

Ana I. Pereira · Florbela P. Fernandes ·
João P. Coelho · João P. Teixeira ·
Maria F. Pacheco · Paulo Alves ·
Rui P. Lopes (Eds.)

Optimization, Learning Algorithms and Applications

First International Conference, OL2A 2021
Bragança, Portugal, July 19–21, 2021
Revised Selected Papers

Editors

Ana I. Pereira 
Instituto Politécnico de Bragança
Bragança, Portugal

João P. Coelho 
Instituto Politécnico de Bragança
Bragança, Portugal

Maria F. Pacheco 
Instituto Politécnico de Bragança
Bragança, Portugal

Rui P. Lopes 
Instituto Politécnico de Bragança
Bragança, Portugal

Florbela P. Fernandes 
Instituto Politécnico de Bragança
Bragança, Portugal

João P. Teixeira 
Instituto Politécnico de Bragança
Bragança, Portugal

Paulo Alves 
Instituto Politécnico de Bragança
Bragança, Portugal

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-91884-2 ISBN 978-3-030-91885-9 (eBook)
<https://doi.org/10.1007/978-3-030-91885-9>

© Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The volume CCIS 1488 contains the refereed proceedings of the International Conference on Optimization, Learning Algorithms and Applications (OL2A 2021), an event that, due to the COVID-19 pandemic, was held online.

OL2A 2021 provided a space for the research community on optimization and learning to get together and share the latest developments, trends, and techniques as well as develop new paths and collaborations. OL2A 2021 had more than 400 participants in an online environment throughout the three days of the conference (July 19–21, 2021), discussing topics associated to areas such as optimization and learning and state-of-the-art applications related to multi-objective optimization, optimization for machine learning, robotics, health informatics, data analysis, optimization and learning under uncertainty, and the Fourth Industrial Revolution.

Four special sessions were organized under the following topics: Trends in Engineering Education, Optimization in Control Systems Design, Data Visualization and Virtual Reality, and Measurements with the Internet of Things. The event had 52 accepted papers, among which 39 were full papers. All papers were carefully reviewed and selected from 134 submissions. All the reviews were carefully carried out by a Scientific Committee of 61 PhD researchers from 18 countries.

July 2021

Ana I. Pereira

Organization

General Chair

Ana Isabel Pereira

Polytechnic Institute of Bragança, Portugal

Organizing Committee Chairs

Florbela P. Fernandes

Polytechnic Institute of Bragança, Portugal

João Paulo Coelho

Polytechnic Institute of Bragança, Portugal

João Paulo Teixeira

Polytechnic Institute of Bragança, Portugal

M. Fátima Pacheco

Polytechnic Institute of Bragança, Portugal

Paulo Alves

Polytechnic Institute of Bragança, Portugal

Rui Pedro Lopes

Polytechnic Institute of Bragança, Portugal

Scientific Committee

Ana Maria A. C. Rocha

University of Minho, Portugal

Ana Paula Teixeira

University of Trás-os-Montes and Alto Douro, Portugal

André Pinz Borges

Federal University of Technology – Paraná, Brazil

Andrej Košir

University of Ljubljana, Slovenia

Arnaldo Cândido Júnior

Federal University of Technology – Paraná, Brazil

Bruno Bispo

Federal University of Santa Catarina, Brazil

Carmen Galé

University of Zaragoza, Spain

B. Rajesh Kanna

Vellore Institute of Technology, India

C. Sweetlin Hemalatha

Vellore Institute of Technology, India

Damir Vrančić

Jozef Stefan Institute, Slovenia

Daiva Petkeviciute

Kaunas University of Technology, Lithuania

Diamantino Silva Freitas

University of Porto, Portugal

Esteban Clua

Federal Fluminense University, Brazil

Eric Rogers

University of Southampton, UK

Felipe Nascimento Martins

Hanze University of Applied Sciences,
The Netherlands

Gaukhar Muratova

Dulaty University, Kazakhstan

Gediminas Daukšys

Kauno Technikos Kolegija, Lithuania

Glaucia Maria Bressan

Federal University of Technology – Paraná, Brazil

Humberto Rocha

University of Coimbra, Portugal

José Boaventura-Cunha

University of Trás-os-Montes and Alto Douro, Portugal

José Lima

Polytechnic Institute of Bragança, Portugal

Joseane Pontes

Federal University of Technology – Ponta Grossa,
Brazil

Juani Lopéz Redondo

University of Almeria, Spain

Jorge Ribeiro	Polytechnic Institute of Viana do Castelo, Portugal
José Ramos	NOVA University Lisbon, Portugal
Kristina Sutiene	Kaunas University of Technology, Lithuania
Lidia Sánchez	University of León, Spain
Lino Costa	University of Minho, Portugal
Luís Coelho	Polytechnic Institute of Porto, Portugal
Luca Spalazzi	Marche Polytechnic University, Italy
Manuel Castejón Limas	University of León, Spain
Marc Jungers	Université de Lorraine, France
Maria do Rosário de Pinho	University of Porto, Portugal
Marco Aurélio Wehrmeister	Federal University of Technology – Paraná, Brazil
Mikulas Huba	Slovak University of Technology in Bratislava, Slovakia
Michał Podpora	Opole University of Technology, Poland
Miguel Ángel Prada	University of León, Spain
Nicolae Cleju	Technical University of Iasi, Romania
Paulo Lopes dos Santos	University of Porto, Portugal
Paulo Moura Oliveira	University of Trás-os-Montes and Alto Douro, Portugal
Pavel Pakshin	Nizhny Novgorod State Technical University, Russia
Pedro Luiz de Paula Filho	Federal University of Technology – Paraná, Brazil
Pedro Miguel Rodrigues	Catholic University of Portugal, Portugal
Pedro Morais	Polytechnic Institute of Cávado e Ave, Portugal
Pedro Pinto	Polytechnic Institute of Viana do Castelo, Portugal
Rudolf Rabenstein	Friedrich-Alexander-University of Erlangen-Nürnberg, Germany
Sani Rutz da Silva	Federal University of Technology – Paraná, Brazil
Sara Paiva	Polytechnic Institute of Viana do Castelo, Portugal
Sofia Rodrigues	Polytechnic Institute of Viana do Castelo, Portugal
Sławomir Stępień	Poznan University of Technology, Poland
Teresa Paula Perdicoulis	University of Trás-os-Montes and Alto Douro, Portugal
Toma Roncevic	University of Split, Croatia
Vitor Duarte dos Santos	NOVA University Lisbon, Portugal
Wojciech Paszke	University of Zielona Gora, Poland
Wojciech Giernacki	Poznan University of Technology, Poland

Optimization of Glottal Onset Peak Detection Algorithm for Accurate Jitter Measurement	123
<i>Joana Fernandes, Pedro Henrique Borghi, Diamantino Silva Freitas, and João Paulo Teixeira</i>	
Searching the Optimal Parameters of a 3D Scanner Through Particle Swarm Optimization	138
<i>João Braun, José Lima, Ana I. Pereira, Cláudia Rocha, and Paulo Costa</i>	
Optimal Sizing of a Hybrid Energy System Based on Renewable Energy Using Evolutionary Optimization Algorithms	153
<i>Yahia Amoura, Ângela P. Ferreira, José Lima, and Ana I. Pereira</i>	
Robotics	
Human Detector Smart Sensor for Autonomous Disinfection Mobile Robot	171
<i>Hugo Mendonça, José Lima, Paulo Costa, António Paulo Moreira, and Filipe Santos</i>	
Multiple Mobile Robots Scheduling Based on Simulated Annealing Algorithm	187
<i>Diogo Matos, Pedro Costa, José Lima, and António Valente</i>	
Multi AGV Industrial Supervisory System	203
<i>Ana Cruz, Diogo Matos, José Lima, Paulo Costa, and Pedro Costa</i>	
Dual Coulomb Counting Extended Kalman Filter for Battery SOC Determination	219
<i>Arezki A. Chellal, José Lima, José Gonçalves, and Hicham Megnafi</i>	
Sensor Fusion for Mobile Robot Localization Using Extended Kalman Filter, UWB ToF and ArUco Markers	235
<i>Sílvia Faria, José Lima, and Paulo Costa</i>	
Deep Reinforcement Learning Applied to a Robotic Pick-and-Place Application	251
<i>Natanael Magno Gomes, Felipe N. Martins, José Lima, and Heinrich Wörtche</i>	
Measurements with the Internet of Things	
An IoT Approach for Animals Tracking	269
<i>Matheus Zorawski, Thadeu Brito, José Castro, João Paulo Castro, Marina Castro, and José Lima</i>	



Optimization of Glottal Onset Peak Detection Algorithm for Accurate Jitter Measurement

Joana Fernandes^{1,2} , Pedro Henrique Borghi^{3,4} , Diamantino Silva Freitas²,
and João Paulo Teixeira⁵

¹ Research Centre in Digitalization and Intelligent Robotics (CeDRI),
Instituto Politecnico de Braganca (IPB), 5300 Braganca, Portugal

`joana.fernandes@ipb.pt`

² Faculdade de Engenharia da Universidade do Porto (FEUP),
4200-465 Porto, Portugal

`dfreitas@ipb.pt`

³ Instituto Politecnico de Braganca (IPB), 5300 Braganca, Portugal

⁴ Federal University of Technology - Parana (UTFPR),
Cornelio Procopio 86300-000, Brazil

`pedromelo@alunos.utfpr.edu.br`

⁵ Research Centre in Digitalization and Intelligent Robotics (CeDRI),
Applied Management Research Unit (UNIAG), Instituto Politecnico
de Braganca (IPB), 5300 Braganca, Portugal

`joaopt@ipb.pt`

Abstract. Jitter is an acoustic parameter used as input for intelligent systems for the diagnosis of speech related pathologies. This work has the objective to improve an algorithm that allows to extract vocal parameters, and thus improve the accuracy measurement of absolute jitter parameter. Some signals were analyzed, where signal to signal was compared in order to try to understand why the values are different in some signal between the original algorithm and the reference software. In this way, some problems were found that allowed to adjust the algorithm, and improve the measurement accuracy for those signals. Subsequently, a comparative analysis was performed between the values of the original algorithm, the adjusted algorithm and the Praat software (assumed as reference). By comparing the results, it was concluded that the adjusted algorithm allows the extraction of the absolute jitter with values closer to the reference values for several speech signals. For the analysis, sustained vowels of control and pathological subjects were used.

Keywords: Jitter · Algorithm · Optimization · Speech pathologies · Acoustic analysis

This work was supported by Fundação para a Ciência e Tecnologia within the Project Scope: UIDB/05757/2020.

© Springer Nature Switzerland AG 2021

A. I. Pereira et al. (Eds.): OL2A 2021, CCIS 1488, pp. 123–137, 2021.

https://doi.org/10.1007/978-3-030-91885-9_10

1 Introduction

Speech pathologies are relatively common and can be found in different stages of evolution and severity, affecting approximately 10% of the population [1]. These pathologies directly affect vocal quality, as they alter the phonation process and have increased dramatically in recent times, mainly due to harmful habits, such as smoking, excessive consumption of alcoholic beverages, persistent inhalation of dust-contaminated air and abuse of voice [2].

There are a variety of tests that can be performed to detect pathologies associated with the voice, however, they are invasive becoming uncomfortable for patients and are time-intensive [3].

Auditory acoustic analyzes, performed by professionals, lack objectivity and depend on the experience of the physician who makes the assessment. Acoustic analysis allows non-invasively to determine the individual's vocal quality. It is a technique widely used in the detection and study of voice pathologies, since it allows to measure properties of the acoustic signal of a recorded voice where a speech or vowels are said in a sustained way [4,5]. This analysis is able to provide the sound wave format allowing the evaluation of certain characteristics such as frequency disturbance measurements, amplitude disturbance measurements and noise parameters.

Jitter is one of the most used parameters as part of a voice exam and is used by several authors Kadiri and Alku 2020 [6], Teixeira et al. 2018 [7], Sripriya et al. 2017 [8], Teixeira and Fernandes 2015 [9] to determine voice pathologies.

Jitter is the measure of the cycle to cycle variations of the successive glottic cycles, being possible to measure in absolute or relative values. Jitter is mainly affected by the lack of control in the vibration of the vocal folds. Normally, the voices of patients with pathologies tend to have higher values of jitter [10].

This work intends to improve the algorithm developed by Teixeira and Gonçalves [10,11] to obtain the jitter parameter, to later be used in a complementary diagnostic system and obtain reliable jitter values. As a means of reference, to ensure reliable values of the algorithm, Praat is used as a reference for comparison, as this software is accepted by the scientific community as an accurate measure and is open software. This software is developed by Paul Boersma and David Weenink [12], from the Institute of Phonetic Sciences at the University of Amsterdam.

This article is organized as follows: Sect. 2 describes the determination of the absolute jitter, the pathologies and the database used, as well as the number of subjects used for the study; Sect. 3 describes some of the problems found in some signals in the Teixeira and Gonçalves algorithms [10,11], as well as the description and flowchart of the adjusted algorithm; Sect. 4 describes the studies carried out, the results obtained and the discussion. Finally, in Sect. 5 the conclusions are presented.

2 Methodology

In Fig. 1, the jitter concept is illustrated, where it is possible to perceive that jitter corresponds to the measure of the variation of the duration of the global periods.

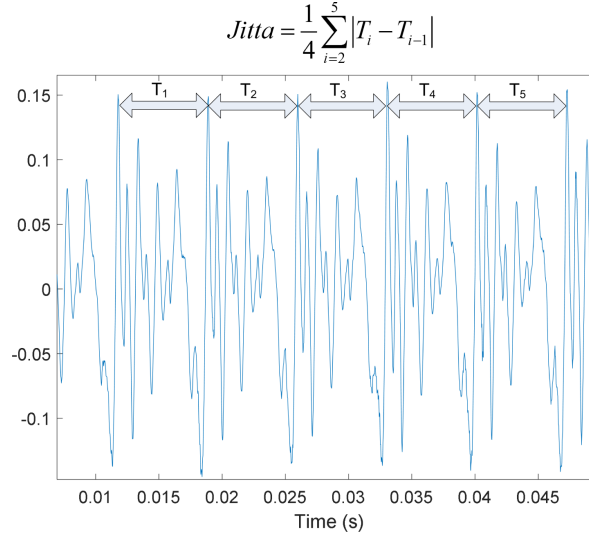


Fig. 1. Jitter representation in a sustained vowel /a/.

2.1 Jitter

Jitter is defined as a measure of glottal variation between cycles of vibration of the vocal cords. Subjects who cannot control the vibration of the vocal cords tend to have higher values of jitter [11].

Absolute jitter (jitta) is the variation of the glottal period between cycles, that is, the average absolute difference between consecutive periods, expressed by Eq. 1.

$$jitta = \frac{1}{N-1} \sum_{i=2}^N |T_i - T_{i-1}| \quad (1)$$

In Eq. 1 T_i is the size of the glottal period i and N is the total number of glottal periods.

2.2 Database

The German Saarbrücken Voice Database (SVD) was used. This database is available online by the Institute of Phonetics at the University of Saarland [13]. The database consists of voice signals of more than 2000 subjects with several disease and controls/healthy subjects. Each person has the recording of phonemes /a/, /i/ and /u/ in the low, normal and high tones, swipe along tones, and the

German phrase “Guten Morgen, wie geht es Ihnen?” (“Good morning, how are you?”). The size of the sound files is between 1 and 3 s and have a sampling frequency of 50 kHz.

For the analysis it was used a sub-sample of 10 control subjects (5 male and 5 female) and 5 patient subjects (2 male and 3 female), of 3 diseases. Of these, 2 subjects had chronic laryngitis, one of each gender, 2 subjects with vocal cord paralysis, one of each gender and 1 female with vocal cord polyps.

3 Development

3.1 Comparative Analysis Between the Signals Themselves

Teixeira and Gonçalves in [11] reported the results of the algorithm using groups of 9 healthy and 9 pathological male and female subject. The authors compared the average measures of the algorithm and Praat software, over the groups of 9 subjects. The comparison shows generally slightly higher values measured by the algorithm, but with difference lower than $20\mu\text{s}$ except for the female pathological group. $20\mu\text{s}$ corresponds to only one speech signal sample at 50 kHz sampling frequency. Authors claim a difference lower than 1 speech sample in average. Anyhow, in a particular speech file it was found very high differences between the algorithm measures and the Praat measures.

In order to try to understand why the jitter measure in some particular speech files of the Teixeira and Gonçalves 2016 [11] algorithm are so different from the reference values, we proceeded to a visual comparison of the signals in the reference software and in the Teixeira and Gonçalves 2016 algorithm.

In this way, we took the signals in which the values for absolute jitter were very different and it was noticed that for a particular shapes of the signal the algorithm do not find the peaks in a similar way as Praat did.

In a female control signal, the reference value for absolute jitter was $10.7\mu\text{s}$ and the algorithm value was $478.6\mu\text{s}$. When the signal was observed, it was realized that one of the reasons that led to this difference was the fact that the algorithm was selecting the maximum peaks to make the measurement, while the reference software was using the minimum peaks. This led us to try to understand why the algorithm is selecting the maximum peaks, instead of selecting the minimums. The conclusion observed in Fig. 2 was reached.

As it can be seen in Fig. 2, the algorithm is finding two maximum values for the same peak. This first search of peaks is taken in a initial phase of the algorithm to decide to use the maximum or minimum peaks. This two close peaks leads the algorithm to erroneously choose the peaks it will use to make the measurement, since it does not satisfy the condition of the algorithm (number of maximum peaks = 10, in a 10 glottal periods length part of speech). This problem occurs in more signals. Correcting this error, for this signal, the absolute jitter value went from $478.6\mu\text{s}$ to $39.3\mu\text{s}$.

In another signal, also from a control subject, in which the reference value for the absolute jitter is $9.6\mu\text{s}$ and the value of the algorithm was $203.1\mu\text{s}$.

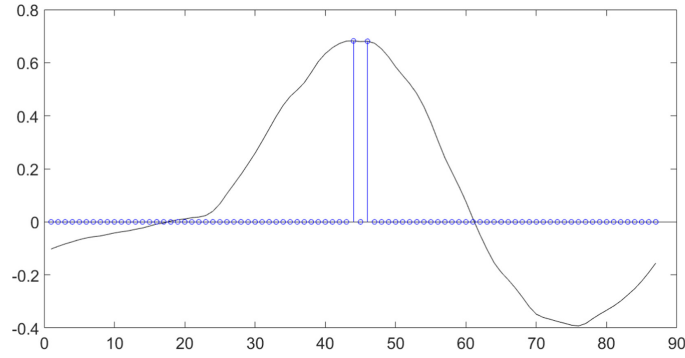


Fig. 2. Problem 1 - finding peaks in a short window for decision for use the maximum or minimum peaks.

When observing the signal, it was also noticed that the algorithm was using the maximum peaks to make the measurement, while the reference software used the minimum peaks. Looking at Fig. 3 (presenting the speech signal and the identification of the right minimum peaks), it is possible to see why that algorithm selected the maximum peaks instead of the minimum peaks.

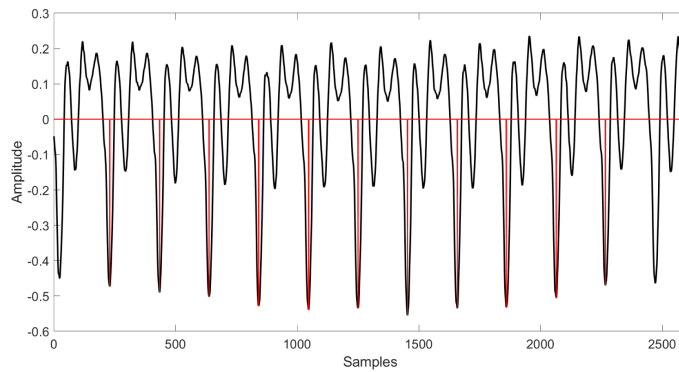


Fig. 3. Problem 2 - selection of maximum/minimum peaks (in red the position of the minimum peaks). (Color figure online)

As it is possible to observe in Fig. 3 the length of the window used, defined in global periods, comprises 11 minimum peaks. Taking into account the conditions of the algorithm to select the maximum or minimum peaks, for later measurement, the fact of having more than 10 minimum peaks leads the algorithm to select the maximum. When the window correction was made, for this signal in question, it started to select only 10 minimum peaks and the algorithm starts to select the minimum peaks for measurement. Thus, the jitter absolute value went from $203.1 \mu\text{s}$ to $32.3 \mu\text{s}$.

In two pathological signals, one female and one male, in which the reference value for absolute jitter is $8.0 \mu\text{s}$ and the value of the algorithm is $68.6 \mu\text{s}$ for female signal. For the male, the absolute jitter in reference software is $28.7 \mu\text{s}$ and that of the algorithm is $108.0 \mu\text{s}$. When the signal of each of them was observed,

it was noticed that the choice between selecting the maximum/minimum was well done, however, this type of signal have two peaks very close in same period that alternatively have the minimum/maximum value. Once the algorithm takes always the minimum/maximum value, the selection of the first and second peak changes along the signal, making the measure of the jitter inconsistent.

In Figs. 4 and 5 it is possible to observe the problems found in these two signals. These figures present the speech signal in blue line, the moving average in red line and the marked peaks with red circle.

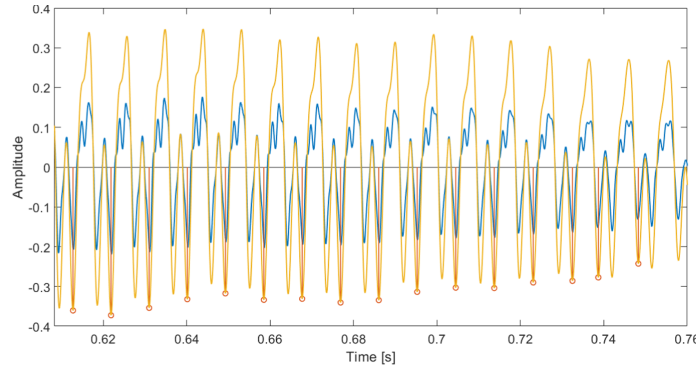


Fig. 4. Problem 3 - variation along the signal of the moving average minimum peak. (Color figure online)

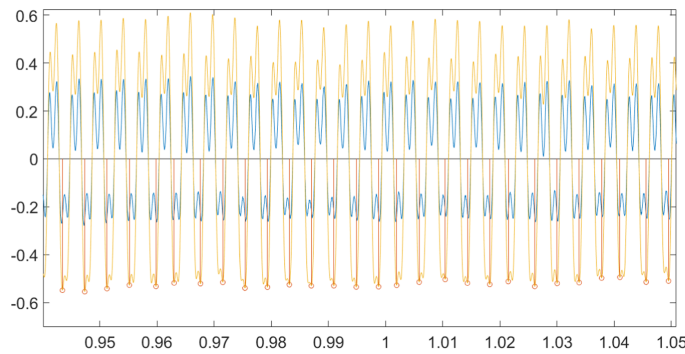


Fig. 5. Problem 4 - variation along the signal of the signal minimum peak. (Color figure online)

As it is possible to observe in Figs. 4 and 5 the fact that there are 2 peaks (minimum peaks in these examples) in the same glottal period, sometimes the maximum peak is the first, other times it is the second, which leads to increase the absolute jitter determined.

3.2 Algorithm Adjustments

The main purpose of the present work is to present the optimizations applied to the algorithm proposed in [11], since it was mainly observed the inability to deal with some recordings, both in the control group and those who have some type of pathology. Thus, a recap of the steps in [11] is proposed in order to contextualize each improvement developed. It must be emphasised the importance of the exact identification of the position of each onset time of the glottal periods, for an accurate measure of the Jitter values. The onset time of the glottal periods are considered as the position of the peaks in the signal.

According to [11], the process of choosing the references for calculating the Jitter parameter can be better performed by analyzing the signal from its central region. There it is expected to find high stationarity. In addition, it was concluded that a window of 10 fundamental periods contains enough information to determine the reference of the length of the fundamental period and the selection between the minimums or maximums. On this window, the moments that occur the positive and negative peaks of each glottal period are determined, as well as the amplitude module of the first cycle. The step between adjacent cycles is based on the peak position of the previous cycle plus a fundamental period. Focusing on this point, the position of the posterior peak is determined by searching for the maximum (or minimum) over a two-third window of the fundamental period. Finally, in an interval of two fifths of fundamental period, around each peak, the search for other peaks corresponding to at least 70% of its amplitude is carried out. The reference determination is made considering the positive and negative amplitudes of the first module cycle and the count of positive and negative peaks in the interval of 10 cycles. Basically, the measure of amplitude directs the decision to the parameter with the greatest prominence, since it is expected to have isolated peaks for the one with the greatest amplitude. On the other hand, the count of cycles indicates the number of points and oscillations close to the peaks, which may eventually prevail over these. The occurrence of these behaviors may or may not be something characteristic of the signal, in a way that it was observed that for cases in which this rarely occurred, a slight smoothing should be sufficient to readjust the signal. In [11], when the peak count exceeded the limit of 10 for both searches, a strong smoothing was applied over the signal that sometimes de-characterized it.

Based on the mentioned amplitude parameters and the peaks count, a decision-making process is carried out, which concludes with the use of maximums or minimums of the original signal or its moving average, to determine all glottal periods. In Fig. 6 the flowchart of the proposed algorithm is presented.

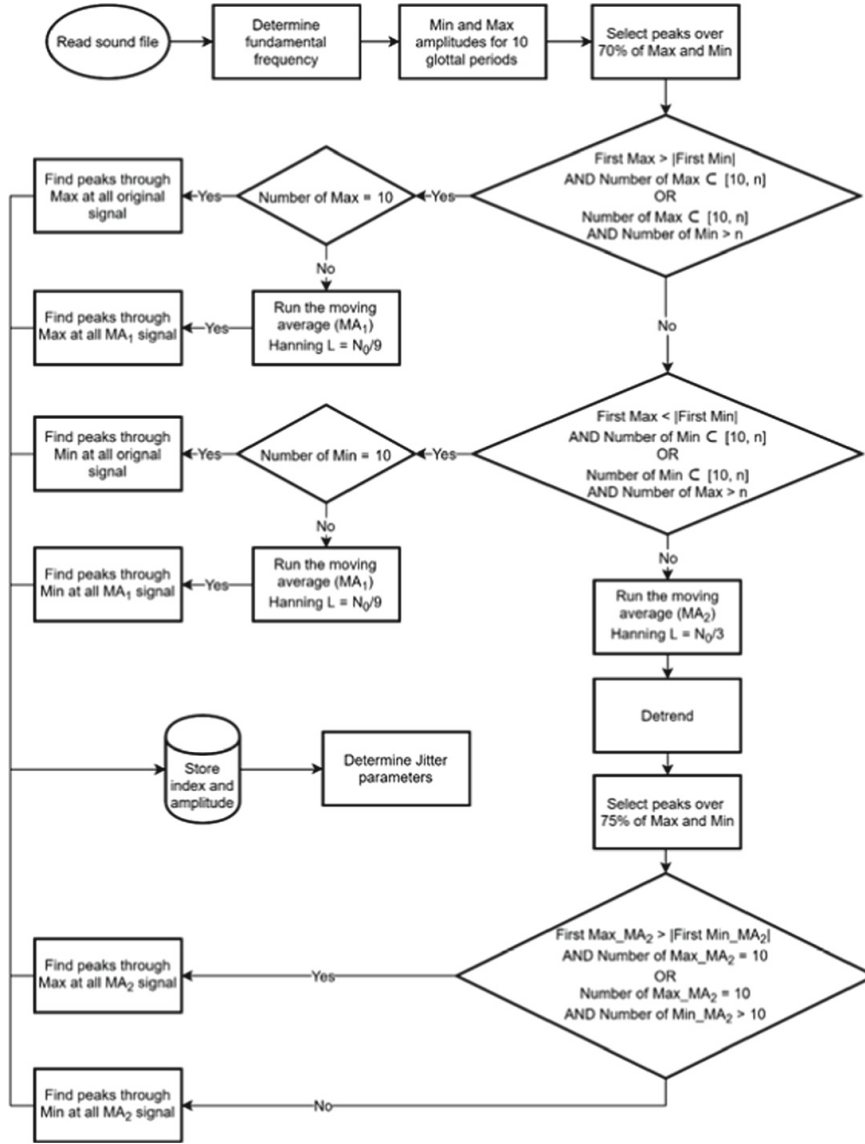


Fig. 6. Algorithm flowchart.

The rules of the decision process are described below, where the limit parameter $n = 13$ is defined as the maximum peak count accepted for the use of the original signal or for the application of light straightening.

- **If** the maximum amplitude is greater than the minimum amplitude in module and the maximum peak count is in the range $[10, n]$, or, the minimum peak count is greater than n and the maximum peak count is in $[10, n]$:
 - **If** the maximum count is equal to 10, the search for peaks is performed on the original signal using the maximum as a cycle-by-cycle search method;
 - **Else**, a moving average with a Hanning window of equal length to a ninth of a fundamental period is applied to the signal and the search for peaks in this signal is carried out through the maximum of each cycle. Figure 7 shows the comparison between the original signals, the moving

average MA_1 and the moving average MA_2 . The peak marking was performed within this rule, that is, starting from MA_1 and using maximum method. It is noted that the slight smoothing applied keeps the peak marking close to the real one and maintains, in a way, the most significant undulations of the signal, whereas MA_2 , in addition to shifting the position of the peaks more considerably, their undulations are composed of the contribution of several adjacent waves. This can reduce the reliability of marking signals that have peaks' side components with high amplitude in modulus.

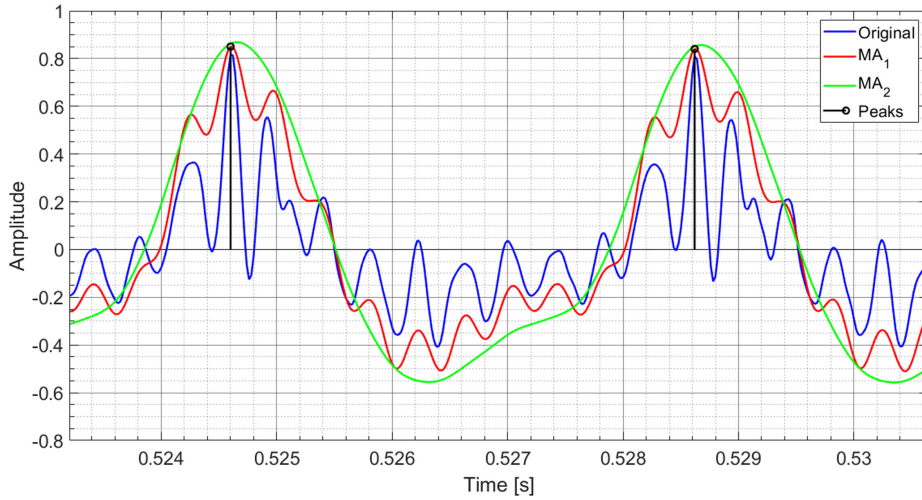


Fig. 7. Comparison between the original signal, the signal after moving average with Hanning window of length $N_0/9$ (MA_1) and the signal after moving average with Hanning window of length $N_0/3$ (MA_2). The peaks of glotal periods are shown in black and were detected through maximum method over MA_1 . Recording excerpt from a healthy subject performing /i/ low. Normalized representation.

- **Else**, if the minimum amplitude in module is greater than the maximum amplitude and the minimum peak count is in the range $[10, n]$, or, the maximum peak count is greater than n and the minimum peak count is in $[10, n]$:
 - **If** the minimum count is equal to 10, the search for peaks is performed on the original signal using the minimum as a cycle-by-cycle search method;
 - **Else**, a moving average with a Hanning window of equal length to a ninth of a fundamental period is applied to the signal and the search for peaks in this signal is carried out through the minimum of each cycle. Figure 8 shows an example of the path taken through this rule. Here, the generated MA_1 signal proves to be much less sensitive than the MA_2 (generated for comparison) since in this one, the wave that contains the real peak is strongly attenuated by the contribution of the previous positive amplitude to the averaging. In this way, the application of smoothing

contributes to eliminates rapid oscillations in the samples close to the minimum, but preserves the amplitude behavior of signals that do not need a strong smoothing.

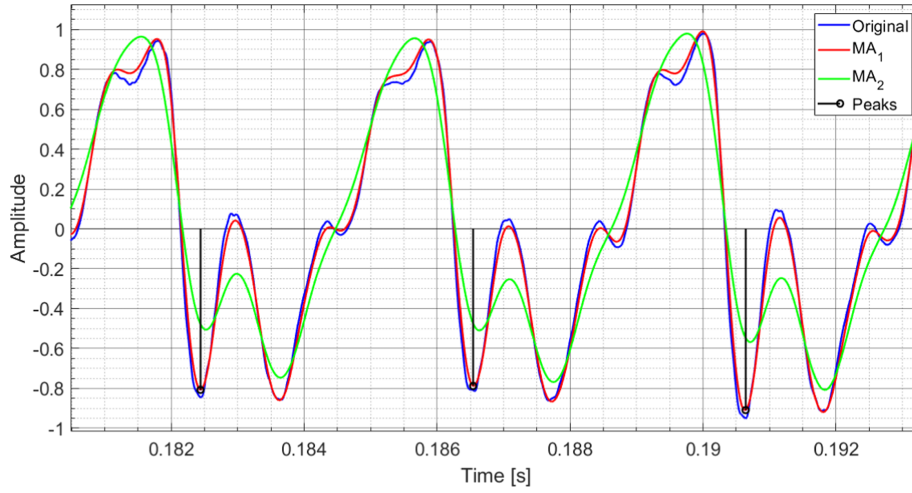


Fig. 8. Comparison between the original signal, the signal after moving average with Hanning window of length $N_0/9$ (MA_1) and the signal after moving average with Hanning window of length $N_0/3$ (MA_2). The peaks of glotal periods are shown in black and were detected through minimum method over MA_1 . Recording excerpt from a pathological subject performing /a/ high. Normalized representation.

- **Else**, a moving average with a Hanning window of equal length to a third of the fundamental period is applied to the signal. On the result, linear trends are removed and a new analysis is performed in the central region of the signal. In this analysis, an interval of 10 fundamental periods is evaluated cycle by cycle about the maximum, minimum peaks and their adjacent peaks greater than 75% of their amplitude:
 - **If**, in the first cycle, the maximum amplitude is greater than the minimum amplitude in module, and the maximum count is equal to 10 or, the minimum count is greater than 10 and the maximum count is equal to 10, the search for peaks is made on the moving average signal using the maximum as a method of inspection cycle by cycle. Figure 9 demonstrates the case in which this rule is followed. It is possible to observe, by comparing the original signal, MA_1 and MA_2 , that the determination of the peaks of the global periods is only possible when a heavy smoothing is applied, since the region with the greatest positive amplitude and energy, can only be highlighted with a single peak approaching a context long enough to avoid the two adjacent peaks in the original signal. It is also noted that MA_1 reduces the high frequencies but keeps the maximum point undefined, varying between the left and right along the signal;

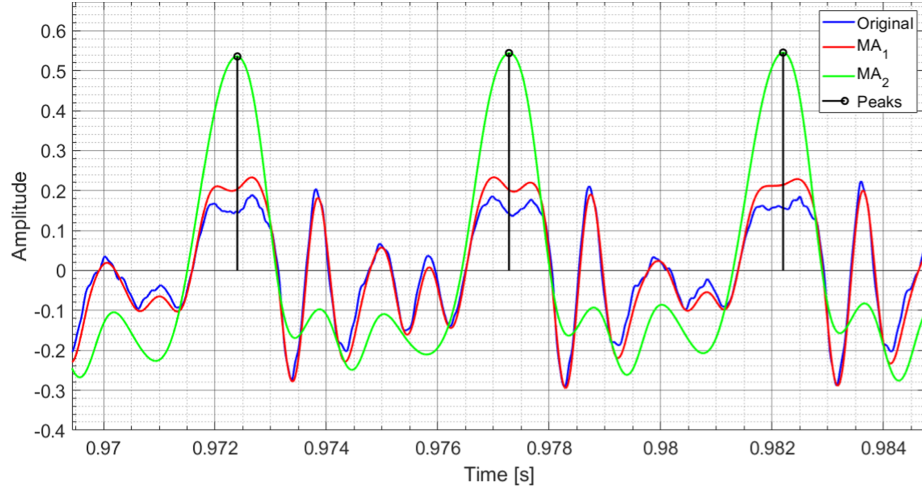


Fig. 9. Comparison between the original signal, the signal after moving average with Hanning window of length $N_0/9$ (MA_1) and the signal after moving average with Hanning window of length $N_0/3$ (MA_2). The peaks of glotal periods are shown in black and were detected through maximum method over MA_2 . Recording excerpt from a pathological subject performing /a/ normal. Normalized representation.

- **Else**, the search for peaks is done on the moving average signal using the minimum as a method of inspection cycle by cycle.

In all cases, the inspection of the peaks cycle by cycle of every signal is carried out with steps of a fundamental period from the reference point (maximum or minimum) and then an interval of one third of the fundamental period is analyzed. The reduction in the length of this interval from two thirds (used in [11]) to one third makes the region of analysis more restricted to samples close to one step from the previous peak. With this it is expected to guarantee that the analysis is always carried out in the same region of the cycles, avoiding the exchange between adjacent major peaks. However, in the cases where there is a shift in the fundamental frequency over time, the use of a fixed parameter as a step between cycles, as done in [11] and in this work, can cause inevitable errors in the definition of the search interval. The maximum or minimum points of each cycle are used as a source for determining the parameters Jitter according to Eqs. 1.

4 Results and Discussion

In this section, the results of some analyzes made will be reported, as well as their discussion.

A comparative analysis was made for the values of absolute jitter between the values obtained by the algorithm developed by Teixeira and Gonçalves 2016 [11], the reference software [12] and the adjusted algorithm, in order to try to understand if the adjusted algorithm values are closer to the values of the

reference software. Praat software was used as a reference, although the exact values of jitter cannot be known.

For 10 control subjects and 5 pathological subjects, the absolute jitter was extracted and averaged by tone and vowel. Table 1 shows the results of this analysis for control subjects and Table 2 for pathological subjects.

Table 1. Average absolute jitter (in microseconds) for each vowel and tone for control subjects.

Vowel	Tone	Praat	Teixeira and Gonçalves 2016	Adjusted
a	High	14.673	26.795	25.890
	Low	27.614	80.084	34.105
	Normal	24.916	50.091	32.016
i	High	16.092	23.621	23.617
	Low	23.805	26.246	26.378
	Normal	18.593	23.909	23.900
u	High	13.361	23.422	23.857
	Low	29.077	60.089	37.140
	Normal	18.376	23.759	23.755

Table 2. Average absolute jitter (in microseconds) for each vowel and tone for pathological subjects.

Vowel	Tone	Praat	Teixeira and Gonçalves 2016	Adjusted
a	High	32.817	56.447	56.379
	Low	39.341	82.316	82.316
	Normal	36.462	139.866	103.326
i	High	39.774	62.280	62.284
	Low	35.512	48.347	48.347
	Normal	25.611	33.921	33.916
u	High	33.606	45.489	45.489
	Low	40.751	65.752	65.745
	Normal	52.168	323.279	103.322

Through the data in Tables 1 and 2 it is possible to see that the adjusted algorithm obtains jitter measures closer to the reference than the Teixeira and Gonçalves 2016 algorithm, since, for example, in the vowel /a/ low tone, for the control subjects, with the algorithm obtained an improvement of 46 μ s.

After the average evaluation an analysis of the individual values per subject was made. Thus, the vowel /a/ normal tone was selected for control and pathological subjects for further analysis. In Fig. 10 it is possible to observe the results

of absolute jitter for the 10 control subjects and in Fig. 11 the results of absolute jitter for the 5 pathological subjects.

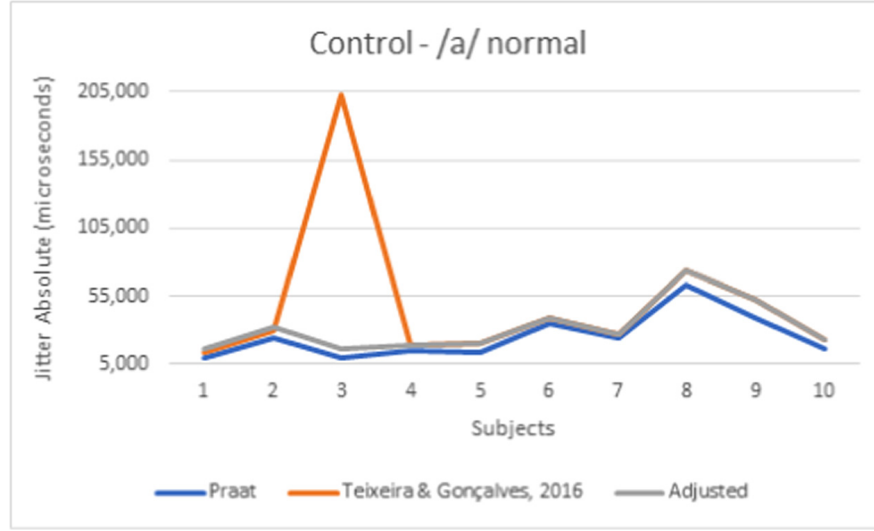


Fig. 10. Comparison of the absolute jitter values for the 10 control subjects.

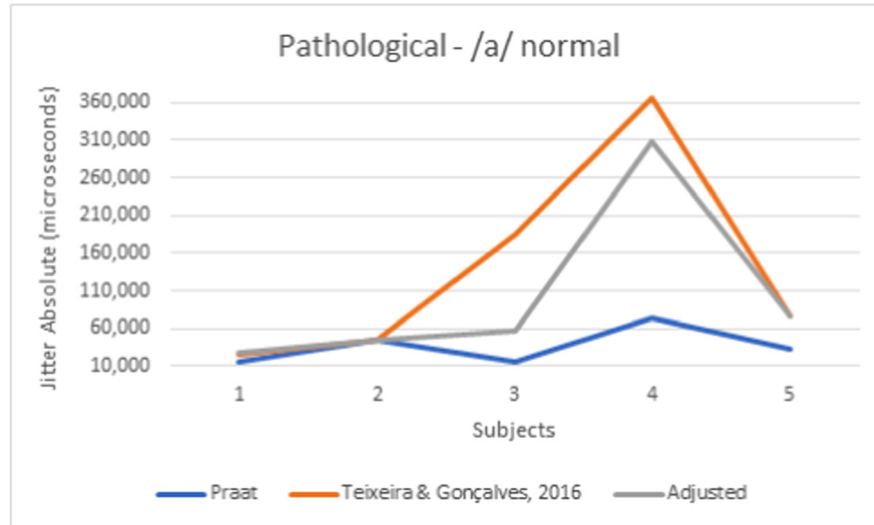


Fig. 11. Comparison of absolute jitter values for the 5 pathological subjects.

Through Figs. 10 and 11 it is possible to observe that comparing the values of the Teixeira and Gonçalves 2016 algorithm with the adjusted algorithm, the adjusted algorithm has closer values to the reference for more subjects. In control subjects, comparing the reference values with the Teixeira and Gonçalves algorithm, 2016, the difference is less than $16.8\mu\text{s}$. Now comparing the reference values with the adjusted algorithm, the difference is less than $7.1\mu\text{s}$. For pathological subjects, the difference between reference values and Teixeira and

Gonçalves algorithm, 2016, is less than $57.9\mu\text{s}$ and the difference with the corrected algorithm is less than $29.4\mu\text{s}$.

Therefore, it can be concluded that the adjusted algorithm obtains values closer to the reference values.

It should also be mentioned that an interpolation procedure of the speech signal was also experimented with interpolations of several order between 2 and 10. The objective was to increase the resolution of the peak position. The interpolation procedure didn't improve the accuracy of the absolute jitter determination, and was discarded.

5 Conclusion

The objective of this work was to improve the algorithm developed by Teixeira and Gonçalves 2016 in some particular signals subjected to inconsistencies finding the onset glottal periods. In order to try to increase the accuracy of the absolute jitter, some adjustments were made in the algorithm, where the detection of the peaks was improved. In order to understand if these corrections obtained improvements in the detection of the absolute jitter, an analysis was made of the averages of the values obtained from the absolute jitter of 10 control subjects and 5 pathological ones. In this analysis it was noticed that the values obtained through the corrections were closer to the reference values. In order to understand how the adjusted algorithm behaves compared to the reference value another analysis was carried out comparing, for one vowel, the 10 control subjects and the 5 pathological subjects. Thus, it was noticed that the adjusted algorithm measures the absolute jitter with values closer to the reference values for more subjects. The algorithm still measure absolute jitter with slightly higher values than Praat. In control subjects, the difference in absolute jitter measurements was reduced from 16.8 to $7.1\mu\text{s}$, and for pathological subjects, the difference was reduced from 57.9 to $29.4\mu\text{s}$.

References

1. Martins, A.I.H.G., Santana, M.F., Tavares, E.L.M.: Vocal cysts: clinical, endoscopic, and surgical aspects. *J. Voice* **25**(1), 107–110 (2011)
2. Godino-Llorente, J., Gomez-Vilda, P., Blanco-Velasco, M.: Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters. *IEEE Trans. Biomed. Eng.* **53**(10), 1943–1953 (2006)
3. Teixeira, J.P., Alves, N., Fernandes, P.O.: Vocal acoustic analysis: ANN versus SVM in classification of dysphonic voices and vocal cords paralysis. *Int. J. E-Health Med. Commun. (IJEHMC)* **11**(1), 37–51 (2020)
4. Sataloff, R.T., Hawkshawe, M.J., Sataloff, J.B.: Common medical diagnoses and treatments in patients with voice disorders: an introduction and overview. In: *Vocal Health and Pedagogy: Science, Assessment and Treatment*, p. 295 (2017)
5. Godino-Llorente, J., Gómez-Vilda, P.: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *IEEE Trans. Biomed. Eng.* **51**(2), 380–384 (2004)

6. Kadiri, S., Alku, P.: Analysis and detection of pathological voice using glottal source features. *IEEE J. Sel. Top. Signal Process.* **14**(2), 367–379 (2020)
7. Teixeira, J.P., Teixeira, F., Fernandes, J., Fernandes, P.O.: Acoustic analysis of chronic laryngitis - statistical analysis of sustained speech parameters. In: 11th International Conference on Bio-Inspired Systems and Signal Processing, BIOSIGNALS 2018, vol. 4, pp. 168–175 (2018)
8. Sripriya, N., Poornima, S., Shivaranjani, R., Thangaraju, P.: Non-intrusive technique for pathological voice classification using jitter and shimmer. In: International Conference on Computer, Communication and Signal Processing (ICCCSP), pp. 1–6 (2017)
9. Teixeira, J.P., Fernandes, P.: Acoustic analysis of vocal dysphonia. *Procedia Comput. Sci.* **64**, 466–473 (2015)
10. Teixeira, J.P., Gonçalves, A.: Accuracy of jitter and shimmer measurements. *Procedia Technol.* **16**, 1190–1199 (2014)
11. Teixeira, J.P., Gonçalves, A.: Algorithm for jitter and shimmer measurement in pathologic voices. *Procedia Comput. Sci.* **100**, 271–279 (2016)
12. Boersma, P., Weenink, D.: Praat: doing phonetics by computer [Computer program]. Version 6.0.48, 15 April 2019 (1992–2019). <http://www.praat.org/>
13. Barry, W., Pützer, M.: Saarbruecken Voice Database. Institute of Phonetics at the University of Saarland (2007). <http://www.stimmdatenbank.coli.uni-saarland.de>. Accessed 15 Apr 2021