

# Vocalization of Heart Rate Variability

Sokol Saliu<sup>1</sup>, Ahmet Birand<sup>2</sup>, Gulmira Kudaiberdieva<sup>2</sup>

<sup>1</sup> Chalmers University of Technology, Lindholmen College, Göteborg, Sweden

<sup>2</sup> Division of Cardiology, Çukurova University, Adana, Turkey

**Abstract**— In this paper, we have proposed vocalization of heart rate variability (HRV) as a perceptual analysis tool. We adapted a phonation-production model to encode external signals and generate audible representations of them. HRV changes, caused by induced perturbations to the autonomous nervous system, could be perceived on vocalized HRV.

**Keywords**— Heart rate variability, voice perturbations, voice synthesis, man-machine interface.

## I. INTRODUCTION

To convey information, numerical data should be appropriately mapped to our sensory world. As a prominent example, computer visualization, which exploits our complex vision system, has successfully helped to 'make sense' of vast amount of numerical data. In the same vein, computer vocalization could engage another subtle physiological tool namely the auditory system.

The perception of perturbations on sustained phonation has provided diagnostic information related to voice disorders. For this, various perceptual voice quality scales, such as GRBAS scale [1](with grades 'hoarseness', 'rough', 'breathy', 'asthenia', and 'strained'), and other voice characteristics, such as 'rough', 'creak', 'fry', 'falsetto', and combinations thereof are being used.

The perceptual characteristics of the phonation are found to be related to the amplitude and pitch of the perturbations and to the power of laryngeal noise [2]. Depending on the characteristics of signal perturbation, pitch perturbations can produce 'natural', 'fry', 'creak' and pitch-varying phonations. If perturbation consists of multiple frequency components, the sound will be perceived as 'polyphonic.' On the other hand, amplitude perturbation is perceived as 'loudness' shimmer whilst 'laryngeal' noise as 'breathiness' or 'hoarseness' [3]. Changes on shape of 'glottal' pulses produces sounds with different scales of 'naturalness' and 'quality'. Modification of vocal tract parameters produces different phonemes or phoneme-like sounds [3].

The HRV is similar to the perturbations of the phonations—both signals are perturbations to physiological rhythms and carry useful clinical information—with a perceptual difference, the HRV is not audible.

Here we have propose vocalization as a process of coding digital signals to the parameter space of a voice-synthesis model. The model is called the vocalization model (VM) and the audible sound generated by it the vocalized signal(VS). The coding scheme needs to consider both the perception characteristics of auditory system and the characteristics of data to be vocalized.

Signals vocalization could add a new perceptual dimension to the biomedical (or other) data. Anyhow, some problems that would need to be addressed are: which VM would

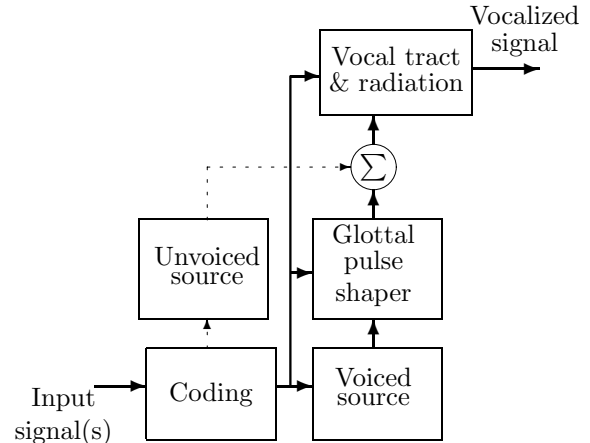


Fig. 1. Vocalization model

be more appropriate for a given class of signals? Which type of coding to choose? Which VM could assist pattern recognition? Which coding could provides better perceptual information?

This paper presents the VM and addresses the last questions by an application from the HRV analysis, which is an important noninvasive tool that sheds light on the autonomous nervous system(ANS) control of heart rate [4] and provides valuable clinical information [5].

We would let the HRV signals to modify the pitch period of the VS in VM. After some practice in listening to the vocalized HRV, we could differentiate amongst various states of HRV induced respectively by the parasympathetic and sympathetic blockade on healthy subjects.

## II. MATERIAL AND METHOD

### A. Vocalization Model

Figure 1 shows the block diagram of the vocalization model. In difference to speech-synthesis models, which usually modifies the vocal tract and switch between voiced and unvoiced sources [6], [7], the VM was basically a sustained-phonation synthesis model, which modified primarily the voiced source—and could also provide for additive unvoiced source.

The voiced source was modelled by a generator of Kronecker delta pulses followed by Rosenberg's glottal pulse[8] shaper(Fig. 1). The vocal tract was modelled by an autoregressive model of order twelve and lips' radiation by a high pass filter ( $H(z^{-1}) = 1 - 0.99z^{-1}$ ). Laryngeal noise or unvoiced source was modelled by a white noise generator. The coding unit coded one or more input signals to time-varying parameters of the other blocks. The VM depicted in Fig. 1 has various degrees of freedoms. The

## Report Documentation Page

<b>Report Date</b> 25 Oct 2001	<b>Report Type</b> N/A	<b>Dates Covered (from... to)</b> -
<b>Title and Subtitle</b> Vocalization of Heart Rate Variability	<b>Contract Number</b>	
	<b>Grant Number</b>	
	<b>Program Element Number</b>	
<b>Author(s)</b>	<b>Project Number</b>	
	<b>Task Number</b>	
	<b>Work Unit Number</b>	
<b>Performing Organization Name(s) and Address(es)</b> Chalmers University of Technology Lindholmen College Goteborg, Sweden	<b>Performing Organization Report Number</b>	
<b>Sponsoring/Monitoring Agency Name(s) and Address(es)</b> US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	<b>Sponsor/Monitor's Acronym(s)</b>	
	<b>Sponsor/Monitor's Report Number(s)</b>	
<b>Distribution/Availability Statement</b> Approved for public release, distribution unlimited		
<b>Supplementary Notes</b> Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.		
<b>Abstract</b>		
<b>Subject Terms</b>		
<b>Report Classification</b> unclassified	<b>Classification of this page</b> unclassified	
<b>Classification of Abstract</b> unclassified	<b>Limitation of Abstract</b> UU	
<b>Number of Pages</b> 3		

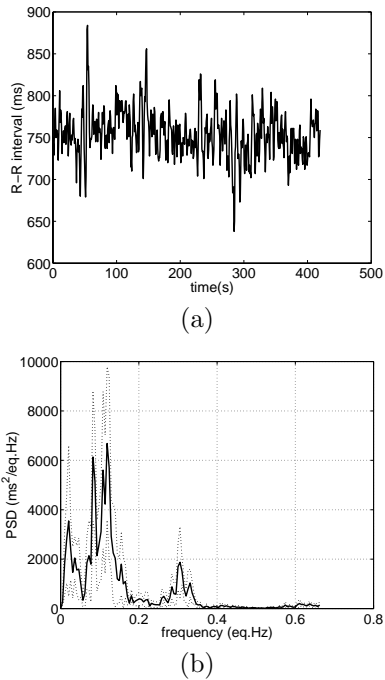


Fig. 2. Short term HRV signal(a) and its power spectral density (b) of a healthy female subject in rest position.

amplitude(i) and the period(ii) of impulses (voiced source block) can be modulated either individually or in combination. Glottal pulse shapes(iii) can be varied parametrically. The unvoiced source could be scaled(iv) and correlated(v) by filters whose parameters could vary smoothly. The formants(vi) (vocal tract block) could also be modified parametrically or otherwise.

Thus, variations of pitch, loudness, and 'laryngeal' noise, and possibly 'glottal' shaping and 'vocal tract' changes, could produce a rich variety of sounds for the trained ear.

### B. HRV Signals

The HRV data were used here were obtained from a group of normal subjects that participated in a study on the influence of the ANS on HRV. The HRV signals with length of about six to seven minutes were obtained from subjects on supine and tilt positions with and without induced parasympathetic and sympathetic autonomic blockade as described in [9].

A typical short-term HRV signal, obtained from a female subject in rest position, is shown in Fig. 2(a); the power spectral density (PSD) of the detrended signal is shown in Fig. 2(b). The PSD consists typically of three main frequency components [10], [11] and has a maximum frequency extent of few Hertz— significantly lower than the lowest frequency on the audible frequency range. Fig. 3 shows the HRV signal and its PSD from the same subject in rest position with induced parasympathetic blockade.

### C. Voiced Source Coding

We confined the discussion to voiced source coding. To construct the VM model we first obtained the mean pitch

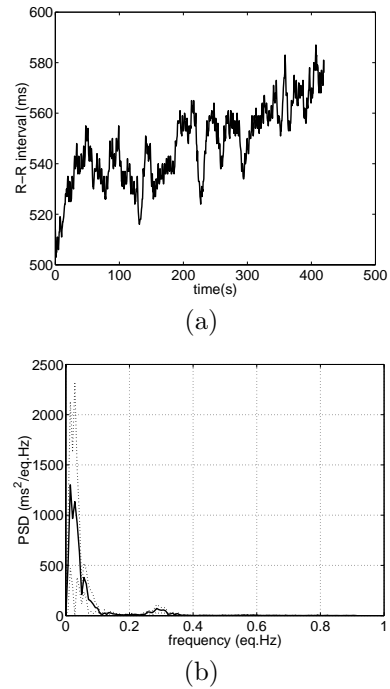


Fig. 3. Short term HRV signal(a) and its PSD(b) of the same subject with induced parasympathetic blockade.

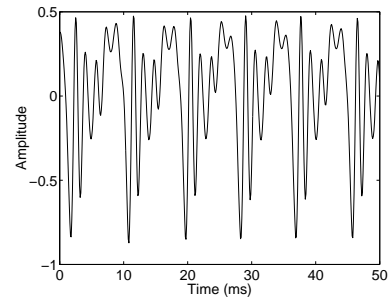


Fig. 4. A segment of vocalized HRV

period of a phonation /a/ (sampled at 10 kHz) from a male subject and estimated the vocal tract parameters as described previously. The voiced source had the pitch above.

Second, the period of voiced source was modulated by the HRV signal. The period variation of the voiced source was related to the HRV signal by

$$\Delta\nu = c \frac{\Delta p}{p_m} \nu_m \quad (1)$$

where  $p_m$  and  $\nu_m$  are respectively the sample means of the HRV signal and the pitch period of the voiced source, whereas  $\Delta p = p - p_m$  and  $\Delta\nu = \nu - \nu_m$  denote respectively deviations from the corresponding means. The constant  $c$  is a subjective scale factor. Passing the source signal through the 'glottal' pulse shaper and the vocal tract generates the VS. A segment of the a vocalized HRV signal is shown in Fig. 4.

### III. RESULTS

We have proposed a vocalization model and applied it to vocalize HRV data obtained from group of normal patients under various experimental settings. The vocalization model was included in a HRV evaluation system [12] where comparative perceptual assessment of HRV was obtained by selectively listening to HRV signals in various windows. After some practice we could discern few auditory patterns on vocalized short-term HRV signals related to postural changes and drugs effects on the parasympathetic-sympathetic balance of the ANS.

### IV. DISCUSSION

The vocalization model was found quite flexible and could be applied generally to various types of signals. Anyhow, at present, its main limitation was its utility. As an example, although the vocalized HRV offered many auditory clues to both short-term and long-term HRV recordings, to make use of them, one should need to establish perceptual categories, like the scales used in pathological voice assessment [1], and train physicians accordingly.

We discussed here a simple coding scheme, that of the single-source VM based on the modulation of a single parameter. Other coding schemes could yield quite different perceptual information whereas the multi-parameters coding of a multi-sources VM could create a high dimensional auditory space.

An advantage we could observe from the application on long-term HRV analysis, was that long-term pitch variation as well as local alterations of the pitch could be perceived concurrently. In difference, global on-screen visualization of such data suffered from spacial aliasing (due to limits on monitor resolutions) whereas on local visual scanning we needed to refer to values on the ordinate axis to follow up long-term variations.

The combined visual and auditory perception of same or related data, could provide a more natural access to such information, probably with a synergistic effect. For example scanning concurrently the HRV data, both visually (HRV signal) and audibly (vocalized signal), could provide a different mental picture of HRV which could facilitate the recognitions of patterns.

The results presented are preliminary. The full extent of the HRV application would need to be further evaluated and the vocalization model to be accordingly adjusted.

In conclusion, the vocalization model, as demonstrated by the application on heart rate variability, provided an additional tool to data analysis. Using it, we were able to perceive auditory patterns on vocalized heart rate variability signals caused by shifts on parasympathetic-sympathetic balance of the autonomous nervous system.

### REFERENCES

- [1] M. Hirano. *Clinical examination of Voice*. Springer Verlag, 1981.
- [2] H. Kasuya and H. K. Masubuchi. Preliminary experiments on voice screening. *Journal of Phonetics*, 14:463–468, 1986.
- [3] D. G. Childers. Vocal quality factors: Analysis, synthesis, and perception. *J. Acoust. Soc. Am.*, 90, 1991.
- [4] R. J. B Macaulay M. Pomeranz and M. A. Caudill. Assessment of autonomic functions in humans by heart rate spectral analysis. *Am. J. Physiol.*, 248, 1985.
- [5] ESC/NASPE Task Force. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Eur. Heart. J.*, 17:354–381, 1996.
- [6] S. Furui. *Digital Speech Processing, Synthesis and Recognition*. Marcel Dekker, 1988.
- [7] L. W. Rabiner. *Digital Processing of Speech*. Prentice Hall, 1978.
- [8] A. E. Rosenberg. Effects of glottal pulse shape on the quality of natural vowels. *J. Acoust. Soc. Am.*, 49(2):582–590, 1971.
- [9] A. Birand, S. Saliu, and Z. G. Kudaiberdieva. Time-frequency analysis of heart rate variability: smoothed pseudo Wigner distribution. *Annals of Non-invasive Electrocardiology*, 1(4):411–418, October 1996.
- [10] B. M. Sayers. Analysis of heart rate variability. *Ergonomics*, 16:17–32, 1973.
- [11] M. Pagani, F. Lombardi, S. Guzzetti, and et al. Power spectral analysis of heart rate and blood pressure variabilities as a marker of sympathovagal interaction in man and conscious dog. *Circ. Res.*, 59:178–193, 1986.
- [12] A. Birand S. Saliu and G. Kudaiberdieva. A computer system for assessment of heart rate dynamics. In *Proc. ICSPAT*, pages 344–348, 1996.