

Automatic decoding of sensor types within randomly-ordered, high-density optical sensor arrays

Keith J. Albert^a, Daljeet S. Gill^b, Tim C. Pearce^{b,*}, and David R. Walt^{a,*}

^aThe Max Tishler Laboratory for Organic Chemistry, Department of Chemistry, Tufts University, Medford, MA 02155 *Corresponding author for sensor work. Fax: 617.627.3443; Email: david.walt@tufts.edu

^bDepartment of Engineering, University of Leicester, Leicester LE1 7RH, United Kingdom *Corresponding author for computational work. Tele: +44(0)116.223.1290 Fax: +44(0)116.252.2619; Email: t.c.pearce@leicester.ac.uk

(authors are listed alphabetically to imply equal credit)

ABSTRACT

In this paper, automatic sensor identification of sensor classes within a high-density randomized array is demonstrated without knowing sensor locations *a priori*. Two different fluorescence-based sensor types, with hundreds of replicates each, were randomly distributed into an optical imaging fiber array platform. The sensor element types are vapor-sensitive microspheres, which have the environmentally-sensitive fluorescent dye Nile Red adsorbed onto their surface. Nile Red undergoes spectral changes when exposed to different microenvironmental polarity conditions, i.e., microsphere surface polarity or odor exposure. These reproducible sensor spectral changes, or sensor response profiles, allow sensors within a randomized array to be grouped into categories via optical decoding methods. Two computational decoding methods (supervised and unsupervised) are introduced and both demonstrated equal classification rates. By comparing sensor responses from a randomized array to those obtained from known (control) arrays, 587 sensors were correctly classified with 99.32% accuracy. Though both methods were equally effective, the unsupervised method, which uses the sensor response changes to odor exposure, is a better decoding model for the vapor-sensitive arrays studied since it relies only on the odor response profiles. Another decoding

technique employed the sensors' emission spectra and is more applicable to other types of multiplexed fluorescence-based arrays and assays. The sensor-decoding techniques are compared to demonstrate that sensors within high-density optical chemosensor arrays can be positionally-registered, or decoded, with no additional overhead in time or expense other than collecting the sensor response profiles.

Keywords

Optical decoding, optical imaging fibers, high-density sensor arrays, fluorescence, microscopy, imaging, cross-reactive vapor sensors, PCA, DFA.

Abbreviations

Principal components analysis (PCA); discriminant function analysis (DFA); charge coupled device (CCD); 1,3-dinitrobenzene (DNB); 2,4-dinitrotoluene (DNT); standard deviation (SD).

INTRODUCTION

Over the last few years, the development of high-density optical chemosensor arrays has become an attractive sensing architecture for many chemical detection tasks [1-11]. These fluorescence-based arrays are assembled by first preparing sensing microspheres (beads) in which a fluorescent indicator is attached to polymeric or porous silica beads. The arrays used for this study employed two different types of porous silica microspheres, which were both removed from HPLC columns and stained with the solvatochromic dye Nile Red [12-15]. Microsphere surface modifications, i.e., amino groups or alkyl chains, alter the dye's excitation and emission spectra, but they are also likely to have an effect on the microsphere's ability to attract and

adsorb vapor molecules, much like thin polymer coatings for surface acoustic wave sensors [16]. When a microsphere sensor array is exposed to a particular vapor, the vapor affects each microsphere type differently. For example, a given vapor may strongly adsorb or partition into or onto some microsphere types but may not interact with other types. The vapor concentration and the vapor's polarity can be monitored by Nile Red via time dependent spectral changes (excitation/emission shifts and/or intensity change), which form unique patterns for each odor [1,17,18]. These fluorescence-based microbeads have proven to be effective sensors for detecting low-level nitroaromatic compound vapors [1,2,9,19,20] and discriminating simple and complex odors [1,21]. As reported herein, these unique and reproducible sensor-odor patterns can be used to identify, or decode, different sensor types when they are combined and randomly distributed on a high-density array platform. This paper does not elaborate on odor discrimination, which has previously been documented for similar sensor types [1,2,4,9,19-22].

High-density arrays prepared with these beads have inherent advantages because they can incorporate hundreds to thousands of individual sensors while maintaining the ability to individually address each one. For example, with the large number of sensor elements and sensor replicates in each array, these platforms allow for detection limit enhancements through multi-sensor signal averaging [2,4]. High-density sensor arrays on optical imaging fiber platforms may be prepared with two different fabrication motifs: (a) single-sensor, homogeneous arrays are fabricated exclusively with one sensor type, i.e., all elements within the array are identical (Figure 1a and 1b); and (b) multiplexed random arrays, where multiple sensor types are combined and distributed on the platform in a randomized fashion [6] (Figure 1c). To prepare an optical fiber array, etched wells [7] on the optical imaging fiber's distal face are filled randomly with the microsphere sensors so that each well contains a complementary-sized sensor element

[6]. These array platforms can consist of 3,000-100,000 individual microbeads but the burden of fabricating identical sensors is alleviated because billions of beads are fabricated at once, i.e., one milliliter of a bead suspension contains $\sim 6 \times 10^9$ microspheres [2,4,6]. Therefore, thousands of reproducible sensors can be deposited in an addressable format within seconds. Each fiber bundle is then positioned onto a microscope-based imaging system with a CCD detector and vacuum-controlled sparging apparatus odor delivery unit [18].

<INSERT FIGURE 1>

The single-sensor homogeneous arrays are effective detection platforms, but to create an array with multiple sensor types, many arrays must be combined, bundled and aligned [1,20]. Depending on the number of arrays one wishes to combine, complications may arise. For example, not all sensor elements in the bundle may be in the field of view and, as one lowers microscope magnification to incorporate more elements, CCD pixel resolution and fluorescence signals decrease. As a promising alternative, one can combine sensors into *a single* array platform (Figure 1c). These arrays are termed *randomized* because the sensor elements are randomly-distributed and not assigned to pre-determined sensor locations. The random array approach is gaining popularity because of its simplicity and effectiveness. Randomized arrays can be fabricated in two simple steps: sensor element preparation and sensor dispersion onto the array platform. Unlike some current array platforms, our randomized arrays do not require micro-robotic manipulation, photolithography, or chemical deposition/treatment for assembly. Due to the array’s randomized nature, as new sensors are developed, they can be readily incorporated into the array with little additional effort (or cost), thus increasing the types or kinds of detection tasks one can perform.

Once a randomized array is fabricated, the different types of sensor elements must be positionally-registered before the array can be employed for sensing tasks. Multiplexed randomized arrays have been used for odor discrimination [9], immunosensing [10], and DNA hybridization [5,23]. In each example, sensor elements within the randomized array were optically decoded via different processes before the detection task was carried out. The spatially and spectrally resolvable sensor elements within these arrays are mapped by using each sensor’s unique properties or inherent optical characteristics. For instance, sensors within a multiplexed-randomized DNA-microchip array can be decoded by successive hybridization steps because each sensor type within the array has a different DNA target. By simply labeling complementary target probe(s) with fluorophore(s), the microsensor locations within the high-density array can be mapped after hybridization [5,8]. Because of their potential applications to high-throughput screening, both DNA and “artificial nose” randomized arrays must be easy to implement. Simply stated, if the different sensor elements within the random arrays cannot be decoded, the array is of limited or no value. Optical decoding schemes are therefore vital to the success and implementation of high-density random arrays. Recent studies have documented, or made reference to, optical decoding processes. These processes include quantum dots [24], fluorescent colloids [25,26], metallic barcodes [27], flow cytometry [28,29], protein- and cell-based assays [11,30], multiplexed assays [3,10,31], and spatial and spectral imaging [32-34], further signifying the importance of optical decoding techniques for a broad range of detection applications.

This paper describes efforts aimed at systematically decoding sensor types and positions within a randomized high-density array by comparing their response patterns to sensors in homogeneous (control) arrays. Each microsphere sensor belongs to a discrete class and by use of

optical decoding processes, these class identifications can be ascertained [4-6,8,9]. *Decoding hundreds of identical sensor elements would be impossible if the sensors did not behave in a reproducible fashion.* Sensor-to-sensor and array-to-array reproducibility for high-density arrays prepared with one microsensor type were recently demonstrated [2,4]. These studies showed consistency in the response pattern to a selected odor exposure from separate homogeneous arrays, which were fabricated on different occasions over the course of several months [2,4]. Array-to-array reproducibility was also examined by Stitzel *et al.*, who showed that sensor response profiles were reproducible when multiplexed arrays were interchanged [9].

There are two parts to this paper. In the first part, two sensor types are used to create single-sensor, homogeneous arrays (Figures 1a and 1b). The sensor types are shown in Table 1. The two homogeneous arrays were exposed to fourteen different odors and time-dependent sensor-odor response profiles were recorded before, during, and after odor exposure. The arrays were exposed to each of the fourteen odor types in independent, separate experiments, *i.e.*, different odor types were not simultaneously combined or delivered. These sensor-odor response profiles are used here as one method to decode the sensor types in the randomized array, which was prepared with the same sensor types (Figure 1c). The second part of this paper discusses another optical technique, which employs the microsensors’s fluorescence emission spectra to decode the randomized array. In addition to recording sensor-odor response data for both homogeneous arrays, an emission spectrum was also acquired via a liquid crystal tunable filter for each sensor element within the array. Recording a sensor’s emission spectrum is similar in some respects to recording the sensor’s response profile upon odor exposure, *i.e.*, response from all sensor elements in the field of view can be extracted from the image data, except that the emission response is not time-dependent and no odor exposure takes place. Odor

scans consist of sixty image frames as a function of time at a single wavelength for each odor exposure, and each frame corresponds to a *time stamp*. In contrast, the emission scans consist of fifty image frames collected at different wavelengths and each frame refers to a *wavelength stamp*. Both types of image data are acquired with a liquid crystal tunable emission filter, which can be tuned to different emission wavelengths without moving any other portion of the detection system. This filter allows the user to collect fluorescence emission data along multiple spectral windows. The array emission scan referred to here is nearly identical to acquiring a dye’s fluorescence spectrum, such as with a fluorescence spectrometer, except each sensor element produces its own emission spectrum. As shown previously, Nile Red-stained microsensors have different emission spectra depending on the microsphere material’s type and polarity [2].

<INSERT TABLE 1>

As discussed in more detail below, the emission spectrum for each known sensor type was used to discriminate between randomly distributed sensor elements by comparing to control (known) emission spectra. This process is useful for discriminating between different sensor types only when the sensors within a given array have different spectral characteristics. Response data from the two homogeneous arrays, with a total of 2,292 sensor elements, were extracted from the odor exposure and emission scan image data. These profiles were used as standards to compare the sensors in the random array. Here we report on new methods for optically decoding sensor elements in a randomized array, which contained 587 sensor elements. Figure 2 briefly summarizes our decoding procedure for solving the sensor identification in randomized arrays. The decoding techniques are discussed and compared below.

<INSERT FIGURE 2>

EXPERIMENTAL SECTION

Materials. The porous silica microsphere packing were obtained from Phenomenex (Torrance, CA). Acetone, ethanol, and toluene were purchased from Fisher (Fairlawn, NJ). Columbian coffee beans (Jim's Organic Coffee) were used as purchased. The carrier gas used in the experiments is ultra zero grade air from Northeast Gas, Inc. (Salem, NH). Nile Red dye, 1,3-dinitrobenzene (DNB), 2,4-dinitrotoluene (DNT) 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 pre-etched, 5.5 μm -core optical imaging fibers were acquired from Illumina Inc. (San Diego, CA).

High-density homogeneous, single-sensor arrays. Single-sensor arrays were fabricated with pre-etched optical imaging fibers (Illumina, Inc. San Diego, CA) with both sensor types (one sensor type per imaging fiber). The two arrays were fabricated as previously described [2]. Briefly, each optical imaging fiber's etched face was rubbed into a prepared microsensor slurry. The sensors spontaneously assemble into the pre-etched high-density microwell array platform (one sensor per well), as previously documented [6]. A burst of air from an air gun was used to remove loose sensors from the fiber's distal face. It is believed that the beads stay in the microwells due to capillary forces. Emission scans for the homogeneous arrays were performed so emission profiles could be compared to those from the randomized array. After each high-density fiber array was positioned on the detection system, an emission scan was collected first and then the array was exposed to 50% saturated methanol vapor (to condition the sensors) prior to acquiring data for subsequent odor exposures. All emission scans were acquired with a liquid

crystal tunable emission filter (see below). The homogeneous arrays were prepared and tested on different days. All arrays were tested two days after they were prepared. The temperature and percent relative humidity over the course of testing ranged from 21.7° to 26.2°C and 39% to 46%, respectively.

High-density randomized array. The two sensor types were combined in approximately equal portions and placed into a 4-mL glass vial. The sensor mixture was then labeled and vortexed. A small amount of sensor stock was removed from the vial with a glass capillary tube and the sensor mixture was smeared into the surface of an optical imaging fiber as discussed above. The random array was prepared and tested on the same day, which was ~21 days after the single-sensor arrays were tested. The temperature and percent relative humidity for the random array was 25.6°C and 21% respectively.

Vapor delivery, emission scan, data collection and image processing. Our sparging vapor delivery apparatus, similar to one developed by Kauer and Shepherd [35], has been well-documented and employed for many detection tasks in our laboratory [1,2,4,9,17-20,22,36]. All solvent or solid analytes were individually placed in sealed flasks for vapor delivery; for ‘air’ (blank) delivery an empty sealed flask is employed in order to purge the carrier (air) gas onto the sensors. Carrier gas is simply the background air used to transport a tested vapor from the source to the sensor array. The non-humidified carrier air flows through a flask containing a test vapor and carries it through the vapor delivery apparatus to the sensors.

Image data for two homogeneous arrays and one randomized array were acquired for fourteen odor exposures and an emission scan. All image data were acquired with 2 x 2 CCD chip binning (to improve signal-to-noise performance) and a 20x microscope objective (N.A. 0.50). All *odor* exposures were acquired with 100 ms CCD exposure, a 560 (band pass 40) nm

excitation filter, and the liquid crystal tunable emission filter (Cambridge Research & Instrumentation, Inc, Cambridge, MA) set to 630 nm. Sixty image frames were collected in the order: 10 before odor exposure, 20 during exposure, 30 following odor exposure. For each exposure, 50% saturated vapor was delivered to the sensor arrays with a total flow of 200 mL/min. The fourteen odors (x 3 replicates) were sampled in this order: acetone, air, benzene, chloroform, Colombian coffee, cyclohexanone, 1,3-dinitrobenzene, 2,4-dinitrotoluene, ethanol, n-propanol, heptane, methanol, toluene, and water. The tunable emission filter was employed to capture fluorescence responses from each sensor array, as $\lambda_{(\text{emission filter})}$ was incremented by 2 nm/steps from 580 nm to 678 nm. The emission profile captured spectra over 100 nm (50 image frames x 2 nm steps). The excitation filter and CCD exposure time employed for the emission scan was 530 (band pass 30) nm and 400 ms respectively. Data extraction from the CCD image stacks were carried out automatically using a Labview BeadFinderVI, previously developed for the purpose [1].

RESULTS AND DISCUSSION

Both single-sensor, homogeneous sensor arrays and a multiplexed, randomized sensor array were employed to show that sensors within the randomized array can be positionally-registered, or decoded, by comparing sensor response profiles to known (control) response profiles. The control response profiles were produced from the homogeneous sensor arrays. Figure 1 shows a cartoon of the three sensor arrays examined in this study and Figure 2 shows an overview of the decoding process. This method for sensor identification was applied to the randomized array's response data. One decoding method was based upon a supervised pattern recognition technique (DFA) and the other one was unsupervised (PCA). Throughout this paper,

standard pre-processing metrics were used to extract information from the time-dependent sensor response profiles. These metrics, which include difference, fractional, relative, array normalization difference, array normalization fractional, array normalization relative, sensor normalization difference, sensor normalization fractional, and sensor normalization relative, are well documented [37,38] and are not discussed further. The type of metric employed in each situation will be identified where necessary.

Three decoding techniques were developed for sensor identification within random arrays – a method relying solely on the emission spectra, another method using both the sensor-odor responses and emission spectra (DFA), and another using only the sensor-odor responses (PCA). The first two decoding techniques are *not* independent. For instance, the supervised (DFA) method required two data sets to perform sensor categorization and each data set consisted of extracted sensor response profiles. The DFA’s input data set consisted of the randomized array’s odor response profiles but since DFA is a supervised technique, this method required another data set for classification. The second DFA data set used to categorize the randomly-dispersed sensors consisted of the emission scan responses to provide the class labels for each bead. In order to perform DFA, we have assumed the emission scans are very accurate for identifying sensor classes, i.e., we do not have *a priori* information about exact sensor identification within the randomized array so the emission scan was used as the gold standard to test our classification algorithm outlined in Figure 2. Because the emission scans are so distinct for each of the two sensor types (*vide infra*), the two sensor types could be easily grouped before DFA was employed. The other decoding method employed only the sensor-odor response profiles for an unsupervised (PCA) classification. As discussed in more detail below, the PCA technique may be ideal for our sensor platform because this method is simple, depends only on the sensor-odor

response profiles and does not require a second data set for classification purposes, obviating the need to acquire emission spectra. The emission scan data used to help decode the randomized sensor elements is much more applicable to other multiplexed fluorescence-based arrays or assays because different fluorescent elements can be categorized by their unique fluorescence spectra. It should be noted that the purpose of this paper is to discuss sensor-decoding strategies for multiplexed randomized arrays and assays. We do not focus on (or discuss) odor discrimination performance for either the homogeneous or random arrays, since we have previously shown that odors can be successfully discriminated using both bundled homogeneous arrays [1,20] and multiplexed randomized [4,9] arrays, proving that both platforms are useful for odor discrimination/detection.

Homogeneous, single-sensor arrays. Response profiles for 2,292 microsensors were examined for 42 odor observations (14 different odors x 3 replicates) and an array emission spectra observation. In all, 98,556 sensor response profiles were extracted from the image data (2292 sensors x 43 observations). Table 1 shows the population and identification for both sensors from the homogeneous arrays. Sensor response profiles from replicate elements were averaged to enhance signal-to-noise ratios, as previously demonstrated [2,4]. The sensor-odor profiles produced from the two homogeneous arrays are compared and discussed as one means to decode the sensor positions within the randomized array.

Sensor-odor response profiles from the homogeneous arrays. For the odor exposures, each sensor type responds in a characteristic and reproducible manner to develop unique radar plots (Figure 3). These unique patterns, or sensor fingerprints, are usually employed to discriminate odors with cross-reactive chemical sensor arrays (or electronic noses) [17,22,37-44]. Now, we demonstrate several new ways in which sensor-odor response patterns can be used

to ‘fingerprint’ each sensor type within a random array. This process indicates the potential for decoding optical sensor elements based upon their characteristic signals during odor stimulus. Each spoke in Figure 3 refers to a different odor, as listed in the caption. Two distinct sensor patterns give a preliminary indication that the sensor types within the multiplexed array can be discriminated on the basis of their odor response. For example, when exposed to chloroform odor (spoke-4 in Figure 3), one sensor type increased by +160% (*sensA*) whereas the other sensor gave no response (*sensF*) (fractional values). Single dashed lines on either side of the mean (thick central line) indicate a confidence interval of 1 standard deviation (SD) and show odor response variations from sensor replicate-to-replicate to the first odor exposure. Since the SDs of the sensor response to almost all the odors are relatively small when compared with the mean, the results suggest the sensor-odor response profiles are highly reproducible among sensor replicates, thus supporting other documented studies with these sensor types [1,2,4,9,20].

<INSERT FIGURE 3>

Sensor discrimination via array emission spectra. By using a single excitation wavelength, each array element’s emission spectrum was simultaneously acquired by increasing the tunable emission filter’s spectral window by 2-nm increments (see Experimental section). Emission scan sensor responses were extracted with the same procedure as used for sensor-odor responses. When the environmentally-sensitive, solvatochromic dye Nile Red is adsorbed to microspheres of different characteristics, its spectral properties are changed [2]. Different microspheres therefore have different excitation and emission spectra because the Nile Red is in a different micro-environment. As the microsphere’s surface polarity increases, i.e., -CN to -OH, the sensor’s spectrum exhibits a bathochromic shift. Depending on the magnitude of Nile

Red’s spectral shift, the results may be comparable to a multi-fluorophore assay because each sensor type may have a unique emission spectrum. Decoding randomized arrays via emission spectra has never been demonstrated before. By simply comparing emission spectra of the sensors within the randomized arrays to those in the homogeneous arrays, the sensors within the randomized array could easily be identified. As stated above, emission scan data are used as the standard to compare the effectiveness of decoding the array with the odor response profiles. To perform sensor classifications, each sensor type’s Z parameter was calculated. For example, the two sensor groups in the randomized array, one with cyano- functionality (*sensF*) and the other with hydroxyl functionality (*sensA*), were identified via their emission spectra, λ_{\max} and equation 1:

$$Z = (V_{645 \text{ to } 650 \text{ nm}}) - (V_{625 \text{ to } 630 \text{ nm}}) \quad (\text{eq. 1})$$

where $V_{645 \text{ to } 650 \text{ nm}}$ and $V_{625 \text{ to } 630 \text{ nm}}$ are the average grayscale values across wavelengths 645 to 650 nm and 625 to 630 nm in the random array. Simply, this equation allows each emission scan to serve as the supervised data input to check the DFA classification results. When $Z > 0$, the sensor in question was labeled as *sensA* (-OH functionality) because it closely matched the result from the single sensor array. When $Z < 0$, the sensor was decoded and labeled as *sensF* (-CN functionality). The sensor emission scan response curve for the single-sensor homogeneous arrays can be compared directly to the emission profiles from the random array’s emission scan (Figure 4). Figure 4a shows 80 emission response profiles for the two homogeneous arrays (40 replicates x 2 sensor types) and Figure 4b shows ~200 emission response profiles for the randomized array. The raw emission spectra in Figure 4 were normalized to $\lambda_{\max} = 1$, so that a qualitative comparison between the control and random arrays could be visualized. Although the

spectra produced by the sensors within the random array show slightly greater dispersion (less reproducible from sensor-to-sensor) than those from the homogeneous arrays, it was still possible to unequivocally assign a class label in each case. The noisy emission curves in Figure 4 appear to be an instrumental artifact. Note that the actual data used for sensor identification within the random array employed the grayscale intensity values in equation 1, and not the normalized emission spectra.

<INSERT FIGURE 4>

Random array sensor discrimination using DFA based upon sensor-odor responses. The DFA decoding technique was carried out with the sensor-odor response profiles. DFA classification was performed with the sensor class identifications learned from the emission scan profiles. This supervised decoding technique produced high classification rates and therefore would be used as a benchmark to compare results for the unsupervised decoding method (discussed below). The results showed a clear separation between the two sensor types except for 4 misclassified sensors out of 587 total sensors (Figure 5). A correct sensor classification rate of 99.32% was achieved using DFA. A method for cross-validating these results was carried out with the maximum posterior probability estimate technique, which is used to accurately estimate the likely error of a particular classifier to unseen data. This estimate is the average posterior group membership probability for all cases for those groups to which the cases were assigned by the discriminant functions. The result of this estimation also produced 99.32% correct sensor classification, further demonstrating that nearly all the sensors were correctly grouped and decoded. Because 583 out of 587 sensor elements were correctly classified, this result gave an indication that the odor response profiles and the emission scan are highly

correlated and represent different ways of measuring the same thing, i.e., sensor type class membership. This supervised technique was necessary to prove our decoding concepts, but such complicated methods were not required for practical sensor categorization. For instance, according to the Z parameter and Figure 4, sensor class could be predicted with (presumably) high accuracy since the emission scans are very different and none of the sensors had Z values close to zero. It is therefore reasonable to argue that the emission scan data could be used for classifying sensor types with high accuracy, which is why the emission data were used as a benchmark to test the efficacy of our less “expensive” sensor decoding algorithms based only on the odor response profiles. Less expensive methods refer to unsupervised approaches (discussed below), which do not require control or training data sets, i.e., known sensor response information, to discriminate between samples or classes.

<INSERT FIGURE 5>

Random array sensor discrimination using PCA based upon sensor-odor responses. PCA was applied to the sensor-odor response profiles to test the possibility for bead type identification when no prior information was made available. Two distinct PC-space clusters resulted, each (presumably) corresponding to one of the sensor types. By using the sensor identification information obtained from the DFA-emission scan, the sensors were labeled on the PCA plot (Figure 6). Figure 6 shows that only four sensors are misclassified out of 587 with unsupervised PCA method. The correct classification rate was 99.32%, which is identical to the supervised technique. This method demonstrates the effectiveness of using an unsupervised technique to correctly identify sensor types within the high-density randomized array. Though the emission scan data are very convincing, the unsupervised PCA technique may be more suitable for field

applications where information, such as emission spectra, may not be available or convenient to collect. The emission scan data were used here only to quantify the accuracy of the unsupervised PCA method. These results show that decoding can be performed with only sensor-odor responses, which would be representative of data collected in the field.

<INSERT FIGURE 6>

After observing sensor classification with DFA and PCA, we observed that all 14 sensor-odor responses were not required to correctly identify the sensors in this study. As mentioned above, Figure 3 shows that the two sensor types have very different values for chloroform odor (spoke 4 in the figure), which signifies that the two sensor types could be classified with only one odor exposure. Although both sensor types could be correctly categorized with less total data, i.e., less odor response profiles, we did not seek to optimize the decoding procedure for these two sensor types. In general, the power of our decoding method depends strongly on the number and types of test odors. The two sensor-odor response based decoding techniques were equally effective for the decoding task studied. Both techniques employed the sensor response profiles extracted from the CCD-based image stacks, essentially creating three-dimensions of information for sensor discrimination: odor exposure (spatial, time, intensity) and emission scan (spatial, spectral, intensity). These types of data are similar to data recorded with hyperspectral imagers or spectrograph/CCD combinations [32]. In our case, the emission scan served as a complementary decoding technique to the odor response information. Regardless, the two decoding techniques were equally effective and their results supported each other. The two techniques could be combined to facilitate sensor decoding when more sensor types are added to a random array.

CONCLUSION

We have shown that it is possible to correctly classify sensor types in a randomized array without prior knowledge of sensor locations. By comparing sensor response profiles from high-density homogeneous (standard) arrays, the sensors within the randomized array were decoded according to their chemistries. We demonstrated that two classes of microbead sensors, which were randomly dispersed on an optical imaging fiber’s distal face, could be decoded via both sensor-odor response profiles and emission spectra. In the case of PCA, it is not necessary to know what the sensor types are on the random array in advance, because by using the sensor-odor fingerprint information, the sensor types can be categorized. Though the emission scan is a very powerful and easy approach to identify multiple fluorophore types in an assay, this process may not be ideal for real-world testing conditions.

The use of randomized arrays is changing the face of sensor array technology because the arrays do not have to be fabricated with defined sensor positions. Instead, combinations of prepared sensor materials can be simultaneously and rapidly added and deposited onto the array platform to form a high-density randomized array. Although sensor locations within the array are unknown, decoding processes allow sensor identifications to be learned. This paper demonstrates that sensors within a randomized array could be decoded with both supervised and unsupervised pattern recognition techniques. The unsupervised PCA technique employing the sensor-odor response profiles is a more ideal decoding procedure, but other decoding techniques might be required (or a combination thereof) when more sensor types are included in high-density randomized arrays. Because of the simple steps required for fabricating randomized arrays and their low associated costs, more detection systems may eventually turn to a

randomized array design. Optically-based multiplexed sensor systems and technologies are being developed more readily and with increased capabilities. One major limitation of these systems, however, relates to the number of fluorophore types that can be simultaneously employed. Because of spectral overlap, some systems cannot resolve fluorophores with similar emission spectra. The decoding method described here, which uses the entire emission spectrum for each sensor element, will potentially allow better resolution between overlapping fluorophore spectra to increase the types or numbers of detection tasks one can carry out with the multiplexed assay. Just like other multiplexed arrays which have had technological impacts in genomics related detection devices [45,46], nano-electronics [47-49], cellular analysis [50,51], electrochemistry [52,53], immunochemistry [54,55], molecular recognition [56,57], and combinatorial drug design [58], randomized arrays are likely to create technological changes (software, array platform architecture) in detection platforms.

ACKNOWLEDGEMENTS

The Tufts work is supported by the Defense Advanced Research Projects Agency and Department of Energy (to DRW) and the Leicester work is funded by the European Commission (to TCP), the Engineering and Physical Science Council (to TCP), and the Royal Society, London (to TCP).

REFERENCES

(see other Word documents—reference texts included in separate files)

FIGURE CAPTIONS

Figure 1. Fabrication of the single-sensor, homogeneous arrays (a, b) and the randomized array requires two simple steps (c). Prepared microsensors are randomly dispersed onto an optical imaging fiber’s etched face to create randomized high-density bead arrays.

Figure 2. Schematic diagram of the random array decoding process. CCD images of the randomized array are shown above. Sensor response profiles are extracted for n odor exposures or for the array emission scan. The extracted sensor profiles are subjected to PCA or DFA and plotted accordingly. The clusters formed from the response profiles are used to identify each sensor type and location within the random array, as shown in the image below. Note: the image with the mapped sensor locations is only used for explanatory purposes, i.e., the mapping is not the actual decoded random array.

Figure 3. Sensor-odor response radar plots for the homogeneous arrays. The plots show the difference in the average fractional response to the first exposure for each odor. The labels for each odor observation (1-14): (1) acetone, (2) air (control), (3) benzene, (4) chloroform, (5) Columbian coffee beans, (6) cyclohexanone, (7) 1,3-dinitrobenzene, (8) 2,4-dinitrotoluene, (9) ethanol, (10) n-propanol, (11) n-heptane, (12) methanol, (13) toluene, and (14) water. In each case the dashed lines indicate ± 1 Standard Deviation (SD) across all the odors showing the variation of odor responses from sensor-to-sensor within the first odor repeat of a particular odor. Positive values indicate an overall increase in grayscale value in response to the odor whilst negative values indicate an overall decrease.

Figure 4. (a) Emission spectra for the two homogeneous sensor arrays. For each sensor type, 40 replicate responses are shown for signal reproducibility. The plot shows a total of 80 beads (40 beads x 2 sensor types). (b) Emission spectra for 193 beads out of 587 from the randomized array.

Figure 5. Plot showing one canonical discriminant variable after DFA was applied to the array normalized difference metric for the randomized array’s emission spectra. The two groups show bead types sensA and sensF being classified into their respective groups. The x-axis shows the reference number of each individual bead in the array. As shown by the dotted line, the first bead is of type F and the last bead (bead 587) is of type A. * indicates beads with uncertain values of Z parameter ($32 \leq Z \leq -21$). These 4 beads are termed deviant beads.

Figure 6. Plot showing classification of the two sensor types (sensA and sensF) for the randomized array using an unsupervised technique (PCA) and the sensor-odor response profiles. The four beads outside the two clusters are deviant beads (*). The letters (A and F) correspond to the assumed class as calculated using the Z parameter in equation 1.

TABLE CAPTION

Table 1. Sensor Materials Employed.

Sensor ID	<i>(sensA)</i>	<i>(sensF)</i>
Sensor Name	<i>3phenosOH</i>	<i>3ibsilCN</i>
Microbead diameter (μm)	3	3
Material Name ^a and Surface Chemistry	Phenosphere (-OH)	Ib-Sil (-CN)
Pore Size (\AA)	80	125
Approximate Sensor Age at Time of Test	20 months	20 months
Number of Sensors Analyzed in Homogeneous Arrays	1077	1215

^aBoth materials named according to distributor (Phenomenex, Inc, Torrance, CA USA).

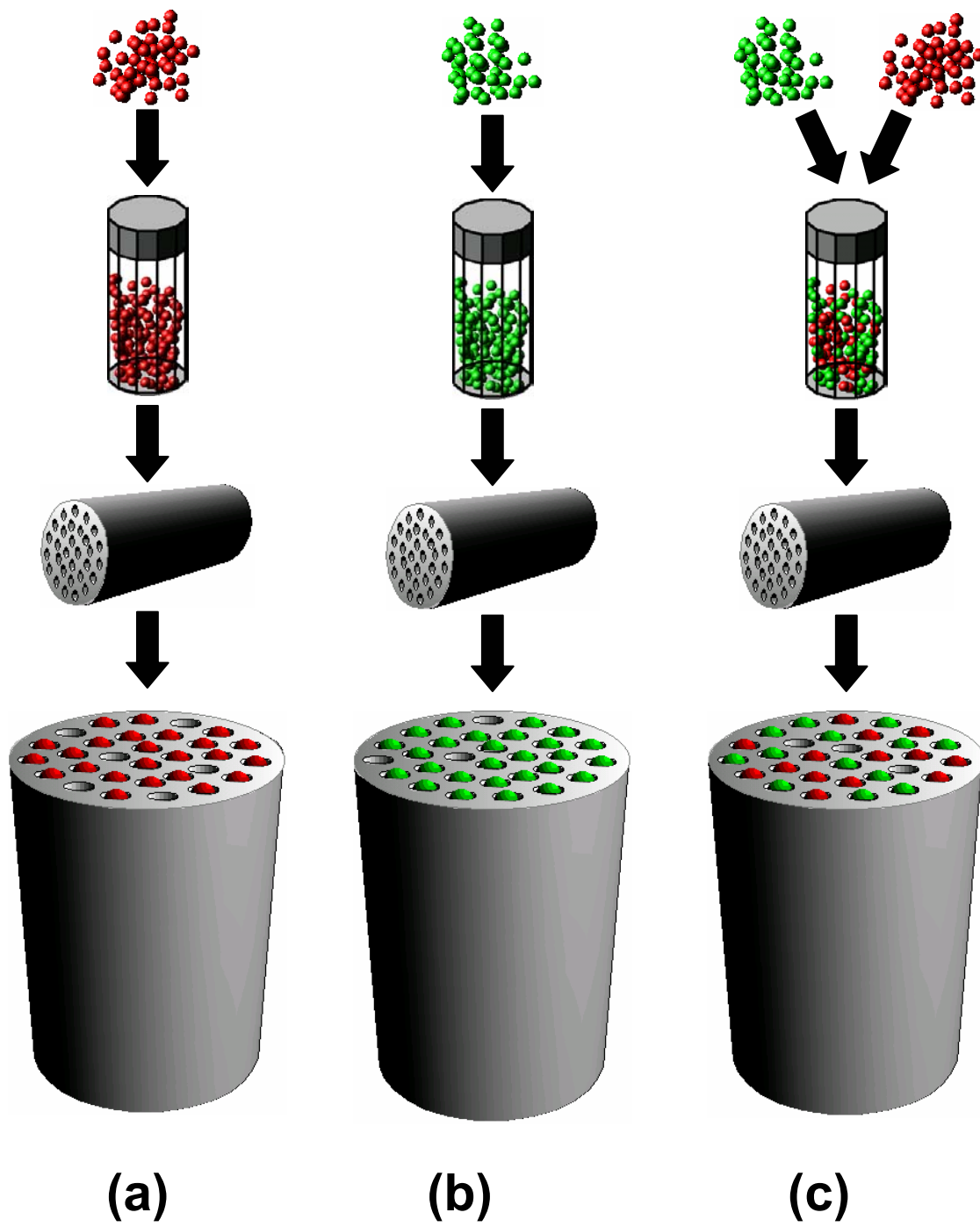


FIGURE 1

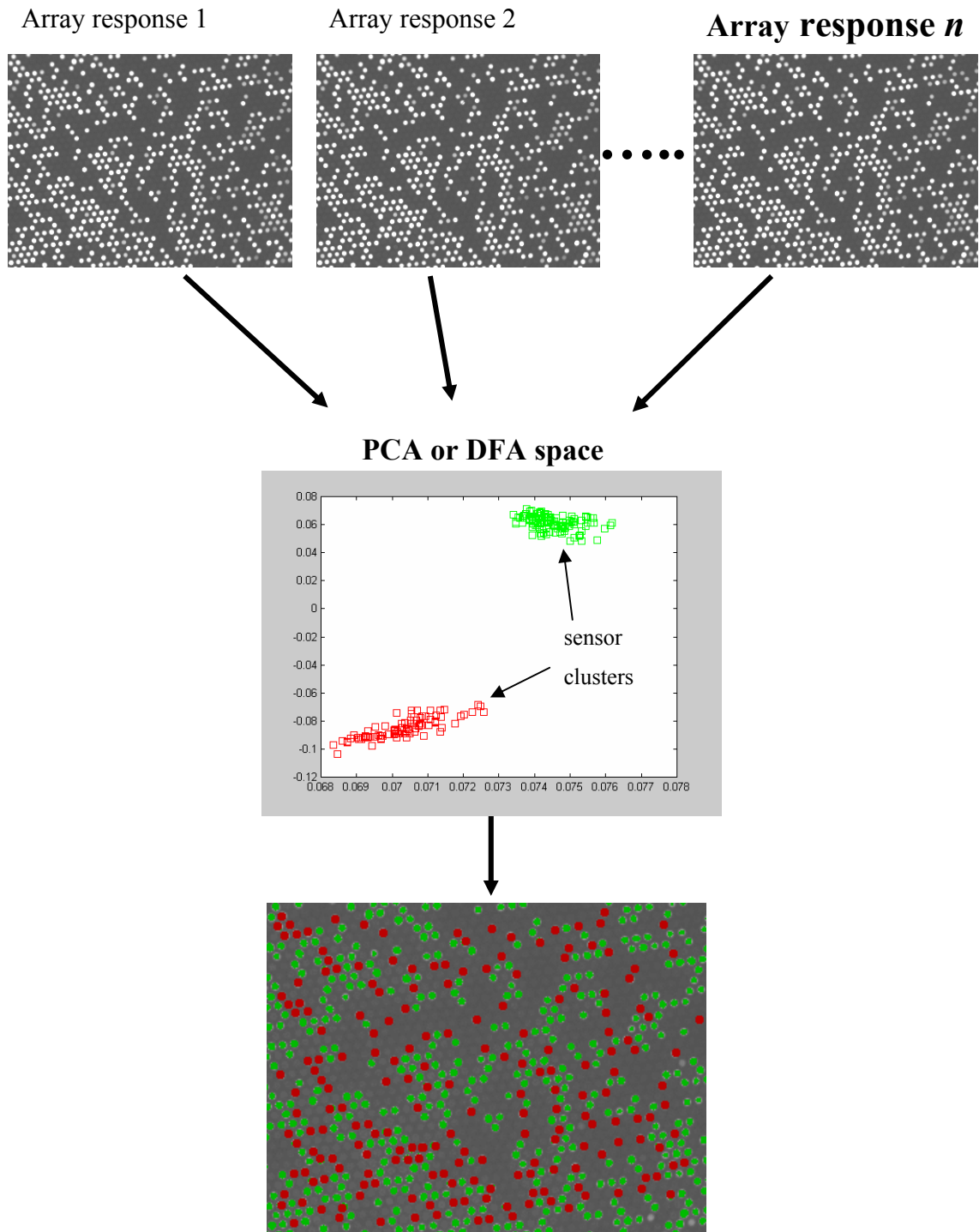


FIGURE 2

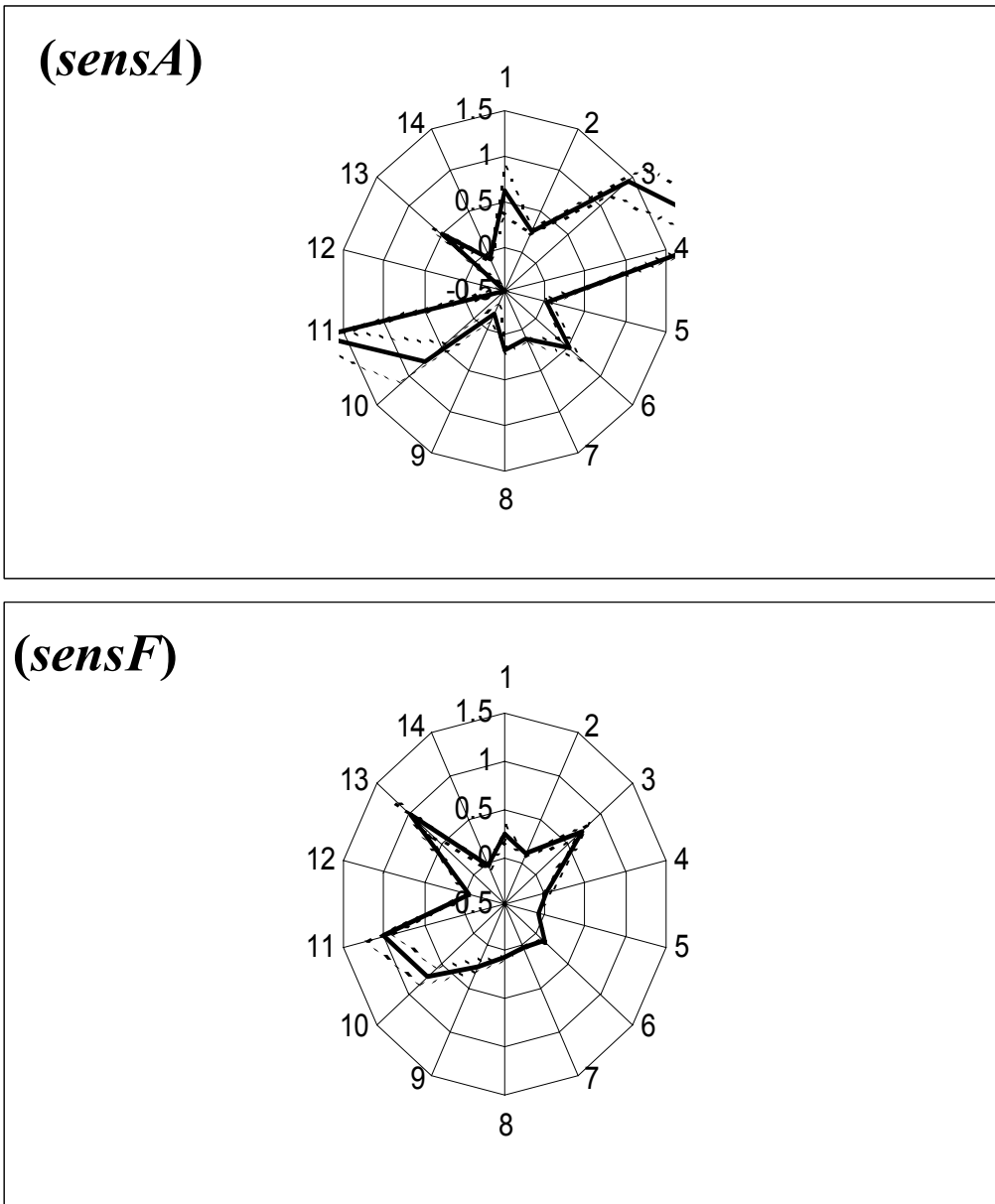


FIGURE 3

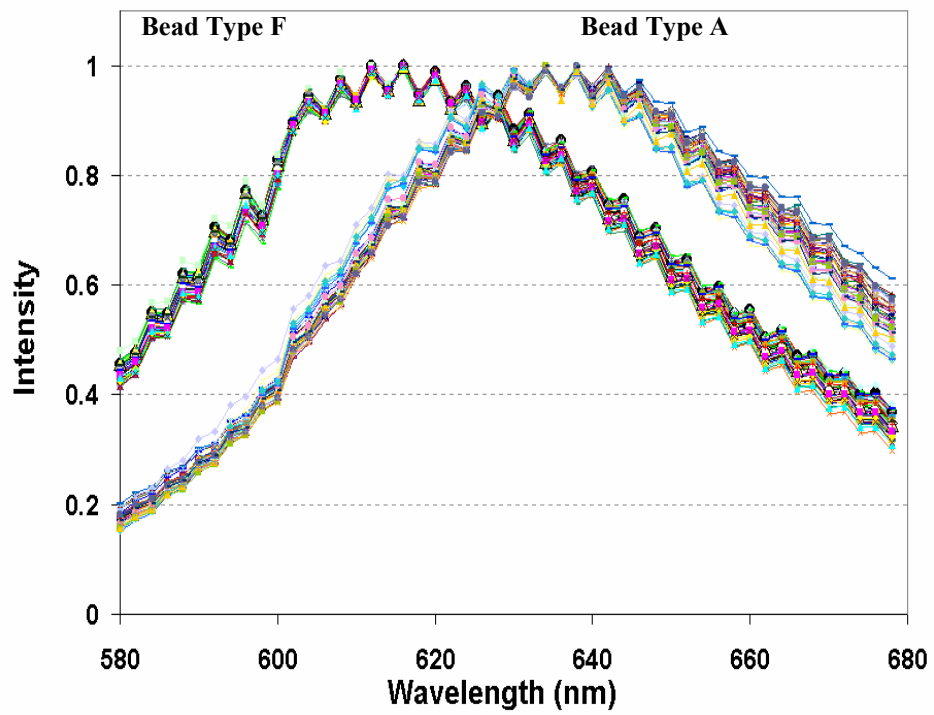
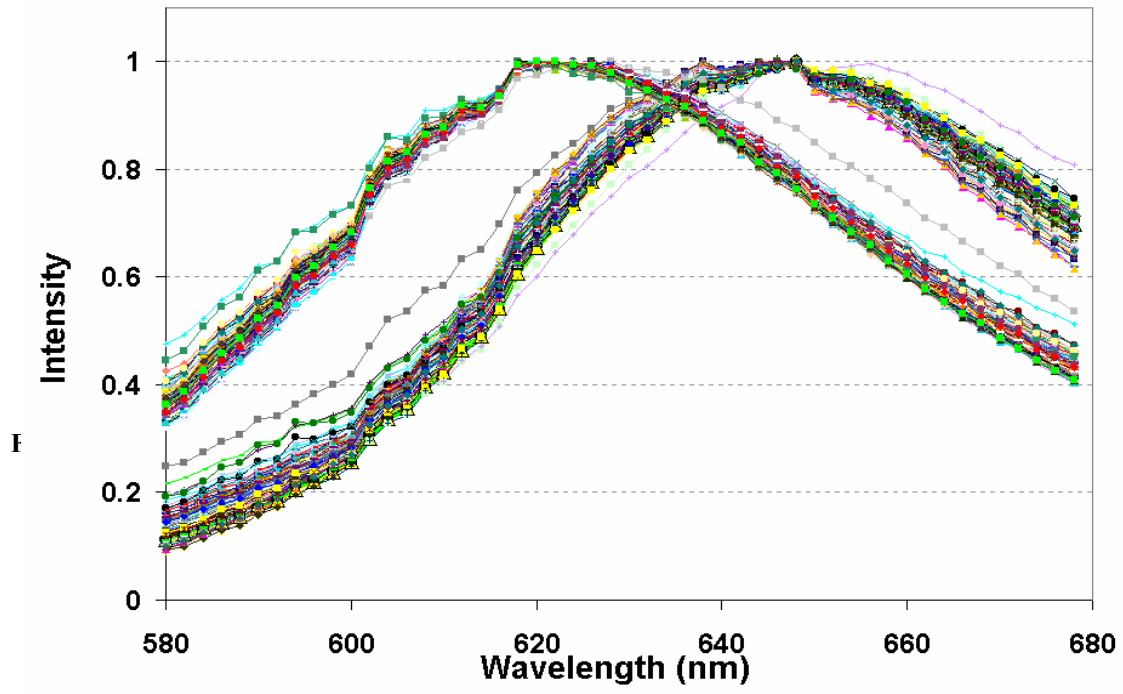


FIGURE 4a



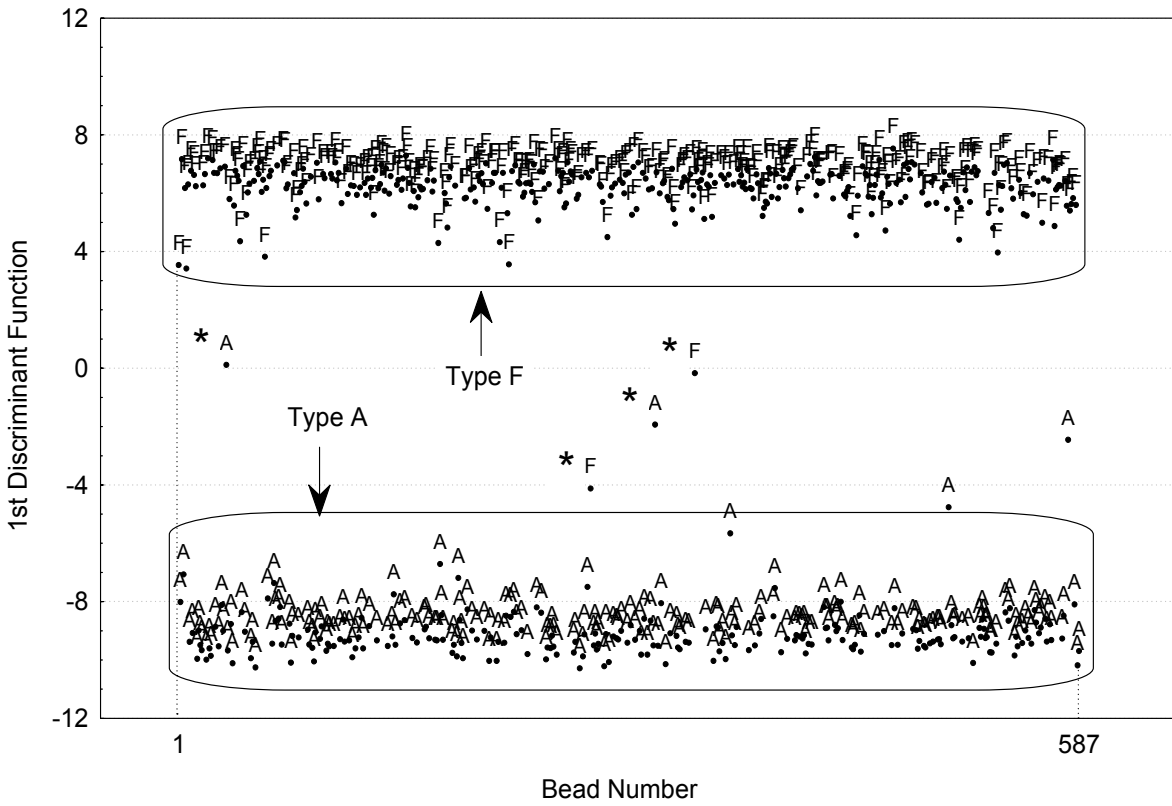


Figure 5

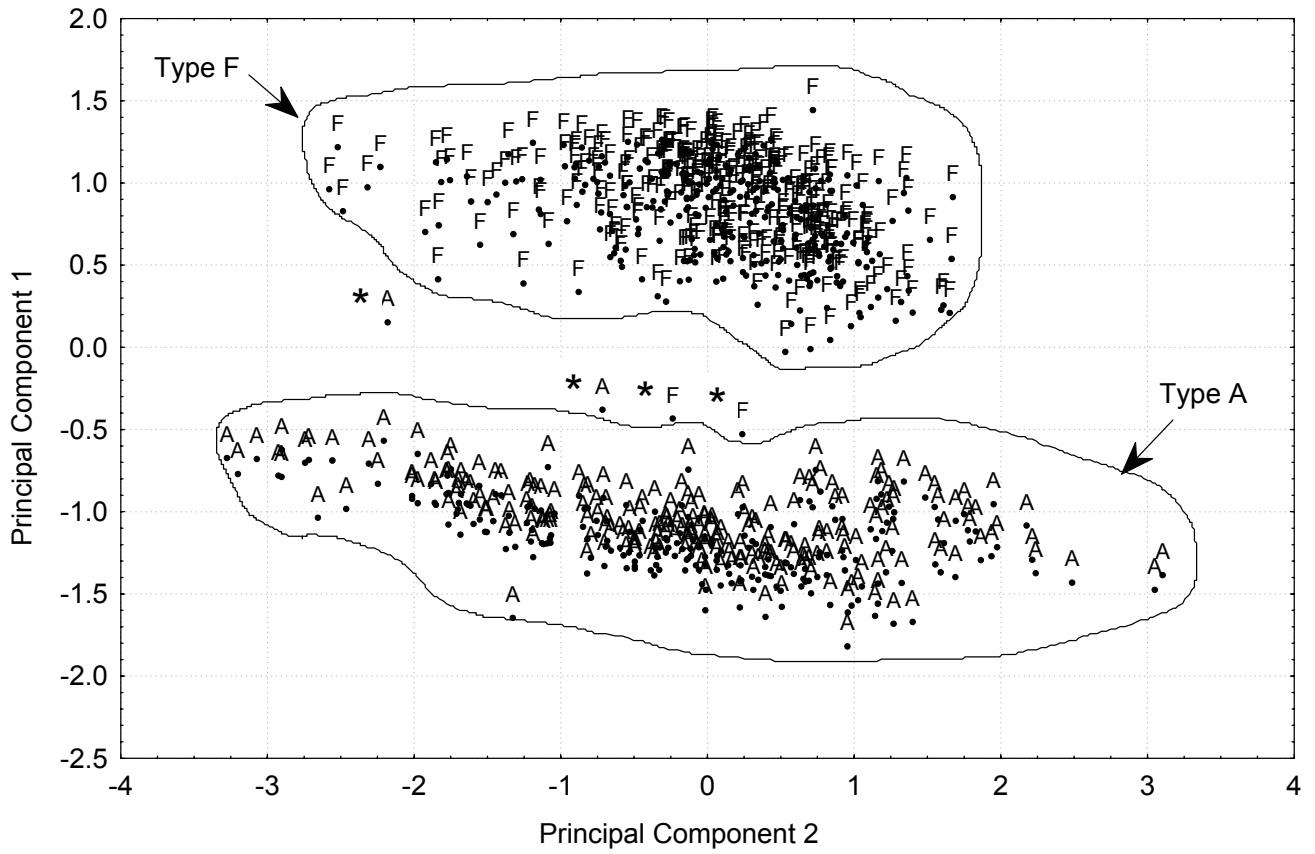


Figure 6