

# Modeling the Relationship between Heart Rate and Features of Vocal Frequency

Motoki Sakai

School of Information Environment  
Tokyo Denki University

## ABSTRACT

Heart rate (HR) is an informative index to assess one's physical condition from the viewpoint of preventive healthcare; it is beneficial if information of HR can easily be obtained without a particular clinical instrument. In this research, voice data was used to estimate HR, because it can easily be recorded using a common device such as smartphone. To evaluate the feasibility of HR estimation using recorded voice, experiments were conducted with two subjects. In the experiment, 60 sets of HRs and voice data were measured per subject. A correlation between HR and the vocal frequency was observed, and the feasibility of HR estimation using voice data was confirmed. The relationship between HR and voice data was modeled using a polynomial function of vocal frequency. Using the proposed model, HR could be estimated with high correlation coefficients and small estimation errors.

## General Terms

Data mining, Biomedical engineering

## Keywords

Heart rate, preventive healthcare, vocal frequency, modeling

## 1. INTRODUCTION

Heart rate (HR) is one of the most important and informative indices used to assess our physical condition from a preventive healthcare perspective. There are many applications available to measure HR for the user's health maintenance, and our physical condition can be determined using these applications. For example, simplified cardio-tachometers can be used for prophylaxis in lifestyle diseases, such as obesity. Obesity is responsible for many symptoms, including diabetes, hypertension or hyperlipidemia, and can result in death [1]. To prevent obesity, appropriate exercise can be effective. However, hard-intensity exercise is not always advised for obese patients, and it is important to exercise under good conditions and appropriate exercise intensities. Therefore, it is important to assess the physical condition and exercise intensity. HR is recommended as one effective way to assess patients' physical condition [2, 3].

Due to recent developments, it is easy to know physical condition required to maintain our health without the use of medical equipment. For example, exercise intensity can be easily evaluated with an acceleration sensor installed in a smartphone. HR can also be measured with a smartphone camera. For health maintenance, it is beneficial to evaluate effectiveness of daily exercise not only with medical equipment, but also with general devices such as smartphones. However, existing cardio-tachometers that use smartphone cameras have drawbacks from a practical perspective. For example, cardio-tachometers using smartphone cameras are easily influenced by the user's body motion or external light.

Therefore, this research proposes a method to more easily estimate HR by measuring the human voice.

The relationship between HR and voice has already been observed; HR and voice have been simultaneously measured in stress assessments and other similar assessments [3]. In previous research, several algorithms to estimate HR from vocal frequency have been presented [4, 5, 6, 7, 8, 9]. Research papers [4, 5], and [6] present HR estimation methods by using machine learning algorithms, such as the support vector machine (SVM). However, these research papers have not revealed the explicit relationship between HR and voice, and it is unclear which features of voice were conducive to estimate HR. For example, the SVM does not require explicit features of voice to estimate HR because learning and regression are conducted in unknown feature space. In [7], HR is expressed as a multivariate function of statistics, including entropy and standard deviation of vocal signal, but the reason why HR is expressed as a multivariate function of statistics is not clear. Research papers [8, 9] focus on fluctuations synchronized with R-waves of the ECG signal in the speech spectrogram. In [8], HR is roughly estimated by image processing for 2D images of speech spectrogram. In [9], researchers have chosen the component of vocal frequency that is most affected by R-waves on one-dimensional scanning of the speech spectrogram; this allowed them to estimate HR with high accuracy. As described above, [4, 5, 6, 7] do not explain the relationship between HR and voice, while [8, 9] present methods for R-wave detection in the speech spectrogram. Furthermore, [10] shows that blood pressure can be estimated from vocal frequency under the assumption that the hardness of blood vessels change according to blood pressure variation, and this influences fluctuations in the vocal tract. It is believed that HR similarly influences vocal frequency through blood vessels in the vocal tract. In this research, we note that a fluctuation transmitted from the blood vessel does indeed influence the voice, and the relationship between HR and vocal frequency is modeled.

The relationship between HR and the vocal frequency is examined, a modeling of their relationship is conducted, and the feasibility of HR estimation with the proposed model is evaluated. If HR estimation from the vocal frequency is feasible, HR can be easily measured during movement without specific clinical instrument, and it will be beneficial from the viewpoint of preventive healthcare. In addition, a study that investigates how HR affects voice can be scholastically beneficial.

## 2. EXPERIMENTS TO MEASURE HR AND VOICE DATA

To model the relationship between HR and vocal frequency, we conducted experiments to measure the HR and voice. Two subjects were studied: a 33-year-old male and a 25-year-old female.

During the experiment, factors that probably affect vocal frequency were controlled. For example, subjects always sat with their head upright in a relaxed manner, because body posture is known to influence the production of sound.

First, the voice was recorded with a voice recorder: an ICD-TX50 produced by SONY. The voice recorder was kept 5 cm from the subject's mouth, and the subject was asked to sustain the sound [ā] (one of the Japanese vowels) for 7 s. The sampling rate of the voice recording was 44.1 kHz with 16-bit resolution. Immediately following the voice recording, HR was measured with the cardio-tachometer: a HEM-1010 produced by OMRON, assuming that the HR does not change suddenly in the absence of external stimulus. Subjects were given no stimulus after measuring their voice.

The above test was repeated 60 times per subject.

## 3. CORRELATION BETWEEN HR AND VOCAL FREQUENCY

To evaluate the relationship between HR and the vocal frequency, the recorded voice signals were expressed as the frequency representation from the FFT, and the correlation coefficients between HR and each component of vocal frequency (120 Hz – 10,000 Hz) were computed.

Fig. 1 shows the correlation coefficient between the HR and the vocal frequency for the male subject. Likewise, Fig. 2 shows the correlation coefficient between the HR and the vocal frequency for the female subject.

As shown in Fig. 1, higher correlation coefficients between the HR and the vocal frequency reached approximately 0.4, and they could be seen in frequency bands between 2,000 Hz and 7,000 Hz. In contrast, Fig. 2 showed that the highest correlation coefficient was approximately 0.5 around 5,500 Hz for the female subject. These results indicated that the HR and the vocal frequency were not non-correlated ( $>0.330104$ ), and non-correlation between the HR and the component of vocal frequency was rejected in various frequency bands.

For example, a change in vocal frequency associated with HR variation is shown in Fig. 3. This demonstrated that formants were shifted to the right with a decrease in HR, and the amplitude of each formant was also changed with HR variation. In this example, the higher the HR, the larger the amplitude of the 5-th formant. However, factors that changed vocal frequency were not limited to HR, and this example

indicated that HR was one factor that influenced vocal frequency, because other factors that changed vocal frequency could be controlled in this experiment.

The aforementioned results suggested that HR could be estimated from the vocal frequency, because the HR and the component of vocal frequency were not non-correlated. Next, we put forth a model to explain the relationship between the HR and vocal frequency. We also test the utility of the model is estimating HR from voice data.

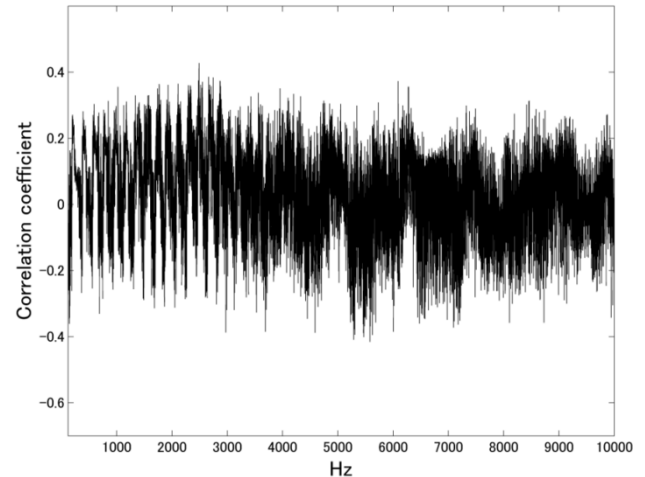


Fig. 1: Correlation coefficient between the HR and his voice for male subject

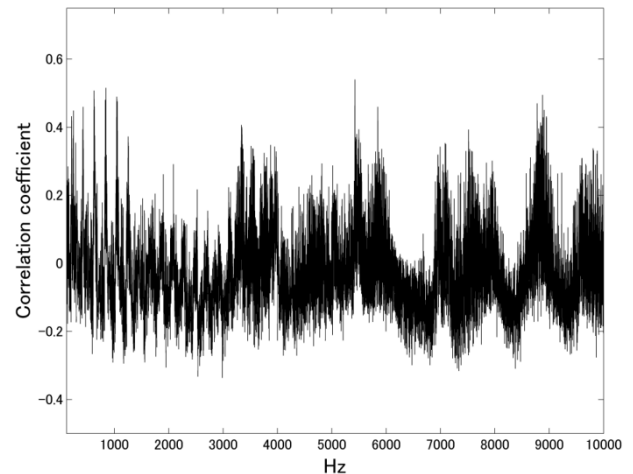


Fig. 2: Correlation coefficient between the HR and her voice for female subject

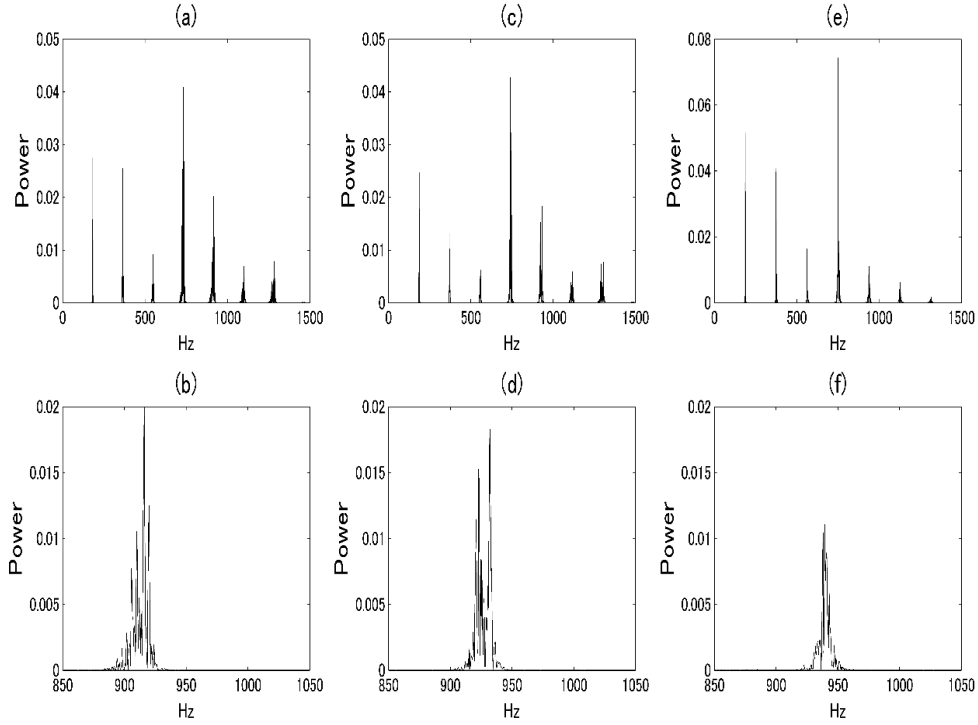


Fig. 3: Example of formants movement associated with change in HR for male subject: (a) Vocal frequency when HR is 105 bpm, (b) Vocal frequency around 950 Hz when HR is 105 bpm, (c) Vocal frequency when HR is 95 bpm, (d) Vocal frequency around 950 Hz when HR is 95 bpm, (e) Vocal frequency when HR is 78 bpm, (f) Vocal frequency around 950 Hz when HR is 78 bpm.

#### 4. MODELING OF THE RELATIONSHIP BETWEEN HR AND THE FEATURES OF VOCAL FREQUENCY

##### 4.1 Empirical model of relationship between HR and spectral components of voice

In Section 3, it was suggested that voice HR and vocal frequency were correlated to a certain extent. From this result, the following hypothesis was adopted again, much like the previous research [10]. First, the hardness of blood vessels vary in synchronism with the heartbeat because the blood pressure is different between the systole and diastole phases. In short, it is contemplated that the hardness of blood vessels in the systole phase increases compared with that in the diastole phase, and this influences fluctuations in the vocal tract. As a result, the spectral characteristics of the voice were altered. In other words, the HR can be regarded as one of the factors that affect the voice frequency. Suppose that the feature of vocal frequency can be expressed as a product of two functions  $F(HR)$  and  $G(\theta)$  as shown in Eq. (1):

$$v_{voice} = F(HR) \cdot G(\theta), \quad (1)$$

where  $v_{voice}$  is a feature of vocal frequency,  $HR$  is the HR value,  $F(HR)$  is a function of  $HR$ , and  $G(\theta)$  is a function of other variables  $\theta$  excluding  $HR$ . In theory,  $G(\theta)$  is a feature value of vocal frequency calculated in the absence of any HR

influence, and can be regarded as a feature calculated from vocal frequencies generated on the basis of existing speech production mechanisms such as the vocal tract filter. In this research,  $G(\theta)$  is treated as a constant, which is set to a feature value of vocal frequency that can be obtained during a low HR (around 70 bpm). To estimate the HR value from the feature of vocal frequency, it is only necessary to find the inverse function of  $F(HR)$  as in Eq. (2):

$$HR = F^{-1}\left(\frac{v_{voice}}{G(\theta)}\right). \quad (2)$$

The feature of vocal frequency is empirically defined as follows:

$$v_{voice} = \prod_f S(f), \quad f = \{f_1, f_2, \dots, f_n\}, \quad (3)$$

where  $f$  means frequencies ranged from  $f_1$  Hz through  $f_n$  Hz, and  $S(f)$  is the component of vocal frequency in  $f$  Hz.  $f = \{f_1, f_2, \dots, f_n\}$  can be regarded as frequencies that are strongly influenced by change in HR, and these frequencies are determined using the sequential feature selection (SFS) method. In summary, the feature of vocal frequency is determined as a product of spectral components related to HR variation.

In this research, it is assumed that the feature of vocal frequency bears a non-linear relationship with HR value, and a model of the relationship between HR and the feature of

vocal frequency shown in Eq. (1) is rewritten with a polynomial function as demonstrated in Eq. (4):

$$v_{Voice} = (\alpha \cdot HR - \beta)^d \cdot G(\theta), \quad (4)$$

where  $\alpha$  is a scale factor, and  $\beta$  is an offset. Finally, an estimated HR value can be obtained with Eq. (5), which is an inverse function of Eq. (4):

$$HR = \delta \cdot \left( \frac{\prod_f S(f)}{G(\theta)} \right)^{\frac{1}{d}} + \varepsilon, \quad (5)$$

where  $\delta = 1/\alpha$  and  $\varepsilon = \beta/\alpha$ .

## 4.2 SFS method to select frequencies related to HR variation

The SFS method was used to find the frequency band strongly influenced by HR variation. SFS is performed according to the following steps. This procedure demonstrates how to find  $f = \{f_1, f_2, \dots, f_n\}$  as shown in Eq. (3).

I. SFS begins with an evaluation of the accuracy of the estimation for each component of vocal frequency  $S(f)$  (120 Hz – 10,000 Hz), identifying the frequency component that contributes to the most accurate estimation. This frequency is labeled as  $f_1$ .

II. SFS is then performed using a paired-component of vocal frequency. The component of vocal frequency identified in the previous step is coupled with each of the other frequency components from the remaining components, selecting the pair with the highest degree of accuracy ( $f_2$ ).

III. SFS next evaluates two components of vocal frequency: the pair from Step II and another selected from the remaining frequency components.

IV. Steps I–III are repeated with four or more spectral components.

Generally, the estimation accuracy must peak when the number of selected components of vocal frequency reaches a given number. The best set of components of vocal frequency can be singled out for maximum accuracy in the SFS procedure. Note that the “accuracy” is determined by the correlation coefficient between the measured and estimated HRs.

Eventually, components of vocal frequency selected by the SFS method are regarded as strongly influenced by HR variation.

## 5. EVALUATIONS AND RESULTS

To accurately estimate HR from vocal frequency, the parameters  $d$ ,  $\delta$ ,  $\varepsilon$  and  $G(\theta)$  shown in Eq. (5) must be effectively selected. For the male subject, the most accurate estimations were obtained when the parameters  $d$ ,  $\delta$  and  $\varepsilon$  were set to 250, 5 and 85, respectively. For the female subject, when the parameters  $d$ ,  $\delta$  and  $\varepsilon$  were set to 200, 5 and 65, respectively, the estimation of HR was the most accurate.

In this paper, two indices of accuracy are adopted. One is a correlation coefficient between measured HR and the estimated one. Another is an estimation error defined in Eq. (6):

$$error = |HR - eHR|, \quad (6)$$

where  $eHR$  is estimated HR value computed with Eq. (5).

Fig. 4 and Fig. 5 show correlation coefficients between HR and vocal frequency components selected by the SFS method for male and female subjects, respectively. These figures illustrate the relationship between accuracies of estimation and the number of components of vocal frequency selected by the SFS. Table 1 shows the highest correlation coefficients for male and female subjects. Table 2 shows a mean estimation error (ME) and standard deviation (SD). Fig. 6 and Fig. 7 illustrate histograms of vocal frequencies selected by the SFS method for male and female subjects, respectively. As shown in Figs. 4 and 5, HRs were estimated with up to 450 vocal frequencies. In Fig. 6 and Fig. 7, the first 220 selected vocal frequencies are shown. (Correlation coefficients reached approximately 0.9 for the male and female subjects when 220 vocal frequencies were selected to compute the features of vocal frequency.)

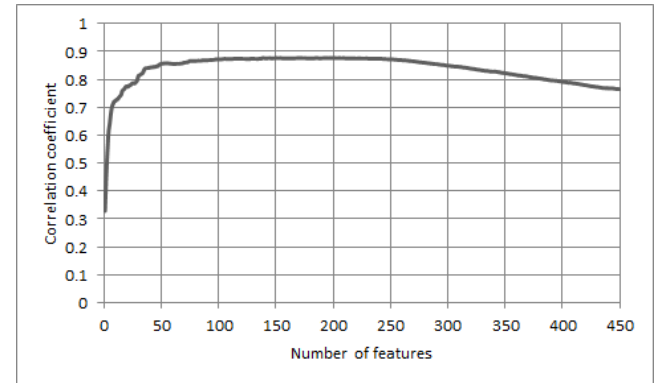


Fig. 4: Correlation coefficient between HR and selected vocal frequency components for the male subject

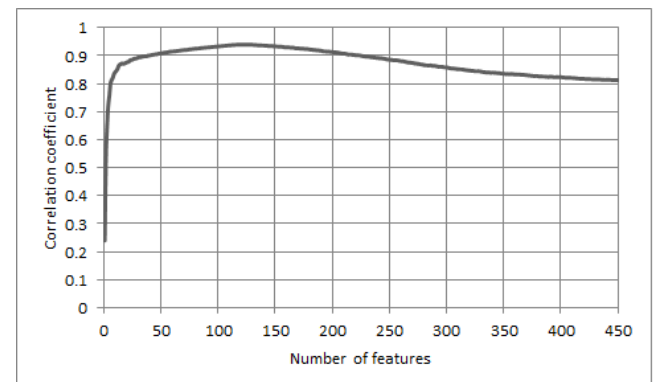


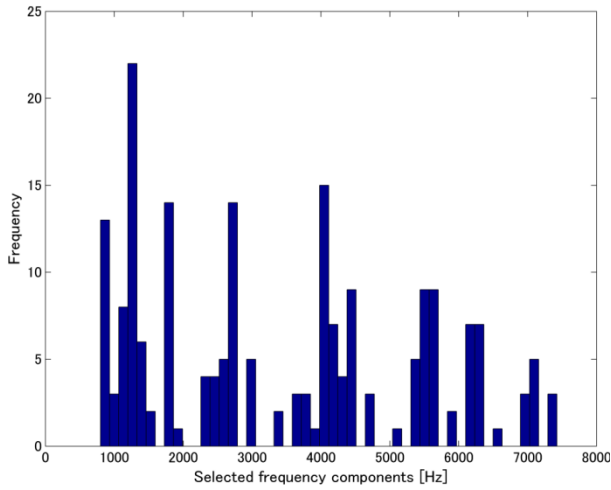
Fig. 5: Correlation coefficient between HR and selected vocal frequency components for the female subject

**Table 1: Correlation coefficients between measured and estimated HR**

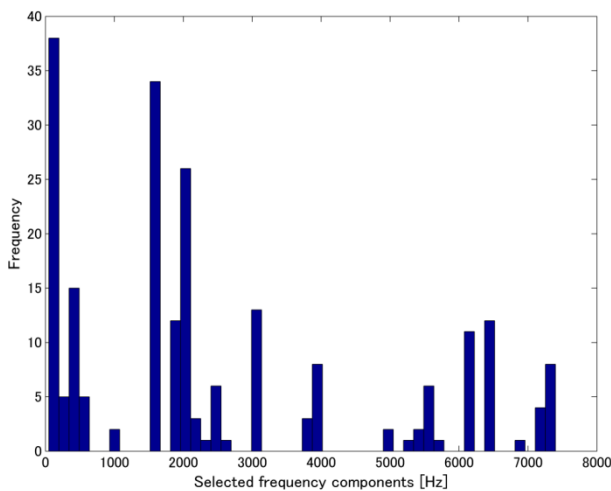
Subject	Correlation coefficient
Male	0.88
Female	0.93

**Table 2: Estimation errors**

Subject	ME (SD) [bpm]
Male	3.8 (2.5)
Female	3.4 (2.9)



**Fig. 6: Histogram of the first 220 frequency components selected by the SFS for the male subject**



**Fig. 7 Histogram of the first 220 frequency components selected by the SFS for the female subject**

## 6. DISCUSSION

The purpose of this research was to model the relationship between the HR and vocal frequency, and then estimate the value of HR from vocal frequency using the proposed model.

Table 1 shows that correlation coefficients between actual HR and the estimated one are around 0.9 for both the male and female subjects. This result indicated that the magnitude of the relationship of feature values of vocal frequency was corresponding to that of HRs.

On the other hand, mean estimation errors of HR were less than 4 bpm for both the male and female subjects. These estimation errors were small to a certain degree, because the measurement error of the instrument used in this research is within  $\pm 5\%$  of the observed value, and the resultant estimation errors were around this measurement error value. (It is likely that measurement error ranges from approximately  $\pm 3.5$  bpm to  $\pm 5$  bpm because normal HR ranges from 70 bpm to 100 bpm.)

Next, frequency components that actually contribute to estimate HR were discussed. Fig. 6 and Fig. 7 illustrate a difference of selected frequency components between the male and female subjects. For the male subject, selected frequency components ranged widely from approximately 800 Hz to 7,400 Hz. On the other hand, frequency components between 3,000 Hz and 5,000 Hz were rarely selected for the female subjects. In addition, frequency components less than 1,000 Hz, which were not selected for the male subject, were culled for the female subject. It is believed that this difference may be derived from gender difference, but a physiological rationale cannot be provided at this time. This result also invites a new perspective. In [9], it is shown that an effect of the heartbeat on vocal frequency can be seen mainly between 1,000 Hz and 6,000 Hz. However, the result shown in this paper indicates that frequency bands wider than ones described in [9] can contribute to estimate HR.

From the above results, it is concluded that the proposed model is effective in estimating the value of HR from vocal frequency, even though a physiological interpretation of the proposed model is not clear yet. However, this research has a drawback in terms of the number of subjects. As for parameters used in the proposed model, suitable parameters for accurate estimation differed between male and female subjects except for parameter  $\delta$ . This may be regarded as gender difference. Conversely, it is well known that the characteristics of vocal frequency are unique to each person. In this case study, individual differences in the vocal frequency have not been sufficiently examined. Individual differences must influence parameters shown in Eq. (4) or (5) such as  $d$ ,  $\alpha$ ,  $\beta$  and  $\delta$ . In future work, individual and gender differences of parameters shown in Equations (4) and (5) will be investigated in a study involving more subjects. Additionally, parameters will be physiologically translated.

## 7. CONCLUSION

The goal of this paper was to estimate the value of HR from vocal frequency. First, HR and voice were measured from male and female subjects, and the correlation coefficients between HR and vocal frequency were evaluated. It was then shown that vocal frequency components at specific frequency bands were not non-correlated to HR. From this result, the relationship between HR and vocal frequency was modeled with the polynomial function of a feature of the vocal

frequency. The feature of vocal frequency was defined as the product of components of the vocal frequency selected by the SFS method. With this model, HR could be estimated with a small estimation error and high correlation between actual and estimated HRs, even though the physiological explanation of the proposed model has been unclear. In future work, individual and gender differences of parameters of the proposed model will be evaluated with more subjects.

## 8. REFERENCES

- [1] A. Halpern, MC. Mancini, ME. Magalhães, M. Fisberg, R. Radominski, MC. Bertolami, A. Bertolami, ME. de Melo, MT. Zanella, MS. Queiroz and M. Nery, "Metabolic syndrome, dyslipidemia, hypertension and type 2 diabetes in youth: from diagnosis to treatment," *Diabetology & Metabolic Syndrome*, 2010.
- [2] MH. Davenport, S. Charlesworth, D. Vanderspank, MM. Sopper and MF. Mottola, "Development and validation of exercise target heart rate zones for overweight and obese pregnant women," *Applied Physiology, Nutrition, and Metabolism.*, vol. 33, no. 5, pp. 984-989, 2008.
- [3] SC. Franckowiak, DA. Dobrosielski, SM. Reilley, JD. Walston and RE. Andersen, "Maximal heart rate prediction in adults that are overweight or obese," *The Journal of Strength & Conditioning Research*, vol. 25, no. 5, pp. 1407–1412, 2011.
- [4] J. Hada and Y. Takeuchi, "Evaluation of mental stress by voice analysis during simulated landing," *The Japanese Journal of Ergonomics*, vol.44, No.3, pp. 171–174, 2008.
- [5] B. Schuller, F. Friedmann and Florian Eyben, "The Munich Biovoice Corpus: Effects of Physical Exercising, Heart Rate, and Skin Conductance on Human Speech Production," *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, pp. 1506–1510, 2014.
- [6] B. Schuller, F. Friedmann and Florian Eyben, "Automatic Recognition of Physiological Parameters in the Human Voice: Heart Rate and Skin Conductance," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 7219–7223 2013.
- [7] J. Kaur and R. Kaur, "Extraction of Heart Rate Parameters Using Speech Analysis," *International Journal of Science and Research*, vol. 3 issue 10, pp. 1374–1376, 2014.
- [8] D. Skopin and S. Baglikov, "Heartbeat feature extraction from vowel speech signal using 2D spectrum representation," *Proceedings of the 4th International Conference Information Technology*, June 2009.
- [9] A. Mesleh, D. Skopin, S. Baglikov and Anas Quteishat, "Heart Rate Extraction from Vowel Speech Signals," *Journal of computer science and technology*, vol. 27, no. 6, pp. 1243–1251, 2012.
- [10] M. Sakai, "Feasibility Study on Blood Pressure Estimations from Voice Spectrum Analysis," *International Journal of Computer Applications*, vol. 109, no. 7, pp. 39 –43, 2015.