Computers in Biology and Medicine 39 (2009) 97 – 105

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Objective characterization of oesophageal voice supporting medical diagnosis, rehabilitation and monitoring

B. García*, I. Ruiz, A. Méndez, M. Mendezona

Department of Telecommunication, Engineering Faculty (ESIDE), University of Deusto, Av. de las Universidades 24, 48007 Bilbao, Spain

ARTICLE INFO

Article history:
Received 22 July 2008
Accepted 24 November 2008

Keywords:
Medical informatics
Rehabilitation engineering
Signal processing
Oesophageal voice
Acoustical measurements
Speech disorder database

ABSTRACT

Otolaryngologists use computational tools in the objective diagnosis of vocal folds pathologies by means of a set of acoustic parameters among others. This can be achieved in the cases of slight pathologies, but it does not exist any commercial software suitable for severe degradations of speech, as they are the oesophageal voice of laryngectomees. The present article shows a high-accuracy algorithm for the detection of the periodicity cycles of both oesophageal and laryngeal voices with low quality which allows the accurate and automatic estimation of pitch, jitter and shimmer. As the proposed algorithm works also with slighter pathologies, it is a useful contribution which allows doctors to perform an objective control during rehabilitation and monitoring stages. Thus, a patient can be controlled during oesophageal voice learning stage and it can also be saved a medical record with the results of the acoustic parameters' measurements in order to detect possible relapses.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Nowadays, many otolaryngology (ORL) specialists take advance of software tools to support their vocal folds pathologies diagnosis by means of objective parameters. These tools complete the information gathered both in visual analysis of stroboscopic images and in several perceptual tests that ORL perform to their patients.

The cancer of the vocal folds needs to pay special attention in its diagnosis, treatment, rehabilitation and monitoring mainly because it can cause death. Once the cancer has been detected, the ORL arranges the vocal folds removal. This implies that patients in such situation will not be able to produce laryngeal voice and hence, they lose the speaking ability.

After the operation and during the rehabilitation, the patient will begin the learning stage of oesophageal speech: the voice produced due to the modulation of the air by means of the oesophagus. This will allow the patient to use oesophageal speech which has a degraded quality but it makes possible to maintain a fluid oral communication. However, the main difficulty is that is not possible to evaluate objectively oesophageal voices during the rehabilitation stage, because there is not any application in the market which can automatically obtain the acoustic parameters for this kind of voices. The quality of the oesophageal voices is such low that the voice periodicity algorithms do not obtain reliable measures.

The algorithms that allow the objective evaluation of oesophageal voices have two applications: in the short term just after the surgery and in the long term for monitoring purposes to prevent from possible relapses. In the monitoring stage, it is very useful to have objective measures in order to represent the evolution curves of those parameters. These graphs are a very clear indicator to see either the positive evolution or stability of the patient, or to detect a possible relapse.

Evidently, the accuracy of the algorithms presented will be also applicable in such cases of less severe pathologies (polyps, nodules, hipomobility of the folds, etc…) where the degradation of voice quality is so high that commercial software applications used by doctors become useless. In fact, it can be very useful in cases of early detection of the cancer where the laryngeal voice of the patients can have a very poor intelligibility and an increased level of noise [1].

The pitch [2] defined as (1) is the property of a sound or musical tone measured by its perceived frequency [2]. Due to de pseudo-periodic nature of the voiced speech, there are variations in the instantaneous frequency f_i so the pitch can be defined as

$$\text{Pitch(Hz)} = \frac{\sum_{i=1}^{N} f_i}{N} \tag{1}$$

being N the number of extracted pitch periods.

Fundamental frequency estimation has consistently been a difficult topic in audio signal processing because is so difficult to define the time instants which define the voice cycles used to obtain their related instantaneous frequency, f_i.

* Corresponding author. Tel.: +34 94139000x2798; fax: +34 94139101.
E-mail address: mbgarcia@eside.deusto.es (B. García).

0010-4825/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2008.11.009

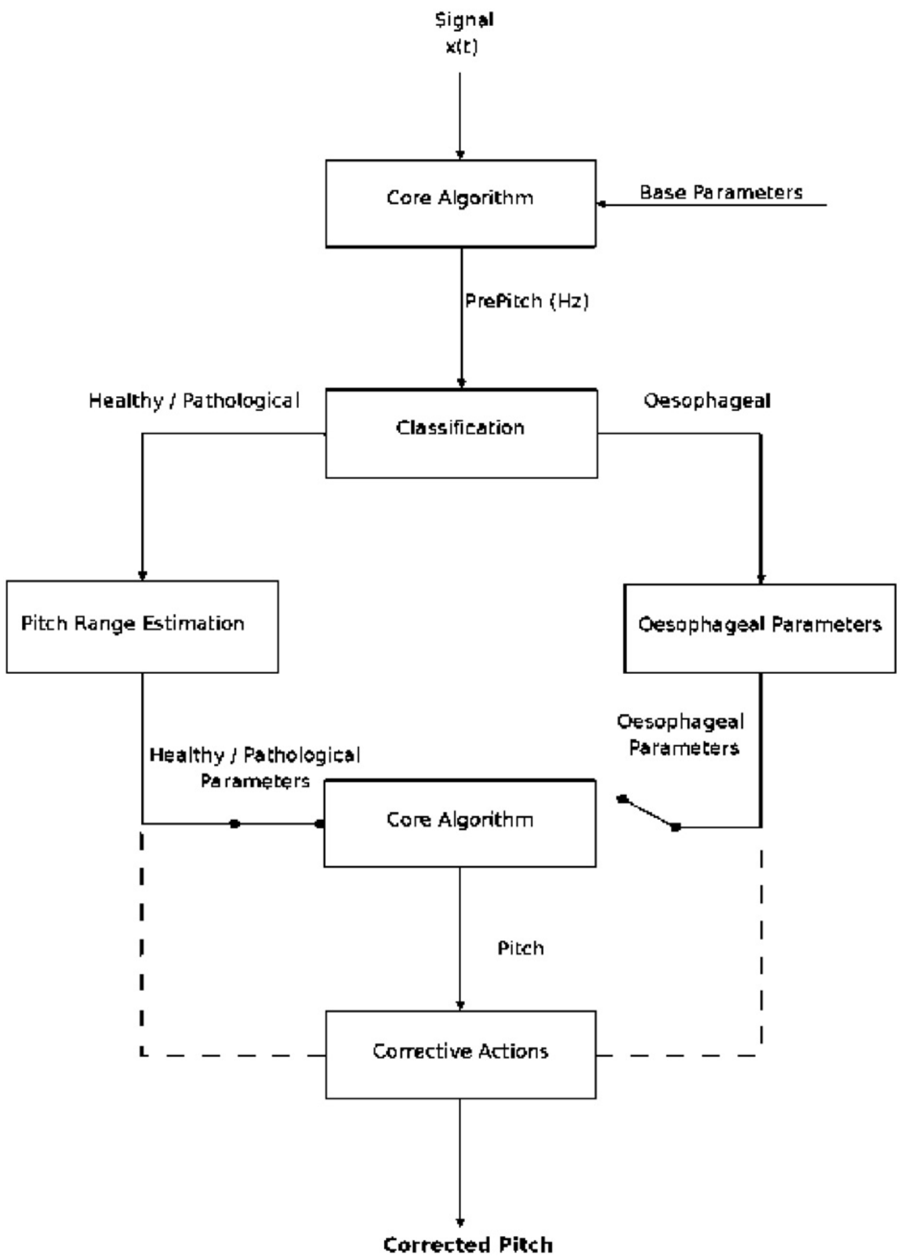

Fig. 1. Block diagram of the algorithm.

Furthermore, in acoustical parameterization it is of capital importance to calculate those instants because they are basic features used in this kind of characterization.

In one hand, jitter is a parameter that represents the variation of the fundamental frequency as it is defined in [2] and, in the other hand, specialists also use the reference of shimmer [2] which is the parameter that represents the amplitude perturbation of the voice signal. The voice produced in vocal folds is supposed to have the ability to maintain its amplitude almost constant, thus an increased value of shimmer may imply a symptom of a voice disorder.

Several authors have reported decent results using voice cycle detection [3,4] and there are many techniques widely detailed in literature: time domain estimators (e.g. zero crossing rate [5]), fundamental frequency estimators [6,7], autocorrelation methods (Yin estimator, [8]), phase space representation [9], cepstrum [10] and statistical methods [11–13]. Some of them define directly the voice cycles [3] while others are used to calculate a numerical approximation [8] to the fundamental frequency value. In these ones, a further step is necessary in order to identify clearly which instants define the voice cycles.

However the results of these works are, none of them have been even tested with oesophageal voices, thus it can be assured that they are not suitable to work with such kind of noisy signals.

This article presents a proposal for medical usage on ORL which consists on an algorithm which has been specifically designed to work with oesophageal voices in order to perform an objective evaluation. For that, the presented algorithm establishes the basic procedure to calculate the acoustical parameters related to speech periodicity both in diagnosis, rehabilitation and monitoring of the patient.

It has been included in a user-friendly software application so that the ORL can take advantage of the algorithm and either measure the parameters or save historical records of pitch, jitter and shimmer of their patients.

Download English Version:

https://daneshyari.com/en/article/505784

Download Persian Version:

https://daneshyari.com/article/505784

Daneshyari.com