31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Development of Statistical Regression Models for Ventilation Estimation

Shaopeng Liu[1], Robert X. Gao[1*], Qingbo He[1], John Staudenmayer[2] and Patty Freedson[3]

Abstract—**Estimation of ventilation volume from dimensional changes of the rib cage and abdomen is of interest to researchers interested in quantifying internal exposure to environmental pollutants in the atmosphere. In this paper, we present different statistical regression models for estimating ventilation volume during free-living activities. The movements of the rib cage and abdomen were measured by piezoelectric sensor belts. Multiple linear regression as the calibration method was applied. Five regression models with different combinations out of thirteen features were developed and the performance of these models was compared through experimental study of 11 subjects. The effect of training approaches – model trained for each subject and for all subjects, and the effect of time intervals for computing features were also investigated. The results indicate that Model 2, combining respiratory features and breathing frequency, with a longer time intervals will lead to a higher accuracy.**

I. Introduction

Continuous monitoring of respiration in free-living environments is potentially important for quantifying exposure to environmental particles in the air. Additionally, respiration measurement is a physiological response that increases with exercise and thus may be useful in estimating an individual's level of physical activity. Unlike a conventional mask or mouthpiece that is burdensome to human test subjects in a free–living environment, non-invasive techniques have been examined. Respiratory inductive plethysmography (RIP), respiratory magnetometer, and piezoelectric sensors are three well known devices that have been used to estimate respiration. Specifically, these methods use sensing belts placed around the rib cage and abdominal region to estimate ventilation from the magnitude of dimensional changes of the belts during breathing [1] - [5].

In 1967, Konno and Mead showed that the respiratory system could be assumed as a simple two degrees-of-freedom system [6]. The volume changes of the whole respiratory system were approximated as the sum of volume changes of rib cage and abdominal compartments, which were related to the dimensional changes of the compartments measured from the strap assembly. To examine the relationship between ventilation volume and the dimensional changes, multiple linear regression has been used by a number of research groups [7]-[9]. However, given that the human respiratory system is not a simple system, providing more variables for the prediction of ventilation may improve prediction accuracy. These variables include breathing frequency, products of breathing frequency and dimensional changes of the rib cage and abdomen respectively, height, weight, circumferences of the rib cage and abdomen.

This study investigates different linear regression models for estimating ventilation volumes from data measured by piezoelectric sensor belts. A total of thirteen features, which have either a direct or indirect relationship with ventilation, were chosen, and five models were developed, based on different combinations of the features. The performance of each model is subsequently assessed and compared with each other through experimental study.

Manuscript received April 7, 2009. This work was supported by NIH under award No. UO1 CA130783.

The authors are with [1]Electromechanical Systems Laboratory, University of Connecticut, Storrs, CT 06269, [2]Department of Mathematics and Statistics, and [3]Exercise Physiology Laboratory, University of Massachusetts Amherst, Amherst, MA 01003, USA (correspondence: 860-486-7110, rgao@engr.uconn.edu).

II. Experimental Parameters

A. Subjects

Eleven test subjects (5 male, 6 female) were recruited from students at the university. After verbal explanation of the experimental procedures, all subjects read and signed an informed consent document approved by the university Institutional Review Board. The characteristics of all subjects are shown in Table I.

TABLE I
SUBJECT CHARACTERISTICS (N = 11)

Characteristics	Mean±SD
Age (years)	24.3±2.7
Mass (kg)	67.7±12.3
Height (cm)	171.2±8.6
Body Mass Index (kg/m^2)	23.2±4.6

B. Experimental Sequence

The subjects performed a continuous activity protocol in the exercise physiology laboratory at the University of Massachusetts Amherst. The test consisted of a rest period with subjects lying, facing up on a bed, three treadmill exercise conditions: slow walking (2.4 km/h), fast walking (4.8 km/h), jogging (7.2 km/h), and two other activities: sweeping floor and moving a 4.5-kg box. Table II presents the details of the test. Subjects performed each activity for 10 minutes, followed by a 2-minute rest. Except for rest which was always performed first, the subjects completed the other five activities in a random order. Prior to the test, subjects were asked to lie down on a bed to rest for 10 minutes, in order to achieve a baseline metabolic rate. All the tests were performed in the morning, and the subjects were not allowed to have food for four hours before the test.

TABLE II
EXPERIMENTAL SEQUENCE

Activity	Test Duration	Description
Resting	10 minutes	Lying down, facing up
Slow Walking	10 minutes	Treadmill, at 2.4 km/h
Fast Walking	10 minutes	Treadmill, at 4.8 km/h
Jogging	10 minutes	Treadmill, at 7.2 km/h
Sweeping Floor	10 minutes	Laboratory
Moving Box	10 minutes	Laboratory, 4.5-kg box

C. Data Acquisition System

Piezoelectric crystal or film sensors have been commonly used for monitoring human respiration. The sensor is usually combined with an elastic belt whose output is proportional to the expansion of the belt. When encircling the rib cage and abdomen snugly at the level of nipples and umbilicus, respectively, the piezoelectric belt sensor assembly measures the chest and abdominal expansion associated with respiratory effort.

Fig. 1 illustrates the configuration of the data acquisition system used. It includes a belt assembly (Piezo Respiratory Effort, Ambu Sleepmate) for ventilation measurement on test subjects, and a respiratory gas exchange system (Oxycon Mobile, Cardinal Health) that serves as the criterion measure. The sensitivity of the piezoelectric sensor is 30 μv/mm. The sensor belts were directly connected to an amplifier (AM503, Tektronix, Inc.) and the amplification gain was set to 1,000. The amplified voltage signals from both sensors were acquired by a 12-bit A/D converter (DAQCard-AI-16E-4, National Instruments), at a sampling frequency of 50 Hz. The volume transducer from the respiratory gas exchange system served as the standard reference. The respiratory gas exchange system is secured to the subject using an adjustable vest. The breath-by-breath respiratory data are collected through a facemask and then transmitted to a host laptop wirelessly. A pre-calibration of the system was performed before experiments.

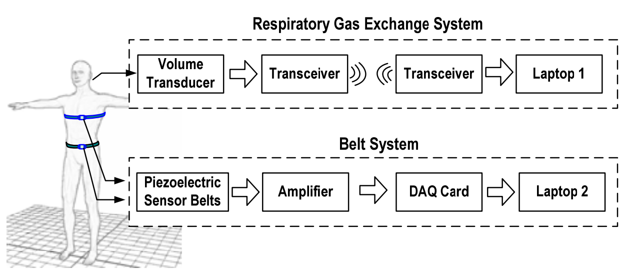

Fig. 1. Schematic of the data acquisition system.

III. DATA ANALYSIS

A. Statistical Methods

In order to estimate the ventilation volume, a relationship was established between the chest and rib displacement data obtained from the piezoelectric sensor belts and the ventilation volume. In general, regression analysis can be used for modeling the relationship between a dependent variable (response) $\boldsymbol{y}$ and one or more independent variables (predictors) $\boldsymbol{x_1}, \boldsymbol{x_2} \ldots \boldsymbol{x_p}$. The general regression equation is expressed as:

$$y = f(\underline{X}, \underline{\beta}) \tag{1}$$

where $\boldsymbol{y}$ is the response in space $\boldsymbol{R^1}$, $y \in R^1$, $\underline{X}$ is a set of independent variables in a space $\boldsymbol{R^P}$, $\underline{X} = \{x_1, x_2, \cdots, x_p\} \subset R^P$, $\underline{\beta}$ are regression coefficients, $\underline{\beta} = \{\beta_0, \beta_1, \beta_2, \cdots, \beta_p\}$, and $\boldsymbol{f}$ represents the regression function which maps from space $\boldsymbol{R^{P+1}}$ to $\boldsymbol{R^1}$, $f:\ R^{P+1} \rightarrow R^1$.

The current study is based on a linear mapping for the estimation of ventilation volume. This is expressed as:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \tag{2}$$

Note that Eq. (2) can be expressed in the matrix form as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}_{n\times(p+1)} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)\times 1} \tag{3}$$

where $\boldsymbol{y}$ is the estimated ventilation volume, $\boldsymbol{x_i}$ is the ith variable data set obtained from the sensors, $\boldsymbol{p}$ is the number of independent variables, and $\boldsymbol{n}$ is the number of data in each variable data set.

B. Features

Application of the linear model as shown above requires the determination of the independent variables. The two degrees-of-freedom model developed in [6] assumed that the changes of ventilation volume are equal to the sum of the rib cage and abdominal volumes, which are related to the dimensional changes of the two body parts. The relationship can thus be expressed as the following:

$$V_T = a \cdot RC + b \cdot AB \tag{4}$$

where $\boldsymbol{V_T}$ represents the ventilation volume, $\boldsymbol{a}$ and $\boldsymbol{b}$ are coefficients, $\boldsymbol{RC}$ and $\boldsymbol{AB}$ are two independent variables that represent the dimensional changes of the rib cage and abdomen, respectively. Based on this assumption, a number of studies [7]-[9] have been performed by using multiple linear regressions. Considering that the human respiratory system is a complex system, multiple variables associated with the ventilation volume are needed in order to improve the accuracy of the ventilation estimation. In this study, the following three types of features are utilized for this purpose:

1) *Respiratory features* (*RF*), which have direct relationship with the ventilation volumes;
2) *Frequency features* (*FF*), which are related to the respiratory frequency;
3) *Body features* (*BF*), which may be related to the ventilation volume.

A total of thirteen features were chosen as shown in Table III. For *RF*, there are six features, including the 10th and 90th percentiles of an abdominal belt signal within a certain time interval (feature F_1 and F_2), 10th and 90th percentiles of a rib cage belt signal within the same time interval (F_3 and F_4), the differences between 90th and 10th percentiles of abdominal signal (F_5) and that of the rib cage signal (F_6). These six features represent stable estimates of the dimensional changes in abdomen and rib cage respectively, which are directly related to ventilation volumes. When computing these features, five different time intervals were used, i.e., 20, 30, 40, 50, and 60 seconds in order to investigate the effect of time intervals on the estimation of ventilation volume. These time intervals were chosen empirically.

Breathing rate may also reflect breathing volume. Thus, three breathing frequency related features consisting of *FF*: breathing frequency (feature F_7) estimated from the sensor belt using spectrum analysis, product (F_8) of breathing frequency and feature F_5, product (F_9) of breathing frequency and feature F_6. The above features were computed after tissue artifact in each of the sensor signals were removed by the method of empirical mode decomposition [10].

It was assumed that features from the subjects' characteristics may be related to the ventilation volume, e.g. a subject's height may be related to ventilatory volume. Thus, four body size features, including the circumferences of abdomen and rib cage (feature F_{10} and F_{11}), height (F_{12}) and weight (F_{13}) were examined.

TABLE III
FEATURES

Feature Type	Feature No.	Description
RF	F_1	10th percentile of abdominal signal
	F_2	90th percentile of abdominal signal
	F_3	10th percentile of rib cage signal
	F_4	90th percentile of rib cage signal
	F_5	F_2 minus F_1
	F_6	F_4 minus F_3
FF	F_7	Breathing frequency
	F_8	F_7 times F_5
	F_9	F_7 times F_6
BF	F_{10}	Circumference of abdomen
	F_{11}	Circumference of rib cage
	F_{12}	Height
	F_{13}	Weight

C. Regression Models

Since some of the features are linear combinations of other features, it is not possible to include all the thirteen features in a single regression model. In this study, the following five models (Table IV) were developed to evaluate the performance of feature combinations on estimating the ventilation volume.

Model M_1 is based on Konno and Mead's model in Eq. (4) and serves as the comparison basis for the other four models. We include breathing frequency – feature F_7 in all the other four models as we assumed that the breathing frequency will be highly related to the ventilation volume.

The least squares algorithm is used to estimate, or train, the multiple linear regression models that estimate ventilation volume as a linear function of the features. The quality of the estimated models is evaluated with root mean square error (RMSE), which is the square root of the average of the squared differences between the model predictions of ventilation volume and the measured volumes.

TABLE IV
REGRESSION MODELS

Model No.	Feature Combination
M_1	F_5, F_6
M_2	F_1, F_2, F_3, F_4, F_7
M_3	F_5, F_6, F_7
M_4	F_7, F_8, F_9
M_5	F_7, F_{10}, F_{11}, F_{12}, F_{13}

The models are trained in two ways: a separate model was trained for each subject using only that subject's data, and a single model was trained for all subjects using all the data. In each approach, the estimation performance is assessed with the leave-one-out cross validation. Cross-validation estimates of RMSE provide valid estimates of how the models would perform if they were applied to different data under similar conditions. In the subject specific models, the model is trained using all but one of the data points for a specific subject, and RMSE is calculated for the left out data point. That procedure is repeated for all the data points and all the subjects, and the average RMSE is reported. In the case of the single model for all subjects, the model is trained using all but one subject's data, and RMSE is calculated for the left out subject. Each subject is left out in turn, and the average RMSE is reported.

IV. RESULTS

Fig. 2 shows the results obtained from the two training approaches for all the five models. As seen in Fig. 2, except for the results from subject No. 10 and No. 3 using model M_5 with 60-second interval, the results using the model trained on one subject are consistently better than that from all subjects. The overall absolute bias is 0.159 L (95% confidence interval: 0.142–0.175 L) using the model trained from one subject, and 0.179 L (95% confidence interval: 0.153-0.206 L) using the model trained on all subjects. And, the average of the overall RMSEs is 0.214 L (95% confidence interval: 0.185-0.244 L) using the model trained on each subject, and 0.269 L (95% confidence interval: 0.234-0.296 L) using model trained from all subjects.

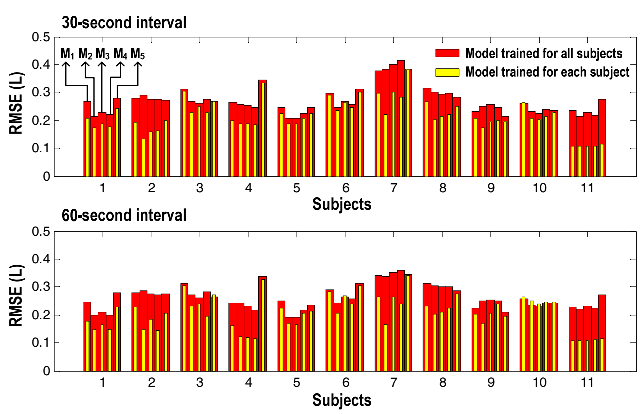

Fig. 2. Performance comparison on training approaches.

Fig. 3 shows the performance comparison of the five types of models. Compared with the traditional model M_1, M_2 reduces the error by about 20% when the model was trained for each subject.

The effect of different time intervals is shown in Fig. 4. The accuracies of ventilation estimation increase when longer time intervals are used. However, the accuracies are not significantly increased: compared with 20-second interval, the error of 60-second interval is reduced by about 8%. The result indicates the impact of selections of time intervals is less significant than the choice of regression models.

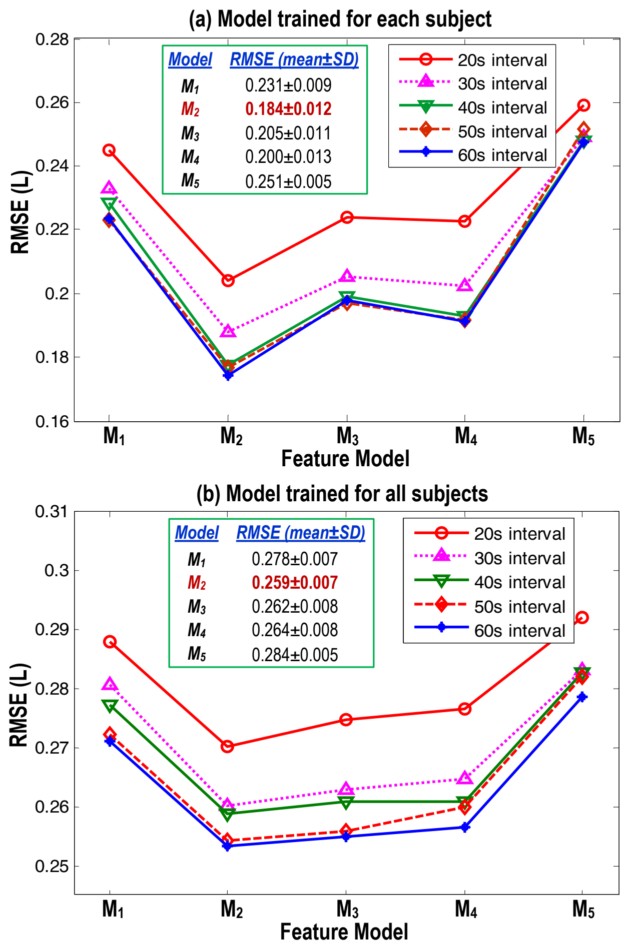

Fig. 3. Performance comparison on feature models.

V. Conclusion

The accuracy of ventilation volume estimation has shown to be affected by various parameters. To enable quantitative analysis, a total of five regression models were developed for the estimation of ventilation volume by using multiple linear regressions. Through experimental study, the model M_2 has the highest accuracy out of the five models, and it reduces error by about 20% as compared with the commonly used model M_1. Our results show that the model trained for all subjects produced a 25% higher error than the models trained for each subject. Also, a longer time interval generally leads to a higher accuracy, although the effect is less significant than that of the regression models.

The linear regression method is computational efficient and thus can be implemented for online monitoring. However, given the complexity and inherent nonlinearity associated with the human respiratory system, further studies are being conducted to investigate the performance of nonlinear models on the estimation of ventilation volume.

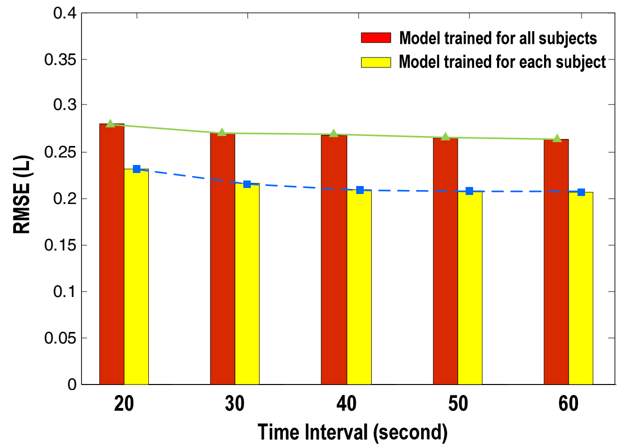

Fig. 4. Performance comparison on time intervals.

Acknowledgment

The authors would like to thank Dr. Todd Hagobian for his assistance in experimental data collection.

References

[1] J. Milledge, F. Stott, "Inductive plethysmograph – a new respiratory transducer," *J. Physiol*., vol. 267, 1977, pp. 4–5.

[2] R. Banzett, S. Mahan, D. Carner, A. Brughera and S. Loring, "A simple and reliable method to calibrate respiratory magnetometers and Respitrace," *J. Appl. Physiol*., vol. 79, 1995, pp. 2169-2176.

[3] B. Pennock, "Rib cage and abdominal piezoelectric film belts to measure ventilatory airflow," *J. Clin. Monit*., vol. 6, 1990, pp. 276-283.

[4] J. Witt, J. Fisher, J. Guenette, K. Cheong, B. Wilson, A. Sheel, "Measurement of exercise ventilation by a portable respiratory inductive plethysmograph," *Respiratory Physiology and Neurobiology*, vol. 154, 2006, pp. 389-395.

[5] K. Cohen, W. Ladd, D. Beams, W. Sheers, R. Radwin, W. Tompkins, and J. Webster, "Comparison of impedance and inductance ventilation sensors on adults during breathing, motion, and simulated airway obstruction," *IEEE Trans. Biomed. Eng*., vol. 44, 1997, pp. 555-566.

[6] K. Konno and J. Mead, "Measurement of the separate changes of rib cage and abdomen during breathing," *J. Appl. Physiol*., vol. 22, 1967, pp. 407–422.

[7] B. Loveridge, P. West, N. Anthonisen and M. Kryger, "Single-position calibration of the respiratory inductance plethysmograph," *J. Appl. Physiol*., vol. 55, 1983, pp. 1031-1034.

[8] N. Stromberg, G. Dahlback and P. Gustafsson, "Evaluation of various models for respiratory inductance plethysmography calibration," *J. Appl. Physiol*., vol. 74, 1993, pp. 1206-1211.

[9] E. Bar-Yishay, A. Putilov and S. Einav, "Automated, real-time calibration of the respiratory inductance plethysmograph and its application in newborn infants," *Physiol. Meas*., vol. 24, 2003, pp. 149-163.

[10] S. Liu, Q. He, R. Gao, and P. Freedson, "Empirical mode decomposition applied to tissue artifact removal from respiratory signal", *Proc. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3624-3627, Vancouver, British Columbia, Canada, August 20-24, 2008.