Design of a Breath Analysis System for Diabetes Screening and Blood Glucose Level Prediction

Ke Yan, David Zhang*, *Fellow, IEEE*, Darong Wu, Hua Wei, and Guangming Lu

***Abstract*—It has been reported that concentrations of several biomarkers in diabetics' breath show significant difference from those in healthy people's breath. Concentrations of some biomarkers are also correlated with the blood glucose levels (BGLs) of diabetics. Therefore, it is possible to screen for diabetes and predict BGLs by analyzing one's breath. In this paper, we describe the design of a novel breath analysis system for this purpose. The system uses carefully selected chemical sensors to detect biomarkers in breath. Common interferential factors, including humidity and the ratio of alveolar air in breath, are compensated or handled in the algorithm. Considering the intersubject variance of the components in breath, we build subject-specific prediction models to improve the accuracy of BGL prediction. A total of 295 breath samples from healthy subjects and 279 samples from diabetic subjects were collected to evaluate the performance of the system. The sensitivity and specificity of diabetes screening are 91.51% and 90.77%, respectively. The mean relative absolute error for BGL prediction is 21.7%. Experiments show that the system is effective and that the strategies adopted in the system can improve its accuracy. The system potentially provides a noninvasive and convenient method for diabetes screening and BGL monitoring as an adjunct to the standard criteria.**

***Index Terms*—Blood glucose level (BGL), breath analysis, chemical sensors, diabetes screening, electronic noses.**

I. Introduction

DIABETES has become a great threat to human health. The timely diagnosis and frequent monitoring are important for managing the disease. To diagnose or monitor diabetes, traditionally, one must draw blood samples to check if his blood glucose level (BGL) falls within the normal range. This method is accurate but painful, invasive, and inconvenient [1]. Therefore, noninvasive diabetes screening and BGL prediction is arousing more and more interest recently. Approaches including reverse iontophoresis, fluorescence technology, bioimpedance spectroscopy, and so on [2], [3] have been studied. These approaches are painless and convenient, but still suffer disadvantages such as lack of specificity, inaccuracy due to subject's movement and sweating, skin irritation, etc. [2], [3].

Breath analysis is a noninvasive approach for clinical applications. By analyzing the concentrations of the biomarkers in breath, we are able to detect disease, monitor disease progression, or monitor therapy [4]. Lots of efforts have been devoted to study the breath biomarkers of diabetes. Acetone [5]–[8] as well as many other volatile organic compounds (VOCs) [9] in breath are proved to either have abnormal concentrations in diabetics or correlate with BGL. Compared to other approaches, breath analysis is readily acceptable and easy to collect samples [10], which makes it an attractive way for noninvasive diabetes screening and BGL prediction [1], [10].

Gas chromatography, mass spectroscopy (GC/MS), and related techniques can be used to analyze components in breath. For example, proton-transfer-reaction mass spectrometry (PTR-MS) has been applied to measure acetone during exercise and sleep [11], [12]. GC/MS-related methods have high accuracy but relatively high cost, low portability, and complex usage, which limits their applications in massive diabetes screening and household BGL monitoring. Another breath analysis method makes use of chemical sensor systems, also known as electronic noses (e-noses), which are generally cheaper, faster, more portable, and easier to operate. With the development in sensor technology, their accuracy has been improving. They have been applied in medicine for bacteria identification [13], diagnosis of renal disease [14], diabetes, airway inflammation [15], and present satisfactory performance.

A few chemical sensor systems have been developed for either diabetes diagnosis [15]–[18] or BGL prediction [19], [20]. However, there are still problems not solved in these systems. First, the chemical sensor array should be further optimized for the specific application. Second, in previous systems, common fluctuations in breath samples were not well taken into account, such as humidity and the ratio of alveolar air. More importantly, considering the intersubject variance of the components in breath, subject-specific BGL prediction models should be built [1]. Furthermore, the number of samples in the experiments of previous studies are small.

In this paper, a novel breath analysis system for both diabetes screening and BGL prediction is proposed. The sensors

Manuscript received February 25, 2014; revised April 6, 2014; accepted April 28, 2014. Date of publication June 9, 2014; date of current version October 16, 2014. This work was supported in part by the Natural Science Foundation of China under Grants 61332011, 61020106004, 61272292, and 61271344, the GRF fund from the HKSAR Government, the central fund from Hong Kong Polytechnic University, The National Basic Research Program of China (973 Program: 2011CB505404), Shenzhen Fundamental Research fund (JCYJ20130401152508661), and Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China. *Asterisk indicates corresponding author.*

K. Yan is with the Department of Electronic Engineering, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China (e-mail: yank10@mails.tsinghua.edu.cn).

*D. Zhang is with the Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China, and also with the Department of Computing, Biometrics Research Centre, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: csdzhang@comp.polyu.edu.hk).

D. Wu and H. Wei are with the Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou 510120, China (e-mail: drdarongwu@gmail.com; zhongxywh@163.com).

G. Lu is with the Biocomputing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China (e-mail: luguangm@gmail.com).

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

Digital Object Identifier 10.1109/TBME.2014.2329753

TABLE I
BASIC PARAMETERS OF THE PROPOSED DEVICE

Device Parameters	Specifications
Size	22 cm × 15 cm × 11 cm
Sampling duration	144 s
Sampling frequency	8 Hz
Injection flow rate	50 mL/s
Chamber volume	100 mL
Number of sensors	11

array is carefully selected with the help of two pilot devices to improve the accuracy. Some of the sensors are under temperature modulation, which is an effective technique to enrich the information content and enhance the selectivity of gas sensors. A temperature-humidity sensor and a carbon dioxide sensor are used to compensate for fluctuations in breath samples. In the algorithm of BGL prediction, subject-specific prediction models are built by incorporating subject identity information into the feature vector. The purpose of these optimization strategies is to enhance the accuracy and robustness of the system.

To evaluate the proposed system, a series of experiments were made. Experiments with simulated samples confirm the system's capability in predicting the concentration of acetone with the presence of interference in breath. In the experiments with real breath samples, a total of 295 healthy and 279 diabetes breath samples were collected. The sensitivity and specificity of diabetes screening is 91.51% and 90.77%, respectively. The mean relative absolute error (MRAE) for BGL prediction is 21.7%. The results prove the effectiveness of the system as well as the optimization strategies used in our system. The hemoglobin A1c (HbA1c) values of 62 diabetic subjects were also predicted. The prediction accuracy for BGL and HbA1c is compared.

The rest of this paper is organized as follows. Section II is the overall description of the system. Section III introduces the optimization strategies in the sensor array and prediction algorithms. The details of the experiments with simulated samples and real breath samples are in Sections IV and V, respectively. Section VI concludes the paper.

II. SYSTEM DESCRIPTION

A. Structure of the Device

The proposed system consists of two parts: a device for breath measurement and a set of algorithms for data analysis. In the device, a vacuum pump and a gas chamber make up the gas route. The pump draws breath or air from outside and injects it into the gas chamber. The gas chamber is a metal container with sensors embedded in its shell. A signal processing circuit is used to magnify the sensors' signals and filter high frequency noises. The processed signals are digitized and transmitted to a computer by a data acquisition card. A fan is placed next to the gas chamber to take away the heat emitted by the sensors. Some parameters of the device are listed in Table I.

It is worth noting that instead of the common box-shaped gas chamber, we designed a column-shaped gas chamber, as shown in Fig. 1. The internal shape of the chamber is cylindrical and the external shape is hexagon. The sensors are embedded on the six facets of the chamber. This design has three advantages: its internal shape allows gases to flow smoothly; its symmetry ensures that the gas concentration in the head space of each sensor is similar; the size of the chamber is miniature.

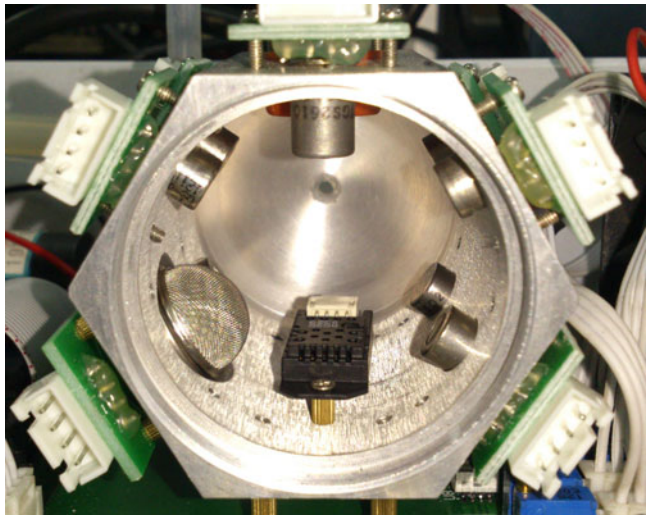


Fig. 1. Snapshot of the column-shaped gas chamber. Sensors are embedded on its wall. Gas enters the chamber from the inlet hole at one end. The outlet end is removed in the figure to show the inside of the chamber.

TABLE II
SUMMARY OF THE SENSOR ARRAY

Channel	Model	Manufacturer	Function
1	TGS4161	Figaro, Inc., Japan	Carbon dioxide
2	TGS822		VOCs, hydrogen, carbon monoxide, etc.
3	TGS826		
4	TGS2610-D00		
5	SP3S-AQ2	FIS, Inc., Japan	
6	GSBT11	Ogam, Inc., Korea	
7	WSP2111	Winsen, Inc., China	
8	TGS2600-TM	Figaro, Inc., Japan	
9	TGS2602-TM		
10	WSP2111-TM	Winsen, Inc., China	
11	HTG3515CH	Humirel, Inc., France	temperature
12			humidity

B. Sensor Array

The device is designed to measure the VOCs, carbon dioxide, humidity, and temperature in breath samples. It is equipped with 11 sensors, including 6 ordinary metal oxide semiconductor (MOS) sensors, 3 temperature modulated MOS sensors, a carbon dioxide sensor, and a temperature-humidity sensor. There are 12 input channels since the temperature-humidity sensor has two input channels. Table II summarizes the model, manufacturer, and function of each sensor. The suffix "-TM" indicates that the sensor is a temperature modulated sensor. This sensor array is specially optimized for the purpose of diabetes screening and BGL prediction. The optimization scheme will be introduced in Section III-A.

C. Sampling Procedure

Similar to [15], when collecting a breath sample, a subject is asked to take a deep breath and exhale into a 600-mL Tedlar gas bag through a disposable mouthpiece. Then, the filled gas bag is plugged onto the connector of the device. The computer software controls the device to complete the breath measurement automatically. All breath samples are measured by the same process, which includes four stages:

1) *Baseline stage (1 s):* The baseline values of the sensors are recorded for future data preprocessing.
2) *Injection stage (7 s):* The pump is ON. Breath is drawn from the gas bag to the gas chamber at a constant speed. The sensors' signals start to respond to the injected breath.
3) *Reaction stage (56 s):* The pump is OFF. The sensors continue reacting with the components in breath. The responses of the MOS sensors without temperature modulation (TM) reach their maximum values.
4) *Purge stage (80 s):* The pump is ON again. Pure air is drawn in to clean the gas chamber for 80 s. The sensors' responses gradually return to their baselines. After the responses remain stable in their baselines, the device is ready for the measurement of the next sample.

After the measurement process, we will get a digitized breath sample represented by 12 response curves (which will also be referred to as a "sample" hereinafter). Each response curve has $144\,\mathrm{s} \times 8\,\mathrm{Hz} = 1152$ data points. The samples will be analyzed with the algorithms in the next section.

D. Data Analysis Methods

1) Signal Preprocessing: For each sensor, we compute its baseline value by averaging its response in the baseline stage. The value is then subtracted from the whole response curve. It is done to eliminate the interference of background noise of the sensors [21]. Humidity compensation is carried out by building a linear humidity response model for each sensor, which will be described in Section III-B. Temperature compensation is not performed since it did not show big significance in our experiments.

2) Feature Extraction: After signal preprocessing, the responses of the ten chemical sensors (see channels 1–10 in Table II) are concatenated into a feature vector. The feature dimension ($1152 \times 10 = 11520$) is very high, so principal component analysis (PCA) is used to extract low-dimensional features. PCA projects high-dimensional data into a low-dimensional subspace while keeping most of the data variance. In the case of BGL prediction, considering the intersubject variance between breath samples, we further add a categorical feature to indicate the subject's identity. The detail of this feature is described in Section III-C.

3) Classification and Regression: Support vector machine (SVM) is among the most popular techniques for classification. It is a kernel-based method suitable for both linear and nonlinear problems. The main idea of the algorithm is to find a maximum margin hyperplane to separate the training samples. It has been proved to generalize well on test samples [22]. SVM has been adopted as the decision algorithm in many chemical sensor systems [13], [23]. We will use it to discriminate between healthy and diabetes samples in the case of diabetes screening. The support vector regression (SVR) [24] algorithm is chosen to solve the BGL prediction problem, since it also has good generalization ability. The details of SVM and SVR can be found in [22] and [24].

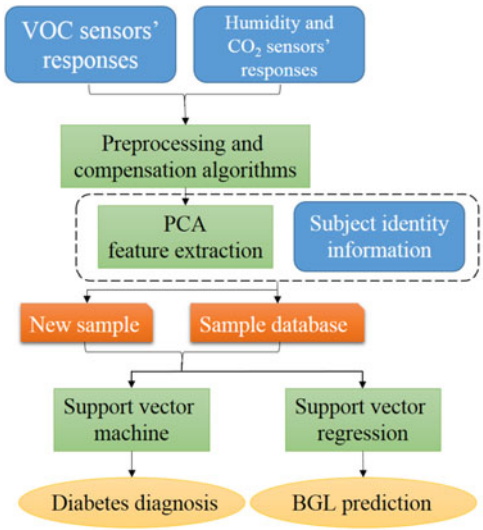


Fig. 2. Framework of the data analysis algorithms. The rounded rectangles in blue are the features. The rectangles in green are the algorithms.

The entire framework of the data analysis algorithms is displayed in Fig. 2.

III. System Optimization

In order to enhance the system's accuracy and robustness for diabetes screening and BGL prediction, several optimization strategies are proposed, including sensor selection, compensation of influential factors, and development of subject-specific prediction models.

A. Sensor Selection

The sensor array is the key part of a chemical sensor system. The sensors should be able to detect the breath biomarkers of diabetes, among which acetone is the most studied one. The concentration of breath acetone of diabetics is higher than that of healthy people [5]–[7]. Furthermore, Wang *et al.* [7] split 30 diabetic subjects into 4 groups and found a linear correlation between the mean concentration of breath acetone and the mean BGL of each group. Turner *et al.* [25] observed that the breath acetone declined linearly with BGL during hypoglycaemic clamps for each volunteer. Besides acetone, compounds such as ethanol [26], carbon monoxide [27], alkanes [28], and methyl nitrate [29] in breath have also been proved to either have abnormal concentrations in diabetics or correlate with BGL. Some researchers have attempted to combine the concentrations of multiple VOCs and achieved good results in diabetes diagnosis and BGL prediction [30], [31].

To detect acetone, one way is to use specially designed acetone sensors [8], [16]. The Si:WO_3 sensor developed by Righettoni *et al.* [8] has high sensitivity and selectivity to acetone. On the other hand, an array of carefully selected cross-sensitive VOC sensors can also have good performance in detecting one or more kinds of gases [21], [32], when pattern recognition algorithms are applied to discriminate different gas "patterns." So we first chose a set of candidate sensors to build pilot e-nose devices, then collected breath samples to evaluate them and select the best sensor combination.

Commercially available sensors are used as candidate sensors because they are easier to acquire, robust, and have good diversity. Some of the candidate sensors can detect and quantify VOCs as low as 0.05 parts per million (ppm) [33], indicating that their precision is satisfactory. Two pilot devices were made with two batches of breath samples collected for sensor selection [34], [35]. Nine sensors were eventually selected to be employed in the final device, i.e., the sensors in channel 2–10 in Table II. Sensors were selected using exhaustive searching experiments, which evaluate the performance of every sensor combination and select the best array in the sense of the highest accuracy in diabetes screening and BGL prediction.

Among the selected sensors, there are three MOS sensors under TM. MOS sensors under TM are believed to provide richer information and have better selectivity than those operated in the ordinary way [36], [37]. In the proposed system, we applied a staircase modulation voltage [36], [37] to three MOS sensors. Experiments with real breath samples show that sensor arrays with these TM sensors have higher accuracy. In the case of diabetes screening, all of the top 50 arrays contain the three TM sensors. To our knowledge, this is the first time that the TM technique is used in breath analysis systems. The results confirm its efficacy.

B. Compensation for Influential Factors

In breath analysis systems, influential factors such as humidity and the proportion of alveolar air in breath samples affect the responses of sensors. Compensation for these factors is important but often neglected. In this section, the influence of these factors will be studied and the methods for compensation will be introduced.

1) Humidity: Humidity compensation is important in breath analysis systems, because human breath contains water vapor, and many VOC sensors are sensitive to humidity. An experiment was made to study the influence of humidity to the sensors. Acetone samples in five concentrations at four humidity levels were provided to the sensors. Results show that if the concentration of acetone is fixed, the maximum value of each sensor rises approximately linearly as the humidity rises. The water vapor in acetone samples has an additive effect to the response of the sensors. Thus, a "humidity coefficient" for each sensor can easily be estimated by linear regression, describing the increase of the sensor's maximum response when humidity increases 1%RH [38]. The humidity compensation model for each VOC sensor is shown in

$$\hat{R}_n^b(t) = R_n^b(t)\left(1 - s_n \frac{\Delta \mathrm{RH}^b}{\max(R_n^b(t))}\right). \tag{1}$$

In (1), $R_n^b(t)$ is the baseline-removed response curve of the *n*th sensor in the *b*th breath sample; s_n is the humidity coefficient of the *n*th sensor; $\Delta \mathrm{RH}^b$ is the difference of humidity between the *b*th breath sample and the environment; $\hat{R}_n^b(t)$ is the compensated response curve. The proportion of magnitude which is considered to be brought by the water vapor in breath is subtracted. Experiment results in Section V-C show that the compensation improves the accuracy of the system.

Temperature in the gas chamber was also measured. Because the temperature was relatively stable among samples, temperature compensation is not applied.

2) Proportion of Alveolar Air: General breath samples consist of two parts: dead-space air from the upper airway and alveolar air from the lungs. VOCs are exchanged between blood and alveolar air. In the case of diabetes screening and BGL prediction, dead-space air is a contaminant and dilutes the concentrations of VOCs in breath samples [15], [39]. So the proportion of alveolar air in a breath sample is an influential factor. This proportion is decided by the phase of the breath. In end-tidal breath, alveolar air is prevailing; whereas breath drawn from the initial phase contains more dead-space air.

Some researchers [19] tried to collect the two parts of breath separately with two cascade gas bags. However, the estimation of the volume of dead-space air may be inaccurate. Moreover, some patients are unable to blow up the two cascade gas bags because of their illness. Another method is to estimate the proportion of alveolar air from the CO_2 concentration in breath samples [39]. Higher CO_2 concentration is an indication of higher proportion of alveolar air. Thereby, we employ a CO_2 sensor in the proposed device. The responses of the CO_2 sensor and the VOC sensors are combined to extract PCA features, which allows the pattern recognition algorithms to learn a better prediction model taking the information of CO_2 concentration into account. The experiment results in Section V-C show that with the information from the CO_2 sensor, better accuracy is acquired.

C. Development of Subject-Specific Prediction Models

Researchers have identified the intersubject variance of the relationship between breath acetone and BGL [9], [25]. As shown in [25], although breath acetone is correlated with BGL for each subject, the baseline values of breath acetone varied among subjects. The author of [1] concluded that calibration of acetone with BGL for each individual is required. However, in previous breath analysis systems aiming at predicting BGL [19], [20], the prediction models are not subject-specific.

To make prediction models subject-specific, an intuitive way is to build a model for each subject with samples from the same subject as training samples. But this method is not applicable when the number of samples from one subject is not enough for an accurate model. In this paper, we propose to add a categorical feature in each feature vector to indicate the subject's identity.

Concretely, for each test sample in the database, a prediction model is trained using all the other samples. For each training sample, the additional categorical feature will be 1 if the training sample is from the same subject with the test sample, or be 0 otherwise. The test sample will also have the additional feature with the value 1. The advantage of this method is that all the training samples can contribute to the prediction model while the training samples from the same subject with the test sample can be emphasized. Thus, the influence brought by the intersubject variance can be reduced. From the results in Section V-C, we find this method can markedly improve the accuracy for BGL prediction.

IV. Experiments With Simulated Samples

An experiment was made to test the system's ability to quantify the main breath biomarker of diabetes, i.e., acetone. According to [5], the concentration of breath acetone in healthy subjects is ranged from 0.22 to 0.80 ppm, while that in subjects with type 2 diabetes is from 1.76 to 3.73 ppm. For subjects with type-1 diabetes, breath acetone could be as high as 21 ppm [25]. So we prepared acetone samples in eight concentrations (0.1, 0.2, 0.5, 1, 2, 5, 10, 20 ppm), with two samples for each concentration. The 16 samples was measured by our device using the sampling procedure in Section II-C. Then, the concentration of acetone in each sample is predicted by leave-one-out cross validation. The data analysis method is preprocessing + PCA + SVR as introduced in Section II-D. The prediction is evaluated by its mean absolute error (MAE) defined in (2), where x_i and $\hat{x}_i$ are the true and predicted concentration of the ith sample, respectively; $n = 16$. In this experiment, the MAE is 0.16 ppm, which indicates the system can predict the concentration of acetone with high accuracy

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i|. \tag{2}$$

Although acetone is among the most abundant VOCs in breath [1], there are many other VOCs in breath that may interfere the measurement of acetone. For example, isoprene in breath has a characteristic concentration of 0.1 ppm [11]. Thus, another experiment was made to test the system's ability to measure acetone with the presence of interfering VOCs. Eight breath samples were collected from each of five healthy volunteers. Then, an addition of acetone was mixed with these 40 breath samples. The eight samples of each volunteer were made to contain an additional acetone of 0, 0.2, 0.3, 0.7, 1.7, 3.3, 5.0, and 6.7 ppm, respectively. These mixed samples are used to simulate the existence of interfering VOCs and the variation of baseline acetone concentrations in real breath samples. Then, the concentration of the additional acetone in each sample is predicted. The leave-one-out strategy and the preprocessing + PCA + SVR algorithm are applied. It is worth noting that the categorical feature described in Section III-C is added to build subject-specific prediction models. In this experiment, the MAE is 0.22 ppm, proving that the system is able to predict the concentration of acetone in the presence of interfering VOCs and the variation of baseline acetone concentrations.

TABLE III
Basic Information of the Diabetic Subjects

Item	Value
Number	87
Male/Female	39/48
Age	39–91
Type 1/Type 2	1/86
Disease duration (years)	0.5–19
Blood glucose level (mmol/L)	4.4–23.1

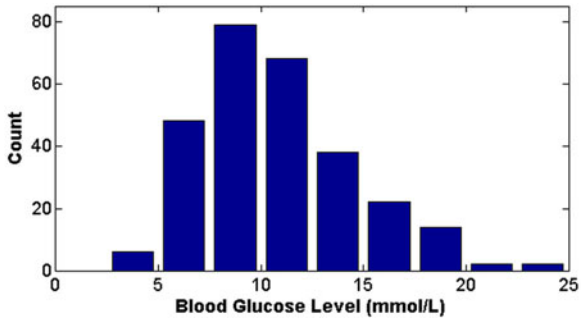


Fig. 3. Distribution of BGLs of the 279 diabetic samples.

V. Experiments With Breath Samples

More than 500 real breath samples were collected to evaluate the system's performance on diabetes screening and BGL prediction. Section V-A is the overview of the samples. Section V-B summarizes the data analysis procedure. Section V-C provides the results and some discussion.

A. Overview of the Breath Samples

A total of 295 healthy samples and 279 diabetes samples were collected from Guangdong Provincial Hospital of Traditional Chinese Medicine (Guangzhou, China). The health states of the healthy subjects were confirmed by physical examinations. The diabetes samples were from 87 inpatient volunteers. For each diabetic subject, several samples were collected at 2 h after meal in different days together with the simultaneous BGLs. The number of samples per subject ranges from 1 to 11. Some information about the diabetic subjects is listed in Table III. Fig. 3 shows the distribution of BGLs of the diabetic samples.

Hemoglobin A1c (HbA1c) is also an important parameter for the diagnosis and monitoring of diabetes. It serves as a marker for average BGL during the preceding 3–4 months with a higher weight over the latest 30 days [40]. Correlation between breath acetone and HbA1c of diabetic subjects has been reported [6], [7]. In this study, 62 out of the 87 subjects had the HbA1c test within the last 13 days. Their HbA1c values range from 5.1% to 15.2%. An experiment was made to predict the HbA1c of the 62 subjects and the accuracy is compared with that of BGL prediction.

B. Data Analysis Procedure

1) Distinguishing Between Healthy and Diabetes Samples: Diabetes screening was achieved by distinguishing between

healthy and diabetes samples. After a digitized breath sample was acquired, it underwent baseline removal, humidity compensation, and PCA feature extraction. The ratio of variance in PCA was set to be 99.98%, extracting about 60 features. They were then scaled to have zero mean and unit variance. SVM [41] with a Gaussian kernel was used for classification. A total of 140 healthy and 140 diabetes samples were randomly selected to train the SVM classifier. Another 139 healthy and 139 diabetes samples were used for testing. We ran this procedure 50 times and computed the average sensitivity and specificity.

2) BGL and HbA1c Prediction: In these two cases, only the diabetes samples were investigated. The data analysis procedures for BGL and HbA1c prediction are mostly the same. Baseline removal, humidity compensation, and PCA feature extraction were applied to the samples. The optimized ratio of variance was set to be 99.1%, extracting about 12 features after PCA. The dimension is lower than that in diabetes screening so as to prevent the regression model from overfitting. The features were further scaled to have zero mean and unit variance. The leave-one-out cross-validation protocol was employed. When the BGL of one sample was predicted, the categorical feature representing the subjects' identity described in Section III-C was added. However, when the HbA1c was predicted, only the first breath sample of each subject was used. There is no need to add the categorical feature since each subject had only one sample. SVR [41] with a linear kernel was adopted to do the prediction.

Three evaluation criteria were utilized to quantify the accuracy of the prediction. The MAE is the average deviation of the prediction from the true target. The MRAE measures the relative error by normalizing the absolute error with the true target. The correlation coefficient r measures the linear correlation between the true target and the predicted value. If we denote x_i as the true target (BGL or HbA1c) of the ith sample, $\hat{x}_i$ as the predicted value, n as the number of samples, $\bar{x}_i$ and $\bar{\hat{x}}_i$ as the mean of all the true and predicted values, then the MRAE and r can be defined as follows:

$$\text{MRAE} = \left(\frac{1}{n} \sum_{i=1}^{n} \left| \frac{x_i - \hat{x}_i}{x_i} \right| \right) \times 100\% \tag{3}$$

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_i)(\hat{x}_i - \bar{\hat{x}}_i)}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2} \sqrt{\sum_{i=1}^{n} (\hat{x}_i - \bar{\hat{x}}_i)^2}}. \tag{4}$$

The aforementioned experiments are used to evaluate not only the performance of the whole system, but also the effectiveness of the system optimization strategies.

C. Results and Discussion

1) Distinguishing Between Healthy and Diabetes Samples: Fig. 4 exhibits the average responses of the healthy samples and the diabetes samples. To observe their differences more clearly, we have made a comparison in Fig. 5. For most VOC sensors (S2-S10), the mean responses of the diabetes samples are larger than that of the healthy samples, showing that the concentration of VOCs in breath of the diabetics is higher than that of the healthy subjects.

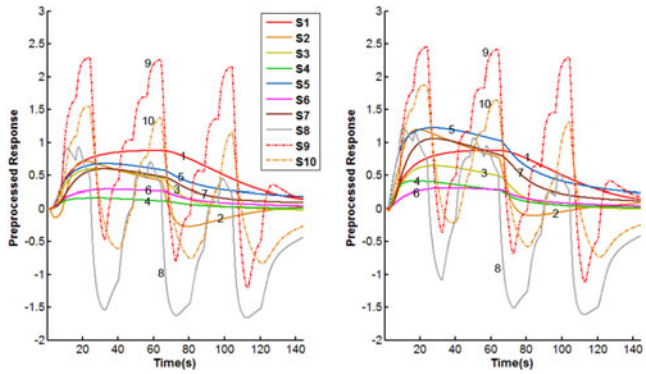


Fig. 4. Average responses of the two classes. Left: healthy; right: diabetes. S1 is a CO_2 sensor; S2–S7 are ordinary MOS sensors; S8–S10 are temperature modulated MOS sensors, so their responses are staircase-shaped.

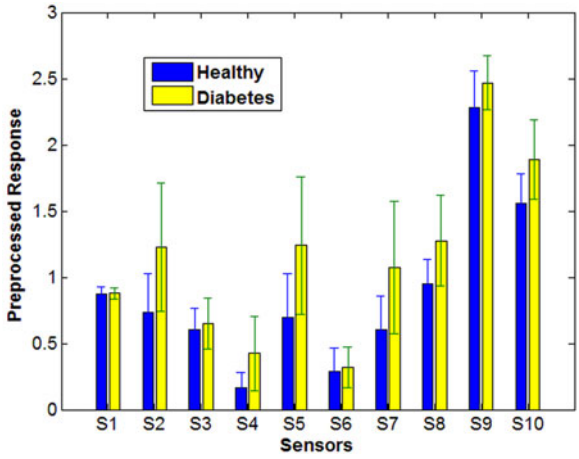


Fig. 5. Average responses of each sensor in the two classes. The coordinates on x-axis are the sensors' indices. S1 is the CO_2 sensor and S2-S10 are VOC sensors. The y-axis is the mean of the maximum value of the preprocessed response. Error bars represent the standard deviations. For VOC sensors, the mean responses of the diabetes samples are larger than that of the healthy ones.

The final sensitivity and specificity for diabetes screening are 91.51% and 90.77%, respectively. The breath analysis system can distinguish between healthy and diabetes samples with a promising accuracy. The accuracy is comparable to previous studies, especially given the fact that the database in this study is larger. The result shows that the system has the potential to be an assistive tool for diabetes screening.

2) BGL and HbA1c Prediction: In order to observe the difference between breath samples from subjects with different BGLs, we divide the diabetes samples in our database into four groups. The BGL thresholds are set to be 7.4, 9.7, and 13.2 mmol/L, so as to make the number of samples in each group close to each other. The mean responses of the VOC sensors in the four groups are shown in Fig. 6. The mean response is ascending from the first to the last group for most sensors except S6 and S9, which is probably because S6 and S9 have higher sensitivity to the interfering components than to acetone. It should also be noticed that the standard deviation in each group is large, which indicates that there are overlaps between groups. This result shows that the prediction task is challenging. The discoveries above are consistent to those in [7].

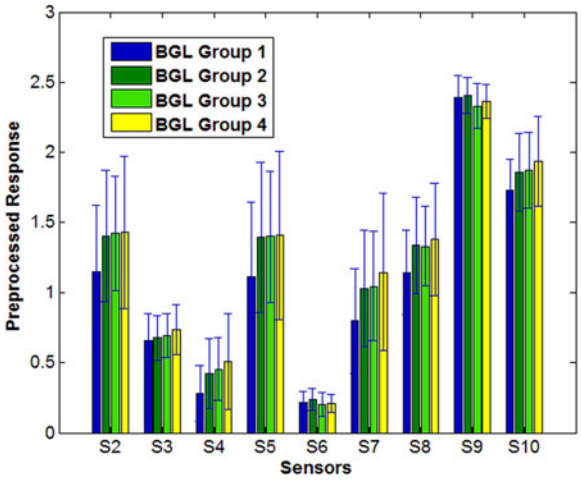


Fig. 6. Mean responses of the VOC sensors in different BGL groups. The diabetes samples are divided into 4 groups according to their BGLs. The coordinates on *x*-axis are the sensors' indices. The *y*-axis is the mean of the maximum value of the preprocessed response. Error bars represent the standard deviations. For most sensors, the mean response is ascending from the first to the last group.

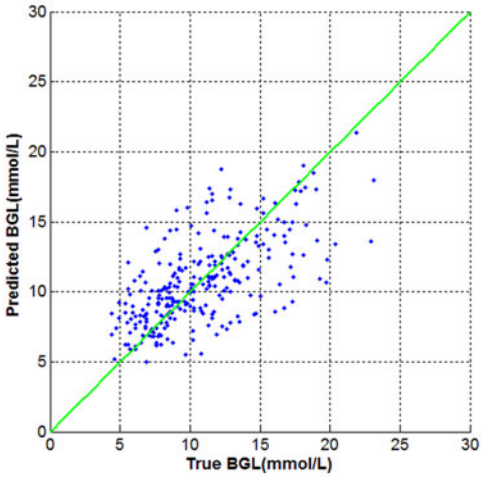


Fig. 7. Scatter diagram for BGL prediction. The x-axis is the true BGL. The y-axis is the predicted BGL. The MAE, MRAE, and correlation coefficient of the prediction are 2.18, 21.7%, and 0.641, respectively.

TABLE IV
COMPARISON OF THE PERFORMANCE ON DIABETES SCREENING AND BGL PREDICTION

Method	Sensitivity (%)	Specificity (%)	MAE	MRAE (%)	*r*
No HC, no CO_2	90.23	88.87	2.40	24.0	0.623
No HC, add CO_2	90.81	89.13	2.25	22.3	0.631
Add HC, no CO_2	90.74	90.14	2.29	23.0	0.636
Add HC, add CO_2	91.51	90.77	2.18	21.7	0.641

BGLs of 279 samples from 87 diabetic subjects are predicted using the leave-one-out protocol. Correlation between the true and the predicted BGLs can be observed from the scatter diagram in Fig. 7. The MAE, MRAE, and correlation coefficient of the prediction are 2.18, 21.7%, and 0.641, respectively. The result is better than a latest study [20], in which a chemical sensor system was designed to predict the BGL of 30 samples with $\mathrm{MRAE} = 25.24\%$.

TABLE V
COMPARISON OF THE PERFORMANCE ON BGL PREDICTION

Method	MAE	MRAE (%)	*r*
Without subject identity information	2.74	27.4	0.350
With subject identity information	2.18	21.7	0.641

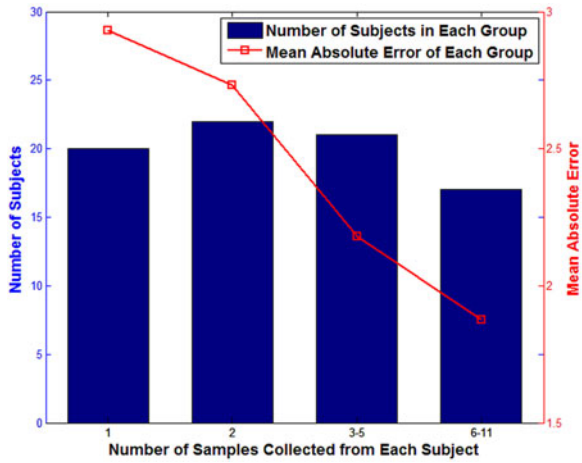


Fig. 8. Relationship between the number of samples collected from a subject and the prediction MAE. Groups 1–4 contain subjects who have 1, 2, 3–5, and 6–11 samples, respectively. The bars show the number of subjects in each group (corresponding to the *y*-axis on the left). The red line shows the MAE of the samples in each group (corresponding to the *y*-axis on the right). The MAE drops as the number of samples increases.

The MAE, MRAE, and correlation coefficient of the HbA1c prediction are 1.86, 21.0%, and 0.56, respectively. The MAE and MRAE are lower than that in the BGL prediction experiment, which is possibly because HbA1cs are more stable and range in a smaller interval. The BGL prediction models are subject-specific and in fact more accurate, so the correlation coefficient of BGL prediction is higher.

3) Effectiveness of the Optimization Strategies: In this section, the effect of the optimization strategies proposed in this paper is assessed. First, the accuracies acquired with or without the compensation algorithms in Section III-B are compared. The results for both diabetes screening and BGL prediction are demonstrated in Table IV. In the tables, HC is short for humidity compensation. CO_2 stands for the information from the CO_2 sensor. For diabetes screening, the addition of HC and CO_2 improves the sensitivity and specificity. For BGL prediction, with the addition of HC and CO_2, the MAE and MRAE are reduced and the correlation coefficient is increased. Therefore, the proposed algorithms aiming at compensating fluctuations of humidity and the proportion of alveolar air are effective.

Table V shows that the BGL prediction accuracy is improved by the strategy of building subject-specific prediction models. We can infer that the influence of the intersubject variance described in [9] and [25] has been reduced. Fig. 8 gives another hint on how much the subject identity information helps the

prediction. The diabetic subjects are grouped according to the number of samples collected from them in the database. Groups are designed so that the number of subjects in each group is close to each other. Then, we compute the MAE of BGL prediction in each group. We find that as the number of collected samples increases, the MAE (the red curve in Fig. 8) decreases. The subjects with the most samples collected has the lowest MAE. This is probably because that with more training samples provided for each subject, the subject-specific prediction model can be more accurate. But this hypothesis still needs further validation using a larger database. To sum up, the subject identity information is important for the training of the prediction model; to improve the prediction accuracy, we can increase the number of training samples for each subject.

VI. Conclusion

This paper proposes a breath analysis system for diabetes screening and BGL prediction. The system includes a breath measurement device and a set of data analysis algorithms. The device has the advantage of being noninvasive, portable, and easy to operate.

To increase the accuracy and robustness of the system, targeted improvements were made on the sensor array, preprocessing, and feature extraction algorithms. The improvements can be roughly categorized into two aspects. In the aspect of medicine, some results in breath analysis studies were consulted. A CO_2 sensor was employed to compensate for the difference of proportion of alveolar air in breath samples. Subject-specific prediction models were built for BGL prediction to reduce the influence of the intersubject variance. In the aspect of sensor technology, an optimal cross-sensitive VOC sensor array was selected with the aid of two pilot devices and two batches of breath sample collection. Temperature modulated MOS sensors were adopted and proved to be useful. The humidity drift of the sensors was compensated. The effectiveness of these improvement strategies were confirmed by experiments. These strategies are expected applicable not only in the proposed system, but in other breath analysis systems as well.

More than 500 breath samples were collected to evaluate the performance of the system. We achieved a promising accuracy in diabetes screening. For BGL and HbA1c prediction, the MRAE is 21.7% and 21.0%, respectively. The BGL prediction result is better than previous breath analysis systems, but still not quite adequate for practical use. One of the error sources is the intersubject variance of the components in breath samples. We have made attempts to reduce the influence of this variance by introducing subject identity feature. With more training samples for each subject and more sophisticated prediction models, the error may be further diminished. Since our experiments were not conducted in strictly controlled environments, there is also intrasubject variance caused by factors such as diet, exercise, and insulin injection. The influence of these factors needs to be further studied to build a prediction model properly taking them into consideration [1], [42]. Larger database should be collected to validate the models.

References

[1] C. Turner, "Potential of breath and skin analysis for monitoring blood glucose concentration in diabetes," *Expert Rev. Mol. Diagn.*, vol. 11, no. 5, pp. 497–503, Jun. 2011.

[2] A. Ramachandran, A. Moses, S. Shetty, C. J. Thirupurasundari, A. C. Seeli, C. Snehalatha, S. Singvi, and J.-P. Deslypere, "A new non-invasive technology to screen for dysglycaemia including diabetes," *Diabetes Res. Clin. Pract.*, vol. 88, no. 3, pp. 302–306, 2010.

[3] S. K. Vashist, "Non-invasive glucose monitoring technology in diabetes management: A review," *Anal. Chim. Acta*, vol. 750, pp. 16–27, Oct. 2012.

[4] T. H. Risby and S. Solga, "Current status of clinical breath analysis," *Appl. Phys. B, Lasers Opt.*, vol. 85, nos. 2/3, pp. 421–426, 2006.

[5] 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," *J. Chromatogr. B*, vol. 810, no. 2, pp. 269–275, 2004.

[6] I. Ueta, Y. Saito, M. Hosoe, M. Okamoto, H. Ohkita, S. Shirai, H. Tamura, and K. Jinno, "Breath acetone analysis with miniaturized sample preparation device: In-needle preconcentration and subsequent determination by gas chromatography–mass spectroscopy," *J. Chromatogr. B*, vol. 877, no. 24, pp. 2551–2556, 2009.

[7] C. Wang, A. Mbi, and M. Shepherd, "A study on breath acetone in diabetic patients using a cavity ringdown breath analyzer: Exploring correlations of breath acetone with blood glucose and glycohemoglobin a1c," *IEEE Sens. J.*, vol. 10, no. 1, pp. 54–63, Jan. 2010.

[8] M. Righettoni, A. Schmid, A. Amann, and S. Pratsinis, "Correlations between blood glucose and breath components from portable gas sensors and PTR-TOF-MS," *J. Breath Res.*, vol. 7, no. 3, 037110, pp. 1–9, 2013.

[9] S. Ghimenti, S. Tabucchi, T. Lomonaco, F. Di Francesco, R. Fuoco, M. Onor, S. Lenzi, and M. Trivella, "Monitoring breath during oral glucose tolerance tests," *J. Breath Res.*, vol. 7, no. 1, 017115, pp. 1–7, 2013.

[10] T. D. C. Minh, D. R. Blake, and P. R. Galassetti, "The clinical potential of exhaled breath analysis for diabetes mellitus," *Diabetes Res. Clin. Pract.*, vol. 97, no. 2, pp. 195–205, 2012.

[11] J. King, A. Kupferthaler, K. Unterkofler, H. Koc, S. Teschl, G. Teschl, W. Miekisch, J. Schubert, H. Hinterhuber, and A. Amann, "Isoprene and acetone concentration profiles during exercise on an ergometer," *J. Breath Res.*, vol. 3, no. 2, 027006, pp. 1–16, 2009.

[12] J. King, A. Kupferthaler, B. Frauscher, H. Hackner, K. Unterkofler, G. Teschl, H. Hinterhuber, A. Amann, and B. Högl, "Measurement of endogenous acetone and isoprene in exhaled breath during sleep," *Physiol. Meas.*, vol. 33, no. 3, pp. 413–428, 2012.

[13] M. Trincavelli, S. Coradeschi, A. Loutfi, B. Soderquist, and P. Thunberg, "Direct identification of bacteria in blood culture samples using an electronic nose," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2884–2890, 2010.

[14] Y.-J. Lin, H.-R. Guo, Y.-H. Chang, M.-T. Kao, H.-H. Wang, and R.-I. Hong, "Application of the electronic nose for uremia diagnosis," *Sens. Actuators B. Chem.*, vol. 76, no. 1, pp. 177–180, 2001.

[15] D. Guo, D. Zhang, N. Li, L. Zhang, and J. Yang, "A novel breath analysis system based on electronic olfaction," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 11, pp. 2753–2763, Nov. 2010.

[16] P. Wang, Y. Tan, H. Xie, and F. Shen, "A novel method for diabetes diagnosis based on electronic nose," *Biosens. Bioelectron.*, vol. 12, no. 9, pp. 1031–1036, 1997.

[17] Q. Zhang, P. Wang, J. Li, and X. Gao, "Diagnosis of diabetes by image detection of breath using gas-sensitive laps," *Biosens. Bioelectron.*, vol. 15, no. 5, pp. 249–256, 2000.

[18] J.-B. Yu, H.-G. Byun, M.-S. So, and J.-S. Huh, "Analysis of diabetic patient's breath with conducting polymer sensor array," *Sens. Actuators B. Chem.*, vol. 108, no. 1, pp. 305–308, 2005.

[19] D. Guo, D. Zhang, L. Zhang, and G. Lu, "Non-invasive blood glucose monitoring for diabetics by means of breath signal analysis," *Sens. Actuators B. Chem.*, vol. 173, no. 6, pp. 106–113, Oct. 2012.

[20] H. M. Saraoglu, A. O. Selvi, M. A. Ebeoglu, and C. Tasaltin, "Electronic nose system based on quartz crystal microbalance sensor for blood glucose and HbA1c levels from exhaled breath odor," *IEEE Sens. J.*, vol. 13, no. 11, pp. 4229–4235, Nov. 2013.

[21] A. Hierlemann and R. Gutierrez-Osuna, "Higher-order chemical sensing," *Chem. Rev.*, vol. 108, no. 2, pp. 563–613, 2008.

[22] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[23] A. Amini, M. A. Bagheri, and G. Montazer, "Improving gas identification accuracy of a temperature-modulated gas sensor using an ensemble of classifiers," *Sens. Actuators B. Chem.*, vol. 187, pp. 241–246, 2013.

[24] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.
[25] C. Turner, C. Walton, S. Hoashi, and M. Evans, "Breath acetone concentration decreases with blood glucose concentration in type I diabetes mellitus patients during hypoglycaemic clamps," *J. Breath Res.*, vol. 3, no. 4, p. 046004, Dec. 2009.
[26] P. R. Galassetti, B. Novak, D. Nemet, C. Rose-Gottron, D. M. Cooper, S. Meinardi, R. Newcomb, F. Zaldivar, and D. R. Blake, "Breath ethanol and acetone as indicators of serum glucose levels: an initial report," *Diabetes Technol. Ther.*, vol. 7, no. 1, pp. 115–123, 2005.
[27] P. Paredi, W. Biernacki, G. Invernizzi, S. A. Kharitonov, and P. J. Barnes, "Exhaled carbon monoxide levels elevated in diabetes and correlated with glucose concentration in blood: A new test for monitoring the disease?," *CHEST*, vol. 116, no. 4, pp. 1007–1011, 1999.
[28] M. Phillips, R. N. Cataneo, T. Cheema, and J. Greenberg, "Increased breath biomarkers of oxidative stress in diabetes mellitus," *Clin. Chim. Acta*, vol. 344, no. 1, pp. 189–194, 2004.
[29] B. Novak, D. Blake, S. Meinardi, F. Rowland, A. Pontello, D. Cooper, and P. Galassetti, "Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes," *Proc. Nat. Academy Sci.*, vol. 104, no. 40, pp. 15613–15618, 2007.
[30] M. Greiter, L. Keck, T. Siegmund, C. Hoeschen, U. Oeh, and H. Paretzke, "Differences in exhaled gas profiles between patients with type 2 diabetes and healthy controls," *Diabetes Technol. Ther.*, vol. 12, no. 6, pp. 455–463, 2010.
[31] J. Lee, J. Ngo, D. Blake, S. Meinardi, A. M. Pontello, R. Newcomb, and P. R. Galassetti, "Improved predictive models for plasma glucose estimation from multi-linear regression analysis of exhaled volatile organic compounds," *J. Appl. Physiol.*, vol. 107, no. 1, pp. 155–160, 2009.
[32] C. Di Natale, R. Paolesse, G. D'Arcangelo, P. Comandini, G. Pennazza, E. Martinelli, S. Rullo, M. Roscioni, C. Roscioni, and A. Finazzi-Agrò, "Identification of schizophrenic patients by examination of body odor using gas chromatography-mass spectrometry and a cross-selective gas sensor array," *Med. Sci. Monit.*, vol. 11, no. 8, pp. CR366–CR375, 2005.
[33] E. J. Wolfrum, R. M. Meglen, D. Peterson, and J. Sluiter, "Metal oxide sensor arrays for the detection, differentiation, and quantification of volatile organic compounds at sub-parts-per-million concentration levels," *Sens. Actuators B. Chem.*, vol. 115, no. 1, pp. 322–329, 2006.
[34] K. Yan and D. Zhang, "A novel breath analysis system for diabetes diagnosis," in *Proc. Int. Conf. Comput. Healthcare*, Hong Kong, China, 2012, pp. 166–170.
[35] K. Yan and D. Zhang, "Sensor evaluation in a breath analysis system," presented at the Int. Conf. Med. Biometrics, Shenzhen, China, May 30–Jun. 1, 2014.
[36] R. Gutierrez-Osuna, A. Gutierrez-Galvez, and N. Powar, "Transient response analysis for temperature-modulated chemoresistors," *Sens. Actuators B. Chem.*, vol. 93, no. 1, pp. 57–66, 2003.
[37] S. Hosseini-Golgoo and F. Hossein-Babaei, "Assessing the diagnostic information in the response patterns of a temperature-modulated tin oxide gas sensor," *Meas. Sci. Technol.*, vol. 22, no. 3, 035201, pp. 1–11, 2011.
[38] K. Kashwan and M. Bhuyan, "Robust electronic-nose system with temperature and humidity drift compensation for tea and spice flavour discrimination," in *Proc. 2005 Asian Conf. Sensors Int. Conf. New Techn. Pharmaceutical Biomed. Res.*, 2005, pp. 154–158.
[39] W. Cao and Y. Duan, "Current status of methods and techniques for breath analysis," *Crit. Rev. Anal. Chem.*, vol. 37, no. 1, pp. 3–13, 2007.
[40] C. L. Rohlfing, H.-M. Wiedmeyer, R. R. Little, J. D. England, A. Tennill, and D. E. Goldstein, "Defining the relationship between plasma glucose and HbA1c analysis of glucose profiles and HbA1c in the Diabetes Control and Complications Trial," *Diabetes Care*, vol. 25, no. 2, pp. 275–278, 2002.
[41] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 27, pp. 1–27, 2011.
[42] P. Španěl, K. Dryahina, A. Rejšková, T. W. Chippendale, and D. Smith, "Breath acetone concentration; biological variability and the influence of diet," *Physiol. Meas.*, vol. 32, no. 8, pp. N23–N31, Aug. 2011.

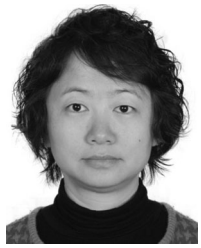
Ke Yan received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China. He is currently working toward the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University, China.

His research interests include biomedical engineering, machine learning, and biometrics.

David Zhang (F'09) received the graduate degree in computer science from Peking University, Beijing, China. He received the M.Sc. degree in computer science in 1982 and the Ph.D. degree from the Harbin Institute of Technology in 1985. He received the second Ph.D. degree in electrical and computer engineering from the University of Waterloo, ON, Canada, in 1994.

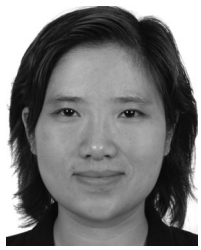
From 1986 to 1988, he was a Postdoctoral Fellow at Tsinghua University and then an Associate Professor at the Academia Sinica, Beijing, China. He is currently the Chair Professor at the Hong Kong Polytechnic University, Hong Kong, where he is the Founding Director of the Biometrics Technology Centre (UGC/CRC) supported by the Hong Kong SAR Government in 1998. He also serves as a Visiting Chair Professor in Tsinghua University, Beijing, China, and an Adjunct Professor in Peking University, Shanghai Jiao Tong University, HIT, and the University of Waterloo. He is the author of more than 10 books and 200 journal papers.

Prof. Zhang is a Croucher Senior Research Fellow, Distinguished Speaker of the IEEE Computer Society, and a Fellow of both IEEE and IAPR. He is the Founder and an Editor-in-Chief of the INTERNATIONAL JOURNAL OF IMAGE AND GRAPHICS; a Book Editor of SPRINGER INTERNATIONAL SERIES ON BIOMETRICS, an Organizer of the International Conference on Biometrics Authentication, and an Associate Editor of more than ten international journals including IEEE Transactions and Pattern Recognition.



Darong Wu received the Master's of Medicine and the Doctor of Medicine degrees from Guangzhou University of Chinese Medicine, Guangzhou, China, in 1999 and 2003, respectively.

From 2005 to 2008, she was a Postdoctoral Fellow at China Academy of Chinese Medical Sciences. She is currently a Chief Physician and a Supervisor of postgraduate in Guangdong Provincial Hospital of Chinese Medicine, Guangzhou, China. She is also the Director of the Office of Clinical Effect Projects. She is also a Part-Time Assistant Professor in the Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada. Her research areas include clinical epidemiology/DME methodology, evaluation of clinical effect of Chinese medicine, evidence-based medicine, and life quality.



Hua Wei received the Master's of Medicine degree from Hubei University of Chinese Medicine, Hubei, China, in 1997.

She is currently a Chief Physician and a Supervisor of postgraduate at Guangdong Provincial Hospital of Chinese Medicine, Shenzhen, China. She is also the Director of the Department of Endocrinology. She is serving as an Executive Director of the Branch of Diabetes of the World Federation of Chinese Medicine Societies. She has been engaged in the clinical work, teaching, and experimental research of endocrinology. She has undertaken or joined more than 20 research projects and published more than 40 papers and ten books around her research area.

Guangming Lu received the B.S. degree in electrical engineering, the M.S. degree in control theory and control engineering, and the Ph.D. degree in computer science and engineering from the Harbin Institute of Technology, Harbin, China, in 1998, 2000, and 2005, respectively.

From 2005 to 2007, he was a Postdoctoral Fellow at Tsinghua University, Beijing, China. He is currently an Associate Professor of the Biocomputing Research Center, Shenzhen Graduate School, HIT, Shenzhen, China. His current research interests include pattern recognition, image processing, and automated biometric technologies and applications.