Beat-to-beat ambulatory blood pressure estimation based on random forest

Rui He[1], Zhi-Pei Huang[1]*, Lian-Ying Ji[1], Jian-Kang Wu[1] , Huihui Li[2] and Zhi-Qiang Zhang[3]*

Abstract—Ambulatory blood pressure is critical in predicting some major cardiovascular events; therefore, cuff-less and noninvasive beat-to-beat ambulatory blood pressure measure-ment is of great significance. Machine-learning methods have shown the potential to derive the relationship between physio-logical signal features and ABP. In this paper, we apply random forest method to systematically explorer the inherent connec-tions between photoplethysmography signal, electrocardiogram signal and ambulatory blood pressure. To archive this goal, 18 features were extracted from PPG and ECG signals. Several models with most significant features as inputs and beat-to-beat ABP as outputs were trained and tested on data from the Multi-Parameter Intelligent Monitoring in Intensive Care II database. Results indicate that compared with the common pulse transit time method, the RF method gives a better performance for one-hour continuous estimation of diastolic blood pressure and systolic blood pressure under both the Association for the Advancement of Medical Instrumentation and British Hyper-tension Society standard.

I. Introduction

As a result of interaction among cardiovascular control mechanics, continuous ambulatory blood pressure (ABP) provided to be an important physiological parameter for stratifying cardiovascular risks and guiding therapy [1]. Therefore, ABP monitoring is useful in clinical situations like disease diagnosis, drugs experiment and surgical procedure.

Since ABP monitoring plays a significant role in many respects, it is necessary to develop beat-to-beat noninvasive ABP measurement methods. Numbers of approaches have been used in solving the problem. The pulse wave velocity (PWV) related methods are most promising due to the reason that they are cuff-less, relatively more convenient than other approaches and comfortable to use in long-term monitoring.

PWV related blood pressure measurement technique is developed based on Moens-Korteweg (M-K) equation as in

$$\mathrm{PWV} = \frac{\mathrm{D}}{\mathrm{PTT}} = \sqrt{\frac{\mathrm{t}\mathrm{E}_0 \mathrm{e}^{\alpha \mathrm{P}}}{\rho \mathrm{d}}} \tag{1}$$

where PTT is pulse transit time, D, t, α, d and E_0 are parameters of the vessel, while ρ stands for blood density. In most studies, a linear relationship is believed to exist between PTT and ABP. Therefore, a calibration procedure could be applied to obtain the parameters in the linear equation. As numerous physiological factors modulate the vessel and blood parameters, it is necessary to run continuous re-calibration, which makes measure devices designed based on those methods uncomfortable to use in daily life.

To avoid the recalibration procedure, one potential idea is estimating parameters in M-K equation from physiological signals such as the pulse and electrocardiograph (ECG) wave. Various researches have shown that waveform properties of those signals are related to equation parameters more or less. Such as K Value extracted from the pulse wave reflects vessel elasticity [2]. Heart rate (HR) derived from ECG waves is also associated with ABP [3]. In our previous study [4], 28 pulse wave time-domain features were proved to have correlations with ABP. Our results thus suggested those features are potentially useful in solving the recalibrating problem of PTT method. In fact, there have been studies focusing on estimating ABP from parameters extracted from various physiological signals based on Machine Learning (ML) techniques [5].

In one representative study [6], researcher collected BP, photoplethysmography (PPG) and ECG from 410 individuals, then tested the performance of several ML techniques for BP estimation. The random forest (RF) technique obtained the best result. However, the blood pressure values in the study were collected by an aneroid sphygmomanometer so they didn't apply the technique to beat-to-beat ABP estimation. Another study concentrated on using common PTT method to calculate beat-to-beat ABP values on the Multi-Parameter Intelligent Monitoring in Intensive Care (MIMIC) database [7]. Due to the lack of available record in database, researcher only use data from 9 patients. By importing data from MIMIC-II database, we were able to apply RF and PTT method to estimate beat-to-beat ABP on dataset from 285 individuals in our study to overcome shortages mentioned above.

This work has two purposes. One is to determine whether more features extracted from ECG and PPG signal could be applied to improve accuracy of PTT-based ABP estimation method. For this purpose, 18 time-domain features and corresponding BPs from each dataset were collected. The other purpose is to check whether RF technique performs better for beat-to-beat SBP and DBP estimation than common PTT method. We estimated beat-to-beat BPs in one-hour period using models based on RF technique, LR technique and typical PTT method respectively. Models were trained by features and BPs value pairs in 30 minutes. The accuracy of each technique was evaluated by percent of sets which reach the Association for the Advancement of Medical Instrumenta-tion (AAMI) and British Hypertension Society (BHS)

1. R. He (e-mail: herui13@mails.ucas.ac.cn), Z.-P. Huang (e-mail: zhphuang@ucas.ac.cn), L.-Y. Ji (e-mail: jilianying@ucas.ac.cn), and J.-K. Wu (e-mail: jkwu@ucas.ac.cn) are with Sensor Networks and Applications Research Center, University of the Chinese Academy of Sciences, Beijing 101407, China.
2. H. Li (e-mail: hh.li@siat.ac.cn) is with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China.
3. Z.-Q. Zhang (e-mail: z.zhang3@leeds.ac.uk) is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9AY, U.K.

standards among all available datasets to come to a conclusion.

II. Method

A. Data collection and feature extraction

The Multi-Parameter Intelligent Monitoring in Intensive Care II (MIMIC-II) database is a massive database recording data from temporal patients in ICU [8]. MIMIC-II contains more subjects than the previous MIMIC database. Currently there exist 4350 records recorded from more than 2000 patients, part of which contains continuous ABP, PPG and ECG signals at the same time, providing an available data source for this study. All the available data were downloaded from MIMIC-II via query builder explorer [9]. 285 records longer than 1.5 hours and contain both ABP, ECG II and PPG signal were selected for feature extraction. WFDB Toolbox[10] was introduced to peak and onset detection in ABP and ECG signal. A derivative based algorithm was applied to PPG signal in the detection procedure. Finally, 18 features were extracted for each heart beat based on the peak and onset position detected. In addition to the 15 features including RRI, PTT, RBW, RBW_z and DBW_x (x from 10 to 75) illustrated in Fig.1, other 3 features including K value, AmBE and RBA were calculated as equation (2), (4) and (5):

$$K = \frac{\overline{ppg_{OO'}} - ppg(t_O)}{ppg(t_O) - ppg(t_{O'})} \tag{2}$$

In which

$$\overline{ppg_{OO'}} = \frac{1}{t_{O'} - t_O} \sum_{i=t_O}^{t_{O'}} ppg(i) \tag{3}$$

$$AmBE = \frac{1}{t_E - t_P} \sum_{i=t_P}^{t_E} \frac{ppg(i) - ppg(t_E)}{ppg(t_P)} \tag{4}$$

$$RBA = \sum_{i=t_O}^{t_P} ppg(i) / \sum_{j=t_O}^{t_{O'}} ppg(j) \tag{5}$$

The branch width features are chosen according to [11], while the AmBE feature is chosen according to patent [12].

C. Feature selection

After extracting all the features and target value groups, a feature importance measurement procedure was imported to rank features following [13]. The RF model based ranking method was chosen due to the reason that it is simple and effective for non-linear relationship description. The principle of this method is to use each feature to train a RF for regression model for ABP evaluation separately. The significance score of each feature is then evaluated as the Pearson correlation between the model's outputs and target values on the test set.

RF are formed by growing a collection of decision tree structured classifiers $\{h(\mathbf{x}, \Theta_k), k=1, \ldots K\}$ which output a vector of class labels y from input $(\mathbf{x}, \Theta_k)$. $\{\Theta_k\}$ are independent identically distributed random vectors with values range from 1 to N, N is splits number. The RF prediction is as in

$$\hat{y}(\mathbf{x}) = \frac{1}{N} \sum_{k=1}^{K} h(\mathbf{x}, \Theta_k) \tag{6}$$

for regression. In practice, a two-step procedure is introduced to generate those trees according to [14]. At step one, bootstrap method is applied to randomly select N samples from all samples. At step two, for each tree in the forest, each node is corresponding to a group of samples. To grow a tree, select m features randomly for each leaf node, then divide the samples in the leaf node to its child nodes according to one feature. The feature could provide greatest information gain among m features. Usually, the preferred m value is the square root of the number of available features [15], so for each feature m is 1. For the concern of train time, use 100 as value of K. A 3-fold cross-validation was applied to avoid over-fitting, during which 70 percent of train set data were randomly selected for training while the left 30 percent were applied for testing for 3 times. The top-ten highest ranking feature vector X_F and corresponded scores were recorded for BPs.

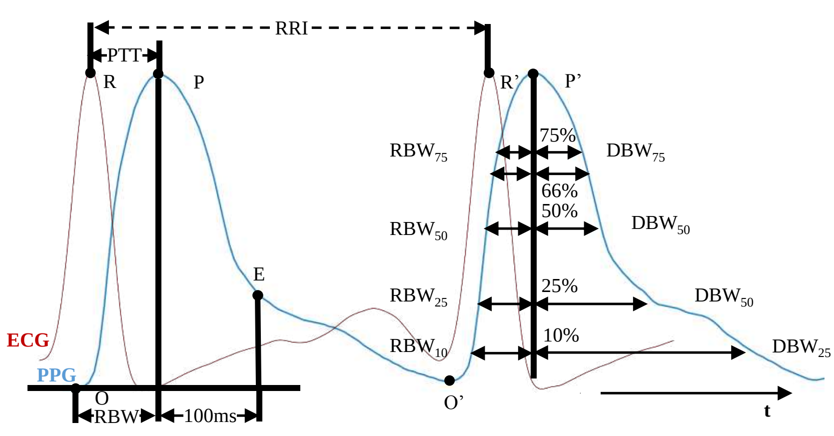

Figure 1. Illumination of pulse wave feature points and features.

Although the relevance of features to ABP could be revealed by the scores, correlation between different features still remained unknown, which may lead to feature redundancy. Therefore, correlation analysis was carried out between all the highest ranking features of SBP and DBP, represented by the correlation coefficient (r) and the p-value (p). Those features which were significantly correlated ($|r| > 0.85$, $p < 0.05$) were selected. For each group of significantly correlated features, only the one which has the highest score were selected. It should be noted that although ABP estimation was applied to each record separately, the feature selection procedure was performed on the whole mixed dataset.

D. ABP estimation

One-hour beat-to-beat estimation of SBP and DBP were run on each record separately. Both train set and test set are all the feature-BP pairs within a specific period of time. 30 minutes was selected as train set length to expand the range of BPs values while one hour was chosen as test set length. Long duration data were partitioned to generate train/test set pairs to ensure all the dataset have approximately same time length with each other. Finally, there were totally 1246 pairs of data for DBP and 1260 pairs for SBP. BPs estimation was then performed for each pair respectively using two techniques:

- LR. The principle of this method is assuming that a linear correlation exists between features and BPs as in the equation

$$y^{bp} = \beta_0^{bp} + \sum_{i=1}^{n} \beta_i^{bp} X_F(i) \tag{7}$$

where bp={SBP, DBP}. By applying plain ordinary least square method to this equation, we get the parameter vector β as

$$\beta = (X_F^T X)^{-1} X_F^T y^{bp} \tag{8}$$

- RF. The principle of this method is mentioned in II.C. Instead of one feature, we use all the selected features for BP estimation. For SBP the feature number is 7, m=2, for DBP the feature number is 7, m=3.

As with each technique, the estimation procedure took two steps. For the first step, a 10-fold cross validation was run on the train set. The purpose of the step was to acquisition specific ML model parameters and avoid over-fitting. For the second step, SBP and DBP were estimated from beat-to-beat features. Then the estimated BP values were compared with measured BP values to determine whether the result reached BHS and AAMI standard.

For comparison purpose, a PTT-based method [7] was imported to estimate beat-to-beat BPs on all the datasets. The fundamental base of this method is the equation

$$y^{bp} = \beta_0^{bp} + \beta_1^{bp} * PTT \tag{9}$$

where bp={SBP, DBP}. Obviously the PTT-based method is an LR method with PTT as the only feature.

E. Statistics

To compare the performance of difference technique, percent of the set which reaches BHS and AAMI standard over all datasets were recorded. Mean square error (MSE) was used as an accuracy interpreter. Then correlation analysis was carried between estimated BP and measured BP to attain the Pearson correlation coefficient (r). For each technique and BP type, those datasets whose estimated values were significantly correlated with measured values were recorded.

It should be noted that all the programs were written in python using APIs provided by scikit-learn module [16].

III. Results And Discussion

The feature ranking results for SBP and DBP are shown in Fig. 2. According to the result, RBW and DBW features have the most important impact on BPs. Correlation study was performed to identify if those impacts come from the correlation between those features and PTT. It turns out that DBW50, DBW33, DBW25 and DBW10 show significant correlation ($r \in [0.8535, 0.8668]$) with each other. The result indicates that the shape of pulse wave descending limb barely changes with time growth. According to Fig. 2, the importance of DBW50 feature ranks higher than other DBW features for SBP. As a result, the DBW10, DBW25, DBW33 were excluded when it came to SBP estimation. No features were removed when estimating DBP. As illustrated in Fig. 3, TABLE I. , TABLE II. and TABLE III. , the RF technique performs better than the LF technique, while the latter shows slight enhancement compared to common PTT method.

Two conclusion could draw from those results. On one side, the common PTT method is based on the linear relationship between PTT and BPs so it is basically an LR method with one feature. As the LR technique with 18 features results better than the PTT method, it is proven that some of those 22 features could provide extra information when it comes to the estimation of beat-to-beat BPs. On the other hand, RF technique shows a higher accuracy than the LR technique, especially for SBP estimation. This result could help validate the conclusion of [6]. It might be the nonlinear characteristic of RF that leads to the result.

To reveal the source of estimation error, correlation analysis was applied to train/test set BP ranges and BP estimation MSEs. Of all results, the correlation between the test set BP range and corresponding MSE is most significant for both SBP and DBP, as illustrated in Fig. 5. This phenomenon indicates that the performance of RF method degrades when applied to subjects with wildly fluctuated BP values. The feature deficiency is a potential cause for the result.

It should be noted that there is still a large scale of datasets which could not meet both BHS and AAMI standards. Many reasons may lead to this result. First of all, we didn't exclude those who had cardiovascular diseases from the database. However, it's known that renovascular disease or use of procaine hydrochloride could result in an abnormal change in blood pressure and reduce the accuracy of estimation.

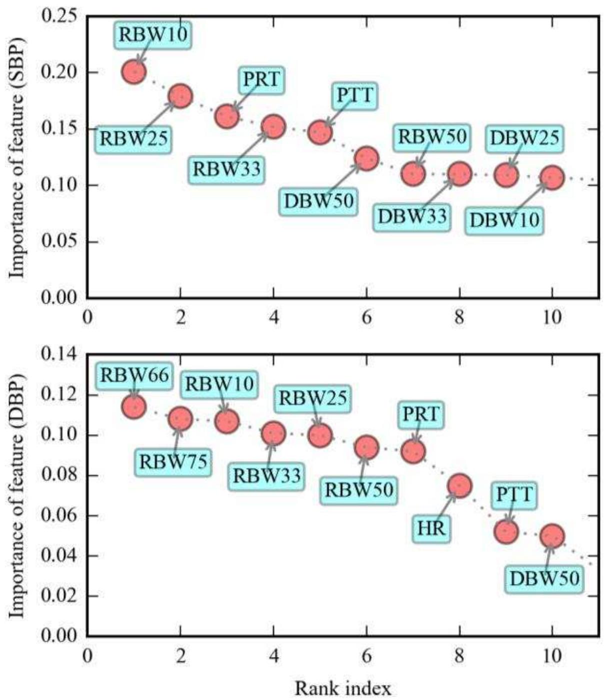

Figure 2. Feature ranking result.

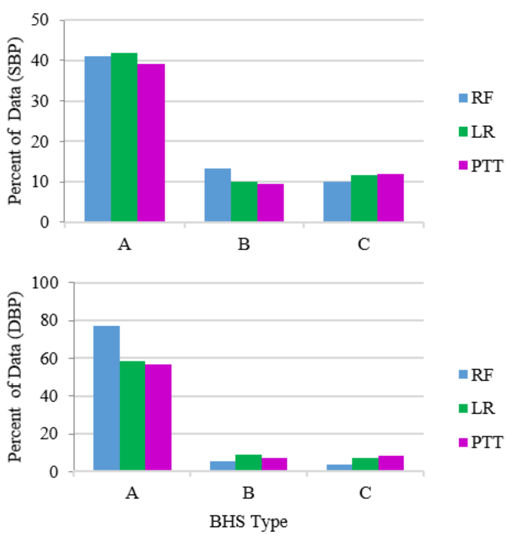

Figure 3. Percent of data reached each BHS standard.

TABLE I. PERCENT OF DATA REACHED AAMI STANDARD

BP Type	Technique Used		
	RF	**LR**	**Typical PTT**
DBP	90.70%	88.21%	83,31%
SBP	66.12%	59.90%	53.65%

TABLE II. MSE (MEAN±STD.)

BP Type	Technique Used		
	RF	**LR**	**Typical PTT**
DBP	4.44±3.72	4.98±4.31	7.12±17.98
SBP	8.29±5.84	9.32±7.65	12.77±41.19

TABLE III. NUMBER OF SIGNIFICANTLY CORRELATED SIGNALS

BP Type	Technique Used		
	RF	**LR**	**Typical PTT**
DBP	297	290	123
SBP	271	271	109

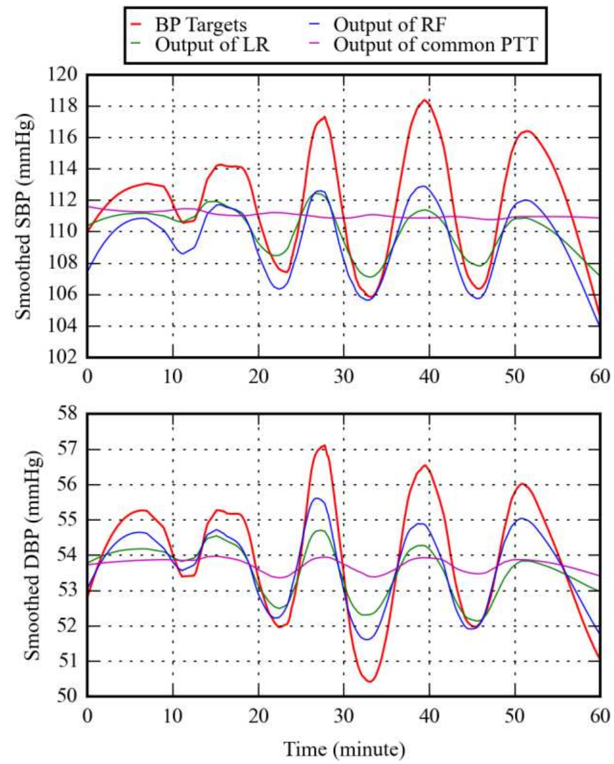

Figure 4. Invasive BPs and prediction BPs output by different methods (MIMIC-II record: a46108_0030).

Although the RF technique has an advantage over other methods, its weakness could not be neglected. Firstly, the estimation speed of RF is much slower than LR though the forest scale is relatively small. On average RF technique takes 20 times longer than LR technique, so using the technique for real-time estimation might be difficult. Secondly, due to the multi decision tree structure of RF method, one feature can appear in more than one tree in the estimation procedure. As a result, the relationships between different features and outputs are not obvious to observe. Therefore, analyzing the result doesn't give us much insight about the physiological meaning of different features.

To summarize, the use of more features, as well as RF technique, leads to higher accuracy for one-hour beat-to-beat SBP and DBP estimation. The technique could be potentially applied to enhance the performance of common noninvasive cuff-less BP monitoring method. Due to the problem of reduced estimated accuracy with target BP range variation, further investigations should be undertaken to obtain more

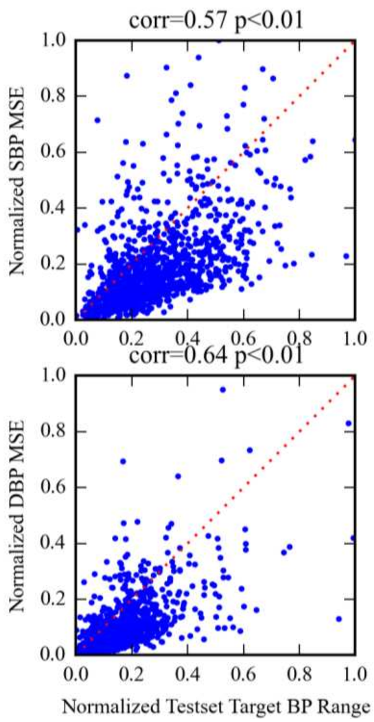

Figure 5. Relationship between test set BP range and estimation MSE.

valid features. In addition, effect of this method should be tested on normal people instead of ICU patients.

REFERENCES

[1] M. E. Ernst and G. R. Bergus, "Ambulatory blood pressure monitoring. (Review Article)," South. Med. J., vol. 96, no. 6, pp. 563–569, 2003.

[2] L. Zhichang, Z. Song, and Y. Wenming, "A research on characteristic information of pulse wave," J. Beijing Polytech. Univ., vol. 22, no. 1, pp. 71–79, 1996.

[3] R. A. Fernandes, I. F. Freitas Júnior, J. S. Codogno, D. G. D. Christofaro, H. L. Monteiro, and D. M. Roberto Lopes, "Resting Heart Rate is Associated with Blood Pressure in Male Children and Adolescents," J. Pediatr., vol. 158, no. 4, pp. 634–637, Apr. 2011.

[4] Y. Su and R. He, "A study on the correlation between arterial pressure, pulse transit time and other time-domain characteristic parameters of radial pulse wave," in ICSIIT, 2015.

[5] P. Li, M. Liu, X. Zhang, X. Hu, B. Pang, Z. Yao, and H. Chen, "Novel wavelet neural network algorithm for continuous and noninvasive dynamic estimation of blood pressure from photoplethysmography," Sci. China Inf. Sci., 2015.

[6] E. Monte-Moreno, "Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques," Artif. Intell. Med., vol. 53, no. 2, pp. 127–138, 2011.

[7] B. Escobar, R. Torres, E. D. I. De Antioquia, and U. Ces, "Feasibility of Non-invasive Blood Pressure Estimation Based on Pulse Arrival Time: a MIMIC Database Study," in 2014 Computing in Cardiology Conference, pp. 1113-1116.

[8] Lee, Joon, Scott, D. J., Villarroel, M., Clifford, G. D., Saeed, M., & Mark, R. G. "Open-access MIMIC-II database for intensive care research." Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE, 2011.

[9] Z. Zhang, "Accessing critical care big data: a step by step approach," J. Thorac. Dis., vol. 7, no. 3, p. 238, 2015.

[10] WFDB Applications Guide, G. Moody, MA, 2012, pp. 165.

[11] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," 2013 IEEE Int. Instrum. Meas. Technol. Conf., pp. 280–283, 2013.

[12] H. Yan, F. Hong, Y. Li, X. Yang, M. Xiao, and M. Wei, "Continuous blood pressure measuring device," CN Patent 102429649 A, May 2, 2012.

[13] R. Genuer, J. Poggi, C. Tuleau-Malot. "Variable selection using Random Forests," Pattern Recognition Letters, vol. 31, no. 14, pp.2225-2236., 2010.

[14] L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001.

[15] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, "Random forest: a classification and regression tool for compound classification and QSAR modeling," J. Chem. Inf. Comput. Sci., vol. 43, no. 6, pp. 1947–1958, 2003.

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine Learning in Python," Machine. Learning., vol. 12, pp. 2825–2830, 2012.

[17] Clifford, G. D., Scott, D. J., & Villarroel, M. User guide and documentation for the MIMIC II database. MIMIC-II database version, 2, 95.,2009.