International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 5
May 2017

Heart Rate Monitoring using Human Speech Features Extraction: A Review

Alampreet Chahal
Department of Electronics &Communication, Punjabi University, Patiala

Silki Baghla
Department of Electronics &Communication, JCDM college of engineering, Sirsa

Dr. Rajbir Kaur
Department of Electronics &Communication, Punjabi University, Patiala

Gautam Kaushal
Department of Electronics &Communication, Punjabi University, Patiala

ABSTRACT

Human speech can be used for monitoring heart rate. Monitoring of the heart can offer important insight with regard to health and wellness. There is a strong correlation between voice signal and heart rate. Several methods such as Mel Frequency Cepstral Coefficient (MFCC), Linear Predictive Coding (LPC), Hidden Markov model (HMM) and Artificial Neural Networks (ANN) are used with a view to identify a straight forward and effective method for voice signal extraction. In this paper we have gathered the contribution made by various researches for heart rate monitoring using human speech feature extraction.

Keywords

Heart rate, HMM, MFCC, LPC, ANN and signal feature extraction

INTRODUCTION

Heart rate indicates the total number of times our heart contracts and relaxes per minute and is expressed as the number of beats per minute (bpm). Heart rate measurement is one of the very important parameters of the human cardiovascular system. The heart rate of a healthy adult at rest is around 72 beats per minute (bpm). Babies have a much higher heart rate at around 120 bpm, while older children have heart rates at around 90 bpm. Athletes normally have lower heart rates than less active people. The heart rate rises gradually during exercises and returns slowly to the rest value after exercise. The rate when the pulse returns to normal is an indication of the fitness of the person. Lower than normal heart rates are usually an indication of a condition known as Brady cardiac, while higher than normal heart rates are known as tachycardia.[1] Heart Rate also depends on the body's need to absorb oxygen and excrete carbon-dioxide. Many factors contribute to variation in heart rate such as level of physical activity, fitness, temperature, body position, emotions, body size and medication. The heart is a muscular pump made up of four chambers. The two upper chambers are called atria, and the two lower chambers are called ventricles. A naturalelectrical system causes the heart muscle to contract and pump blood through the heart to the lungs and the rest of the body [2].Heart rate is measured by detecting arterial pulsation. Virtually all sciences contribute to the maintenance of human health and the practice of medicine. Medical physicists and biomedical engineers are the professionals who develop and support the effective utilization of this medical science and technology as their responsibilities to enhance human health care with the new development of the medical tools such as electrocardiogram (ECG). Electrocardiogram is a method used to measure the rate and regularity of heartbeats to detect any irregularity of the heart. An ECG translates the heart electrical activity into wave-line on paper or screen. The spikes and dips in the line tracings are called P wave, QRS complex, T wave, ST segment and RR interval (period

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 5
May 2017

between two sequences of R and R waves of ECG signal) [3].The clinical method of heart beat monitoring is ECG waveform. The ECG waveform is different for the different patient to extend that they are unlike to each other and at the same time a like for different types of beats. Fig.1 shows the typical

ECG signal with three indicated waves: the P, QRS and T. The P wave is the result of slow moving

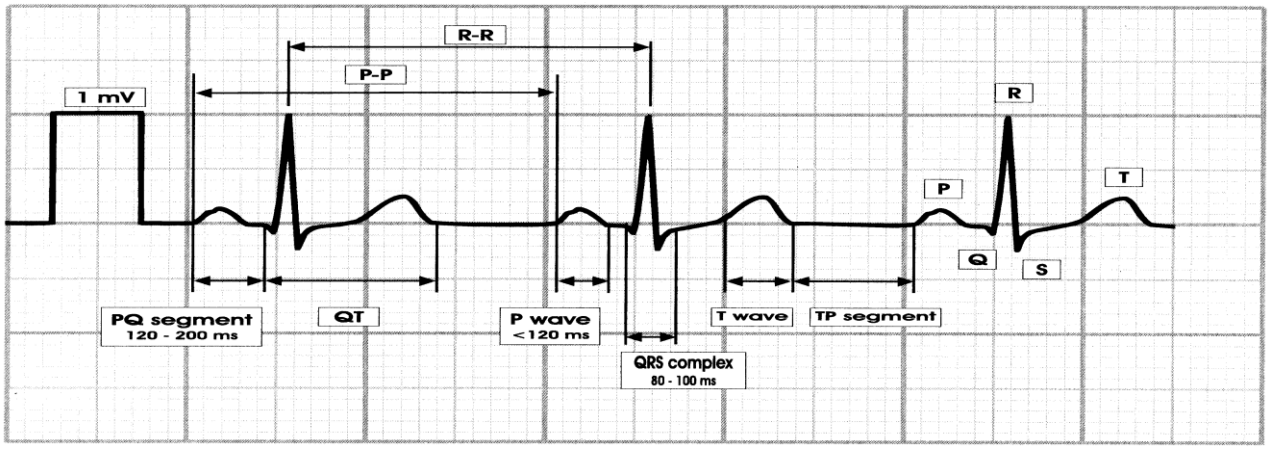

Figure.1 Typical ECG Waveform

Depolarization of the atria. The wave of stimulus spreads rapidly from the apex of the heart upwards, causing rapid depolarization of the ventricles. This results in the QRS complex of the ECG. The Plateau part of action potential of about 100-120ms after the QRS is known as the ST segment. The depolarization of the ventricles causes the slow T wave with amplitude of 0.1-0.3mV. Between T and P waves there is a long plateau part of small amplitude is known as TP segment[4]. The human heart rate is in the range of 60-100bpm but they may change according to age, sex and height of the person. The earlier models of heart rate monitors used electrode leads that are attached to the chest but the modern heart rate monitors use chest strap transmitter and a wrist receiver. Strapless heart rate monitors allows the user to simply touch two sensors on a wrist watch to display the heart rate. The detection of heart rate from human speech was also done based on the modeling of vowel speech signal. The heart rate is calculated as the time interval between the two QRS complexes per unit time. The heart rate is obtained from recorded ECG's using 1500 rule. Count the number of small squares between two neighboring R-waves and divide that number by 1500.

$$H = \frac{1500}{Numberofsmallboxesbetween2R-Waves}$$

The demand for contactless heart rate monitoring has increased lately, especially for long duration monitoring and for patients with particular condition such as burn victims or infants at the risk of sudden infant syndrome. In this paper different methods of contactless heart rate monitoring introduced in the literature are complied and presented.

LITERATURE REVIEW

In[5],Atal and Fumitada Itakura formulated the fundamental concepts of Linear Predictive Coding (LPC) for estimation of the vocal tract response from speech waveforms.The total log prediction residual of an input signal is minimized by optimally registering the LPC onto the input autocorrelation coefficients using the dynamic programming algorithm (DP). An algorithm is used to find the best match between the input pattern and the reference pattern is derived. The dynamic programming technique is used in conjunction with a

sequential decision scheme. A sequential decision procedure is used to reduce the amount of computation in dynamic programming algorithm. The recognition time is about 22 times real time.

In [6], Rollin McCraty, Mike Atkinson, William Tiller, Glen Rein, and Alan D. Watkinsextend the previous findings by demonstrating that anger produces a sympathetically dominated power spectrum. The appreciation produces a power spectral shift toward Medium frequency (MF) and high frequency (HF) activity. The results suggest that the positive emotion lead to alteration in Heart Rate Variability (HRM), which may be beneficial in treatment of hypertension and reducing the death in patients with congestive heart failure.

In [7], Kamal Nigam, Andrew Kachites Mccalum, Sebastian Thrun and Tom Mitchell used an algorithm for learning from labeled and unlabeled documents based on the combination of Expectation Maximization (EM) and a naive Bayes classifier. The Expectation Maximization procedure works well when the data conform to the generative assumptions of the model. There are two extensions to the algorithm that improve classification accuracy, a weighting factor to modulate the contribution of unlabelled data and use of multiple mixture components. The use of unlabelled data reduces the classification error by 30%.

In [8], K. H. Kim, S. W. Bang and S. R. Kim developed a novel emotion recognition system based on the processing of physiological signals. The system consisted of preprocessing, feature extraction and pattern classification stages. Preprocessing and feature extraction methods were derived so that emotion-specific characteristics could be extracted from short-segment signals. A support vector machine was utilized as a pattern classifier to overcome the difficulty in pattern classification due to the large amount of within class variation of features and the overlap between classes. Correct classification ratios for 50 subjects were 78.43% and 61.76%, for the recognition of three and four categories.

In [9], Sigurdur Sigurdsson, Kaare Brandt Petersen and Tue Lehn-Schiøler have discussed the influence of MP3 coding for the Mel frequency cepstral coefficients (MFCCs). The widely used subset of the MFCCs is robust at bit rates equal or higher than 128 kbit/s. The results show that the different MFCC implementations are correlated for approximately the first 15 MFCCs. MFCCs were shown to be very robust at bit rates of 320 and 128 kbit/s at a fixed sampling rate of 44.1 kHz. The robustness decayed rapidly that included higher frequencies in the Mel filter bank. This paper shows that MFCC features are very robust to MP3 encoding and applicable in MIR tasks.

In [10], Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, and Sung Hyon Myaenghave suggested that naive Bayes was quite effective in various data mining tasks; it showed a disappointing result in the automatic text classification problem. Based on the observation of naive Bayes for the natural language text, a serious problem in the parameter estimation process was found, which caused poor results in text classification domain. In this paper they proposed two methods: per-document text normalization and feature weighting method. The proposed naïve Bayes text classifier performed well in standard benchmark collections, competing with state-of-the-art classifiers.

In [11], Dmitriy E. Skopin and Sergey U. Baglikov used a method to measure the rate and regularity of heartbeats to detect any irregularity in a heart. The heartbeat information may be extracted using proposed method of 2D spectrum representation that is extension of Short Time Fourier transform and can be used for AM, FM, ultrasound, Doppler monograph, and other signals analysis.

In [12], Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi presented the viability of MFCC to extract features and DTW to compare the test patterns. The digital signal processes such as Feature Extraction and Feature Matching is introduced to represent the voice signal. The non-parametric method for modeling the human auditory perception system, Mel Frequency Cepstral Coefficients (MFCCs) were utilized as extraction techniques. The principle of DTW is to compare two dynamic patterns and measure its similarity by calculating a minimum distance between them. The voice is a signal of infinite information. Direct analysis and synthesizing the complex voice signal.

In [13], Ibrahim Patel and Dr. Y. Srinivas Rao proposed an approach to the recognition of speech signal using frequency spectral information with Mel frequency for improvement of speech feature representation in a

Hidden Markov Model (HMM) based recognition approach. The Mel frequency approach exploits the frequency observation for speech signal in a given resolution which results in resolution feature overlapping. The two features passed to the HMMnetwork result in better recognition compared to existing MFCC method. They have better efficiency for accurateclassification.

In [14], Abdelwadood Mesleh, Dmitriy Skopin, Sergey Baglikov, and Anas Quteishat used a novel non contact heart rate extraction method from vowel speech signals. The proposed method is based on modeling relationship between speech production of vowel speech signals and heart activities for humans where it is observed that the movement of heart beat causes a short increment of vowel speech formants. The short-time Fourier transform (STFT) is used to detect the formant maximum peaks so as to accurately estimate the heart rate.

In [15], Abdallah Meraoumia, Salim Chitroub and Ahmed BouridaneintegratedFP and FKP in order to construct an efficient multi-biometric recognition system based on matching score level and image level fusion. Finger prints has one of the highest levels of reliability. Biometrics is an effective technology for personnel identity recognition, but uni-modal biometric systems which uses a single trait for recognition suffered from problems like noisy sensor data, non-universality, lack of distinctiveness of the biometric trait, and spoof attacks.

In [16], Media Anugerah Ayua, Siti Aisyah Ismaila, Ahmad Faridi Abdul Matina, Teddy Mantorobrecommended that many mobile phones have been equipped with sensors to enable the delivery of advanced features to the users. Earlier research has shown that data from mobile-phone embedded accelerometer can be used for activity recognition purpose. Overall accuracy rate for classifier training managed to exceed 96% and exceeded 90% for classifier testing.

In [17], Mitar Milacic, A.P.James, Sima Dimitrijev demonstrated the use of formant based features because a better formant extraction method is to increase the effectiveness of formants. The automatic speech recognition system implemented with five formant based features outperforms the conventional features such as MFCC and PLP when noisy levels are high. Combining formant features with other features such as MFCC and PLP can improve the rubustness of ASR systems. The disadvantage of formant feature is lower recognition performance which can be overcome by combining features.

In [18], Waseem Shahad, Salman Asad and Muhammad Asif Khan proposed a technique that used J Rip classifier and association rule mining to select the most relevant features from a data set. J Rip extracts the rules from a data set and then association rules mining technique is applied to rank the features.Select the relevant features increases the accuracy and decreases the computational cost.

In [19], ChenpingHou, FeipingNie, Dongyun Yi and Yi Wuproposed a novel multiple rank regression model (MRR) for matrix data classification. Unlike traditional vector based methods, they employed multiple-rank left projecting vectors and right projecting vectors to regress each matrix data set to its label for each category. Traditional classification methods are vector based. MRR achieves higher accuracy and hence lower computational complexity. Compared with traditional supervised tensor-based methods, MRR performs better for matrix data classification.

In [20], Esratjahan, TilottomaBarua and Umme Salma monitored the heart rate and blood oxygen saturation is a couple of biometrics to provide information regarding the health of the body. Oxygenated blood has different light absorption characteristics than deoxygenated blood under red and infrared wavelengths. This device is able to produce highly reliable results for both heart rate and SPO2 level. ThePrinciple of pulse oximeter is based on the red and infrared light absorption characteristics of oxygenated and deoxygenated hemoglobin. The main advantage is that it can be used by non professional people at home to measure the heart rate and SPO2 level easily and safely.

In [21], AsselDavletcharova, SherinSugathan, Bibia Abraham and Alex Pappachan James proposed a novel non contact heart rate extraction method from vowel speech signals. The proposed method was based on modeling relationship between speech production of vowel speech signals and heart activities for humans

where it was observed that the movement of heartbeat caused a short increment of vowel speech formants. The short-time Fourier transform (STFT) is used to detect the formant maximum peaks so as to accurately estimate the heart rate.

In [22], AibekRyskaliyev, SanzharAskaruly and Alex Pappachan James explained a method for estimating the heart rate from human speech using voice signal analysis and by development of empirical linear predictor model. There is a correlation between voice signal and heart rate established by classifiers and prediction of the heart rates. MFCC was used to extract feature distances for voice spectrum analysis.

CONCLUSION

Several techniques for voice recognition such as Mel-Frequency Cepstral Coefficient (MFCC), Linear Predictive Coding (LMC), Hidden Markov Model(HMM) and Artificial Neural Networks(ANN) has been used in literature to correlate the features of voice with that of human heart rate. These techniques do not consider artificial hearts and transplanted hearts. The methods are based on human speech which cannot give accuracy as compared to traditional ECG but can be improved to provide linear relationship between heart rate measured by clinical process (ECG) and calculated by simulation. In this paper we have compiled the contributions of researchers regarding heart rate detection using human speech feature extraction.

REFERENCES

[1] Meda Sai Kheerthana, Manjunath A 2015 A Survey on Wearable ECG Monitoring System Using Wireless Transmission of Data. International Journal of Advanced Research in Computer and Communication Engineering.
[2] Mesleh A, Skopin D, Baglikov S, Quteishat A 2012 Heart rate extraction from vowel speech signals. Journal of Computer Science and Technology.
[3] Skopin D, Baglikov S 2009 Heartbeat feature extraction from vowel speech signal using 2D spectrum representation. InProceedings of IEEE International Conference of Information Technology.
[4] Linh T.H, Osowski S 2003 On-line heart beat recognition using hermite polynomials and neuro-fuzzy Network. IEEE Transactions on Instrumentation and Measurements.
[5] Itkara F 1975 Minimum Prediction Residual Principle Applied to Speech Recognition. IEEE Transaction on acoustics, speech and signal processing.
[6] McCraty R, Atkinson M, Tiller WA, Rein G, Watkins AD 1995the effects of emotions on short-term power spectrum analysis of heart rate variability. American Journal of Cardiology.
[7] K. Nigam, A.K. McCallum, S. Thrun, and T.M. Mitchell 2000Text Classification from Labeled and Unlabeled Documents Using EM. Machine Learning.
[8] K.H. Kim, S.W. Bang and S.R. Kim 2004 Emotion recognition system using short-term monitoring of physiological signals. Medical and biological engineering and computing.
[9] Sigurdsson S, Petersen K, Lehn-Schiøler T 2006 Mel frequency cepstral coefficients: An evaluation of robustness of mp3 encoded music. Proceedings of the Seventh International Conference on Music Information Retrieval.
[10] Kim S-B, Han K-S, Rim H-C, Myaeng SH 2006 Some effective techniques for naive bayes text classification. IEEE Transactions on Knowledge and Data Engineering.
[11] Skopin D, Baglikov S 2009 Heartbeat feature extraction from vowel speech signal using 2D spectrum representation. In Proceedings International Conference Information Technology.
[12] Muda L, Begam M, Elamvazuthi I 2010 Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. Journal of Computing.
[13] Patel I and Srinivas Rao Y 2010 Speech recognition using HMM with MFCC. Signal and Image Processing: An International Journal.
[14] Mesleh A, Skopin D, Baglikov S, Quteishat A 2012 Heart rate extraction from vowel speech signals. Journal of Computer Science and Technology.

[15] Meraoumia A, Chitroub S, Bouridane A 2012 Multimodal biometric person recognition system based on fingerprint & finger-knuckle-print using correlation filter classifier. International Conference of Communications.
[16] Ayu MA, Ismail SA, Matin AFA, Mantoro T 2012 A comparison study of classifier algorithms for mobile phone's accelerometer based activity recognition. Procedia Engineering.
[17] Milacic M, James AP, Dimitrijev S 2013 Biologically inspired features used for robust phonemeon recognition. International Journal of Machine Intelligence and Sensory Signal Processing.
[18] Shahzad W, Asad S, Khan MA 2013 Feature subset selection using association rule mining and j-rip classifier. International Journal of physical sciences.
[19] Hou C, Nie F, Yi D, Wu Y 2013 Efficient image classification via multiple rank regressions. IEEE Transactions on Image Processing.
[20] Jahan E, Barua T and Salma U 2014 An overview on heart rate monitoring and pulse oximeter system. International Journal of latest research in science and technology.
[21] Davletcharova A, Sugathan S, Abraham B, James AP 2015 Detection and analysis of emotion from speech signals. Procedia Computer Science.
[22] Ryskaliyev A, Askaruly. S and James A.P 2016 Speech signal analysis for the estimation of heart rates under different emotional states.In Proceedings of International Conference on Advances in Computing, Communications and Informations.