International Journal of Applied Engineering and Technology ISSN: 2277-212X (Online)
An Open Access, Online International Journal Available at http://www.cibtech.org/jet.htm
2014 Vol. 4 (3) July-September, pp.13-16/Kaur and Kaur
Research Article

EXTRACTION OF PHYSIOLOGICAL PARAMETERS USING AUDIO SIGNALS

***Jaswinder Kaur and Rupinder Kaur**
DIET, Kharar, Punjab, India
**Author for Correspondence*

ABSTRACT
It is observed by the study that audio recordings of sustained vowels may determine the breath, whether a person's pulse is high or low. However, good results for HR and SC values always required a speaker's training input, which limits spontaneous recognition as well as analysis of audio recordings without physiological reference data in possible applications. To improve the performance, it is aimed to increase recognition accuracy by selection of features, and to compare the importance of different feature groups. Recognition accuracy could be further increased by enlargement of the data set and development of a predictor able to recognize predictions out of an expected range of values. It has also been observed that a threshold for separation of high pulse and low pulse of HR is required and has been taken from the user input. However, this should be computed from the speech signal automatically by combining the speech signal frequency and amplitude values. In the presented work, a self adaptive threshold selection is proposed based on the input speech signal amplitude, speed and frequency. Further a better correlation is established between the speech signal level ad HR and skin conductance.

Keywords: *HR, Speech Synthesis, Zero Crossing*

INTRODUCTION
The audio signals or voice may be used for physiological parameters like heart rate and skin conductance. The variety of affordable and portable medical devices allowing a person to actively contribute to diagnosis and treatment is permanently increasing. Particularly for persons whose mobility is limited or who are living remotely this trend may significantly improve quality of life. There are devices for measuring blood pressure, heart rate, body core temperature, respiration rate and many other physiological parameters autonomously, but yet they are still rather expensive and inconvenient for an everyday use. Ideally, monitoring of vital signals should require a minimal effort by a user and cause minimal disturbance; a user should not have to spend time thinking about the use of a monitoring device or even notice monitoring particularly.
Monitoring should be easy enough to perform it in emergency situations, for example when calling a hospital, and monitoring should also be carried out over longer periods of time so that it could spontaneously react on emergencies or collect vital data for creating a health profile. For these reasons, physiological data should be recorded by sensors that are best unnoticed' in terms of intrusiveness. Disturbance of daily life would be minimal, if not a separate device had to be carried, but if monitoring of physiological parameters could be performed by computers or mobile phones which provide computational power in reach most of the time already. These considerations draw attention on signal types which can be recorded by mobile phones and computers easily: video and audio.
In addition to the advantages mentioned above, video- and audio-based recognition can also be performed on past recordings, e.g., movies, songs and other voice recordings. A major advantage of audio-based over video-based recognition is that a microphone does not have to be directed towards a user's face or skin. It can also be employed as a complimentary technology in any situation in which a video camera is not available or able to record, e.g. in the dark.

Related Works
Heart rate variability (HRV) provides important information about the development of the cardiovascular system in fetuses. This paper presents a new measure of fetal HRV that can be estimated using Doppler ultrasound techniques. This measure employs the multiple signal characterization (MUSIC) algorithm

which is a high-resolution method for estimating the frequencies of sinusoidal signals embedded in white noise from short-duration measurements. We show that the product of the square-root of the estimated signal to noise ratio (SNR) and the variance of the frequency estimates is independent of the noise level in the signal. Since variations in the angle of incidence of the Doppler ultrasound beam effectively changes the input SNR, this measure of HRV is robust to the input noise as well as the angle of incidence. Presented analysis results validate the robustness properties and the usefulness of the HRV measure (Kumari and Mathew, 2003)

A research project on stress assessment is running at De@ University of Technology since 1992. One of the aims of the project is to develop an instrument: for automated stress assessment. The underlying system is based on the analysis of facial expressions, voice analysis and the analysis of physiological signals such as heart rate and blood pressure. Analysis of these multi-medial data takes place in parallel and are based on Artificial Intelligence technology. In each of the parallel subsystems, corresponding to sensor, image and sound data, the functionally is split up into a number of layers: filtering and reduction layer, pre-processing layer, processor layer, application layer and ou/put layer. The results of the analysis are combined by a central interpreter, resulting in an overall stress measure. In /his paper the stress assessment is used to monitor vigilance levels of car drivers with a focus on voice analysis (Rothkrantz *et al.,* 2004).

We discuss a procedure that exploits the strengths of autoregressive (AR) and minimum variance distortion less response (MVDR) methods to circumvent their respective weaknesses, in the context of estimating the power spectral measures from cardiovascular signals. While the AR power spectrum density (PSD) has high resolution, it suffers from a greater degree of variability at low orders and potential spurious peaks at high orders. The MVDR method (of lower resolution) provides the power-spectral measures at lower variability over low filter orders. The weaknesses of the two methods can be circumvented by estimating the power by Low-order MVDR filters at the frequencies estimated by AR PSD. Results of experiments with simulated and real data are supportive (Galigekere and Shoemaker, 2004).

The focus of this paper is mental tension detection in speech to assist control the tension in day-to-day business such as conferences and operations in a call centre. It is difficult to use classical techniques for mental tension detection in day to- day business because those techniques require invasion body by electrodes or squirts and tied up by cables. In order to achieve a non-invasive, non-contact and low-restricting method, this proposed technique uses acoustic features in the speech. The technique uses the vocal tract model which represents the shape and the tightness of throat muscle. The Gaussian Mixture Model (GMM) classifies two mental tension states: high-tension and non-tension. The experiment result shows high recognition rate of mental tension detection (Michiaki *et al.,* 2007).

This paper presents a novel non-contact heart rate extraction method from vowel speech signals. The proposed method is based on modelling the relationship between speech production of vowel speech signals and heart activities for humans where it is observed that the moment of heart beat causes a short increment (evolution) 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. Compared with traditional contact pulse oximeter, the average accuracy of the proposed non-contact heart rate extraction method exceeds 95%. The proposed non-contact heart rate extraction method is expected to play an important role in modern medical applications (Abdelwadood *et al.,* 2012).

Algorithm

The human speech production system is capable of conveying an abundance of information with regards to the text of a sentence, identity of a speaker, as well as emotion and stress. Mood, emotion, personality and other pragmatic information about the state of the speaker are present in every spoken utterance. At present, interest in this area of research is increasing as the number of potential applications is increasing and vocal emotions have also tended to be studied in isolation. About 25% of information contained in a speech signal refers to the speaker. These phonetically irrelevant speaker characteristics make speech recognition less effective but can be used for speaker identification and classification of the speaker's internal state. Voice analysis is based on the way of speaking and on the verbal content of messages. Several stress features can he extracted from voice signals and given below:

FO mean: fundamental frequency of the voice signal

FOO Perturbation: slight variations in glottal cycles

Intensity Mean: energy values for a voice signal

Speech Rate: length of an utterance in time

Zero Crossings: number of times a voice signals crosses averaged over an utterance the zero line based on the recorded values of these parameters, the heart rate can be accessed.

Noise filtering is very important as background noise can completely disguise the original voice signal, Noise filtering is the most important task of the filtering and reduction layer. The other important task is feature extraction. The pre-processor layer takes care of storing samples and average features over speech utterances. The processor layer has to compute the relation between stress and voice features.

Ideally, the STFT spectrogram includes speech formants without any noise. But practically, the heart activities cause distortion to all formants, and the duration of the associated distortion is approximately equal to 0.2 seconds with 100 Hz magnitude.

Below shows an ideal STFT spectrogram for a vowel speech signal. It represents a 40 beats per minute (bpm) heart activity which is the lowest heart rate in real situations. In order to extract the relevant heart rate information from the corresponding STFT spectrogram, a searching algorithm is proposed to horizontally scan the STFT spectrogram starting from the top (Nyquist frequency) to the bottom (DC) of the original speech signal. Each time the algorithm scans the spectrogram of the speech signal horizontally, a one-dimensional (1-D) signal is obtained. Generally, there are two possible cases of the horizontal scanning when a scanning line passes through a part of the spectrogram beyond the bounds of the formants, it contains background information only and it is not able to extract a useful 1-D signal.

When a scanning line passes through any part of a speech formant, it is able to extract a useful 1-D signal.

Each extracted 1-D signal (useful 1-D signal) passes through a 5th-order FIR low-pass filter to suppress high frequency components. Finally, a discrete Fourier transform (DFT) is applied to the filtered 1-D signals.

The amplitude of the 4th harmonic is the maximum. Based on Fourier transform properties, the harmonic number four of a six-second speech signal corresponds to 0.67 Hz frequency and to 40 bits per second heart rate (which is exactly the heart rate of the original signal).

As a result, it is concluded that heart rates can be extracted using a number of harmonics with maximum magnitudes.

RESULTS AND DISCUSSION

The presented work is based on extraction of body physiological parameters work is based on extraction of body skin conductance and heart rate. The accuracy primarily depends upon the authentic recording of speech signal and uniformity of the text material. The different text material may give rise to a large variation in the results as the work is considerably new and the data for the same is not available for validation or comparison. The data or the same has to be generated and text material has to be more specific and not the general one. A heart surgeon may suggest in this case when deciding on text material that may give enough information about the heart variation.

REFERENCES

Galigekere RR and Shoemaker JK (2004). On Utilizing the Strengths of AR and MVDR Methods to Circumvent Their Weaknesses in Application to HRV Analysis. *IEEE-Institute of Electrical and Electronics Engineers. 02766547/04* $20.00 0.

Kumari Fernando L and John Mathew V (2003). Robust Estimation of Fetal Heart Rate Variability Using Doppler Ultrasound. *IEEE-Institute of Electrical and Electronics Engineers* **8**. 0-7803-7663-3/03/$17.00 82003.

Mesleh A, Skopin D, Baglikov S *et al.*, (2012). Heart rate extraction from vowel speech signals. *Journal Of Computer Science And Technology* **27**(6) 1243-1251. DOI 10.1007/s11390-012-1300-6.

Michiaki Ariga, Yoshikazu Yano, Shinji Doki and Shigeru Okuma (2007). Mental Tension Detection in the Speech Based on Physiological Monitoring. *IEEE-Institute of Electrical and Electronics Engineers*. 1-4244-0991-8/07/$25.00/©2007.

Rothkrantz LJM, Van Vark FLJ and Datcu D (2004). Multi-medical Stress Assessment. *IEEE-Institute of Electrical and Electronics Engineers*. 0-7803-8566-7/04/$20.00 Q.