Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 100 (2016) 271 – 279

Procedia Computer Science

Conference on ENTERprise Information Systems / International Conference on Project MANagement / Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN / HCist 2016, October 5-7, 2016

Algorithm for jitter and shimmer measurement in pathologic voices

João Paulo Teixeira*, André Gonçalves

Polytechnic Institute of Bragança, Campus de Sta. Apolónia, 5300-253 Bragança, Portugal

Abstract

An algorithm to measure the jitter (jitta, jitter, rap and ppq5) and shimmer (ShdB, Shim, apq3 and apq5) parameters was developed. These parameters can be used in an intelligent system to diagnose voice pathologies. The algorithm is sensitive to the fundamental frequency and determines the parameters based on maximum and minimum functions applied to each glottal period of the signal. The algorithm was previously tested with synthesized speech signals with very high accuracy, but several improvements had to be included for the analysis of pathologic voices. A comparison using real, control and pathologic voices was made between the developed algorithm and other software program to evaluate the consistency of measures between different algorithms. Based on this comparison it can be said that the algorithm improved the accuracy with synthetic speech signal and showed consistent measures for pathologic speech signal.

Peer-review under responsibility of the organizing committee of CENTERIS 2016

Keywords: Jitter; Shimmer; Voice analysis; Pathologic voice diagnose; Voice clinics.

1. Introduction

The present work has as long term objective the development of an automatic diagnosis of laryngeal pathologies system. From the voice signal it is possible to extract a set of parameters that combined in an intelligent way can give a strong help for diagnose of voice pathologies.

* Corresponding author. Tel.: +351 273303129; fax: +351 273303051.
E-mail address: joaopt@ipb.pt

Peer-review under responsibility of the organizing committee of CENTERIS 2016
doi:10.1016/j.procs.2016.09.155

From the medical point of view various techniques have been used to assess the patient's voice quality. The simpler consists in the auditory perceptual analysis. However, the auditory perceptual analysis may lead to different results depending on the experience of the doctor. Thus, this practice is subjective and leads to the lack of consensus among professionals. Therefore, it became necessary to search for an objective assessment, in which the voices are analyzed by devices which are capable of measuring several acoustic parameters.

The most common signs that may indicate changes in the larynx relate hoarseness, breathiness and roughness. The transient hoarseness may result from abuse of the voice or the casual flu. But when the hoarseness persists and becomes a characteristic voice, is indicative of pathology of the larynx. Hoarseness can also be an early symptom of cancer of the larynx, Teixeira, et al. [1]. The most common pathologies affecting voice are vocal nodules, the laryngitis, the paralysis, polyps, cysts and Reinke's Edema. Other pathologies of the larynx that may lead to dysphonic speech are ulcers of contact, as Lopes [2].

The parameters obtained by the acoustic analysis have the advantage of describing the voice objectively. With the existence of normative databases characterizing voice quality or using intelligent tools combining the various parameters, it is possible, using artificial intelligent tools, to distinguish between normal and pathological voices or even identify or suggest the pathology. These tools allow the monitoring of clinical standpoint and/or employment and reduce the degree of subjectivity of perceptual analysis, as Teixeira, et al. [1].

Currently, acoustic parameters commonly used in applications of acoustic analysis as well as the most referenced in the literature, are the fundamental frequency, jitter, shimmer and HNR. The fundamental frequency (F0), measured in Hertz, is defined as the number of times a sound wave produced by the vocal cords repeats during a given time period. It is also the number of cycles of opening/closure of the glottis. There is a typical range of values of this frequency for the different genders and ages. But these values are not stationary since F0 is also used to convey prosody. Besides, it also vary with sex and age, thought to depend on factors such as the state of mind of the person, the time of day that fit the lifestyle and professional use of voice [1].

Other parameters extracted from the glottal signal were also used recently in [3] besides jitter and shimmer. These new parameters are: closing phase, opening phase, open quotient, closed quotient, amplitude quotient, normalized amplitude quotient, quasi open quotient, speed quotient, difference between harmonics and harmonics richness factor.

Anyhow, the most commonly used parameters as part of a comprehensive voice examination are the voice frequency (jitter) and amplitude (shimmer) variations [4]. Jitter is the measure of the cycle-to-cycle period variations of the successive glottal cycles and shimmer is the cycle-to-cycle amplitude variations of the successive glottal cycles. These measures can be determined using absolute or relative values. All of these parameters have been largely used for description of pathological voice quality [5, 6]. Both parameters are obtained by analysis of a record speech of sustained vowel phonations [7-11].

The jitter is affected mainly by the lack of control of vibration of the cords. The voices of patients with pathologies often have higher values of jitter. The shimmer changes with the reduction of glottal resistance and mass lesions on the vocal cords and is correlated with the presence of noise emission and breathiness [7]. It is expected that patients with pathologies have higher values of shimmer. The articulatory oral space measures [12] may contribute with some new information to explain the higher jitter and shimmer values.

The aim of the work reported here is the accurate determination of the jitter and shimmer measures by the developed algorithm. This is based on a join of the maximum/minimum functions and moving energy over the speech signal that will find automatically the glottal pulses amplitude and position.

The quality of the algorithm was evaluated in previous publication [13], using synthetic speech with controlled values of jitter and shimmer analytically determined. Although the higher accuracy with synthetic speech, the real speech signal impose new challenges, because several differences may appear for different voice persons. For instance the peak searched to identify the position of the glottal pulse, can now be a positive or a negative pulse, or it can be positive in some part of the signal and negative in other parts of the same signal giving rise to a wrong measures of jitter. This paper describes the improvement to the algorithm to deal with this situation for real voices whether they be normal or pathologic voices.

The behavior of this algorithm was compared using control and pathological voice signals. Once this real speech signal have unknown values of jitter and shimmer, the measure obtained by the algorithm were compared with measures given by the Praat program just to evaluate the consistency of the measures in both algorithms.