A Novel Breath Analysis System Based on Electronic Olfaction

Dongmin Guo, David Zhang, *Fellow, IEEE,* Naimin Li, Lei Zhang, *Member, IEEE,* and Jianhua Yang

***Abstract*—Certain gases in the breath are known to be indicators of the presence of diseases and clinical conditions. These gases have been identified as biomarkers using equipments such as gas chromatography (GC) and electronic nose (e-nose). GC is very accurate but is expensive, time consuming, and non-portable. E-nose has the advantages of low-cost and easy operation, but is not particular for analyzing breath odor and hence has a limited application in diseases diagnosis. This article proposes a novel system that is special for breath analysis. We selected chemical sensors that are sensitive to the biomarkers and compositions in human breath, developed the system, and introduced the odor signal preprocessing and classification method. To evaluate the system performance, we captured breath samples from healthy persons and patients known to be afflicted with diabetes, renal disease, and airway inflammation repectively and conducted experiments on medical treatment evaluation and disease identification. The results show that the system is not only able to distinguish between breath samples from subjects suffering from various diseases or conditions (diabetes, renal disease, and airway inflammation) and breath samples from healthy subjects, but in the case of renal failure is also helpful in evaluating the efficacy of hemodialysis (treatment for renal failure).**

***Index Terms*—Breath analysis, electronic olfaction, therapy monitoring, chemical sensor, disease identification.**

I. Introduction

In recent years, there were increasing concerns about the applications of breath analysis in medicine and clinical pathology both as a diagnostic tool and as a way to monitor the progress of therapies [1], [2]. Comparing with other traditional methods such as blood and urine test, breath analysis is non-invasive, real-time, and least harmless to not only the subjects but also the personnel who collect the samples [3]. The measurement of breath air is usually performed by gas chromatography (GC) [4] or electronic nose (e-nose) [5]. GC is very accurate but is expensive and not portable, its sampling and assaying processes are complicated and time consuming (about one hour for one sample), and its results require expert interpretation [6]. A less expensive and more portable alternative is e-nose. It is cheaper and faster (requiring only 30 minutes for one sample) and is often used outside of medicine, in fields related to food, chemistry, fragrances, security, and environment [7]. Recently, e-nose has gradually been used in medicine for the diagnosis of renal disease [8], diabetes [9], lung cancer [10], and asthma [11]. While all of these methods work satisfactorily, they can each identify only one particular disease. One reason for the limited applications of e-noses in breath analysis might be the design of commercial e-noses for broad applications rather than for breath analysis specifically. We thus propose a new specific breath analysis system in this paper in order to extend the applications in medicine.

The system makes use of chemical sensors that are particularly sensitive to the biomarkers and compositions in human breath to trigger responses to a patient's breath sample. In contrast to the broad panel of nonspecific sensors used in commercial e-noses, the sensors of our system were specifically selected for their responses to known components of human breath. The sample is injected into the system using an auto-sampler at a fixed injection rate to guarantee all samples are sampled under the same criterion. The chemical sensors sense the sample and accordingly form a kind of 'odorprint' that is typically associated with a given disease or condition. The 'odorprint' is then sent to computer for signal processing and pattern recognition. We evaluated the system in two experiments. In the first we classified subjects with renal failure before and after hemodialysis. In the second we applied the system to distinguish between healthy subjects and subjects suffering from three types of diseases/conditions (diabetes, renal disease, and airway inflammation). The experimental results show that our system can fairly accurately measure whether hemodialysis has been effective and can identify the three conditions/diseases with quite a high level of accuracy.

The remainder of this paper is organized as follows. Section II describes the composition of the human breath and the certain diseases that may be associated with certain gaseous compounds. Section III describes how a subject's breath is sampled, the setup of the sensor array, and how data is processed. Section IV explains the experimental details. Section V gives the experimental results and discussion. Section VI offers our conclusion.

II. Breath Analysis

Human breath is largely composed of oxygen, carbon dioxide, water vapor, nitric oxide, and numerous volatile organize compounds (VOCs) [12]. The type and number of the VOCs in the breath of any particular individual will vary but there

The work is partially supported by the RGF fund from the HKSAR Government, the central fund from Hong Kong Polytechnic University, and the NSFC fund under contract No. 60803090 in China, the Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China.

D. Guo and L. Zhang are with the Biometrics Research Centre, Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong.

D. Zhang is with Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China and the Biometrics Research Centre, Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: csdzhang@comp.polyu.edu.hk).

N. Li is with the Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001, P.R. China.

J. Yang is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, P.R.China.

TABLE I
TYPICAL COMPOSITIONS FROM THE ENDOGENOUS BREATH OF THE HEALTHY PERSONS

Concentration(v/v)	**Molecule**
percentage	oxygen, water, carbon dioxide
parts-per-million	acetone, carbon monoxide, methane, hydrogen, isoprene, benzenemethanol
parts-per-billion	formaldehyde, acetaldehyde, 1-pentane, ethane, ethylene, other hydrocarbons, nitric oxide, carbon disulfide, methanol, carbonyl sulfide, methanethiol, ammonia, methylamine, dimethyl sulfide, benzene, naphthalene, benzothiazole, ethane, acetic aide

is nonetheless a comparatively small common core of breath which are present in all humans [13]. The molecules in an individual's breath may be exogenous or endogenous [14]. Exogenous molecules are those that have been inhaled or ingested from the environment or other sources such as air or food and hence no diagnostic value [15]. Endogenous molecules are produced by metabolic processes and partition from blood via the alveolar pulmonary membrane into the alveolar air. These endogenous molecules are present in breath relative to their types, concentrations, volatilities, lipid solubility, and rates of diffusion as they circulate in the blood and cross the alveolar membrane [16]. Changes in the concentration of the molecules in VOCs could suggest various diseases or at least changes in the metabolism. Table I summarizes the typical compositions found in the endogenous breath of healthy persons [13], [15].

Some molecules such as nitric oxide, isoprene, pentane, benzene, acetone, and ammonia may indicate specific pathologies [17]–[19]. To take a few examples, nitric oxide can be measured as an indicator of asthma or other conditions characterized by airway inflammation [20]. Breath isoprene is significantly lower in patients with acute respiratory exacerbation of cystic fibrosis [21]. Increased pentane and carbon disulfide have been observed in the breath of patients with schizophrenia [22]. The concentration of VOCs such as cyclododecatriene, benzoic acid, and benzene are much higher in lung cancer patients than in control groups [23]. Acetone has been found to be more abundant in the breath of diabetics [24]. Ammonia is significantly elevated in patients with renal disease [25]. Table II lists some breath compounds and the conditions that research has found to be associated with them. The compounds and conditions listed in Tables I and II were the focus of the work being described in this paper.

III. DESCRIPTION OF THE SYSTEM

The proposed system operates in three phases (Fig. 1), gas collection, sampling, and data analysis, with a subject first breathing into a Tedlar gas sampling bag. This gas is then injected into a chamber containing a sensor array where a measurement circuit measures the interaction between the breath and the array. The signals are then filtered and amplified and sent to computer for further analysis. Fig. 2 shows our system (left) and its laptop interface.

TABLE II
SOME BREATH COMPOUNDS AND ASSOCIATED CONDITIONS

Breath compounds	**Associated conditions**
acetone	diabetes [24]
carbonyl sulphide, carbon disulphide, isoprene	liver diseases [16]
naphthalene,1-methyl-, 3-heptanone, methylcyclododecane, etc.	pulmonary tuberculosis [26]
nonane, tridecane, 5-methyl, undecane, 3-methyl, etc.	breast cancer [27]
benzene,1,1-oxybis-, 1,1-biphenyl,2,2-diethyl, furan,2,5-dimethyl-, etc.	lung cancer [28]
ammonia	renal disease [25]
octane,4-methyl, decane, 4-methyl, hexane, etc.	unstable angina [29]
propane,2-methyl, octadecane, octane, 5-methyl, etc.	heart transplant rejection [30]
pentane, carbon disulfide	schizophrenia [22]
pentane	acute myocardial infarction [31]
pentane	acute asthma [32]
pentane	rheumatoid arthritis [33]
ethane	active ulcerative colitis [34]
nitric oxide	asthmatic inflammation [35]
nitric oxide, carbon monoxide	bronchiectasis [36], [37]
nitric oxide	COPD [38]
ethane, propane, pentane, etc.	cystic fibrosis [39]

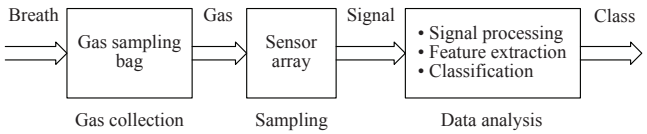


Fig. 1. The working flow defined in our system.

A. Breath Gas Collecting

Fig. 3 shows how the subject's breath is collected using a 600ml Tedlar gas sampling bag (A), an airtight box (B) filled with disposable hygroscopic material to absorb the water vapor from the breath, and, the last component, a disposable mouthpiece (C). The hygroscopic material is silica gel. It is stable and only reacts with several components, such as fluoride, strong bases, and oxidizers. None of them is involved in the breath components showed in Table I and II. Our previous experiments had shown there was no obvious effect on the disease identification by using silica gel as a hygroscopic material. In any case, the mouthpiece is equipped

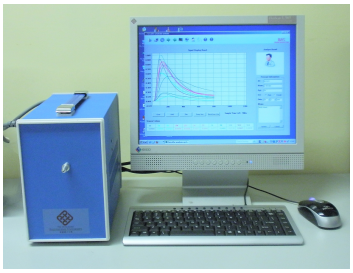


Fig. 2. Breath analysis system and the working interface.

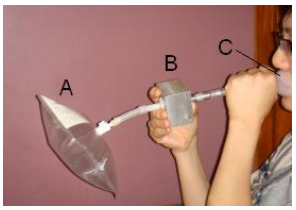


Fig. 3. Exhaled air is collected with a gas sampling bag.

TABLE III
DETAILED INFORMATION WITH RESPECT TO COLLECTED DISEASES

Compounds	Breath sampling locations	Conditions
acetone	alveolar air	diabetes
ammonia	dead-space air	renal disease
nitric oxide	dead-space air	airway inflammation

with an anti-siphon valve that prevents inhalation of the gel.

Subjects are required to give their breath sample in one of two different ways depending on whether the condition under consideration typically exhibits its biomarkers (compounds) in what are known as, dead-space air from the upper airway, or alveolar air from the lungs. Depending on the type of biomarkers and on the breath test tracks, dead-space air may be either a necessity or a contaminant [12]. Alveolar air is required where a condition is typified by biomarkers that are found when there is an exchange from circulating blood. In contrast, dead-space air is required when the biomarkers are released into the airways, and thus into the dead-space air. Table III lists some of the compounds, conditions, and breath sampling locations used in this work [25], [40].

Alveolar and dead-space airs are collected in different ways. Alveolar air is collected by having the subject take a deep breath before breathing into the bag. The first 150 ml of the collected breath is discarded because it may be contaminated [17]. This method would be applied to a subject with, for example diabetes. Dead space air is collected with a pump that draws the breath from the subject's mouth into a sampling bag. The pump is shown as component (B) of the apparatus in Fig. 3. This method would be applied with subjects with, for example, airway inflammation and renal disease. There is no need for the subject to exhale in this process.

B. Signal Sampling

The second phase is signal sampling, which involves acquiring dynamic responses to the interactions between a breath sample and the sensing elements, chemical sensors which form a sensor array in the signal measurement module in the hardware framework. These sensors sense gas particles and generate measurable electronic signals. The signals are then filtered, amplified, and digitized, and then sent to the computer for feature extraction, pattern analysis and classification.

1) Chemical Sensors: The function and the performance of our system highly depend on the capabilities of the chemical sensors. In our system, each sensor in the array has a unique 'odorprint' corresponding to the compounds listed in Tables I and II. Most of the compounds are VOCs but some are inorganic compounds such as ammonia, nitric oxide, carbon dioxide, and hydrogen. Table IV summarizes the main disease biomarkers and compositions in human breath and the type of sensor required. Table V lists the types of sensors used in our system, the gases they are sensitive to and at what sensitivity. These sensors used in our work are metal oxide semiconductor gas sensors from FIGARO Engineering Inc. This kind of sensors is very sensitive, robust, and resistant to humidity and ageing [40]. Seven of the sensors are each sensitive to VOCs. One sensor detects only carbon dioxide. One sensor is sensitive to ammonia, which is associated with renal disease. One sensor is sensitive to nitric oxide which is associated with bronchiectasis, airway inflammation, and COPD. One sensor is sensitive to sulphides, what are associated with liver diseases. Finally, one sensor is sensitive to hydrogen. These sensors are able to sensitive to most of biomakers and compositions in human breath, therefore they have better responses than those commercial e-noses.

TABLE IV
COMPOUNDS DETECTED AND SENSORS REQUIRED

Main compounds in human breath	Requisite sensors
acetone, isoprene, pentane, benzene etc.	VOC sensor
ammonia	NH_3 sensor
nitric oxide	NO sensor
carbonyl sulphide, carbon disulphide	sulphide sensor
hydrogen	H_2 sensor
carbon monoxide, carbon dioxide	CO and CO_2 sensor

TABLE V
TYPE OF SENSORS AND CORRESPONDING SENSITIVE GAS

No.	Sensors	Gases	Sensitivities (ppm)
1	TGS2600	H_2, CO and VOCs	1 - 30
2	TGS2602	VOCs	1 - 30
3	TGS2611-C00	VOCs	500 - 10000
4	TGS2610-C00	VOCs	500 - 10000
5	TGS2610-D00	VOCs	500 - 10000
6	TGS2620	VOCs and CO	50 - 5000
7	TGS825	H_2S	5 - 100
8	TGS4161	CO_2	350 - 10000
9	TGS826	NH_3	30 - 300
10	TGS2201	NO and NO_2	0.1 - 10
11	TGS822	VOCs	50 - 5000
12	TGS821	H_2	10 - 1000

2) Signal Measurement: The framework of the system consists of three modules: signal measurement, signal conditioning, and signal acquisition. The signal measurement module contains a sensor array, temperature control circuit, and measurement circuit (Fig. 4). The temperature control circuit provides negative feedback voltage to the heater of sensors so as to guarantee that the sensors are at stable temperature. The measurement circuit is responsible for transforming odor signals into electronic signals.

The sensor array is composed of 12 sensors set in a 600 ml stainless steel chamber. Breath samples from subjects are

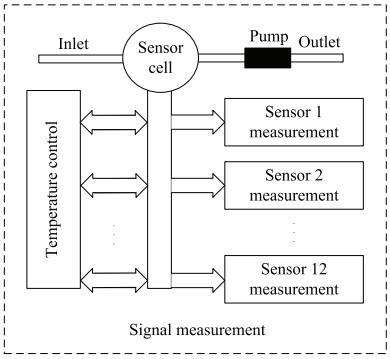


Fig. 4. Basic structure of sensor array used in our system.

TABLE VI
FUNDAMENTAL PERFORMANCE PARAMETERS OF PROPOSED SYSTEM

System parameters	Specifications
Working temperature	$25 \pm 10°C$
Carrier flow	10 ml/s
Chamber volume	600 ml
Sampling injection rate	120 ml/s
Sampling frequency	9 Hz
Sampling time	100 seconds

collected with a 600 ml Tedlar gas sampling bag and then injected into the chamber through an auto-sampler at 120 ml/s. Since the capacity of the sampling bag is 600 ml, the total injection time for one sample is $t = 600/120 = 5s$. The resistances of the sensors change from R_0 to R_s when they are exposed to sampled gas. The output voltage is

$$V_{Out} = \frac{1}{2} V_{CC}(1 - \frac{R_s}{R_0}), \quad (1)$$

where V_{CC} is the transient voltage crossing the sensor and V_{Out} is the transient output voltages of the measurement circuits.

The signal measurement module measures these voltages and converts them into analog electrical signals. The analog signals are subsequently conditioned by signal filtering and amplifying. Finally, the signals are sampled at a 9 Hz sampling frequency and transmitted through a USB interface to a computer. This component is controlled by a 16 bit microprocessor. After data collection, a pump works at a rate of 10 ml/s to purge the chamber. Table VI summarizes the fundamental performance parameters of the proposed system.

3) Sampling Procedure: The sampling procedure, program controlled by the system to ensure all samples are sampled under the same criterion, involves two sub-procedures, a purge cycle and a sampling cycle. In the purge cycle, a pump pulls and purges the air over the sensor array, supplying background air to the array for the baseline measurement as well as refreshing it after sampling. In the sample cycle, the analyte is injected into the chamber. When the sensor array is exposed to the analyte, changes in resistances are measured and recorded. The action of the system is different in each time-slice (Fig. 5). The following explains this in detail.

1) $-10 \sim 0s$ (baseline stage): The chamber is purged and the sensor returns to a steady state. The baseline value is measured and recorded for data manipulation and normalization;
2) $1 \sim 5s$ (injection stage): Sampled gas is injected into the chamber at an invariable rate. Particles of sampled gas inside the chamber accrue during injection, producing a changing of resistance of the sensor and causing the amplitude of the signal to rise;
3) $6 \sim 10s$ (reaction stage): Particles in the chamber continue to accumulate on the sensors but the accumulation rate is decreasing. The resistance of the sensor monotonically increases at a decreasing rate, as does the amplitude of the signals;
4) $15 \sim 90s$ (purge stage): The chamber is purged again. The pump quickly draws out the remaining analyte, thereby shortening the sampling time as well as refreshing the air for the next use.

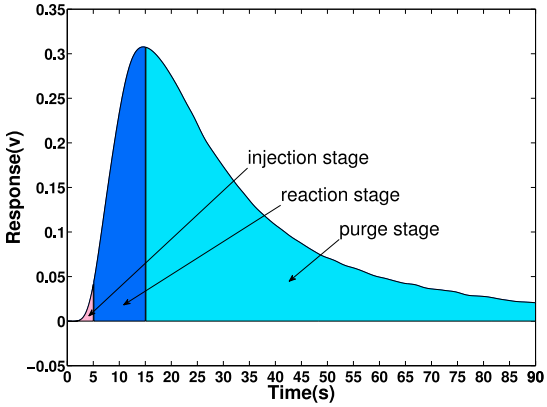


Fig. 5. A typical sensor response curve which undergoes three stages.

In our database, the characteristic curve of one sample is taken from the data for the period from 1 s to 90 s. Since the sampling frequency is 9 Hz, one sensor in one sample creates a $90 \times 9 = 810$-dimension feature vector.

C. Data Analysis

The system measures changes in voltage across each sensor and converts the raw signal into a digital value that can be applied to future analysis. This analysis involves three steps: signal preprocessing, feature extraction, and classification.

1) Signal Preprocessing: The purpose of signal preprocessing is to compensate for drift and eliminate irrelevant information so to improve the performance of the subsequent pattern recognition and classification. It involves baseline manipulation and normalization. Baseline manipulation is implemented for drift compensation, contrast enhancement, and scaling. Its basic idea is to subtract the baseline of each sensor from the sensor response. We assume that one data set has e samples, where $e = 1, ..., N_e$. Each sample consists of s sensor transients, where $s = 1, ..., N_s$. There are k dimensions per transient, where $k = 1, ..., N_k$. The dynamic response of one sample at time t_k is denoted as $R_{e,s}(t_k)$. There are b dimensions in baseline stage, where $b = 1, ..., N_b$. The baseline response of this sample is $B_{e,s}(t_b)$. The relative change for a particular sensor is defined as the preprocessed response

$$R^B{}_{e,s}(t_k) = R_{e,s}(t_k) - \frac{1}{N_b}\sum_{t_b=1}^{N_b} B_{e,s}(t_b), \quad \forall e, s, k, b. \quad (2)$$

Normalization is used to compensate for sample-to-sample variations caused by analyte concentration and pressure of oxygen (PO_2). $R^B{}_{e,s}(t_k)$ is the response of the sensor N_s to the N_e sample in data set, which has been processed by baseline manipulation. The normalized response is defined as

$$R^{BN}{}_{e,s}(t_k) = \frac{R^B{}_{e,s}(t_k)}{max(R^B{}_{e,s}(t_k))}, \quad \forall e, m. \quad (3)$$

2) Feature Extraction: The purpose of feature extraction is to find a low-dimensional mapping $f : x \in R^N \mapsto y \in R^M (M < N)$ that preserves most of information in the original feature vector x. In this paper, we employed principal components analysis (PCA) to extract characteristic features of samples from m classes. We calculated the eigenvectors and eigenvalues of the training set and sort the eigenvectors, i.e., principal components of PCA, by descendant eigenvalues, then projected both test data and training data onto the PCA subspace spanned by selected principal components. The criteria for principal component selection is

$$r_\lambda = \sum_{k=1}^{s} \lambda_i \Big/ \sum_{k=1}^{n} \lambda_i \ , \quad (4)$$

where r_i is the eigenvalue, s is the number of selected principal components, and n is the total number of eigenvalues. Assume $r_\lambda > 99\%$ counts for enough variability in the dataset, $s = 10$ eigenvectors as features. We hence formed a s-dimensional training vector space and test vector space respectively for data classification.

3) Classification: K nearest neighbor voting rule (KNN) was used as a classifier for the features that extracted by PCA. Basically, it classifies an unlabeled test sample by finding the K nearest neighbors in the training set using Euclidean distance and assigning the label of that class represented by a majority among the K neighbors [41]. There are many voting rule to decide which class the unlabeled sample belongs to. In our experiment, we used the following vote rule: assuming there are m classes and one sample has $K_1, K_2, ..., K_m$ nearest neighbors for the m classes, where $\sum_{i=1}^{m} K_i = K$, the classification result is given by

$$c = \arg\max_{i=1,...,m}\{\frac{K_i}{K}\}, \quad (5)$$

where c is the label of the predicted class. The training vectors were classified in advance into m classes, labeled as either healthy or diseased. The test vector was then predicted using Equation 5.

IV. Experiments

In the first experiment, we used our system to distinguish between pre- and post-treatment breath samples from 52 subjects with end-stage renal failure, a kind of condition associates with the accumulation of urinary waste products in the blood because the kidneys are not working effectively (Table VII). A standard treatment for the condition is hemodialysis to help patient remove more urea and creatinine from the blood. There is a reduction in the ammonia concentration in expired breath of patients as hemodialysis proceeds [42]. The results for these experiments are given in Section IV-A.

TABLE VII
Composition of the renal failure database

Type of subjects	Number	Male/Female	Age
Subjects with renal failure	52	33/19	34-70

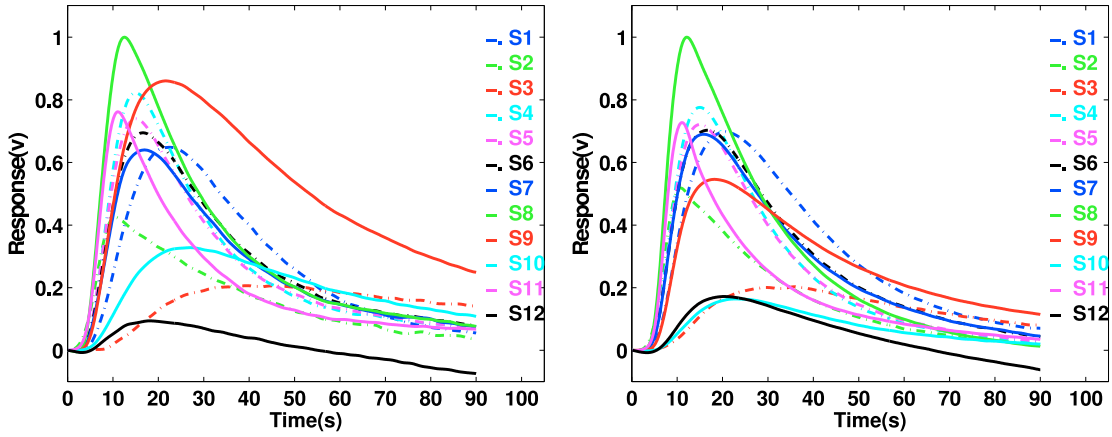
TABLE VIII
Composition of the subject database

Type of subjects	Number	Male/Female	Age
Healthy subjects	108	58/50	23-60
Subjects with diabetes	117	65/52	32-70
Subjects with renal disease	110	63/47	28-70
Subjects with airway inflammation	110	54/56	16-62

In the second set of experiments, we tested the ability of the system to distinguish between subjects assumed to be healthy on the basis of recent health check and subjects known to be afflicted with either diabetes, renal disease, or airway inflammation. Totally, We collected 108 healthy samples, 117 diabetes samples, 110 renal disease samples, and 110 airway inflammation samples using the gas collection and signal sampling procedures described in Sections III-A and III-B. Table VIII details the composition of the subject database. All patients were inpatient volunteers from the Harbin Hospital. Their conditions were confirmed and correlated by comparing their levels with standard clinical blood markers for the relevant diseases and conditions. In each case, these diseases and conditions are associated with characteristic molecules in the breath. Diabetes arises when the glucose produced by the body cannot enter cells and so cells have to use fat as an energy source. One of the by-products of metabolizing fat for energy is ketones. When ketones accumulate in the blood, there is ketoacidosis, which is characterized by the smell of acetone on the patient's breath [43]. Renal disease arise from the inability of the kidneys to effectively filter the blood. This results in an accumulation of nitrogen-bearing waste products (urea), which accounts for the odor of ammonia on the breath of patients [44]. As for airway inflammation, it has been shown that exhaled nitric oxide levels are higher when there is airway inflammation, especially asthmatic airway inflammation [45].

A. Evaluating Outcomes of Hemodialysis

Fig. 6 shows the responses of the twelve different sensors (S1-S12) to the samples of renal failure patients over the 90 s sampling period. Fig. 6(a) shows a typical response of one patient before hemodialysis and Fig. 6(b) shows a response of the same patient after hemodialysis. The curves represent the output of each sensor, S1-S12. These curves have been preprocessed by baseline manipulation and normalization according to Equations 2 and 3. As shown in Table II, the dominant compound marking renal disease is ammonia. From



(a) (b)

Fig. 6. Typical responses from the same patient: (a) before treatment, (b) after treatment. The horizontal axis stands for the sampling time (0-90 s) and the vertical axis shows the amplitude of the sensor output in volts.

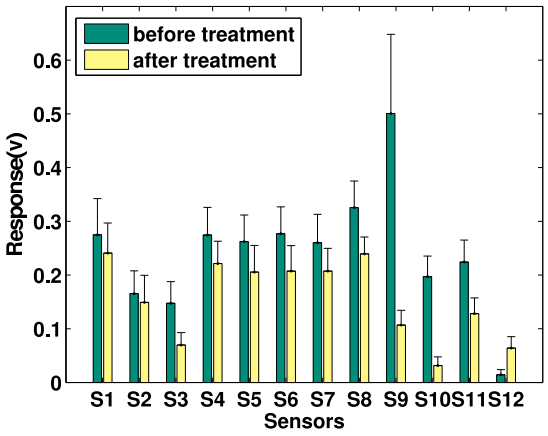


Fig. 7. Mean response of twelve sensors to two cases: before treatment and after treatment. The error bar represents the standard deviation (only the upper bar is drawn). The horizontal axis denotes the twelve sensors and the vertical axis stands for the mean value of each normalized response.

these figures, it is very clear that before hemodialysis (Fig. 6(a)), the amplitude of the ninth sensor (red solid line) is very high, which indicates that the concentration of ammonia in the breath is quite large. However, after treatment (Fig. 6(b)), the amplitude of the ninth sensor clearly decreases, indicating the concentration of ammonia in the subject's breath has fallen.

Fig. 7 presents the mean responses of the twelve sensors showing the response of each sensor to two kinds of samples. The error bar represents the standard deviation, showing the difference between the responses of all samples in one classes and their mean. The mean response is defined as

$$MeanR_{e,s} = \frac{1}{N_k} \sum_{t_k=1}^{N_k} R_{e,s}(t_k), \quad \forall e, s, k. \tag{6}$$

The definitions of these variables are the same as Section III-C1. For each sensor, the left bar presents the class before hemodialysis and the right bar is the class after hemodialysis. After the treatment, the values of several responses clearly fall, especially the ninth sensor, which is sensitive to ammonia.

B. Distinguishing between Subject Breath Samples

Fig. 8 shows the responses of twelve different sensors (S1-S12) to the four different air samples over the 90 s sampling period. Fig. 8(a) is a typical response to a healthy sample. Fig. 8(b) is to a diabetes sample. Fig. 8(c) is to a renal disease sample. And Fig. 8(d) is to an airway inflammation sample. The curves represent the output of each sensor.

Fig. 9 gives the mean responses of the twelve sensors in the four types of air samples as defined by Equation 6. The definition of error bar is the same as that in Section IV-A. In each of the four categories it is possible to find the combinations of sensors which could unambiguously identify each of the four conditions. Thus, the strongest responses to healthy samples came from the sixth, seventh, and eighth sensors while the strongest response to diabetes came from the second, fourth, fifth and twelfth sensors. It is worth mentioning that the twelfth sensor gave a very significant response, though it is not used for VOCs detection. In China, the special diet recommended for diabetics features large amounts of fermentable dietary fiber, which leads to colonic fermentation of indigestible carbohydrates [46]. One product of colonic fermentation is hydrogen [47], which is absorbed into the bloodstream and excreted through the breath. Therefore, the breath air of diabetics we have collected would include hydrogen. The strong response to the renal disease samples came from the first, third, ninth and eleventh sensors, especially the ninth sensor, which is particularly sensitive to ammonia. The largest response to airway inflammation came from the tenth sensor, which is used to detect nitric oxide.

V. Results and Discussion

The outcomes of both sets of experiments were evaluated using PCA coupled with KNN, as introduced in Sections III-C2 and III-C3.

A. Results Evaluating Outcomes of Hemodialysis

Fig. 10 shows a two-dimensional PCA analysis of the responses with the first principal component (PC1) plotted

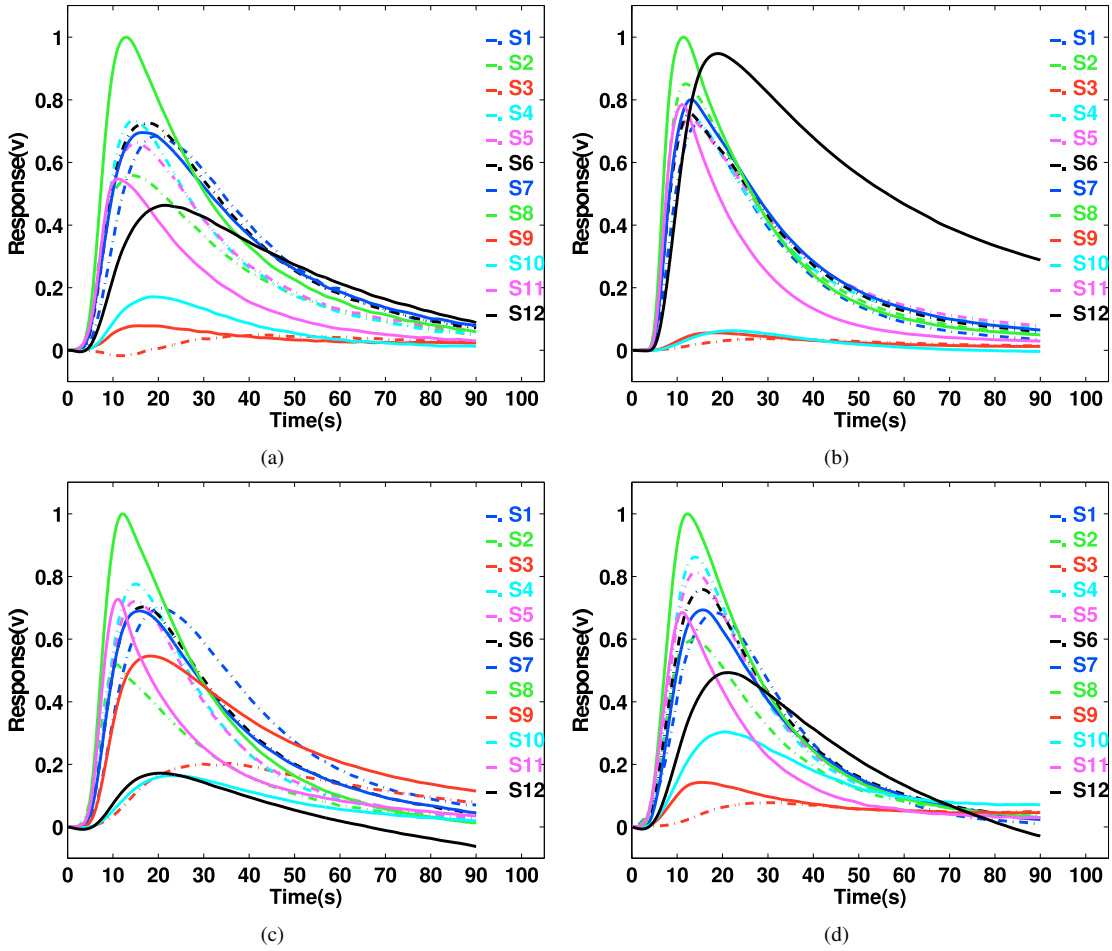


Fig. 8. Typical responses from four subject categories: (a) healthy subjects, (b) subjects with diabetes, (c) subjects with renal disease, and (d) subjects with airway inflammation. The horizontal axis stands for the sampling time (0-90 s) and the vertical axis denotes the amplitude of sensor output in volts.

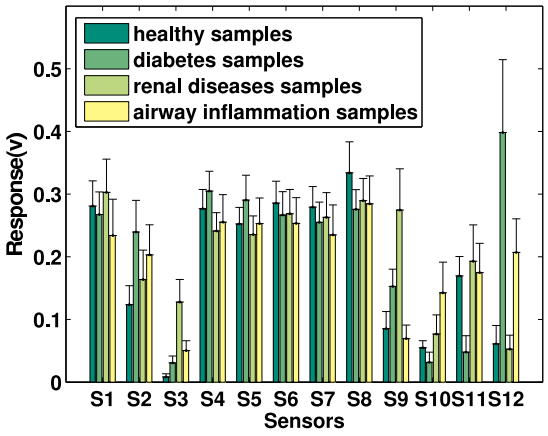


Fig. 9. Mean response of twelve sensors to four classes: healthy, diabetes, renal disease, and airway inflammation. The error bar represents the standard deviation (only the upper bar is drawn). The horizontal axis denotes the twelve sensors and the vertical axis stands for the mean value of each normalized response.

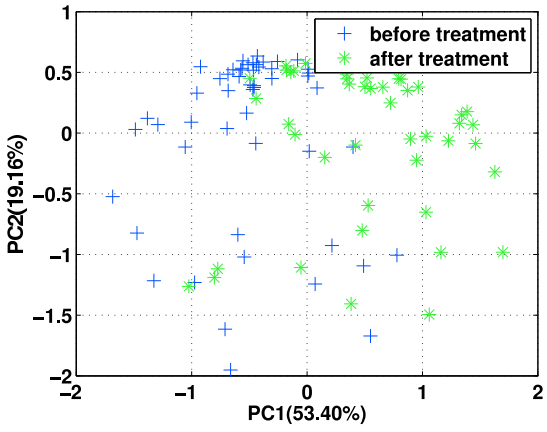


Fig. 10. PCA two-dimensional plot of the sensor signals corresponding to two classes: (a) renal failure samples before treatment($green*$), (b) renal failure samples after treatment($blue+$).

against the second (PC2). $green*$ stands for the samples before treatment and $blue+$ for the samples after treatment. The two dimensions explained 73.01% of the variation in the data, 53.4% for PC1 and 19.61% for PC2. The two classes are discriminative even though some samples overlap.

To measure the classification accuracy of system, we randomly selected a training set of 26 samples from each disease class of 52 samples. The remainder was used as the test set. PCA was used to extract characteristic features of the samples. We calculated the eigenvectors and eigenvalues of the training

TABLE IX
CLASSIFICATION RESULTS OF TWO CLASSES: RENAL FAILURE SAMPLES BEFORE TREATMENT AND AFTER TREATMENT

Actual group member	Number of samples		Predicted group member (average)		Accuracy
	Training set	Test set	Before treatment	After treatment	
Before treatment	26	26	20.84	5.16	80.15%
After treatment	26	26	21.32	4.68	82%

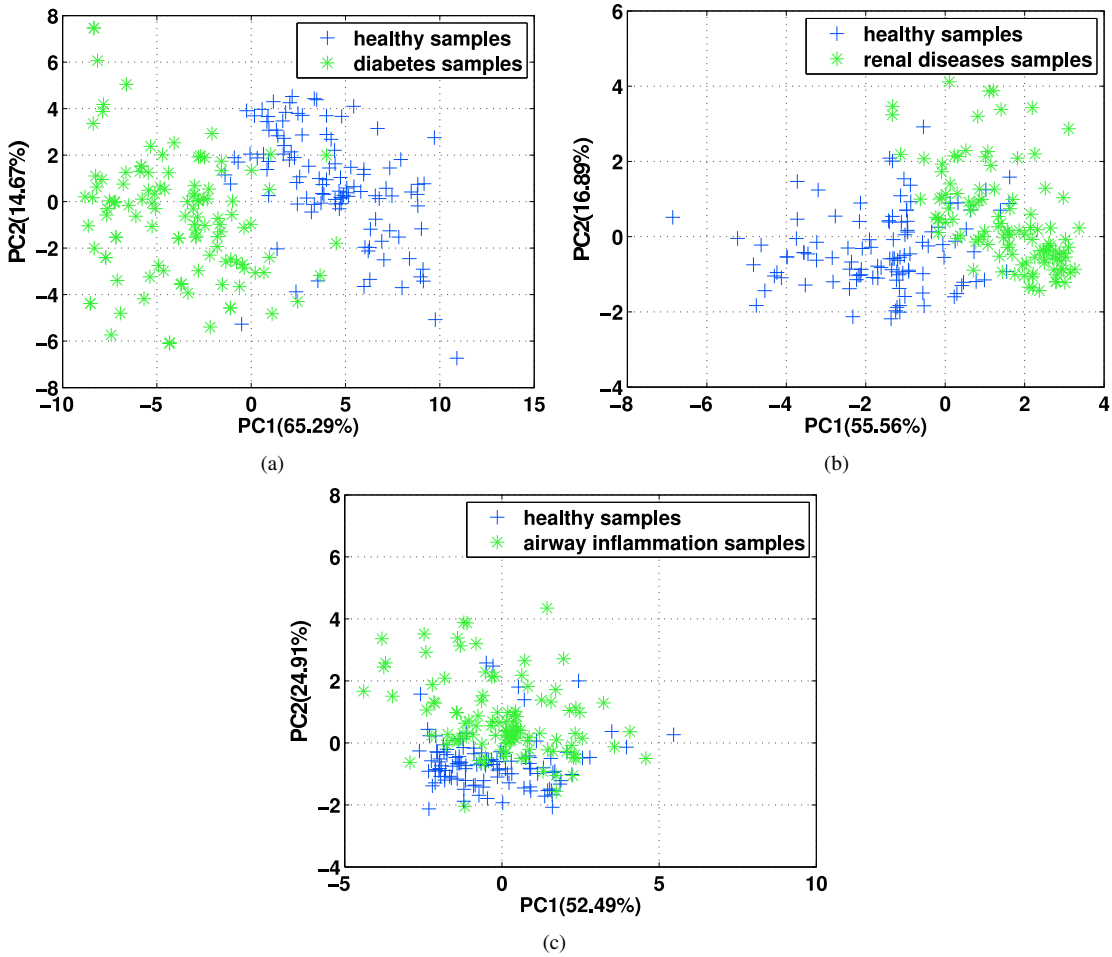


Fig. 11. PCA two-dimensional plot of the sensor signals corresponding to two classes: (a) healthy samples ($blue+$) and diabetes samples ($green*$), (b) healthy samples ($blue+$) and renal disease samples ($green*$), and (c) healthy samples ($blue+$) and airway inflammation samples ($green*$).

TABLE X
THE DEFINITION OF SENSITIVITY AND SPECIFICITY

		Test outcome		Sensitivity	Specificity
		Positive	**Negative**		
Actual condition	**Positive**	tp	fp	$\frac{tp}{tp+fn}$	$\frac{tn}{tn+fp}$
	Negative	fn	tn		

set and sorted the eigenvectors by descendant eigenvalues. We then used Equation 4 and the condition $r_\lambda > 99\%$ to select the first 12 eigenvectors as principle components. Next, we projected all samples onto the PCA subspace spanned by principal components. Then, KNN (K = 5) classifier defined by Equation 5 predicted the class that a test sample belonged to. We ran this procedure 50 times and computed the average classification rate over all 50 runs.

Table IX shows the classification results. In the 26-sample pre-treatment test set, an average of 20.84 samples were classified correctly and 5.16 samples were classified incorrectly, an overall accuracy of 80.15%. In the 26-sample post-treatment test set, an average of 21.32 samples were classified correctly and 4.68 samples were classified incorrectly, an overall accuracy of 82%. Clearly, the proposed system would have some value in evaluating the efficacy of hemiodialysis, and may take the place in some cases of blood tests, given that it is simple, low-cost, and non-invasive.

B. Results Distinguishing between Subject Breath Samples

The classifications of the four types of breath samples were evaluated with PCA coupled with KNN and the results were measured by sensitivity and specificity. The samples from patients with diabetes, renal disease, and airway inflammation

TABLE XI
THE CLASSIFICATION RESULTS DEFINED BY SENSITIVITY AND SPECIFICITY

		Training set	Test set	Test outcome (average)		Sensitivity	Specificity
				Positive	Negative		
Diabetes	**Positive**	57	60	52.6	7.4	86.97%	87.57%
	Negative	48	60	7.88	52.12		
Renal failure	**Positive**	50	60	51.94	8.06	83.96%	86.14%
	Negative	48	60	9.92	50.08		
Airway inflammation	**Positive**	50	60	42.12	17.88	73.79%	71.58%
	Negative	48	60	14.96	45.04		

and the healthy samples were formed three groups. One group contained healthy subjects and subjects with diabetes, the second contained healthy subjects and subjects with renal disease, and the third group contained healthy subjects and subjects with airway inflammation.

Fig. 11 shows the PCA two-dimensional plot of the responses from the two classes with the first principal component (PC1) plotted against the second (PC2). The $green*$ stands for the samples classified as being from patients and $blue+$ for healthy subjects. In the PCA plot of diabetes samples and healthy samples, the two dimensions explained 79.96% of the variation in the data, 65.29% for PC1 and 14.67% for PC2. In the PCA plot of renal disease samples and healthy samples, the two dimensions explained 72.45% of the variation in the data, 55.56% for PC1 and 16.89% for PC2. In the PCA plot of airway inflammation samples and healthy samples, the two dimensions explained 77.4% of the variation in the data, 52.49% for PC1 and 24.91% for PC2.

To compare the test results, we randomly selected 60 samples from each class as the test set and the remainder formed the training set. PCA was used to extract characteristic features of samples. Equation 4 and the condition $r_\lambda > 99\%$ were used to select the first 10 eigenvectors in all classes in every group. The KNN (K = 5) classifier as defined by Equation 5 was then used to determine which class each test sample belonged to.

In medicine, the reliability of a diagnosis is measured in terms of sensitivity and specificity, with the outcome being either positive (unhealthy) or negative (healthy). In the classification, the number of genuine sick subjects is denoted tp; misidentified healthy subjects is fp; genuine healthy subjects is tn; the misdiagnosed sick subjects is denoted as tn [10]. Sensitivity and specificity are thus defined as in Table X. Table XI shows the classification results of all groups.

In the diabetes experiment, out of 60 samples in the test set, the system correctly diagnosed an average of 52.6 samples as diabetes and incorrectly diagnosed 7.4 samples as healthy. In the 60 healthy samples in the test set, an average of 52.12 samples were correctly diagnosed as healthy and 7.88 were incorrectly diagnosed as diabetes. The sensitivity of this diagnosis was thus 86.97% and the specificity was 87.57%.

In the renal disease experiment, an average of 51.94 disease samples were correctly diagnosed as renal disease and 8.06 disease samples were incorrectly diagnosed as healthy. In the 60 healthy samples in the test set, an average of 50.08 healthy samples were correctly diagnosed as healthy; while an average of 9.92 healthy samples were incorrectly diagnosed as renal disease. Consequently, the sensitivity and specificity of this diagnosis were 83.96% and 86.14% respectively.

Same as above, in the experiment of airway inflammation diagnosis, there were averagely 42.12 disease samples diagnosed correctly as airway inflammation and 17.88 disease samples diagnosed incorrectly as healthy, and there were averagely 14.96 healthy samples diagnosed incorrectly as airway inflammation and 45.04 healthy samples diagnosed correctly as healthy. The sensitivity of this diagnosis was thus 73.79% and the specificity was 71.58%.

VI. CONCLUSION

This article proposed a breath analysis system that has a broad application in medicine, such as detecting diseases and monitoring the progress of related therapies. The system structure, working procedure, odor signal preprocessing, and pattern recognition method were introduced. To evaluate the system performance, breath samples were captured and two experiments were conducted on medical treatment evaluation and disease identification. The results showed that the system was not only able to distinguish between breath samples from subjects suffering from various diseases or conditions (diabetes, renal disease, and airway inflammation) and breath samples from healthy subjects, but in the case of renal failure was also helpful in evaluating the efficacy of hemodialysis.

Although the current pattern recognition method produced satisfactory results when we used integral data, it should still be possible to further improve the classification accuracy and speed by selecting proper features. Typically, the performance of an electronic olfaction system depends heavily on the features being provided to the odor classification algorithm. Therefore, in future work we will investigate how to select the most proper features for effective pattern classification. We also intend to extend the number of diseases/conditions that the system can analyze.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments.

REFERENCES

[1] F. Di Francesco, R. Fuoco, M. Trivella, and A. Ceccarini, "Breath analysis: trends in techniques and clinical applications," *Microchemical Journal*, vol. 79, no. 1-2, pp. 405–410, 2005.

[2] R. Dweik and A. Amann, "Exhaled breath analysis: the new frontier in medical testing," *Journal of Breath Research*, vol. 2, pp. 1–3, 2008.

[3] J. Van Berkel, J. Dallinga, G. Möller, R. Godschalk, E. Moonen, E. Wouters, and F. Van Schooten, "Development of accurate classification method based on the analysis of volatile organic compounds from human exhaled air," *Journal of Chromatography B*, vol. 861, no. 1, pp. 101–107, 2008.

[4] M. Phillips, "Method for the collection and assay of volatile organic compounds in breath," *Analytical biochemistry*, vol. 247, no. 2, pp. 272–278, 1997.

[5] E. Thaler and C. Hanson, "Medical applications of electronic nose technology," *Expert Review of Medical Devices*, vol. 2, no. 5, pp. 559–566, 2005.

[6] A. Amann, G. Poupart, S. Telser, M. Ledochowski, A. Schmid, and S. Mechtcheriakov, "Applications of breath gas analysis in medicine," *International Journal of Mass Spectrometry*, vol. 239, no. 2-3, pp. 227–233, 2004.

[7] F. Rock, N. Barsan, and U. Weimar, "Electronic nose: Current status and future trends," *Chem. Rev*, vol. 108, no. 2, pp. 705–725, 2008.

[8] Y. Lin, H. Guo, Y. Chang, M. Kao, H. Wang, and R. Hong, "Application of the electronic nose for uremia diagnosis," *Sensors & Actuators: B. Chemical*, vol. 76, no. 1-3, pp. 177–180, 2001.

[9] J. Yu, H. Byun, M. So, and J. Huh, "Analysis of diabetic patient's breath with conducting polymer sensor array," *Sensors & Actuators: B. Chemical*, vol. 108, no. 1-2, pp. 305–308, 2005.

[10] R. Blatt, A. Bonarini, E. Calabro, M. Della Torre, M. Matteucci, and U. Pastorino, "Lung Cancer Identification by an Electronic Nose based on an Array of MOS Sensors," in *Neural Networks, 2007. IJCNN 2007. International Joint Conference on*, 2007, pp. 1423–1428.

[11] S. Dragonieri, R. Schot, B. Mertens, S. Le Cessie, S. Gauw, A. Spanevello, O. Resta, N. Willard, T. Vink, K. Rabe *et al.*, "An electronic nose in the discrimination of patients with asthma and controls," *The Journal of Allergy and Clinical Immunology*, vol. 120, no. 4, pp. 856–862, 2007.

[12] W. Cao and Y. Duan, "Current status of methods and techniques for breath analysis," *Critical Reviews in Analytical Chemistry*, vol. 37, no. 1, pp. 3–13, 2007.

[13] M. Phillips, J. Herrera, S. Krishnan, M. Zain, J. Greenberg, and R. Cataneo, "Variation in volatile organic compounds in the breath of normal humans," *Journal of Chromatography B: Biomedical Sciences and Applications*, vol. 729, no. 1-2, pp. 75–88, 1999.

[14] W. Miekisch and J. Schubert, "From highly sophisticated analytical techniques to life-saving diagnostics: Technical developments in breath analysis," *Trends in Analytical Chemistry*, vol. 25, pp. 665–673, 2006.

[15] T. Risby and S. Solga, "Current status of clinical breath analysis," *Applied Physics B: Lasers and Optics*, vol. 85, no. 2, pp. 421–426, 2006.

[16] S. Sehnert, L. Jiang, J. Burdick, and T. Risby, "Breath biomarkers for detection of human liver diseases: preliminary study," *Biomarkers*, vol. 7, no. 2, pp. 174–187, 2002.

[17] A. DAmico, C. Di Natale, R. Paolesse, A. Macagnano, E. Martinelli, G. Pennazza, M. Santonico, M. Bernabei, C. Roscioni, G. Galluccio *et al.*, "Olfactory systems for medical applications," *Sensors & Actuators: B. Chemical*, vol. 130, no. 1, pp. 458–465, 2007.

[18] J. Schubert, W. Miekisch, K. Geiger, and G. Nöldge-Schomburg, "Breath analysis in critically ill patients: potential and limitations," *Expert Rev. Mol. Diagn*, vol. 4, no. 5, pp. 619–629, 2004.

[19] W. Miekisch, J. Schubert, and G. Noeldge-Schomburg, "Diagnostic potential of breath analysis-focus on volatile organic compounds," *Clinica Chimica Acta*, vol. 347, no. 1-2, pp. 25–39, 2004.

[20] A. Deykin, A. Massaro, J. Drazen, and E. Israel, "Exhaled nitric oxide as a diagnostic test for asthma: online versus offline techniques and effect of flow rate," *American journal of respiratory and critical care medicine*, vol. 165, no. 12, pp. 1597–1601, 2002.

[21] L. McGrath, R. Patrick, P. Mallon, L. Dowey, B. Silke, W. Norwood, and S. Elborn, "Breath isoprene during acute respiratory exacerbation in cystic fibrosis," *European Respiratory Journal*, vol. 16, no. 6, pp. 1065–1069, 2000.

[22] M. Phillips, M. Sabas, and J. Greenberg, "Increased pentane and carbon disulfide in the breath of patients with schizophrenia," *Journal of clinical pathology*, vol. 46, no. 9, pp. 861–864, 1993.

[23] M. Phillips, N. Altorki, J. Austin, R. Cameron, R. Cataneo, J. Greenberg, R. Kloss, R. Maxfield, M. Munawar, H. Pass *et al.*, "Prediction of lung cancer using volatile biomarkers in breath," *Cancer Biomarkers*, vol. 3, no. 2, pp. 95–109, 2007.

[24] C. Deng, J. Zhang, X. Yu, W. Zhang, and X. Zhang, "Determination of acetone in human breath by gas chromatography-mass spectrometry and solid-phase microextraction with on-fiber derivatization," *Journal of Chromatography B*, vol. 810, no. 2, pp. 269–275, 2004.

[25] S. Davies, P. Spanel, and D. Smith, "Quantitative analysis of ammonia on the breath of patients in end-stage renal failure," *Kidney international*, vol. 52, no. 1, pp. 223–228, 1997.

[26] M. Phillips, R. Cataneo, R. Condos, G. Ring Erickson, J. Greenberg, V. La Bombardi, M. Munawar, and O. Tietje, "Volatile biomarkers of pulmonary tuberculosis in the breath," *Tuberculosis*, vol. 87, no. 1, pp. 44–52, 2007.

[27] M. Phillips, R. Cataneo, B. Ditkoff, P. Fisher, J. Greenberg, R. Gunawardena, C. Kwon, F. Rahbari-Oskoui, and C. Wong, "Volatile markers of breast cancer in the breath," *The breast journal*, vol. 9, no. 3, pp. 184–191, 2003.

[28] M. Phillips, R. Cataneo, A. Cummin, A. Gagliardi, K. Gleeson, J. Greenberg, R. Maxfield, and W. Rom, "Detection of Lung Cancer With Volatile Markers in the Breath," *Chest*, vol. 123, no. 6, pp. 2115–2123, 2003.

[29] M. Salazar, "Breath Markers of Oxidative Stress in Patients with Unstable Angina," *Heart Disease*, vol. 5, no. 2, pp. 95–99, 2003.

[30] M. Phillips, J. Boehmer, R. Cataneo, T. Cheema, H. Eisen, J. Fallon, P. Fisher, A. Gass, J. Greenberg, J. Kobashigawa *et al.*, "Heart allograft rejection: detection with breath alkanes in low levels (the HARDBALL study)," *Journal of Heart and Lung Transplantation*, vol. 23, no. 6, pp. 701–708, 2004.

[31] Z. Weitz, A. Birnbaum, P. Sobotka, E. Zarling, and J. Skosey, "High breath pentane concentrations during acute myocardial infarction," *Lancet*, vol. 337, no. 8747, pp. 933–935, 1991.

[32] C. Olopade, M. Zakkar, W. Swedler, and I. Rubinstein, "Exhaled pentane levels in acute asthma," *Chest*, vol. 111, no. 4, pp. 862–865, 1997.

[33] S. Humad, E. Zarling, M. Clapper, and J. Skosey, "Breath pentane excretion as a marker of disease activity in rheumatoid arthritis," *Free Radical Research*, vol. 5, no. 2, pp. 101–106, 1988.

[34] S. Sedghi, A. Keshavarzian, M. Klamut, D. Eiznhamer, and E. Zarling, "Elevated breath ethane levels in active ulcerative colitis: evidence for excessive lipid peroxidation," *American Journal of Gastroenterology*, vol. 89, no. 12, pp. 2217–2221, 1994.

[35] E. Baraldi and S. Carraro, "Exhaled NO and breath condensate," *Paediatric Respiratory Reviews*, vol. 7, pp. 20–22, 2006.

[36] S. Kharitonov, A. Wells, B. O'connor, P. Cole, D. Hansell, R. Logan-Sinclair, and P. Barnes, "Elevated levels of exhaled nitric oxide in bronchiectasis," *American journal of respiratory and critical care medicine*, vol. 151, no. 6, pp. 1889–1893, 1995.

[37] I. Horvath, S. Loukides, T. Wodehouse, S. Kharitonov, P. Cole, and P. Barnes, "Increased levels of exhaled carbon monoxide in bronchiectasis: a new marker of oxidative stress," *British Medical Journal*, vol. 53, no. 10, pp. 867–870, 1998.

[38] W. Maziak, S. Loukides, S. Culpitt, P. Sullivan, S. Kharitonov, and P. Barnes, "Exhaled nitric oxide in chronic obstructive pulmonary disease," *American journal of respiratory and critical care medicine*, vol. 157, no. 3, pp. 998–1002, 1998.

[39] M. Barker, M. Hengst, J. Schmid, H. Buers, B. Mittermaier, D. Klemp, and R. Koppmann, "Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis," *European Respiratory Journal*, vol. 27, no. 5, pp. 929–936, 2006.

[40] A. Turner and N. Magan, "Electronic noses and disease diagnostics," *Nature Reviews Microbiology*, vol. 2, no. 2, pp. 161–166, 2004.

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

[42] L. Narasimhan, W. Goodman, and C. Patel, "Correlation of breath ammonia with blood urea nitrogen and creatinine during hemodialysis," *Proceedings of the National Academy of Sciences*, vol. 98, no. 8, pp. 4617–4621, 2001.

[43] R. Holt, N. Hanley, and C. Brook, "Essential endocrinology and diabetes," 2006.

[44] A. Greenberg and A. Cheung, *Primer on kidney diseases*. WB Saunders Co, 2005.

[45] K. Ashutosh, "Nitric oxide and asthma: a review," *Current opinion in pulmonary medicine*, vol. 6, no. 1, pp. 21–25, 2000.

[46] F. Brighenti, L. Benini, D. Del Rio, C. Casiraghi, N. Pellegrini, F. Scazzina, D. Jenkins, and I. Vantini, "Colonic fermentation of indigestible carbohydrates contributes to the second-meal effect," *American Journal of Clinical Nutrition*, vol. 83, no. 4, pp. 817–822, 2006.

[47] L. Le Marchand, L. Wilkens, P. Harwood, and R. Cooney, "Breath hydrogen and methane in populations at different risk for colon cancer," *International Journal of Cancer*, vol. 55, no. 6, pp. 887–890, 2006.

Dongmin Guo received the B.S. degree in Automation and the M.S. degree in Precise Instrument and Mechanics from Northwestern Polytechnical University, Xi'an, China, in 2003 and 2006, respectively. She is currently working toward the Ph.D. degree with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong.

From 2006 to 2008, she was a software engineer at UMTS RNC department of Alcatel-Lucent, Shanghai. Her research interests include medical diagnosis, pattern recognition, and bioinformatics.

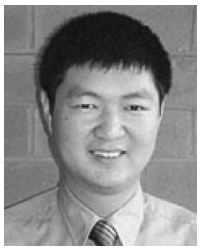
David Zhang (F'09) received the degree in Computer Science from Peking University, Beijing, China, the M.Sc. degree in Computer Science and the Ph.D. degree from Harbin Institute of Technology (HIT), Harbin, China, in 1982 and 1985, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Ontario, Canada, in 1994.

From 1986 to 1988, he was a Postdoctoral Fellow with Tsinghua University, Beijing, and then an Associate Professor with the Academia Sinica, Beijing. He is currently the Head of the Department of Computing and a Chair Professor with the Hong Kong Polytechnic University, Hong Kong, where he is the Founding Director of the Biometrics Technology Centre (UGC/CRC) supported by the Hong Kong SAR Government in 1998. He also serves as a Visiting Chair Professor with Tsinghua University and an Adjunct Professor with Peking University; Shanghai Jiao Tong University, Shanghai, China; HIT; and the University of Waterloo. He is the Founder and Editor-in-Chief for the International Journal of Image and Graphics, a Book Editor for the Springer International Series on Biometrics (KISB), and an Associate Editor for more than ten international journals, including IEEE TRANSACTIONS and Pattern Recognition. He is the author of more than ten books and 200 journal papers.

Prof. Zhang is a Croucher Senior Research Fellow, a Distinguished Speaker of the IEEE Computer Society, and a Fellow of the International Association of Pattern Recognition. He is the Organizer of the International Conference on Biometrics Authentication.

Naimin Li graduated from the Department of Medical Treatment in the Traditional Chinese Medicine (TCM) training class, Shenyang, China, in 1962.

He is a Fellow Professor at the Harbin Institute of Technology, Harbin, China. He began the study and exploration of the application of tongue diagnosis in modern disease diagnosis and treatment since 1965, and put it forward to the fields of internal medicine, pediatrics, epidemiology and gynecology. During 1970 and 1972, he was appointed as a Leader of Experts Group of TCM by the Chinese government in the Middle-Europe and Albania. In 1989, the first tongue image laboratory in China was set up under his guidance. So far, he has authored and coauthored over 250 papers and 14 books around his research areas.



Lei Zhang (M'04) received the B.S. degree from Shenyang Institute of Aeronautical Engineering, Shenyang, China, in 1995 and the M.S. and Ph.D. degrees in Control Theory and Applications from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

From 2001 to 2002, he was a Research Associate with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong. From January 2003 to January 2006, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. Since January 2006, he has been an Assistant Professor with the Department of Computing, The Hong Kong Polytechnic University. His research interests include image and video processing, biometrics, pattern recognition, multisensor data fusion, optimal estimation theory, etc.

Jianhua Yang received the B.S. degree in Computer Software from Xi'dian University, Xi'an, China, in 1989 and the M.S. degree in Underwater Acoustic Engineering and Ph.D. degrees in Detection Technique and Automation Device from Northwestern Polytechnical University, Xi'an, China, in 1992 and 2002, respectively.

From 1992 to 1999, she was a Lecture with the School of Automation, Northwestern Polytechnical University, Xi'an, China. From 1999 to 2001, she worked as a Visiting Researcher in Multigassensors GmbH, Jena, Germany and then was a Research Fellow with Optoelectronics Research Centre, The University of Southampton, Southampton, United Kingdom. In 2008, she was a Research Fellow with the Department of Computing, The Hong Kong Polytechnic University. She is currently the Associate Dean and Professor of the School of Automation, Northwestern Polytechnical University, Xi'an, China. Her research interests include electronic nose technique, integration testing technique, embedded computer and its application, etc.