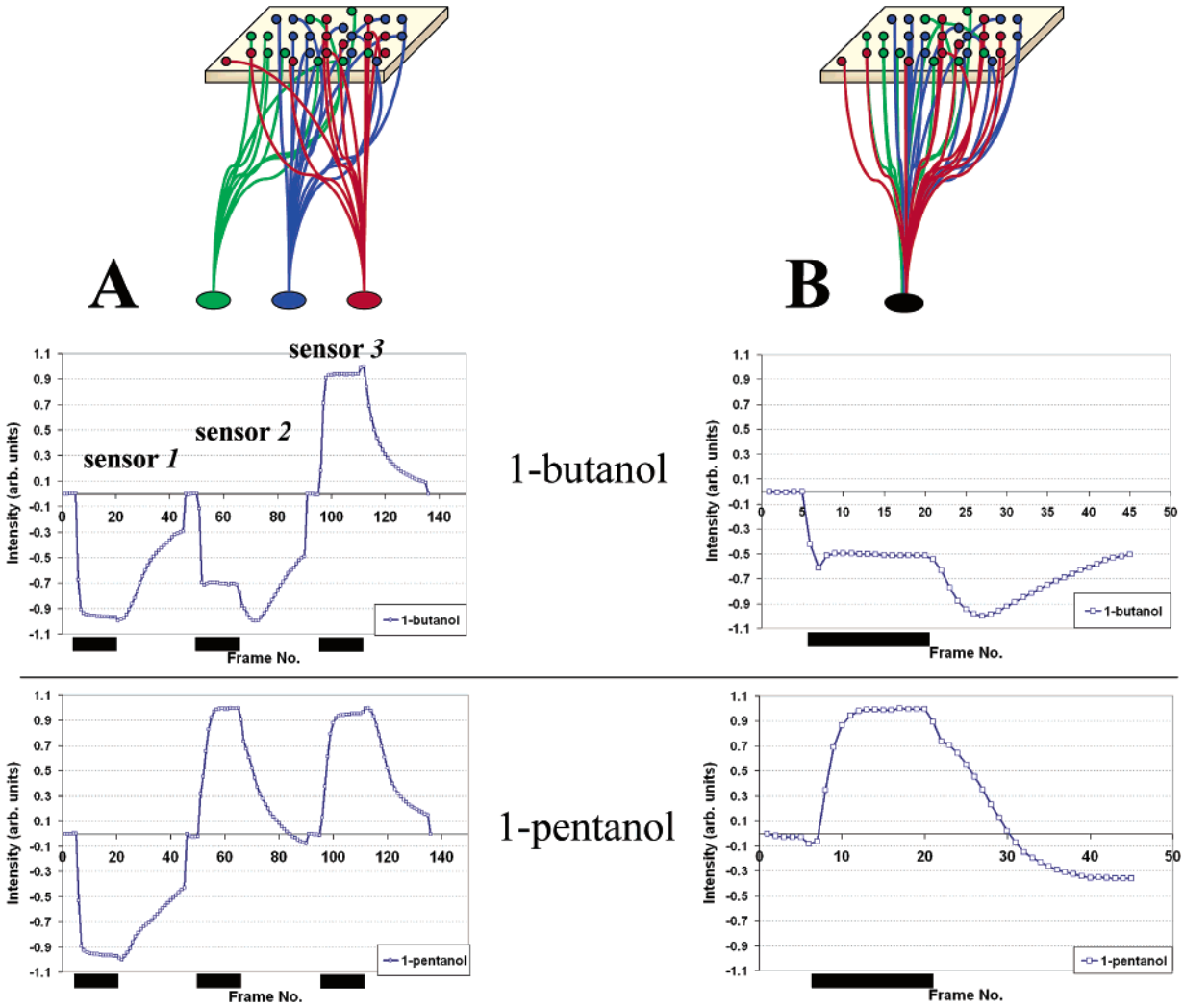

Figure 1. Response profiles. Response data are reduced when the response profiles from three different sensor types (1−3) (A) are combined during signal processing (B). Response profiles from sensor replicates were first averaged for each of the three sensor types and then concatenated (decoded array) (A). Response profiles from all three sensor types, regardless of the number of replicates for each, were averaged (B) to reduce the total amount of information (nondecoded array). Response profiles for 1-butanol and 1-pentanol exposures are shown. *Y*- and *X*-axes are intensity (arbitrary units) and frame number, respectively. All black bars indicate the start and end of the 1.8-s odor stimulus.

ing Information section for sample PC plots for the live/dead bacteria discrimination with the 06-bead array, with separate plots for the decoded and nondecoded arrays.

In addition to saving time, the ensemble summing provides an intrinsic signal-to-noise advantage. Using the 06-bead array as a model, we compared the signal-to-noise differences for one odor exposure's response profile generated from the ensemble approach to a response profile generated from the decoded approach. The decoded array response was generated by combining one response profile from each of the six different sensor types, i.e., responses from one replicate of each of the six different sensor types (see Supporting Information for the plot of the two response profiles). The decoded response profile has more noise compared to the response generated via the ensemble approach. As shown in previous studies by this group, signal-to-noise enhancements are achieved via multisensor summing and signal averaging schemes,[13,15] which is also achieved by the ensemble approach described here.

To test the limits of the system, we performed additional experiments in which we progressively added more sensor types (up to 18) to the array (Figure 2). In contrast to the previous experiment, we extracted only the combined signals from the

Table 2. Odor Discrimination Accuracy for Decoded vs Nondecoded Arrays[a]

	classification rate (%)	no. of PCs	no. of mistakes
3-Bead Array and 4 Odors			
decoded array	100	5	0
nondecoded array[a]	94	3	1
6-Bead Array and 11 Odors			
decoded array	85	15	9
nondecoded array[a]	80	5	12

[a] The odors for the three-bead array consisted of two bacteria headspace (*E. coli* and *K. pneumoniae*) and two controls (water and medium). The odors for the six-bead array consisted of five *live* bacteria types in agar medium (*Micrococcus luteus*, *E. coli*, *Salmonella typhimurium*, *K. pneumoniae*, and *Acinetobacter radioresistens*), the same five bacteria species after being autoclaved, and medium (sampled before and after being autoclaved). PCs, principal components. [b] The response profiles for the nondecoded arrays were extracted via the 1-ROI/array method.

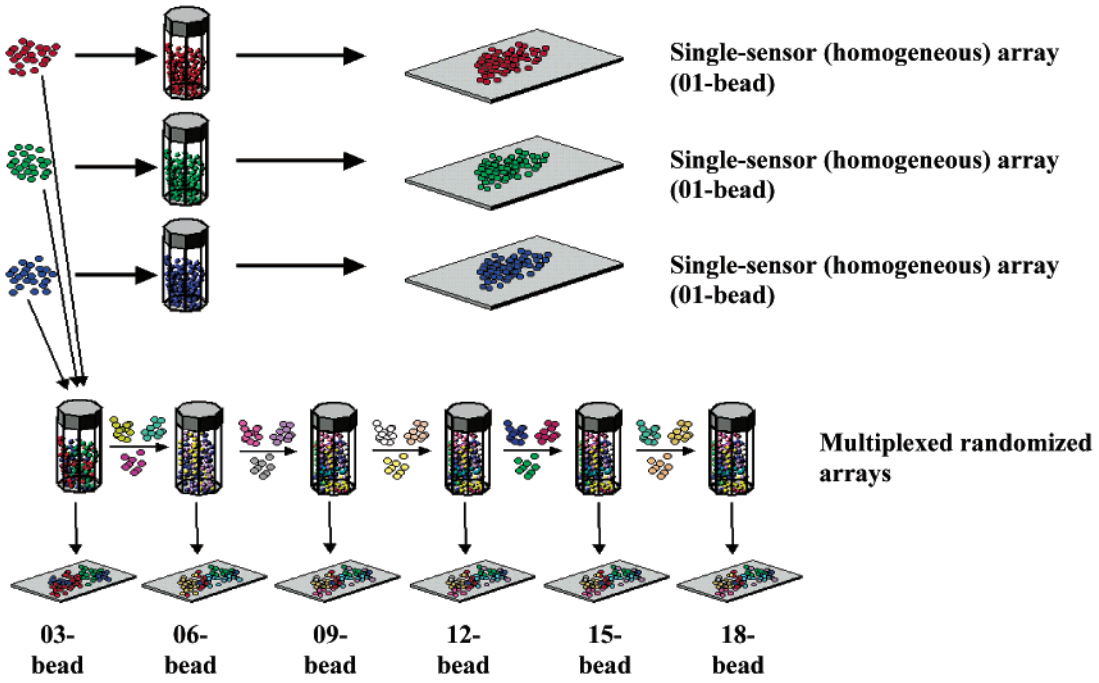

Figure 2. Array fabrication protocol for experiments I and II. Eighteen different sensor stocks were employed to prepare the nine high-density microbead arrays for each of the two experiments.

microsensor ensemble and did not compare the performance to the decoded arrays. The sensor arrays were exposed to a 1.8-s odor stimulus for 100 observations (20 different odor conditions × 5 replicates). The odor set consisted of various volatile organic compounds and binary odor mixtures as listed in Table 3. To ensure ourselves that we were obtaining statistically significant information, different combinations of microspheres were combined into two different sensor array sets (experiments I and II). Both array sets in experiments I and II consisted of nine different array platforms. Arrays were prepared with 1 (three arrays), 3, 6, 9, 12, 15, and 18 microsphere types (2 experiments × 9 arrays = 18 total arrays). Figure 2 shows the detailed fabrication method for one of the array sets. Although experiments I and II employed the same 18 sensor stocks, the 2 array sets were fabricated and tested three months apart and the order of microsensor stock additions for each set was completely different. The 18-bead arrays in experiments I and II were the only arrays with the same set of beads; the other arrays had only some overlap in sensor identity. For example, the 03-bead array in experiment I and the 03-bead array in experiment II were different; i.e., both arrays were composed of three different microsensor types. Likewise, the 06-, 09-, 12-, and 15-bead arrays in experiment I were fabricated with a different combination of sensor stocks compared to the 06-, 09-, 12-, and 15-bead arrays in experiment II. See Supporting Information for a detailed list of array composition, microsensor types employed, and two PC plots showing the 18-bead array results for experiments I and II.

As more sensors were added to the arrays in experiment I, the discrimination performance improved (Table 3, left). It is interesting to note that a single sensor performed better than the combined array until the diversity reached six bead types. In experiment II, where different single-bead sensors were employed, the performance was quite good throughout the testing with even a single-bead sensor achieving excellent odor classification (Table 3, right). These experiments demonstrate that, in some odor recognition problems, one sensor may dominate the response and therefore may be adequate for performing the classification. When poorly discriminating sensors are used, summing multiple sensors helps provide better classification. In some cases, a single-sensor array outperformed multisensor arrays for specific odor discrimination tasks. The fact that any given sensor may be better or worse for a given discrimination task argues for the use of multisensor arrays. These analysis methods were extended to different sensor

Table 3. Odor Discrimination Accuracy for 100 Odor Exposures When All Sensor Responses Are Combined (Nondecoded Arrays)[a]

	classification rate (%) [no. of PCs]	
	experiment I	experiment II
single-sensor (01-bead) array	74 [11]	74 [7]
single-sensor (01-bead) array	86 [7]	98 [7]
single-sensor (01-bead) array	76 [9]	94 [6]
03-bead random array	80 [8]	98 [6]
06-bead random array	86 [13]	97 [7]
09-bead random array	93 [7]	94 [8]
12-bead random array	95 [7]	94 [8]
15-bead random array	85 [8]	98 [7]
18-bead random array	97 [6]	96 [6]

[a] All arrays in experiments I and II employed different microsensor types, even for the 01-bead arrays. The only arrays with the same sensor composition for I and II were the 18-bead arrays (see text). Twenty different odor exposures (× 5 replicates each): (1) air carrier gas, (2) acetone, (3) *n*-heptane, (4) ethanol, (5) toluene, (6) water, (7) ethanol/heptane mixture 1:1 (v/v), (8) methanol/ethanol mixture 1:1 (v/v), (9) benzene, (10) 1-propanol, (11) aqueous 90 ppb 1,3-dinitrobenzene, (12) 1,3-dinitrobenzene (s), (13) methanol/1-propanol mixture 1:2 (v/v), (14) methanol, (15) 1-butanol, (16) 3-pentanol, (17) *p*-xylene, (18) ethanol/1-pentanol mixture 1:3 (v/v), (19) cyclohexanone, and (20) 1-pentanol. Since there were 100 observations, the number of misclassifications is apparent from the classification rate (97% = 3 mistakes; 86% = 14 mistakes).

arrays and different odor detection tasks showing the versatility and utility of our platform in solving multiple (simple and complex) problems with the same analysis approach.

By combining the responses from multiple different sensors into a single response profile, this simplified data extraction process does not result in degradation in system performance. We believe that the decoded sensor arrays perform slightly better than the nondecoded arrays because of "signal canceling", which likely removes, or decreases, sensor response features that would ordinarily add to an array's ability to discriminate (refer to Figure 1, where some response features are washed-out when the sensor responses are combined). Sensor arrays with orthogonal responses/sensors add to an array's discrimination ability. While it is true that adding nonorthogonal sensors decreases discrimination, we believe that an 18-sensor array provides sufficient diversity to solve a multitude of problems as demonstrated by the results. If only three sensors were selected and employed, the results were always worse. This approach is one way to obtain diversity and to compress resources required to perform analysis and data collection. Although the CCD-based detection system employed is relatively expensive and somewhat large compared to other systems used for odor discrimination, it has many inherent advantages. Along with the advantages discussed in the introduction, this system also allows us to rapidly explore (streamlined) data extraction methods that may benefit high-throughput screening. For example, by extracting the collective sensor ensemble response for the 18 sensor types in experiments I and II instead of extracting the individual microsensor responses, more than 1990 MB of computer memory and 44 h of data extraction time were avoided. This streamlined data extraction process optimized system resources without significantly diminishing system performance. We have demonstrated that our system can employ simple data collection strategies to conserve time, labor, and resources *without compromising system performance*. Overhead costs are therefore reduced because the entire system can be streamlined. The methods may open the door to other simple, high-throughput analysis techniques to streamline other tasks where information is needed immediately, such as risk assessment in a biological/chemical warfare exposure.

The results reported here demonstrate that combining sensor responses maintains information content while significantly reducing the amount of data that needs to be collected. The data extraction method proved to be computationally simple, yet was powerful enough to maintain performance even when response information was reduced dramatically. One concern is that we worked with a relatively small set of odorants, and if we had to discriminate between many more odorants, the advantages of working in the higher-dimensional space would become more apparent compared to summing sensor responses. Even with the simplified problem, however, we have demonstrated a significant improvement in performance for an artificial nose microsensor array system, which utilized multiple sensor types and a simple data extraction process to successfully demonstrate odor discrimination.

The use of efficient data extraction schemes may also relate to biological sensing systems. Our optical array system's architecture parallels some aspects of the mammalian olfactory system and allows us to objectively explore information coded in the sensor−odor response profiles. For instance, for each sensor type present in the array, there are hundreds to thousands of replicate sensors, which parallels the replicates[24−26] and randomly positioned[26,27] olfactory cells in the mammalian olfactory system. Responses from these olfactory receptor cells, of which there are many replicates, converge onto glomeruli.[28,29] While the actual processes of signal convergence are not completely understood, mammals are known to have rather large dynamic ranges and are capable of low-level odor perception because the receptor responses are summed to enhance odor recognition.[3,30]

Reich et al. recently reported that individual neurons in the primary visual cortex (V1) contain more response, or spike, information compared to average responses from neuronal clusters.[31] They showed that, for certain types of stimuli, individual neurons provided higher information rates compared to neuron cluster averages because some of the individual neurons produced unique spike patterns. The signal convergence and averaging process, however, removes neuron-to-neuron differences. Information in neuronal spikes diminishes as spike trains are grouped, or summed, from clusters of neurons;[31] i.e., by averaging across neurons in a population, potentially useful information is discarded because averaging spike patterns across neurons underestimates the total available information.[32] On the other hand, by monitoring individual neurons, many more (and potentially costly) resources must be devoted to data collection. These results may complement Reich et al.'s findings because we monitored and quantified the degradation in array performance when response profiles from different types of sensors were combined. Although the microbead sensors have quite different odor response profiles compared to neurons, the results reported in this paper suggest that signal-summing techniques improve the efficiency of processing without compromising the fidelity of classification.

ACKNOWLEDGMENT

The work was supported by the Defense Advanced Research Projects Agency. A special thanks goes to Dr. Sergei Ignatov (State Research Center for Applied Microbiology, Moscow Region, Russia) for bacteria sample preparations, Dr. Gregory Bakken (Pennsyvania State University) for developing the confusion matrix classification program for odor discrimination, and Shannon Stitzel (Tufts University) for helping to both collect odor observations and decode the three-bead and six-bead arrays.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in the text (two tables and three figures). This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review December 31, 2002. Accepted May 27, 2003.

AC0264776

(24) Axel, R. *Sci. Am.* **1995**, *273*, 154.
(25) Buck, L. B. *Annu. Rev. Neurosci.* **1996**, *19*, 517−544.
(26) Buck, L.; Axel, R. *Cell* **1991**, *65*, 175−187.
(27) Vassar, R.; Ngai, J.; Axel, R. *Cell* **1993**, *74*, 309−318.
(28) Ressler, K. J.; Sullivan, S. L.; Buck, L. B. *Cell* **1994**, *79*, 1245−1255.
(29) Mori, K.; Nagao, H.; Yoshihara, Y. *Science* **1999**, *286*, 712−715.
(30) Buck, L. B. *Cell* **2000**, *100*, 611−618.
(31) Reich, D. S.; Mechler, F.; Victor, J. D. *Science* **2001**, *294*, 2566−2568.
(32) Richmond, B. *Science* **2001**, *294*, 2493−2494.