International Journal of Healthcare and Medical Sciences
ISSN(e): 2414-2999, ISSN(p): 2415-5233
Vol. 3, No. 12, pp: 103-116, 2017
URL: http://arpgweb.com/?ic=journal&journal=13&info=aims

Analysis of Heart Rate Variability Via Health Care Platform

Yong-Hong Hsu* Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan

Yang-Yi Chen Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan

Chun-Liang Lin Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan

Changchen Zhao Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan
School of Automation Science and Electrical Engineering, Beihang University, Beijing

Abstract: This research task develops a mobile healthcare analysis system (PHAS) which combines both easy ECG signal measurement and reliable analysis of heart rate variability for home care purpose. The PHAS is composed by a health care platform (HCP) and a data system analysis (DSA) module. The HCP consists of a self-developed two pole electrocardiography (ECG) measuring device and the DSA a data processing unit for detection and analysis of heart rate variability. For the DSA module, the adaptive R Peak detection algorithm is proposed to reliably detect the R peak of ECG for HRV analysis. A number of features are extracted from ECG signals. A data mining method is employed for HRV analysis to exploit the correlation between HRV and these features. Experiments are conducted by establishing a database of ECG signals measured from 29 subjects under rest and exercise condition. The results show the PHAS's significant potential in mobile applications of personal daily health care.

Keywords: Electrocardiography; Heart rate variability; Health care platform; Data processing; Home care application.

1. Introduction

There exists an increasing demand for mobile health care devices for home care purpose over the last few years. The development of tele-medicine and tele-homecare plays an important role in computer science and healthcare application [1-3]. The healthcare system provides abundant of contextual information and alerting mechanisms for both users and professionals. Also, the wearable health system is an emerging technology for continuous monitoring by using biomedical signal such as electrocardiography (ECG), blood pressure, or oxygen.

The ECG signal measures the electrical activity of heart beat by using several electrodes. An accurate measurement of ECG can provide important diagnostic information for doctors' reference. ECG can be used to diagnose a wide range of diseases such as heart attack, pulmonary embolism, perceived cardiac dysrhythmias, etc. Recent researches also focus on developing various kinds of health care devices that accurately measure and monitor ECG signals [4, 5].

The heart rate variability (HRV) based on the ECG is another diagnostic basis related to human health. HRV is a physiological phenomenon which is the variation of time interval between heart beats. HRV is proven to be regularized by the autonomic nervous system and often be used to reflect mental and psychological stress [6, 7]. Analysis methods can be taken under time-domain and frequency-domain. Some also employ non-linear methods. The analysis of HRV can be useful to diagnose cardiovascular and noncardiovascular diseases such as diabetic neuropathy, myocardial dysfunction, or Liver cirrhosis.

However, diagnosis based on ECG and HRV analysis is usually inconvenient for the home care purpose. For medical diagnosis, the measurement of ECG signal has to be accurate enough. ECG signals are typically recorded by placing ten electrodes on the patient's limbs and on the surface of the chest for a short period of time. However, patients at home may be inconvenient to go to the hospital due to a variety of reasons. Hence, development of a mobile ECG measurement device becomes necessary. Besides, patients are usually hard to decide what the illness they have due to the lack of medical experience even if accurate ECG singles are available. This has led to the demand for developing a health care platform with measurement and analysis abilities for the home care purpose.

In this research task, we build up a mobile health-care platform with easy ECG measurement and HRV analysis capabilities for the home care purposes. The platform combines ECG measurement module and data system analysis

*Corresponding Author

module. The objective of this research is to integrate the data system analysis (DSA) into the mobile health care platform so that patients at home can perform a preliminary, near-professional self-diagnose of their illness.

The contributions of this paper are summarized as follows: 1) we integrate HRV analysis into a health care platform so that diagnoses for patients at home become easier, 2) the adaptive R peak detection algorithm is proposed to reliably detect R peak of ECG for HRV analysis, and 3) we propose an HRV analysis method to exploit the correlation between HRV and features derived from ECG signals.

The remainder of this article is organized as follows. Section 2 briefly reviews some of the related works. Section 3 gives an overview of the proposed platform and details of the DSA module are presented in Section 4~6. Section 7 reports the experimental results and Section 8 concludes the whole article.

2. Related Work

This section mainly focuses on two types of works, i.e., health care platform design and ECG and related signal analysis.

2.1. Health Care Platform Design

Health care platform design has long been an active research topic. Considerable health care platforms that are easy to use and reliable have been developed. In, Alemdar and Ersoy [8] advocated the idea of wireless network for healthcare which include: body area network, personal area network, gateway to the wide area networks, wide area network, and end-user healthcare monitoring application. In. Darwish and Hassanien [9] presented several examples of body area network system for continuous monitoring of patients to explain the importance of body sensor networks in medicine and caregivers. In Baig, *et al.* [10] focused on reviewing wearable ECG monitoring system in the form of wireless, mobile, and remote technologies related to older people. Afterwards, there are some examples to establish ECG monitoring system, such as Yang, *et al.* [11] developed a wearable patch-style heart activity monitoring system (HAMS) which was used for recording ECG signal,. [5] designed an ECG processor with HRV analysis in a real-time potable health system, and Wu et al. [6, 7] focused on the real-time ECG monitoring system to estimate heart rate variability and added wireless transmission to the self-developed biofeedback system. More recently, [12] performed ambulatory EEG recording and real-time healthcare applications by using some miniaturized wireless electroencephalography (EEG) sensors. They proposed a scalable real-time energy-efficient EEG compression scheme for the wireless body area sensor network.

2.2. ECG and Related Signal Analysis

With regard to QRS complex detection, Yeh and Wang [13] focused on development of the difference operation method [14] to detect QRS wave of ECG signals. Wu and Lee [15] proposed an improved methodology which provides a total solution from hardware design to software development for HRV analysis, including design of an ECG acquisition platform, QRS signal processing, and HRV parameter calculation. Besides, the regularization algorithm of RR interval can be traced back to [16-18].

For HRV analysis, in 1996, Task Force of the European Society of Cardiology [19] published the standards of HRV analysis, and application with clinical use. [6] described the procedures of HRV analysis, such as time-domain, frequency-domain, and nonlinear analysis. Also, it depicted the relationship with HRV and diseases. In. Niskanen, *et al.* [20] proposed the software for HRV analysis in Matlab environment. After that, the algorithms of frequency analysis via HCP that involves resampling RR sequence have been proposed in [21], and computing the spectrum with fast Fourier transform [22] and autoregressive model [23, 24].

In Milley [25] proposed the idea of healthcare and data mining. In. Cios and Moore [26] analyzed the uniqueness of medical data mining. In Pecchia, *et al.* [27] proposed the data mining application with healthcare. So far, the data mining application appeared in many aspects. Mougiakakou, *et al.* [28] exploited it to medical image acquisition and data management, Pecchia, *et al.* [27] used data mining for early detection of the worsening in patient's condition to enhance effectiveness and efficiency of the health monitoring system, and Jain, *et al.* [29] analyzed several mining physiological conditions to verify HRV analysis. I Wang [30] punished the feature selection with kernel class so as to provide a variety of selection modes with different search strategies. With the extensive experiments for feature selection. Molina, *et al.* [31] conducted distinct feature selection algorithm, and conducted the experimental evaluation. With regard to establish the system model, the application notions come from the Bayesian theorem [14].

3. System Overview

3.1. Health Care Platform (HCP) Module

The HCP module here is implemented by adopting a patented handheld device [32], developed by the authors, to easily collect ECG signals from two conductive electrodes. See Figures 1 and 2 for the description of the module.

Fig-1. Structure of the HCP

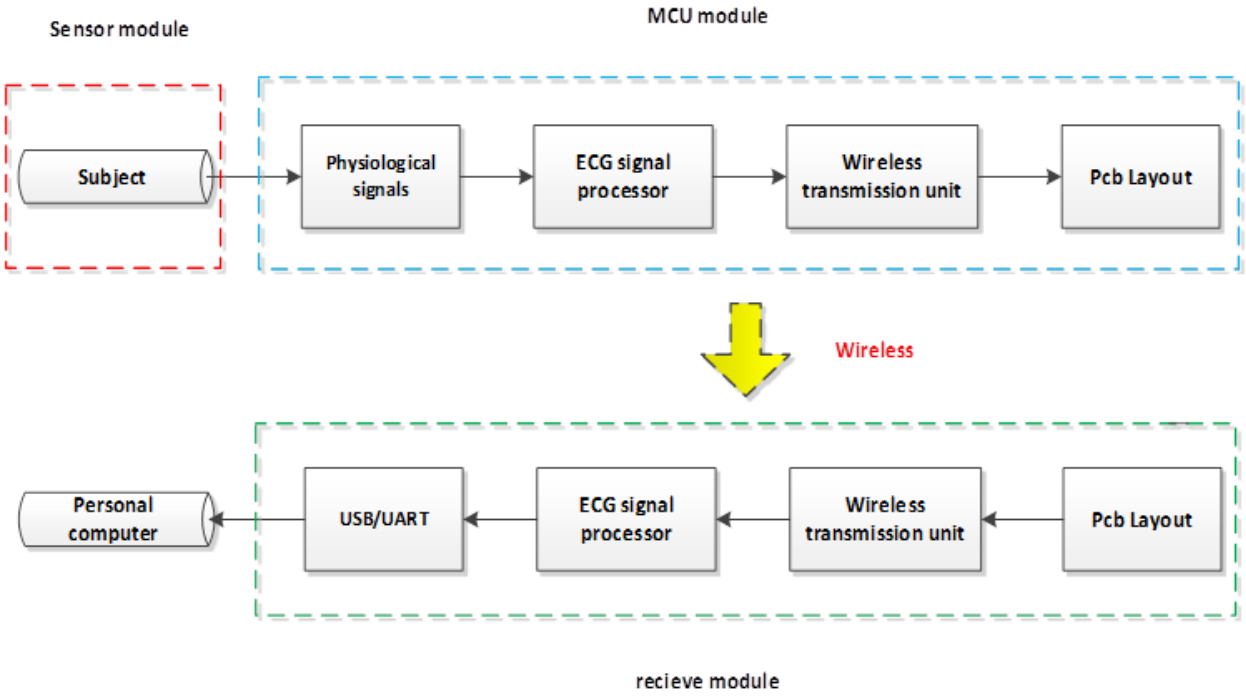


Fig-2. Self-developed two pole ECG acquisition device.

The measurement of ECG signals is implemented by a handheld device using a pair of electrode patches. The signal process is processed by an analog filter/amplifier unit with the frequency range 0.5 to 40 Hz.

The receiver module consists of a PCB-based antenna, an RF communication unit, a microcontroller, and an USB/UART interface. The ECG signal processing algorithm adopts the QRS wave detection method proposed in Pan and Tompkins [33]. The ECG signals are displayed on the monitor with the signal sampling rate at 125 Hz.

3.2. Data System Analysis (DSA) Module

The DSA works with three functions: digital signal processing, database analysis, and establishment of the DSA predictive model. The digital signal processing unit includes QRS complex detection. Database analysis includes feature generation, data mining, and feature selection.

For the digital signal processing, the essential parts are the QRS complex detection and construct RR interval and EDR signal. Afterwards, the data processing analysis is conducted by including feature generation, data mining, feature selection, and establishing the system model. In the data processing analysis, we obtain essential information from time-domain, frequency-domain, and perform nonlinear analysis for decision making. In the time domain analysis, we collect data from a bio-signal processor module obtaining SDNN (standard deviation of RR intervals), RMSSD (root mean square of successive difference RR intervals) and NN50 (the number of successive RR intervals differing more than 50ms). For the frequency analysis, the spectrum distribution of HF (high frequency band), LF (lower frequency band), VLF (very Lower frequency band), and the ratio of HF and LF are obtained. In addition, nonlinear analysis methods also provide the characteristic messages from Poincare plot, approximate entropy, and derived parameters. Then, it passes through a data mining procedure to exhume the unknown informatics from database.

4. Digital Signal Processing

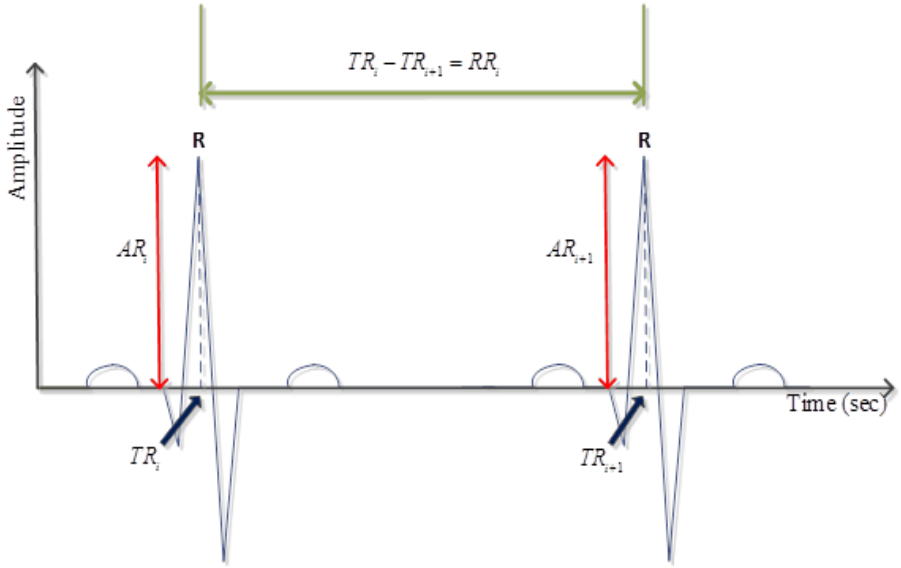
4.1. QRS Complex Detection

The digital signal processing of DSA is designed to extract key information. We derive time series sequences from the original ECG signal, i.e. R-R interval time series and EDR time series. The QRS complex detection consists of adaptive R-peak detection and generation of EDR signal. R-R interval sequence can be established with the adaptive R peak detection algorithm, and EDR time sequence can be established for the EDR signal.

4.2. Adaptive R Peak Detection Algorithm

The acquired ECG signals are preprocessed via high-pass filtering, low-pass filtering, and band-pass filtering processes, to minimize noise corruption while preserving the meaningful part. Figure 3 shows the schematic figure of R-peak detection.

Fig-3. Schematic figure of R peak detection from ECG

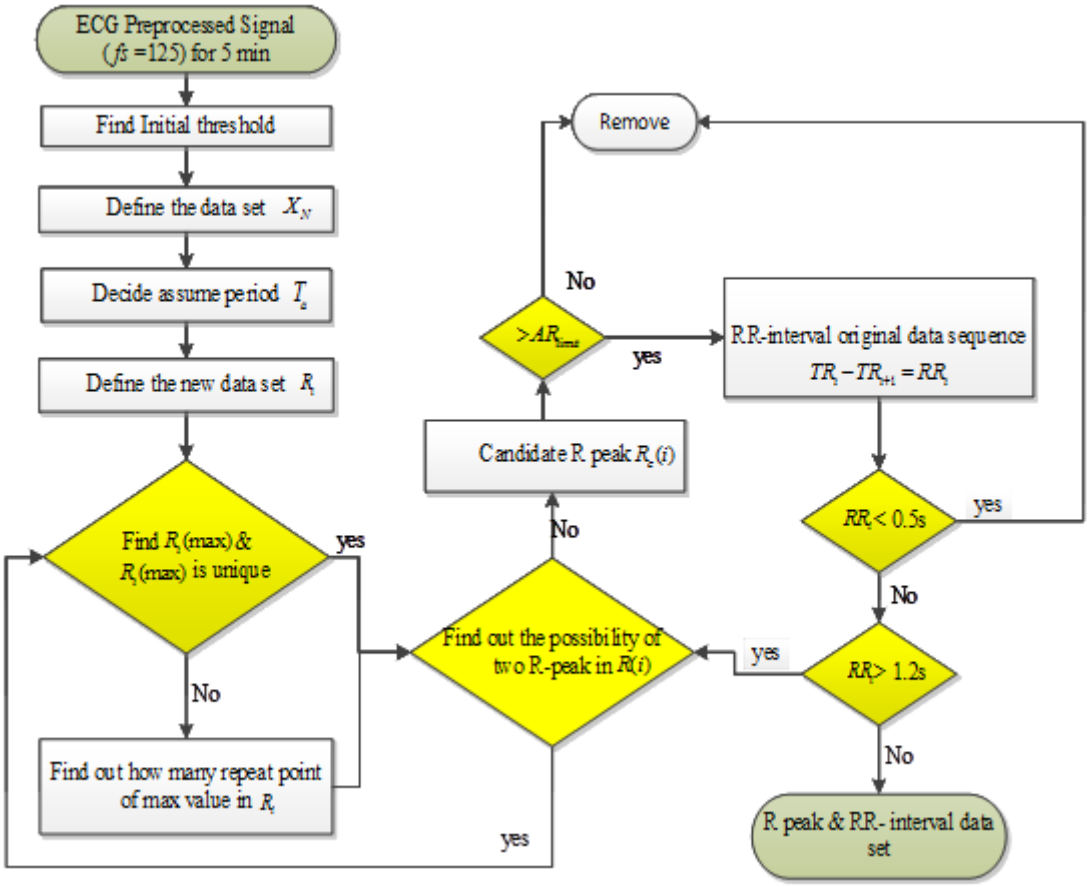


In Figure 3, AR_i is the amplitude of the ECG signal, RR_i represents the time series in ECG, and $TR_i - TR_{i+1} = RR_i$. The operational steps of adaptive R peak detection algorithm can be generated by the steps depicted in Figure 4. This algorithm involves detecting the candidate R peak and constructing the R-R interval data set. As for the candidate R peak, the first step is to find the initial threshold from the ECG database. Because the original data set is saved at 125 samples per second, there might be the case that R peak appears twice in one sec. Secondly, we assume the optimal period T_a of the heart rate to obtain R peak which appears once in the data set R_i. The third step is to find $R_i(max)$ which is unique. If $R_i(max)$ is not unique, we proceed to find the repeating point and tune the ECG data. There is only one specific $R_i(max)$ needed to be considered.

Before conducting the decision process for the R-R interval from the data set $R_c(i)$ which includes the information of all R peaks in ECG, we define the lower limit of amplitude AR_{limit}. The strategy of this step is to remove the unsuitable R peaks in $R_c(i)$, whether a detected event exhibits a suitable R peak or not. If the detected RR_i is large, for example, higher than 1.2 sec (40 bpm) which may relates to a false detection; likewise, if RR_i is less than 0.5 sec (120 bpm), there is probably a missing detection. Finally, we obtain the R-peak data set L(i) and R-R interval data set RR(i).

Different from the usual R-peak detection algorithm (such as Pan-Tompkins' QRS detection algorithm), we have developed an adaptive R-peak detection algorithm shown as in Figure 4 which uses the lower limit of ECG amplitude, instead of the upper and lower limits of amplitude, to simplify the analysis procedure. In addition, we have adopted the optimal period to increase accuracy of the R-peak detection.

Fig-4. Flowchart of the adaptive R peak detection algorithm



4.3. Generating EDR Signals

With periodic breathing, human's chest movement reflects in the R peak of an ECG signal. A surrogate respiratory signal of respiration can be referred as the EDR. We extract the R-peak data set L(i) as the reference point. EDR signal is obtained by calculating the area enclosed by the baseline corrected ECG in the region 100 ms beyond the R-peak. The transmission rate of ECG original data is set at 125 Hz, which means that the interval between a complete P, Q, R, S, T waveform is 8 ms. The first step is to set each R peak as the datum point. Next, we search for 13 complete ECG waveforms located after the datum point. As for the EDR signal, the sum of the 13 ECG data is to substitute the area enclosed by 100 ms after the datum points (i.e. 13 ECG waveforms spanning 104 ms).

4.4. Preprocessing R-R interval

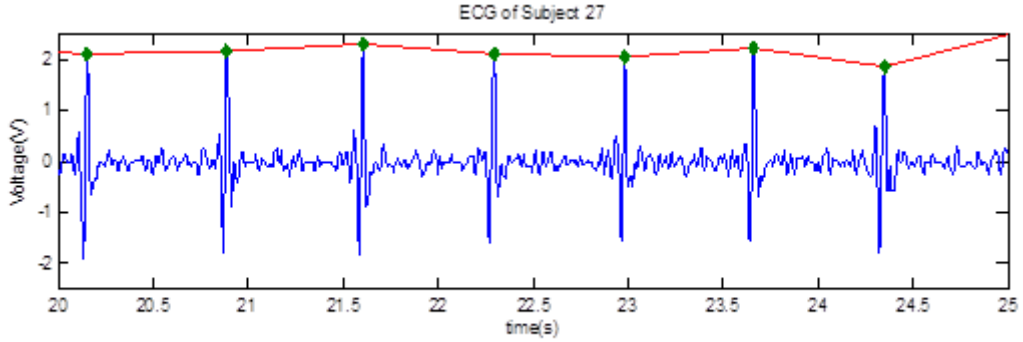
The preprocessing of R-R interval deals with the issue of smoothness prior regularization. We adopt the method presented in Tarvainen, *et al.* [16] to eliminate the base-line drift in R-R interval.

4.5. Signal Processing Verification

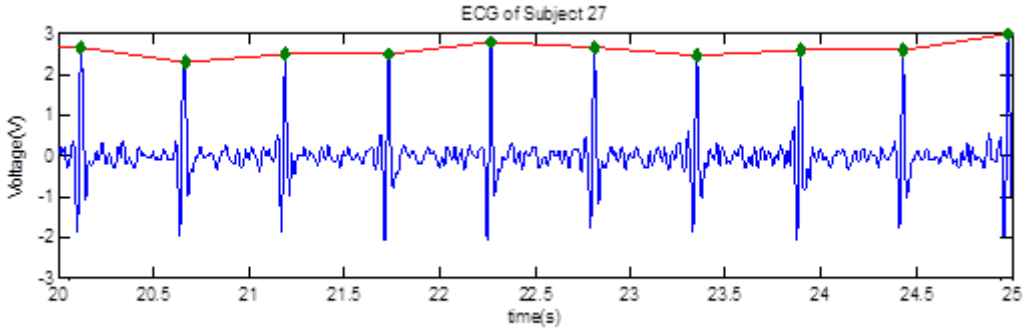
Verification of the digital signal processing consists of four parts: verification of the adaptive R peak algorithm, reconstruction of ECG signal to generate EDR signal, generation of R-R interval and EDR signal, and signal processing of the R-R interval.

With the measurement of ECG signal by the HCP, one can directly observe the differences among all subjects' ECG signals under the rest and the exercise conditions. In addition, it verifies accuracy of ECG measurement via the HCP. Figure 5 shows differences of the subject's ECG signals under the rest and exercise conditions. In this figure, the exercise condition shows higher heart rate than that in the rest condition, thereby affecting composition of the R-R interval. Generation of R-R interval and EDR signal relies on the adaptive R peak detection algorithm and reconstruction of ECG signals. The generated R-R interval and EDR signals are displayed in Figures 6 and 7. One can observe slight variation of R-R interval and EDR signal in different experimental conditions, such as the R-R interval variation infers to HRV, and EDR variation is referred to the respiratory activity under the rest and exercise conditions.

Fig-5. R-R interval of the subject A. (a) rest (b) exercise condition

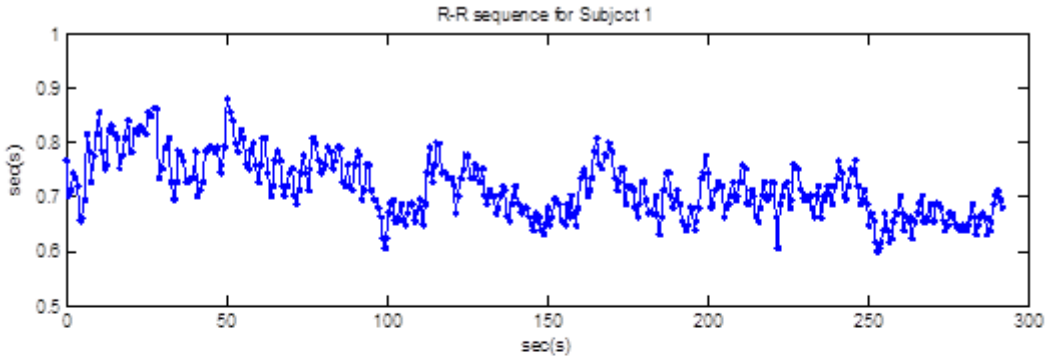


(a)

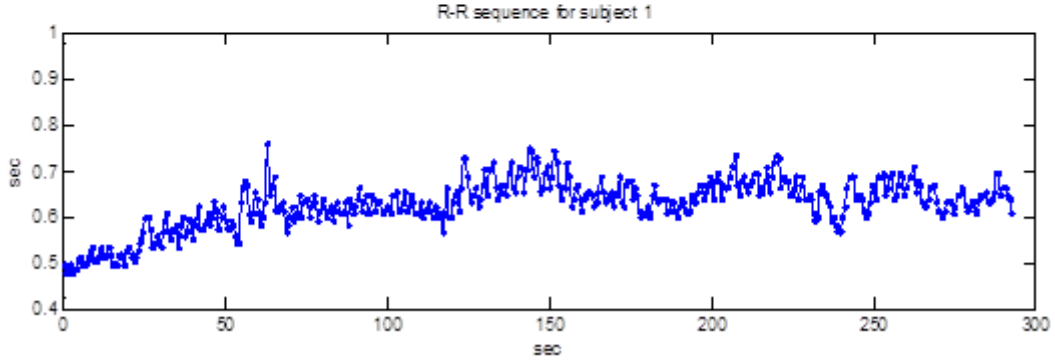


(b)

Fig-6. R-R interval of the subject A. (a) rest condition (b) exercise condition

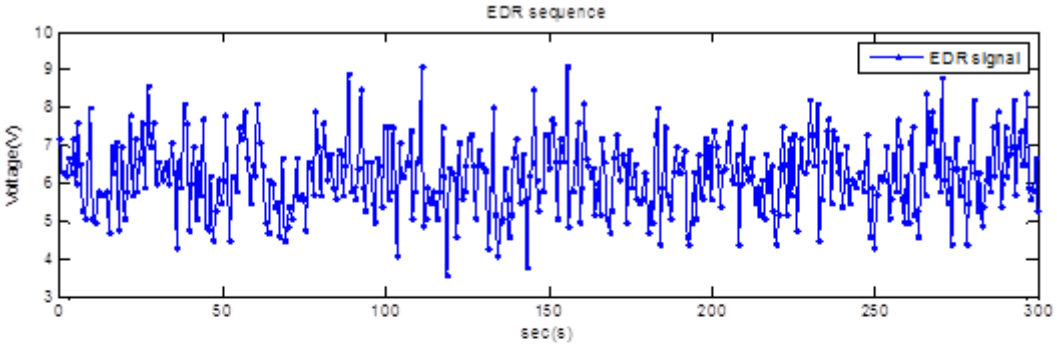


(a)

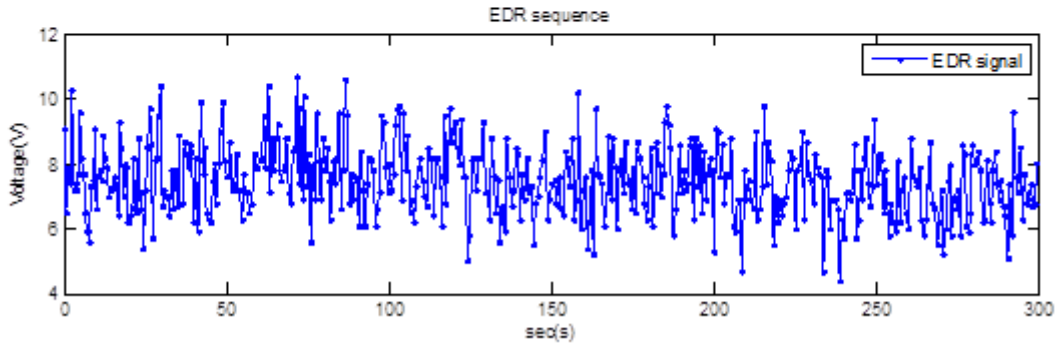


(b)

Fig-7. EDR signal of the subject A. (a) rest condition (b) exercise condition



(a)



(b)

To eliminate the baseline drift of R-R interval, we utilize a smoothness prior analysis mentioned before. As the result of signal processing, one can distinguish the difference between RR-original and RR-preprocessed. The result of comparison can be seen from Figures 8 and 9.

Fig-8. Signal processing of R-R interval for the subject A (exercise).

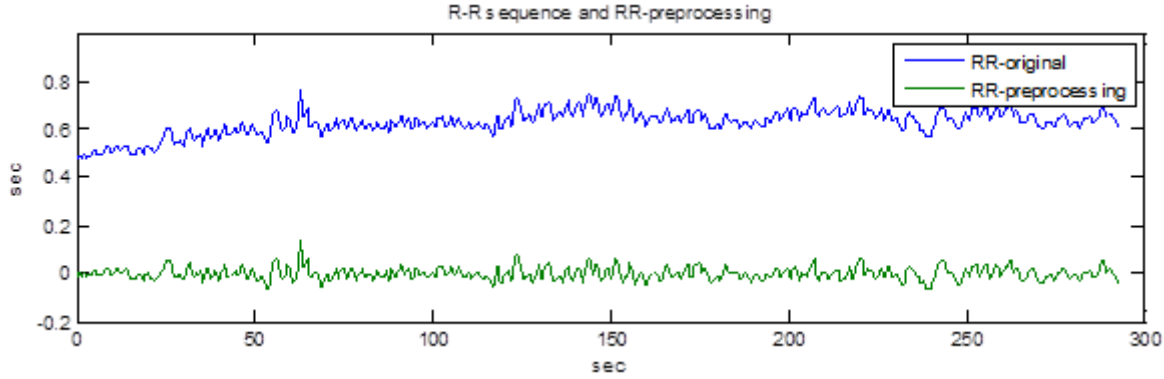
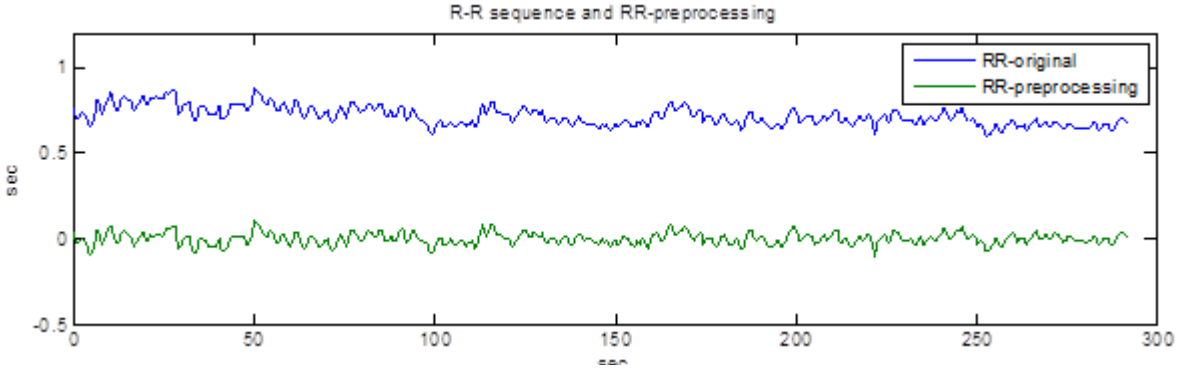


Fig-9. Signal processing of R-R interval for the subject A (rest)



5. Feature Generation

The regulation mechanisms of HRV originated from sympathetic and parasympathetic can be an indicator reflecting the physical function of human's heart. We investigate variability of heart rate via the HCP. There are three major approaches for analysis: time-domain, frequency-domain, and non-linear analysis.

5.1. Time-Domain Analysis

For the time-domain method, we directly use the raw R-R interval time series to calculate related indices. In the time-domain analysis, $\overline{RR}$, $\overline{HR}$, SDNN, SDSD, RMSSD, and pNN50 are major indices of HRV time-domain analysis [4]. $\overline{RR}$ means the mean value of R-R interval and the mean heart rate is $\overline{HR}$. In the experiments, we select 19 characteristic values from the HRV time series.

5.2. Frequency-Domain Analysis

While the time-domain approach is computationally simpler, it lacks the ability to discriminate between sympathetic and parasympathetic contribution to HRV. Note first that an R-R interval time series is an irregularly time-sampled signal. However, in the frequency domain, the feature should not be ignored. Frequency-domain analysis involves spectral analysis. Power spectral density (PSD) of the sampled signal describes distribution of the signal power over the interested frequency range. We refer to Lee, *et al.* [4] for the key frequency-domain indices over the frequency bands of very low frequency (VLF) (0.003-0.04 Hz), low frequency (LF) (0.04-0.15 Hz), and high frequency [2] (0.15-0.4 Hz). In addition, we consider LF/HF, LF norm (LF power in normalized unit), and HF norm. Moreover, we impose the preprocessed R-R interval as another signal source and utilize different methods to examine the spectrum distribution for the purpose of improving accuracy of the frequency analysis. Among different types of R-R interval, the influence on the spectrum needs to be taken into account. Totally, 32 indices were obtained from the frequency analysis.

5.3. Nonlinear Analysis

Poincaré plot serves as a graphical representation of the correlation between the samples RR[n] and RR[n + 1] of R-R intervals [34] with data placed on the plot to fit an ellipse. $SD1=\sqrt{0.5SDSD^2}$ and $SD2=\sqrt{2SDNN^2 - 0.5SDSD^2}$ denote the standard deviations of the data of the major and minor axes of the ellipse, respectively. The former is closely related to the short-term variability which is mainly caused by respiratory sinus arrhythmia (RSA) reflecting the activity of parasympathetic. The latter is related to the long-term variability reflecting the equilibrium of autonomic nervous system [28]. SD1/SD2 reflects the activity of sympathetic [34]. The two indices can be used to describe ANS activity from the Poincaré plot: the cardiac vagal index ($CVI=\log_{10} 16SD1SD2$) and cardiac sympathetic index (CSI=SD2/SD1).

Pincus [35] proposed approximate entropy (ApEn) as a solution to observe the changes in underlying episodic behavior not reflected in peak occurrences or amplitudes. As for the value of ApEn, the larger one indicates higher complexity or irregularity, and the smaller one shows higher regularity. Denote the $x(j) = [u(j)\ u(j+1) \dots u(j+m-1)]$ be a time series of data u(j). The approximate entropy is obtained as

$$ApEn(k, r, N) = \Phi^k(r) - \Phi^{k+1}(r) \quad (1)$$

where $\Phi^k(r) = \frac{1}{N-M+1}\sum_{i=1}^{N-m+1} \log(C_i^m(r))$ with $C_i^m(r)$ being {number of x(j) such that $\frac{d[x(i),x(j)]}{N-m+1} < r$ }where $d[x, x^*] = \max_a ||u(a) - u^*(a)||$ with u(a) being the m scalar components of x. Selection of r here is $r = 0.2SD_x$ where SD_x is standard deviation of the R-R interval. The information can be derived from the distribution of Poincaré plot and the derived parameters from Poincaré, SD1, SD2, CVI, and CSI. In addition, the approximate entropy is designed for the values of r with r = 0.1, 0.15, 0.2 and 0.25.

5.4. Feature Generation

The feature generation process aims at integration of HRV analysis and EDR analysis parameters. For the DSA parameters of HRV analysis, it consists of 95 features with 19 features based on time-domain analysis, 32 features for the frequency-domain analysis, 10 features for the nonlinear analysis, and 34 features for the frequency analysis [36].

The EDR series is derived from ECG signals; it utilizes the QRS wave of ECG to calculate key characteristic parameters. For composition of the indices, the spectral analysis can be used to explode meaningful messages linked to the activity in EDR series. The spectrum is estimated in a similar manner to the R-R interval series.

6. Data Mining

6.1. Classification

Support vector machine (SVM) is adopted here to extract key features embedded in database. The SVM minimizes structural risk as the development objectives. After performing comparison among different decision functions: linear, polynomial and radial basis function (RBF), the RBF was adopted because of its better capability of discrimination among the test subjects. The decision function RBF is given by

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{l} \alpha_{o,i} y_i K(x, x_j) + b\right) \quad (2)$$

where the kernel function $K(x, x_j) = \exp(-\frac{||x-x_j||^2}{2\sigma^2})$ with σ being the width and b is the bias. The coefficients of HRV and EDR time series were used as the input data to the SVM model, which serves as the key factors in the search process to exploit the unclear relationship between HRV and EDR.

6.2. Cross Validation

The jackknife method is a leave-out one cross validation, which is designed to search for the correlation among different features [37]. First, we assume that there are n original data, where the vector of samples $X = (x_1, x_2, \ldots, x_n)$. The jackknife samples $x_{(i)}$ are

$$\begin{aligned} x_{(1)} &= (x_2, x_3, x_4, \ldots, x_n) \\ x_{(2)} &= (x_1, x_3, x_4, \ldots, x_n) \\ &\vdots \\ x_{(i)} &= (x_1, x_2, x_3, \ldots x_{i-1}, x_{i+1}, \ldots, x_n), \quad i = 1,2, \ldots, n \end{aligned} \quad (3)$$

The definition of the jackknife samples is taken out from the sample vector X. The rest of the dataset $m = n - 1$ is to compute a series of jackknife samples. After finishing n times of leave one out, we come to access a group of jackknife samples $Q = \{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$. Estimation of the jackknife samples $\hat{\theta}_{(i)}$ is the jackknife replication defined as

$$\hat{\theta}_{(i)} = S(x_{(i)}) \quad (4)$$

6.3. Statistical Analysis

The bootstrap method [38] involves taking the original data set of N objects and sampling from it to form a new sample (i.e. bootstrap sample), also of the size N. The bootstrap sample is taken from the original data using sampling with replacement. This process is repeated a significant number of times and, for each of these bootstrap samples, its mean is computed to get bootstrap estimates. The method is designed here to estimate the empirical distribution. The aim is to assess accuracy of a given value for a given sampling distribution, which can be used to estimate the unknown empirical distribution.

The original data set X are assumed to be independent. The bootstrap sample θ is generated by repeating random sampling replacement, where $\theta^* = (\theta_1^*, \theta_2^*, \ldots, \theta_n^*)$ and θ_1^* represents one of the original data set, such as $\theta_1^* = x_3$, $\theta_2^* = x_5$, $\theta_3^* = x_1$. Constructing the bootstrap sample to estimate the statistic values by average, correlation function, and standard deviation is called the bootstrap replications $S(\theta^*)$. The empirical distribution function $\hat{G}$ is the discrete distribution. The empirical distribution $\hat{G}$ is computed and considered as an estimator of the same statistics for the entire distribution G.

6.4. Feature Selection

The sequential forward selection (SFS) method aims at enhancing accuracy of the results in data mining by removing noisy and inappropriate features. It iteratively adds features to the initial data subset, according to the data mining results for extracting one dimensional feature at one time.

6.5. Establishment of DSA Model

We exploit the optimal feature subset U(i) as the basis of the DSA model Y(i). The weights of the DSA model, denoted $W_i, i = 1, \ldots, r$, can be considered as the probability of accuracy in each feature. The optimal feature subset U(i) is

$$U(i) = (x_1, x_2, \ldots, x_n), i = 1,2, \ldots, r \quad (5)$$

The prediction analysis T_i is obtained by

$$T_i = \frac{1}{i} \sum_{i=1}^{r} Y(i)^T \quad (6)$$

where $Y(i)^T = W_i U(i)^T$ with $W_i = p_i I_i$ and p_i being the probability of accuracy in each feature.

7. Results

Twenty-nine subjects including ten males and nineteen female ones aged from 18 to 61 joined in this experiment. Each subject was asked to hold the two handlebars to record his/her ECG signals. The measurement involves the rest condition for 5 minutes, riding the exercise bike for another 5 minutes then measuring ECG for 5 minutes. The purpose is to verify that the DSA procedure works correctly with this HCP. We select 95 characteristic indices from HRV and EDR analyses as the feature generation database.

7.1. Classification

We make use of the kernel function from SVM for classification with the width of the Gaussian function 1/2000 and the regulation parameter 100. According to the results of classification with 29 participated subjects, they can be divided into three groups to separate the collected results in one of three categories as Type I, Type II, and Type III. For Types I and III, variation of the heart behaviors can be distinguished directly. Type II is for the cases, the results collocated in certain area which are difficult to be determined. The situation of Type II would depend on the kernel function chosen. Comparison of performance between linear and nonlinear kernel functions can be referred to Lin, *et al.* [32].

7.2. Leave-one-out Validation

Secondly, leave-one-out cross validation (L1O CV) is adopted to pick the abnormal subjects. The L1O CV uses one observation as the validation set and the remaining observations as the training set [39]. The process is repeated to cut the original sample on a validation set of one observation and a training set. We considered the classification result of testing data as the reference to find out the subjects belonging to Type III. The leave-one-out validation was conducted 10 times. By this way, we try to identify the possibility of abnormal subjects. The category of exceeding or equal to three unusual features D_{3-dem} involves subjects 3, 5, 6, 7, 15, 17, 21. The category of exceeding or equal to five abnormal features D_{5-dem} involves subjects 3, 5, 21. See Table 1 for the summary.

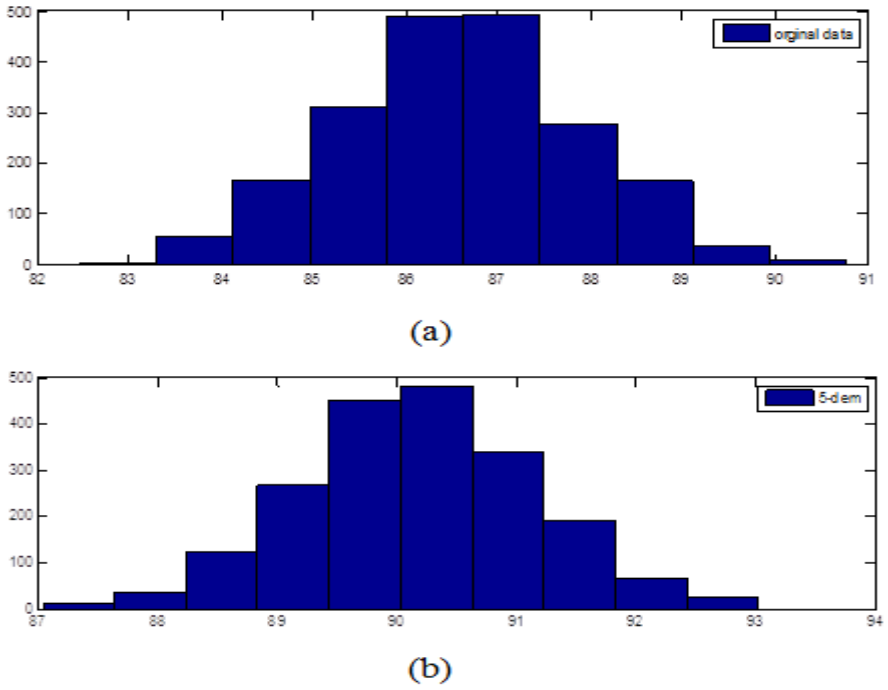
Table-1. Summary of the abnormal test subjects

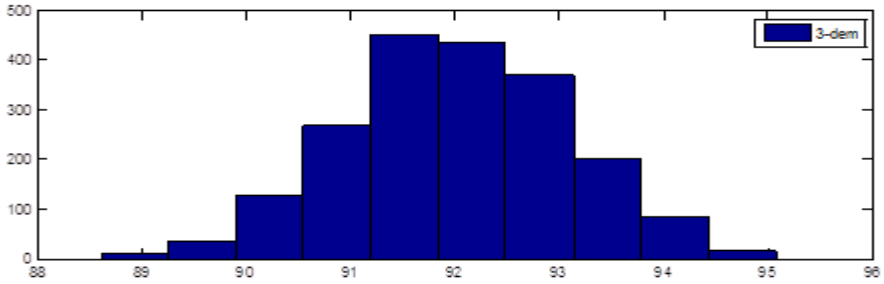
Case 1: SDNN (Feature 7), Case 2 : RMSSD (Feature 8), Case 3: SDSD (Feature 9), Case 4: SD1 (Feature 54), Case 5: Poincare (Feature 55), Case 6: CVI (Feature 56), Case 7: CSI (Feature 57), Case 8: VLF (Feature 20), Case 9: HF (Feature 22), Case 10: Total (Feature 23)			
Subject	**Case**	**Subject**	**Case**
1	5	12	10
3	1,2,3,5,6,7,9	13	10
5	1,4,6,7,9	15	1,4,6
6	5,6,10	17	1,4,6
7	1,7,9	19	1
10	1	21	1,2,3,4,5,6,8,9,10
11	1		

7.3. Statistical Analysis

Three categories ($D_{original}$, D_{3-dem}, D_{5-dem}) were used to conduct bootstrap analysis to evaluate the empirical distribution. Figure 10 shows empirical distribution with the bootstrap method, with 2000 bootstrap replications for each category. From the comparison of sub-figures, we should be confident to pick abnormal subjects. D_{3-dem} shows the best experimental results, which are located in 91-93%. We can thus consider it as the base dataset for further research.

Fig-10. Empirical distribution with (a) $D_{original}$, (b) D_{3-dem} , (c) D_{5-dem}



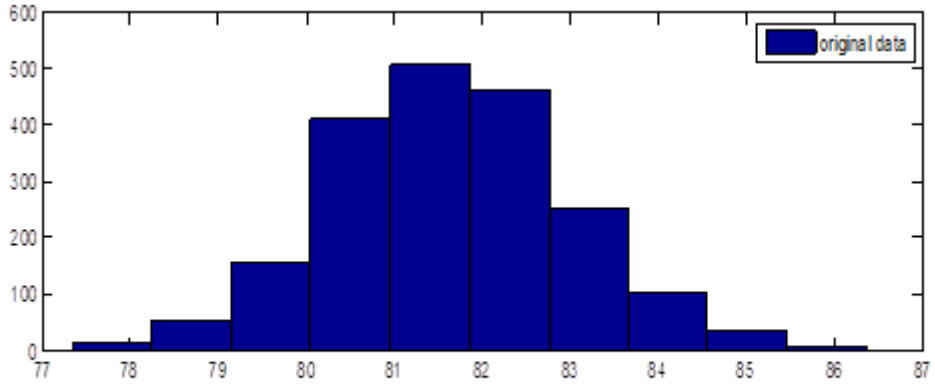


(c)

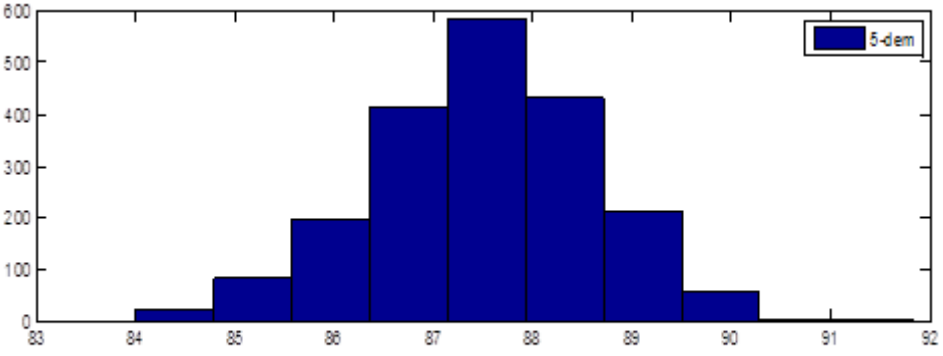
7.4. Feature Selection

The feature selection result was performed with 20 features by observing the empirical distribution. Figure 11 shows empirical distribution with feature selection for $D_{original}$, D_{5-dem} and D_{3-dem}. The result of selection shows closeness while compared with the original dataset; the difference is within 5 %. The feature selection can thus be used to reduce the computational load.

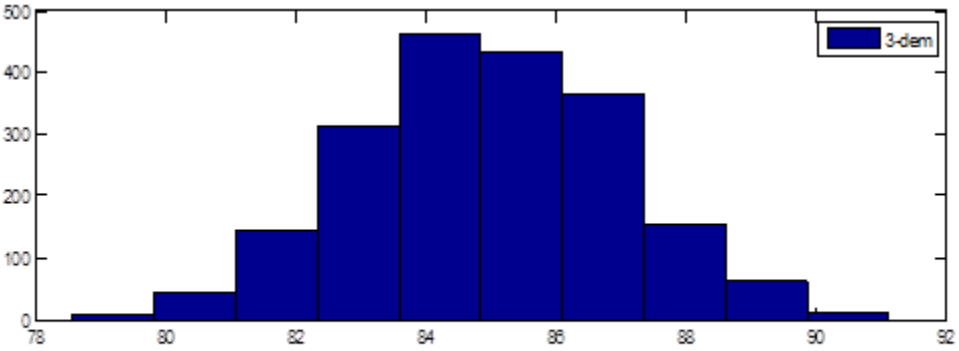
Fig-11. Empirical distribution with feature selection for (a) $D_{original}$, (b) D_{3-dem} , (c) D_{5-dem}.



(a)



(b)

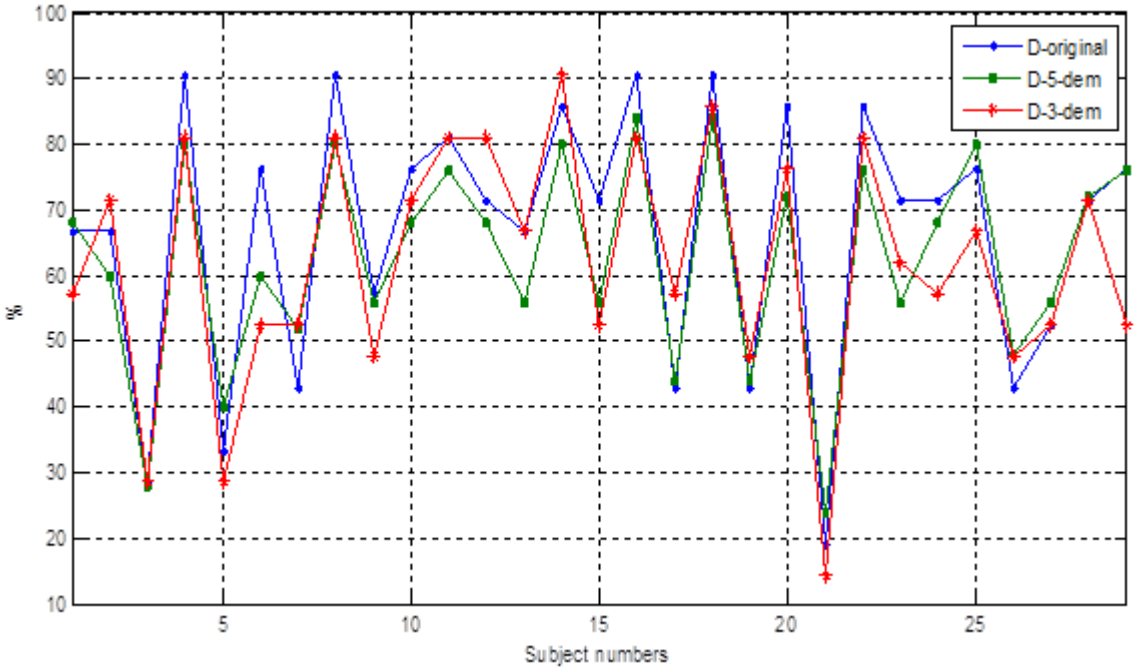


(c)

7.5. Results of DSA Model Prediction

According to the results of feature selection, we established DSA models to predict the physiological information for each subject. Figure 12 shows the result of the predictive analysis for 20 features.

Fig-12. Results of the predictive analysis (20 features).



7.6. Discussions

The percentage of RBF-based classification belonging to Type II is smaller than that of the linear one. Indices considered here show a higher classification results (70-90%) in RBF-based classification. It meant that the situation between the rest and after the exercise condition have no effect in those situation. In other words, the classification with the RBF method exhibit relatively ideal performance.

Leave-one-out validation shows that better performance of classification in the modified results revealing that Type III may exist abnormal subjects. From Figures 6-9, we are confident to eliminate the abnormal subject. D_{3-dem} shows the best experimental result, which located in 91-93%. Therefore, one can consider D_{5-dem} as the base data set for further research works.

From Figures 10 and 11, the mean values of the corresponding features are significantly different, which can be used for the purpose of identification. The results of selection show that they are close to the original dataset; the error is less than 5%. Results presented here show potential of the developed platform for the purpose of future mobile health care.

8. Conclusions

The major aim of this research task is to develop an easily applicable healthcare analysis system which can be used in conjunction with the current trend of mobile health care applications. The proposed design shows HCP's significant potential in mobile applications of personal daily health care. The platform has been shown to be a convenient tool for collecting and analyzing physiological information from human body. The proposed procedure in data analysis, DSA allows one to generate ECG patterns that may affect HRV parameters. Extensive experiments demonstrate the use of this platform.

Acknowledgement

This research was sponsored by Ministry of Science and Technology, Taiwan under the grant MOST 104-2622-E005-011.

References

[1] Lee, S. C. and Chung, W. Y., 2014. "A robust wearable u-healthcare platform in wireless sensor network." *Journal of Communications and Networks*, vol. 16, pp. 465-474.

[2] Baig, M. M., Hosseini, H. G., Connolly, M. J., and Kashfi, G., 2014. "A wireless patient monitoring system for hospitalized older adults: Acceptability, reliability and accuracy evaluation," IEEE-EMBS." *International Conference on Biomedical and Health Informatics*, pp. 330-333

[3] Zhang, L. L., Aldanondo, M., and Kumar, A., 2014. "Healthcare platforming for healthcare service development in hospitals." *IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 321-324.

[4] Lee, C. Y., Lai, Y. T., and Hsu, S. Y., 2015. *A sensor-fusion solution for mobile health-care applications, smart sensors and systems*. Springer International Publishing, pp. 387-397.
[5] Available: https://patient.info/health/electrocardiogram-ecg
[6] Wu, W., Lee, J., and Chen, H., 2009. "Estimation of heart rate variability changes during different visual stimulation using non-invasive real-time ECG monitoring system" *International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing*. pp. 344-347.
[7] Wu, W., Gil, Y., and Lee, J., 2011. "A prototype of wireless electrocardiogram biofeedback platform and its preliminary evaluation on HRV alteration in cognitive unconscious status." *IEEE International conference on Healthcare Informatics, Imaging and Systems Biology*, pp. 17-24.
[8] Alemdar, H. and Ersoy, C., 2010. "Wireless sensor networks for healthcare: A survey." *Computer Networks*, vol. 54, pp. 2688-2710.
[9] Darwish, A. and Hassanien, A. E., 2011. "Wearable and implantable wireless sensor network solutions for healthcare monitoring." *Sensors*, vol. 11, pp. 5561–5595.
[10] Baig, M. M., Gholamhosseini, H., and Connolly, M. J., 2013. "A comprehensive survey of wearable and wireless ECG monitoring systems for older adults." *Medical and Biological Engineering and Computing*, vol. 51, pp. 485-495.
[11] Yang, H. K., Lee, J. W., Lee, K. H., Lee, Y. J., Kim, K. S., Choi, H. J., and Kim, D. J., 2008. "Application for the wearable heart activity monitoring system: analysis of the autonomic function of HRV." *IEEE International Conference on Engineering in Medicine and Biology Society*, pp. 1258-1264.
[12] Ramy, H., Mohamed, A., and Alghoniemy, M., 2015. "Scalable real-time energy-efficient EEG compression scheme for wireless body area sensor network." *Biomedical Signal Processing and Control*, vol. 19, pp. 122-129.
[13] Yeh, C. Y. and Wang, W. J., 2008. "QRS complex detection for ECG signal: The Difference Operation Method." *Computer Methods and Program in Biomedicine*, vol. 91, pp. 245-254.
[14] Domingos, P. and Pazzani, M., 1997. "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss." *Machine Learning*, vol. 29, pp. 103-130.
[15] Wu, W. and Lee, J., 2009. "Improvement of HRV methodology for positive/negative emotion assessment." *IEEE International Conference on Collaborative Computing Networking, Applications and Worksharing*, pp. 1-6.
[16] Tarvainen, M. P., Ranta-aho, P. O., and Karjalainen, P. A., 2002. "An advanced detrending method with application to HRV analysis." *IEEE Transactions on Biomedical Engineering*, vol. 49, pp. 172-175.
[17] Karjalainen, P., 1997. "Regularization and bayesian methods for evoked potential estimation." *Univ. Kuopio, Dept. Appl. Phys.*
[18] Gersch, W., 1991. "Smoothness priors." *New Directions in Time Series Analysis, Part II*, pp. 113-146.
[19] Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996. "Heart rate variability-standards of measurement, physiological interpretation, and clinical use." *European Heart Journal*, vol. 17, pp. 354-381.
[20] Niskanen, J. P., Tarvainen, M. P., Ranta-Aho, P. O., and Karjalainen, P. A., 2004. "Software for advanced HRV analysis." *Computer Methods and Programs in Biomedicine*, vol. 76, pp. 73-81.
[21] West, B. J. *Fractal physiology and medicine: studies of nonlinear phenomena in life, Science 1*. Singapore: World Scientific.
[22] Cooley, J. W., Lewis, P. A., and Welch, P. D., 1969. "The fast fourier transform and its applications." *IEEE Transactions on Education*, vol. 12, pp. 27-34.
[23] Kay, S. M. and Marple, S. L., 1981. "Spectrum analysis-a Modern perspective." *Proceeding of the IEEE*, vol. 69, pp. 1380-1419.
[24] Brodman, A., Schlindwein, F. S., Rocha, A. P., and Leite, A., 2002. "A study on the optimum order of autoregressive models for heart rate variability." *Physiology Measurement*, vol. 23, pp. 325-336.
[25] Milley, A., 2000. "Healthcare and data mining." *Health Management Technology*, vol. 21, pp. 44-47.
[26] Cios, K. J. and Moore, G. W., 2002. "Uniqueness of medical data mining." *Artificial Intelligence in Medicine*, vol. 26, pp. 1-24.
[27] Pecchia, L., Melillo, P., and Bracale, M., 2011. "Remote health monitoring of heart failure with data mining via CART method on HRV features." *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 800-804.
[28] Mougiakakou, S. G., Valavanis, I. K., Mouravliansky, N. A., and Nikita, A., 2009. "DIAGNOSIS: a telematics-enabled system for medical image archiving, management, and diagnosis assistance." *IEEE Transactions on Instrumentation and Measurement*, vol. 58, pp. 2113-2120.
[29] Jain, A. J., Duin, R. P. W., and Mao, J., 2000. "Statistical pattern recognition: a review." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 4-37.
[30] Wang, L., 2008. "Feature selection with kernel class separability." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1534-1546.
[31] Molina, L. C., Belanche, L., and Nebot, A., 2002. "Feature selection algorithms: a survey and experimental evaluation." *IEEE International Conference on Data Mining*, pp. 306-313.
[32] Lin, C. L., Chen, Y. Y., and Chiang, C. T., 2013. "Identity identifier based on the mixed ECG and fingerprint." *Taiwan Patent M*, p. 453912.

[33] Pan, J. and Tompkins, W. J., 1985. "A real-time QRS detection algorithm." *IEEE Transactions on Biomedical Engineering*, vol. 32, pp. 230-236.

[34] Brennan, M., Palaniswami, M., and Kamen, P., 2001. "Do existing measures of Poincaré plot geometry reflect nonlinear feature of heart rate variability." *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 1342-1347.

[35] Pincus, S. M., 1991. "Approximate entropy as a measure of system complexity." *National Academy of Sciences*, vol. 88, pp. 2297-2301.

[36] Yu, Y. H., 2014. "Heart rate variability analysis in health care platform, Master Thesis." National Chung Hsing University. Taichung.

[37] Browne, M. W., 2000. "Cross-validation methods." *Journal of Mathematical Psychology*, vol. 44, pp. 108-132.

[38] Efron, B., 1979. "Bootstrap methods: Another look at the jackknife." *The Annals of Statistics*, vol. 7, pp. 1-26.

[39] Devijver, P. A. and Kittler, J., 1982. *Pattern recognition: A statistical approach*. London: Prentice-Hall.