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

Cuffless Blood Pressure Estimation by Error-Correcting Output Coding Method Based on an Aggregation of AdaBoost with a Photoplethysmograph Sensor

Satomi Suzuki and Koji Oguri

Abstract— **This paper presented a novel cuffless and non-invasive technique of Blood Pressure (BP) estimation with a pattern recognition method by using a Photoplethysmograph (PPG) sensor instead of a cuff. Error-Correcting Output Coding (ECOC) method was adopted as a multi-classifier machine based on an aggregation of general binary classifiers. AdaBoost was applied as binary classifier machine. 368 volunteers participated in the experiment. The estimated Systolic Blood Pressure (SBP) was calculated from their individual information and several features of their Pulse Wave (PW). As a result of the comparison between measured SBP, estimated SBP, *MD*=-1.2 [mmHg] and *SD*=11.7 [mmHg] were obtained. Hence, this technique would be helpful to the advance the development of continuous BP monitoring system, since the only one device to monitor BP is smaller than traditional measurements.**

I. Introduction

BLOOD Pressure (BP) is a quite important factor for assessing a cardiovascular condition. In particular, Systolic Blood Pressure (SBP) is considered as a better predictor of major cardiovascular adverse events compared to Diastolic Blood Pressure. In general, non-invasive BP measurements need a cuff. However, such measurement techniques give loads to patients mentally and physically. Moreover, these can hardly monitor BP continuously. The realization of gauge for monitoring BP at all times is eagerly anticipated.

Recently, it has been easy to measure daily biological signals since sensor technologies have been developed and all kinds of measurement instruments have become much smaller and consumed less power. Therefore, this study suggests a method of SBP estimation with a wearable sensor instead of a cuff.

The relation between biological signals and BP has been concerned for more than 30 years. Most studies of biological signals with BP have focused on Pulse Transit Time (PTT) [1]-[5]. In these studies, patients have to keep using at least two sensors, since two devices (e.g. electrocardiogram, Photoplethysmograph (PPG)) are necessary for obtaining PTT. Hence, patients would feel uncomfortable or inconvenient. As other previous study without PTT, Park et al. [6] have predicted mean arterial pressure with physiological characteristics from pressure that measured by a tonometry sensor. A device for BP monitoring should be small-scale due to patients' loads. Therefore, our study depends on only Pulse Wave (PW) signal which detected by a PPG sensor, since PPG sensor is a very small gauge (Fig. 1). Moreover, subject just wears the sensor on the surface of the body to measure the signal. We aim to estimate SBP by analyzing PW.

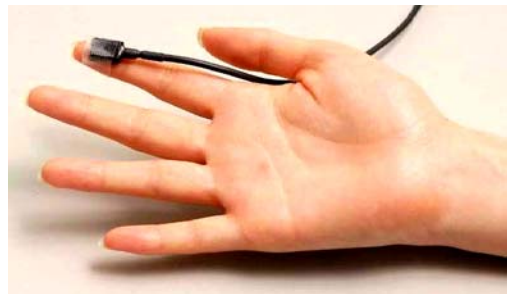

Fig. 1. Photoplethysmograph sensor.

This paper proposes a novel algorithm to estimate SBP by means of a multi-class recognition method based on an aggregation of general binary classifiers. We classify SBP value into multiple classes according to arbitrary thresholds of SBP. And then, estimated SBP is calculated by using the learning SBP estimation models that are constructed from training examples.

II. Photoplethysmograph

Oxygenated hemoglobin in blood has very high absorption of blue and green light. In addition, surface of the body can hardly reflect the light which has wavelengths in 550 [nm] and under regions. Therefore, the signal that obtained by PPG sensor with the light has the wavelengths in this region represents capacity changing of blood vessel according to volume of blood flowing. The signal is a kind of PW.

Individual features of PW would reflect the blood ejection from heart, the shape of blood vessels and the condition of blood vessel walls. Although the signal of PW is a bit different depending on measured areas, PW is basically constructed by the ejection wave and the returning reflected wave from blood vessel walls. In this study, several features are calculated from PW to estimate SBP. Figure 2 shows the characteristics of PW.

S. Suzuki is with Graduate School of Information Science and Technology, Aichi Prefectural University, Aichi, Japan (e-mail: s.suzuki@bme.ist.aichi-pu.ac.jp)

K. Oguri is with Faculty of Information Science and Technology, Aichi, Prefectural University, Aichi, Japan (e-mail: oguri@ist.aichi-pu.ac.jp)

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

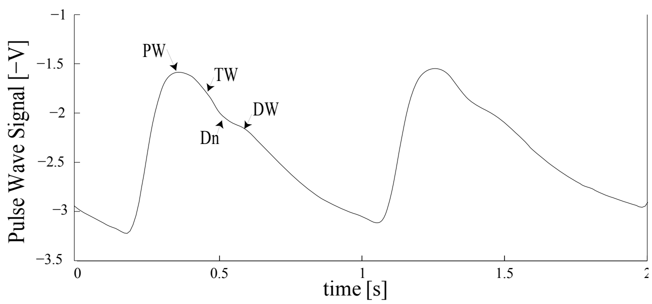

Fig. 2. The feature of pulse wave.

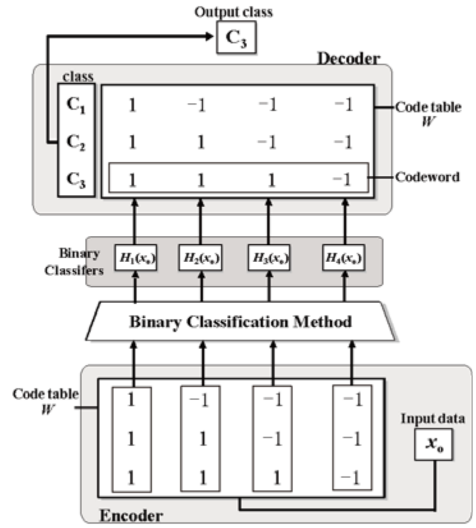

Fig. 3. A diagram of multi-class recognition method based on ECOC method by means of binary classifier.

III. METHOD

In this study, SBP is estimated by a multi-class recognition based on a combination of two-class recognition methods. Two main types of a multi-class predictor have been proposed. One is classical combinations and other predictors such as Neural Network and multi-Support Vector Machine. Another is an exhaustive combination of binary classifiers such as Pairwise and Error Correcting Output Coding (ECOC) method. Many easily-implemented and high-speed binary predictors have been presented. Moreover, we can choose optimal method and parameters for each class. Therefore, a promising high accuracy multi-class predictor would be constructed.

Our study applies ECOC method which suggested by Dietterich and Bakiri [7]. ECOC method is defined by following algorithm.

1. Encoding: Multiple lavels are encoded to binary data with a code table.
2. Binary Classification: Binary classification problem is preformed for each binary sequence.
3. Decoding: Output multi-class are calculated by decoding method with the code table.

In this study, the multi-class predictor comprises an aggregation of AdaBoost. We focus on a Humming distance to decode. Figure 3 depicts a diagrammatic illustration of a multi-class recognition method based on ECOC method by means of an aggregation of binary-class predictors.

A. Code Table

Error rate in the encoding should be as low as possible for high accuracy estimation. To that end, we have to create an applicable code table. This study suggested Global Data Sequence Code Table (GDS-CT) as an opportune code table to encode. The elements of GDS-CT have "+1" or "-1". In particular, only each adjacent element in GDS-CT have "+1".

The number of class is defined G and p denotes the number of binary problem. The number of binary predictor increases exponentially by up to $\sum_{i=1}^{pG-1}(i+1)$ with an increase in G. For given G = 4, Equation (1) denotes an example of code table $W_{GDS\text{-}CT}$.

$$W_{GDS-CT} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & -1 \end{pmatrix} \tag{1}$$

B. Encoding Method

As preprocessing step of multi-class recognition, SBP value is classified into binary class by means of several thresholds (*BPThreshold*). To specify a certain range of SBP, the maximum *BPThreshold* (*maxThreshold*) is defined 180[mmHg] and the minimum *BPThreshold* (*minThreshold*) is defined 80[mmHg]. And then, *BPThreshold* is calculated with Threshold Interval for SBP (*BPTI*). *BinaryClass* as an indicator function consists of the following formula.

$$BinaryClass = \begin{cases} -1 & (BP < BPThreshold) \\ +1 & (BP \geq BPThreshold) \end{cases} \tag{2}$$

C. Binary Classification

We focus on AdaBoost which is one of boosting algorithms in ensemble learning methods [8]. AdaBoost has been applied to many tasks for pattern classification due to its high precision rate and very high speed for the binary classification problem. The key idea of AdaBoost is to give a reweighing of examples in accordance with the best performance tuned in each step, and to succeed to the next step in terms of a renewal error rates by reweighing.

Let us consider a classification problem where for a given feature vector $\boldsymbol{x}$, the corresponding label $\boldsymbol{y}$. Figure 4 shows a diagrammatic illustration of AdaBoost algorithm as binary classifier for given n examples. This study adopted a Classification And Regression Trees (CART) algorithm as a weak learner machine.

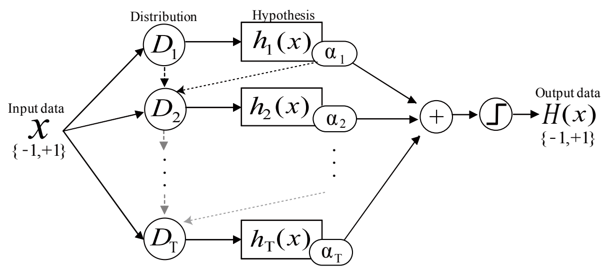

Fig. 4. A diagram of AdaBoost algorithm for binary problem. The number of weak learner is T. For weak learner t, a weight of example i is defined $D_t(i)$. For initialization, $D_1(i)=1/n$. Final decisions $\boldsymbol{H}$ as the majority vote from both h_t denoting a result of weak learner t and α_t denoting the reliability of h_t.

D. Decoding Method

Humming decoding is applied as a decoding method. The Humming decoding algorithm in this study consists of following steps:

1. Humming distance between results obtained with binary predictors and a codeword C_k indicating an optional class for each row was calculated.
2. Multiple classes (*OutputClass*) are defined a codeword which has a minimum Humming distance. *OutputClass* is obtained by following formula,

$$OutputClass = \arg\min_{1\le k\le G} \sum_{m=1}^{p} \frac{1 - H_m C_k(m)}{2} \quad (3)$$

where $\boldsymbol{H}$ is the results which obtained by means of binary predictors, p is the number of binary predictor.

E. Evaluation Index

The estimated SBP (*estSBP*) is calculated with *OutputClass, minThreshold* and *BPTI*.

$$estSBP = OutputClass \times (BPTI - 1) + minThreshold \quad (4)$$

A correlation coefficient (r), a Mean Difference (MD) and a standard deviation of difference (SD) between measured SBP (*measSBP*) and estimated SBP are applied as evaluation indexes. These are obtained by following formulas,

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (5)$$

$$MD = \frac{1}{n}\sum_{i=1}^{n} e_i \quad (6)$$

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - \bar{e})^2} \quad (7)$$

TABLE I
THE NUMBER OF SUBJECTS FOR EACH AGE GROUP

Age [years]	-19	20-29	30-39	40-49	50-59	60-69	70-
N [subjects]	14	29	0	4	17	184	119

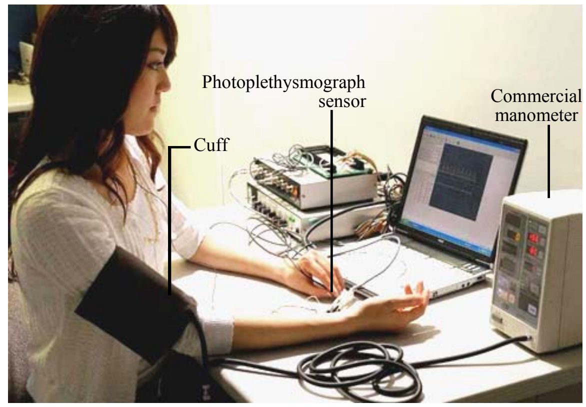

Fig. 5. Experimental environment.

where n is the number of data, e_i is the difference between measured SBP and $estSBP_i$.

IV. EXPERIMENT

160 males and 208 females participated in our experiment. Table I shows the number of subjects for each age group. All volunteers kept their posture for 5 minutes.

SBP was measured on the right brachial by a commercial BP meter (TM-2540R, A&D Corp.) with a cuff every minute. PPG signal was measured on the left finger by PPG reflection mode sensor (DENSO Corp.) continuously. The sampling frequency is 100 [Hz]. The PPG sensor used wavelengths in 520 [nm] regions. Figure 5 shows the experimental environment. In addition, the subjects described their personal information such as their age, height and weight.

V. RESULT

Database was needed for a balance of subjects' age. Subject data are randomly selected for the database in terms of subjects' age and SBP. And then, training set and test set were produced (Table II).

Specifically, we constructed the learning models with the training set to estimate SBP. And then, estimated SBP was calculated by using the models in terms of the test set.

As a result of the comparison between *measSBP* and *estSBP*, MD=-1.2 [mmHg] and SD=11.7 [mmHg] were obtained (Table III). Figure 6 displays a distribution of *measSBP* and *estSBP*.

TABLE II
DETAILED INFORMATION OF EACH DATASET

	training set	test set
N [subject]	195	21
Age [year]	51.1±21.4	46.8±21.1
SBP [mmHg]	118.6±18.4	112.0±17.7
Male [%]	37.8	50.0

TABLE III
COMPARISON OF MEASURED SBP AND ESTIMATED SBP

	r	*MD* [mmHg]	*SD* [mmHg]
training set	0.96	1.3	5.1
test set	0.75	-1.2	11.7

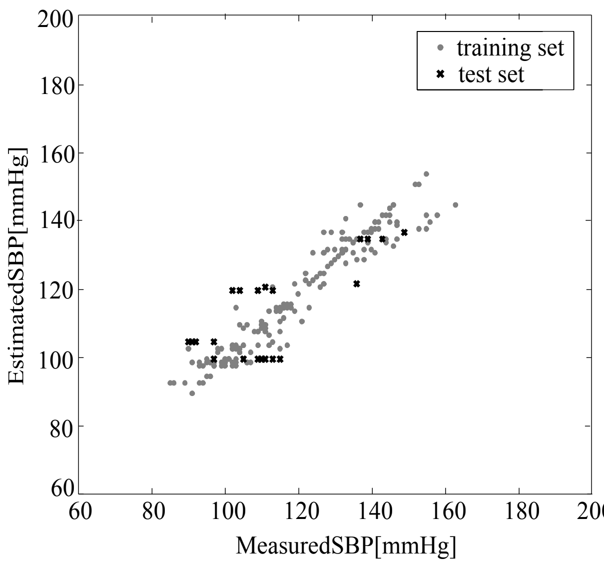

Fig. 6. Distribution of *measSBP* and *estSBP*.

VI. DISCUSSION

According to American National Standard (ANS) for automated sphygmomanometers, *MD* should be less than 5.0 [mmHg] and *SD* should be less than 8.0 [mmHg]. For training set, calculated *MD* and *SD* fully satisfy ANS. For test set, calculated *MD* fully satisfies ANS, but *SD* does not. This reason is that AdaBoost for test set cannot find the value which does not exist in training set. At that time, the error in test set is produced. Therefore, the database needs the sufficient number of sample data to realize cuffless BP monitoring system by proposed method.

In case the number of sample data set is not sufficient, proposed method should be combined a linear analysis such as a multiple regression analysis.

VII. CONCLUSION

This study presented a novel non-invasive technique of SBP estimation without a cuff. We focused on PPG sensor which has been extensively used for the evaluation of the vascular system in recent years. SBP was estimated with a multi-classifier machine. Specifically, the machine comprised a coupling ensemble Adaboost due to its high precision rate and extremely high speed. Humming decoding was applied as a decoding method.

Two data sets to create the learning models and to test were chosen due to the imbalance of subjects' age-group. As a result of the comparison between measured SBP and estimated SBP, *MD* calculated for test set fully satisfies ANS. Cuffless BP monitoring would become more convenient and helpful through this technique since the only one device for monitoring is smaller than traditional measurements. Overall, this study represents a great advance in the field of the cuffless continuous monitoring system development for BP.

In the future, the data of middle age will be increased further. Therefore, it is intended that the estimation accuracy for test set will fully satisfy ANS.

VIII. ACKNOWLEDGMENTS

Authors gratefully acknowledge classes of Aichi Silver College.

REFERENCES

[1] W. Chen, T. Kobayashi, S. Ichikawa and Y. Takeuchi : "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent caliburation", Medical & Biological Engineering &Computing, Vol.38, pp.569-574 (2000)

[2] Carmen C.Y. Poon and Yuan-Zhang : "Using the Changes in Hydrostatic Pressure and Pulse Transit Time to Measure Arterial Blood Pressure", 29th IEEE EMBC, Lyon, France (2007)

[3] Yuan-ting Zhang, Carmen C.Y. Poon, Chun-hung Chan, Martin W.W. Tsang and Kin-fai Wu : "A Health-Shirt using e-Textile Materials for the Continuous and Cuffless Monitoring of Arterial Blood Pressure", 3rd IEEE EMBS, pp.86-89, MIT, Boston, USA (2006)

[4] Mikyoung Park, HeeJung Kang, Young Huh and Kyung-Chul Kim : "Cuffless and Noninvasive Tonometry Mean Arterial Pressure Measurement by Physiological Characteristics and Applied Pressure", 28th IEEE EMBC, New York, USA (2006)

[5] Javier Espina, Thomas Falck, Jens Muehlste and Xavier Aubert : "Wireless Body Sensor Net-work for Continuous Cuffless Blood Pressure Monitoring", 3rd IEEE EMBS, MIT, Boston, USA(2006)

[6] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes", Journal of Artificial Intelligence Research, Vol.2, pp263-286 (1995)

[7] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences Vol.55, No.1, pp119-139 (1997)

[8] Schapire, R., Freund, Y., Bartlett, P., and Lee, W., "Boosting the margin: a new explanation for the effectiveness of voting methods", Proc. 14th International Conference on Machine Learning, pp.322-330, (1997)