

RESEARCH

Open Access

Correlation analysis of the speech multiscale product for the open quotient estimation

Wafa Saidi*, Aicha Bouzid and Nouredine Ellouze

Abstract

This article proposes a multiscale product (MP)-based method for estimating the open quotient (OQ) from the speech waveform. The MP is operated by calculating the wavelet transform coefficients of the speech signal at three scales and then multiplying them. The resulting MP signal presents negative peaks informing about the glottis closure, and positive ones informing about the glottis opening. Taking into account the shape of the speech MP close to the derivative of electroglottographic (EGG) signal, we proceed to a correlation analysis for the fundamental frequency and OQ measurement. The approach validation is done on voiced parts of the Keele University database by calculating the absolute and relative errors between the OQ estimated from the speech and the corresponding EGG signals. When considering the mean OQ over each voiced segments, results of our test show that OQ is estimated within an absolute error from 0.04 to 0.1 and a relative error from 8 to 21% for all the speakers. The approach is not so performant when the evaluation concerns the OQ frame-by-frame measurements. The absolute error reaches 0.12 and the relative error 30%.

Keywords: speech, open quotient, multiscale product, crosscorrelation

1. Introduction

According to the source-filter theory of the speech production [1], voiced speech is represented as the response of the vocal tract filter to the glottal voice source. The glottal source consists of quasi-periodic pulses which are created by the vocal folds oscillations. It is characterised by two crucial moments; the glottal closure (GCI) and opening instants (GOI). GCIs and GOIs are required to be estimated accurately for many applications in various speech areas, such as voice quality assessment [2], speech analysis and coding [3], speaker identification [4] and glottal source estimation [5].

A glottal source parameter widely related to the GCI and GOI is the open quotient (OQ). It is defined as the ratio between the glottal open phase duration and the speech period. The open phase is the proportion of the glottal cycle during which the glottis is open. Thus, it is the duration between one GOI and the consecutive GCI. The speech period is the interval limiting two successive GCIs.

OQ is of considerable interest as it has been reported to be related to voice quality such as “breathy” and “pressed” voices [6,7]. A breathy voice happens when the vocal folds do not completely close during a glottal cycle and thus the OQ is large. A pressed voice is produced with constricted glottis and it corresponds to a small OQ. Vocal quality is studied with more details in [8].

In [9], the OQ changes with vocal registers were analysed using high-speed digital imaging and electroglottography (EGG). The work presented in [10] proposes the OQ measurements from the EGG signal and studies the relationship between the OQ and the perception of the speaker’s age. The correlation between the OQ and the fundamental frequency has been studied for male and female speakers in [11,12]. Henrich [13] provides an overview of the OQ variations with the vocal intensity and the fundamental frequency.

The EGG signal was the easiest way to measure the OQ as it is a direct representation of the glottal activity. In this context, Henrich et al. [13-15] suggested a correlation-based method called DECOM for automatic measurement of the fundamental frequency (F0) and the OQ using the derivative of electroglottographic (DEGG)

* Correspondence: saidiwafa@yahoo.com

Signal, Image and Pattern Recognition Lab., National School of Engineers of Tunis, ENIT Le Belvédère, B.P.37. 1002 Tunis, Tunisia

signals. Bouzid and Ellouze [16] used the multiscale product (MP) of the wavelet transform (WT) for detecting singularities in speech signal caused by the opening and the closing of the vocal folds. But no quantitative results were given.

For estimating the OQ and other glottal parameters from the speech signal only, many approaches have been proposed to estimate the glottal source signal. These methods are based on the digital inverse filtering using linear prediction or vocal-tract deconvolution [17-19]. A recent study done in [20] uses the zeros of the z -transform with a general model of the glottal flow to compute the OQ and the asymmetry quotient on speech signal of various voice qualities.

In this article, we are inspired by the approach presented in [14] where the OQ is estimated from the EGG signal using a correlation-based algorithm. Knowing that the speech MP provides a signal having a shape strongly close to the DEGG signal, we apply the Henrich correlation approach on the newly obtained signal and not on the EGG one. Therefore, we can give an estimation of the pitch period and the OQ from the speech signal over frames of a fixed length.

This rest of the article is organised as follows. Section 2 presents the MP analysis of the speech signal. Section 3 describes the proposed approach to estimate the OQ over a given frame. The method is divided into three stages. The first one operates the speech MP consisting of making the WT coefficients at three scales. The second step consists of windowing the MP signal and then split it into positive and negative parts. The third step computes the crosscorrelation function between the obtained two parts for estimating the open phase duration, and the autocorrelation of the negative part for estimating the pitch period. Evaluation results are presented in Section 4. Conclusion is drawn in Section 5.

2. MP for speech analysis

WT is a multiscale analysis widely used in image and signal processing. Owing to the efficient time-frequency localisation and the multiresolution characteristics, the WTs are quite suitable for processing signals of transient and non-stationary nature. Mallat and Zhong [21] have shown that multiscale edge detection is equivalent to find the local maximum of its wavelet representation. Several wavelet-based algorithms have been proposed to detect signal singularities [22-24]. GCIs and GOIs are such events characterising the speech signal. The peak displaying the discontinuity in the WT is often damaged by noise when the scale is so fine or smoothed when the scale is large.

To improve edge detection using wavelet analysis, the MP method is proposed. It consists of making the

product of the WT coefficients of the acoustic signal over three scales. It enhances the peak amplitude of the modulus maxima line and eliminates spurious peaks due to the vocal tract effect.

The product of the WT of a function $f(n)$ at scales is

$$p(n) = \prod_j W_{s_j} f(n) \quad (1)$$

where $W_{s_j} f(n)$ represents the WT of the function $f(n)$ at scale s_j .

The product $p(n)$ shows peaks at signal edges, and has relatively small values elsewhere. An odd number of terms in $p(n)$ preserve the edge sign.

The MP was first related to the edge detection problem in image processing [25,26]. Besides, the MP is proposed by Bouzid and Ellouze [16,27] to extract crucial information concerning the vocal source such as glottal opening and closure instants, the fundamental frequency, the OQ and the voicing decision. In previous studies, we proved that the MP is a robust and efficient method for determining the GCI from both clean and noisy acoustic signal [28,29].

Figure 1 illustrates a frame of a voiced speech signal followed by its MP and the DEGG signal. The MP shows minima marking the instants of glottis closing with a high precision and maxima denoting the glottis opening with less precision.

Figure 2 shows the EGG signal followed, respectively, by its derivative and MP. The MP of the EGG signal presents only one peak even when these peaks are imprecise or doubled on the DEGG. In this example, we clearly see the effect of the MP on cancelling the noise and giving accurate peaks.

The strength of the MP of the EGG signal compared to the DEGG signal is profoundly studied by Bouzid and Ellouze [16]. This study attempts to measure the voice source parameters using the MP of the EGG signal.

3. Proposed method for OQ estimation

3.1. Overview of the method

Our proposed approach for the OQ estimation from the speech signal follows three stages as shown in Figure 3.

First stage: consists of computing the MP of a voiced speech signal and then the signal is divided into frames of a fixed length. To compute the MP, we multiply the WTs of the speech signal at scales 2, 5/2 and 3 using the quadratic spline function.

To divide the MP signal into frames of a length N , we multiply it by a sliding rectangular window $w[N]$. The MP over a window of index i is given by

$$MP_{wi}[k] = MP[k - iN]w[k] \quad (2)$$

where k is within $[1, N]$ and i is the frame index.

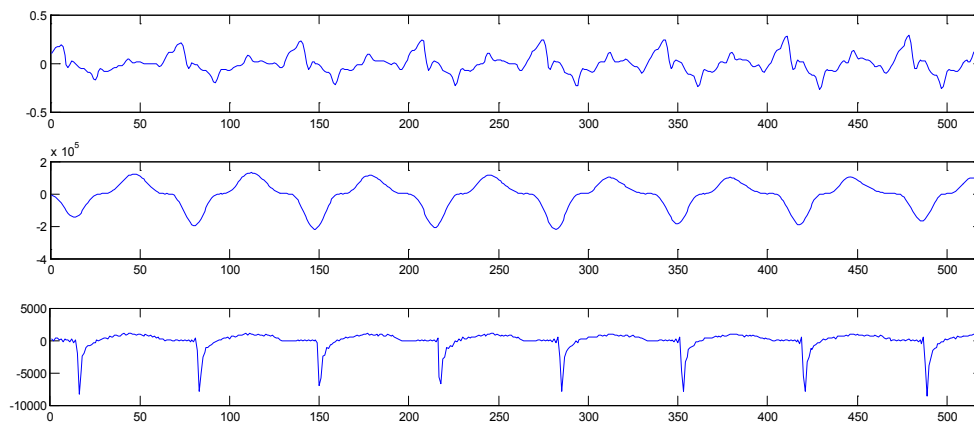


Figure 1 Speech signal followed by its MP and the DEGG signal.

Second stage: consists of separating the speech MP into two parts: a negative part MP^c which contains information concerning glottal closure peaks, and a positive part MP^o which contains information about glottal opening peaks. The MP^c signal is derived from the original signal by replacing any positive value by 0. In the same way, the MP^o signal is derived from the original signal by replacing any negative value by 0.

Figure 4 depicts the speech signal of the vowel/o/pronounced by the female speaker f1 followed by its MP, the MP^o and the MP^c . Minima of the MP negative part correspond to the GCI and peaks of the positive part fit with GOI.

Third stage: concerns the calculation of the cross-correlation function between the positive and negative parts (MP^o and MP^c) for estimating the open phase, and the autocorrelation function of MP^c to estimate the fundamental frequency over each frame. The open phase and the fundamental frequency are, respectively, given by the non-null index matching with the

first maximum of the crosscorrelation and autocorrelation functions. The OQ is then deduced by calculating the ratio between the open phase and the pitch period.

The crosscorrelation function between MP^o and MP^c over a frame i is calculated as follows

$$R_o(k) = \sum_{l=1}^N MP_{w_i}^o(l) MP_{w_i}^c(k+l) \quad (3)$$

By the same way, the autocorrelation function of MP^c over a frame i is calculated as follows

$$R_c(k) = \sum_{l=1}^N MP_{w_i}^c(l) MP_{w_i}^c(k+l) \quad (4)$$

3.2. Frame selection

Assuming that the fundamental frequency value is approximately known, the frames length is chosen to

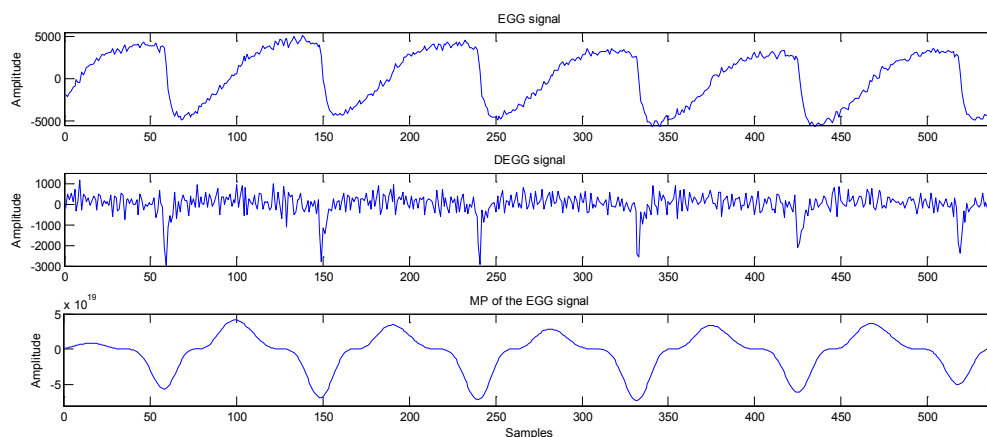
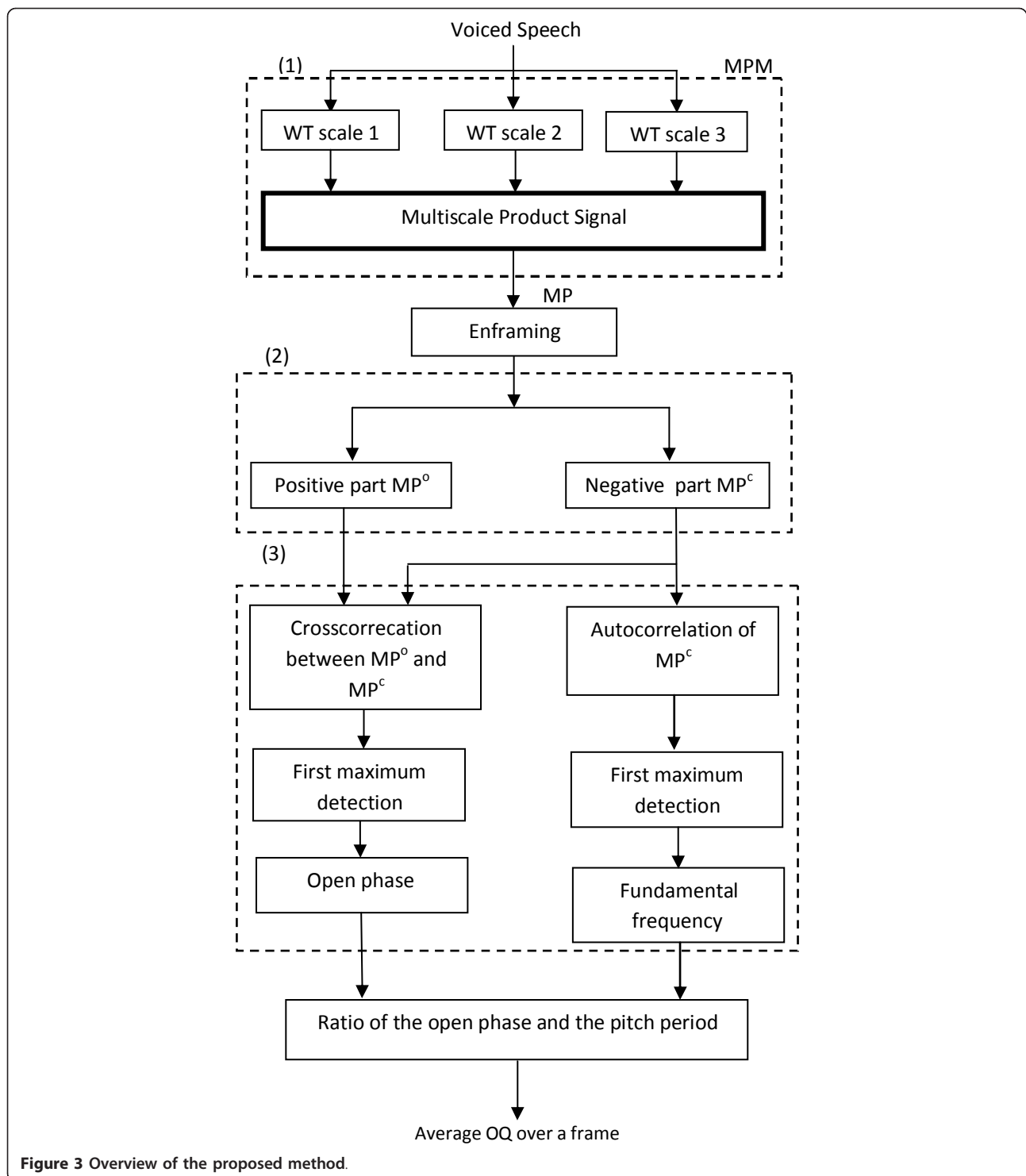


Figure 2 EGG signal, DEGG signal and the MP of the EGG signal.



be no less than four periods and no longer than eight periods. We chose these limits for the frame because on running speech, the fundamental frequency varies by a significant amount over eight periods of pitch. So, we use a rectangular window with a fixed length of

25.6 ms for female speakers and 51.2 ms for male speakers.

Figure 5 illustrates the instantaneous fundamental frequency of each glottal cycle over a voiced segment of 97 periods long. F0 is extracted from both the EGG and

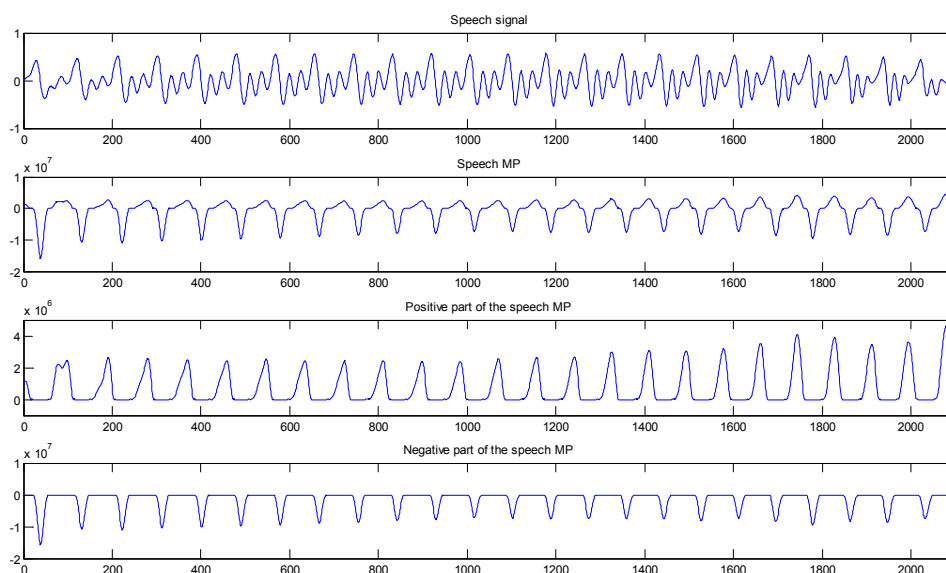


Figure 4 Speech signal, the MP of the speech signal, MP^o and MP^c .

speech signals by detecting GCIs manifested as minima of the MP. This example shows the variation sustained by F0 over running speech. F0 varies significantly when exceeding eight glottal cycles.

3.3. MP autocorrelation for the fundamental frequency estimation

Autocorrelation analysis is a well-known method for fundamental frequency estimation. This technique was

firstly used by Rabiner [30] as a pitch detector. Henrich et al. [14] applied this approach to estimate the fundamental frequency from the EGG signal.

For us, we focus on applying the autocorrelation technique to calculate the fundamental frequency from the speech signal. In fact, we calculate the speech MP of the speech over a frame, and then we compute the autocorrelation function of its negative part. The non-null index of the first maximum corresponds to the mean

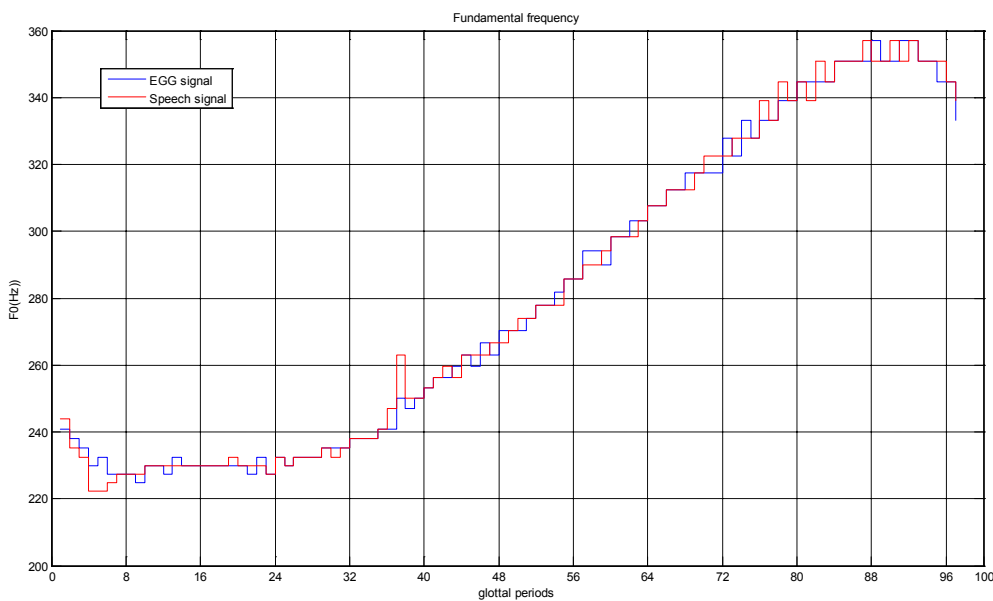


Figure 5 F0 from EGG signal, F0 from speech signal over a voiced segment.

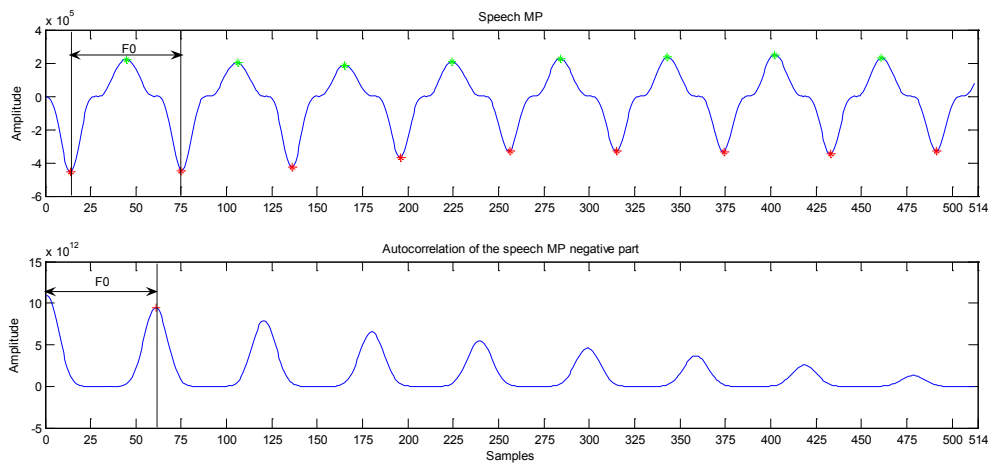


Figure 6 Speech MP and the autocorrelation function of the speech MP negative part.

value of the duration between two successive GCIs. Figure 6 gives an example where the fundamental period is estimated using the proposed approach.

In [14], Henrich et al. discuss the problems of double or imprecise peaks happening on the DEGG signal at the opening and the closing of the glottis and how to handle them. This glottal behaviour is observed by Anastalpo and Karnell [31]. These problems are overcome using the MP of the EGG signal as proposed in [16]. For real speech, typical cases are absent for closing peaks and are seldom observed for opening peaks.

Figure 7 represents an example of a noisy DEGG signal. Peaks are imprecise and double on the DEGG but they are unique not on the MP of the EGG. We note the ability of the MP to eliminate spurious peaks. In this case, we see that peaks indicating the glottis closing are weak and difficult to detect especially at the beginning of the frame. We also note the efficient role of the

autocorrelation function to give a distinguishable maximum indicating the average value of the fundamental frequency over a given frame.

Figure 8 represents the F0 estimated from the speech and the EGG signals using the autocorrelation technique over voiced frames spoken by a female speaker (f_3). F0 extracted from the speech signal is often near to the reference one and they are confused for many frames.

3.4. MP crosscorrelation for open phase estimation

To calculate the glottis open phase duration of the speech signal, we calculate its MP at first. Then, we operate the crosscorrelation between its positive and negative parts. The first maximum index is considered as the open phase.

Figure 9 shows the speech MP followed by the crosscorrelation calculated between its negative and

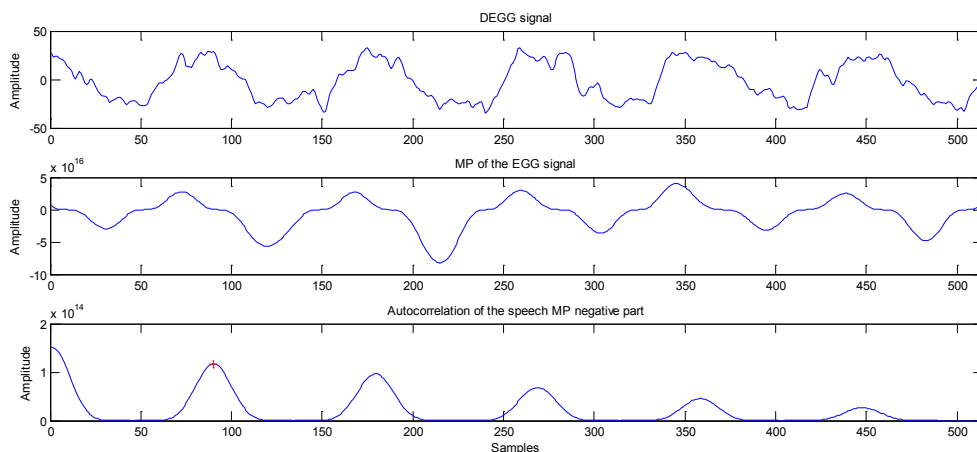


Figure 7 DEGG signal, MP of the EGG signal, autocorrelation of the speech MP negative part.

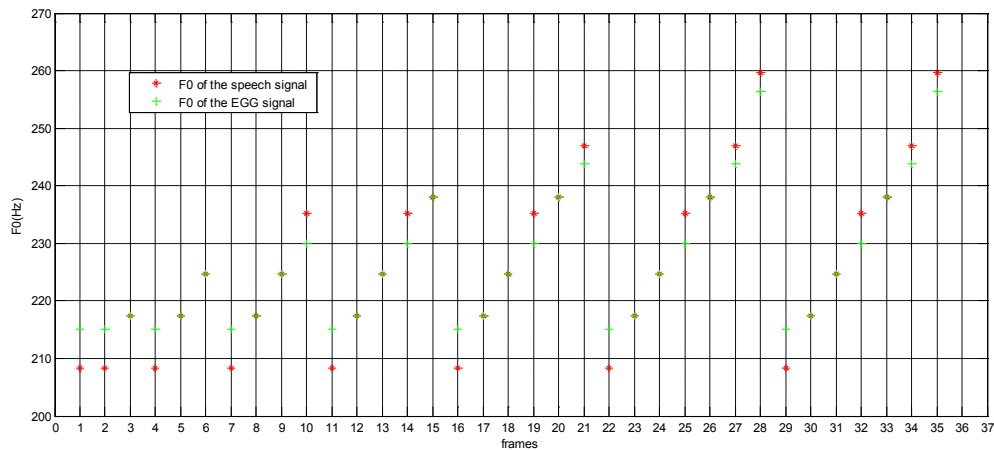


Figure 8 The F0 estimated from the speech signal and the F0 estimated from the EGG signal.

positive parts. The non-null index matching with the first maximum of the crosscorrelation function corresponds to the time between an opening peak and the consecutive closing peak which is termed as the open phase.

However, we note the cases where the speech MP produces more than one positive peak during a period. This behaviour induces double peaks on the crosscorrelation function. So, we consider the mean value of the two maxima. Our solution gives the nearest value to the open phase measured by the EGG signal as it is considered as the ground truth.

Figure 10 illustrates a problematical case where the opening peaks are double and have very weak amplitude on the MP. On the crosscorrelation function, these peaks are also double but with reinforced amplitude. The middle of the two peaks coincides well with the unique peak given by the EGG signal.

3.5. OQ estimation

Since the fundamental frequency and the open phase are given, it is possible to estimate the OQ.

Figure 11 illustrates the OQ measured from the reference EGG signal and the OQ estimated from the speech signal for the voiced segments uttered by the female speaker f4. In Figure 12, we draw the OQ estimation accuracy by computing the standard deviation of the error calculated between OQ measured from the EGG signal and OQ estimated from the real speech over each voiced segment. We effectively note a good coherence between the estimation from the speech signal and the reference from the EGG signal.

Figure 13 depicts the results of the OQ estimation from both the speech and the reference EGG signals for the frames contained in all the voiced segments corresponding to the speaker f4. Figure 14 shows the OQ accuracy over the whole frames.

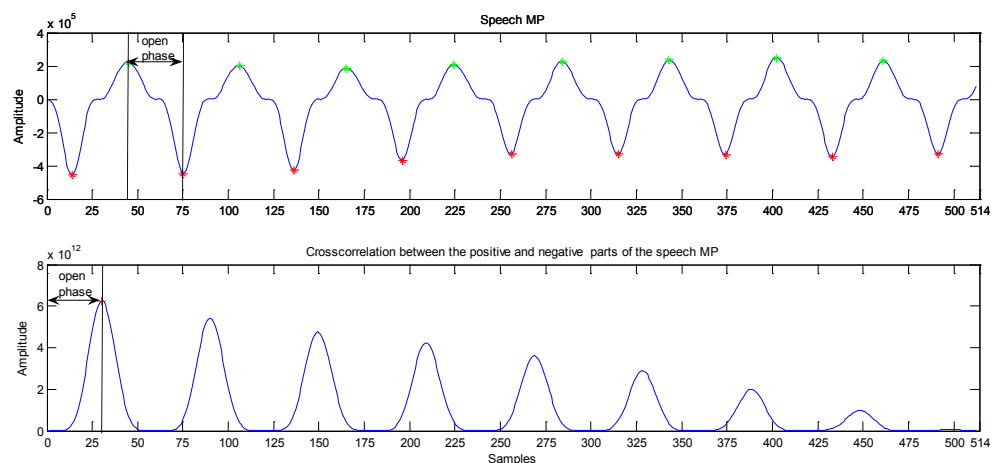


Figure 9 Speech MP and the crosscorrelation of the negative and positive parts of the speech MP.

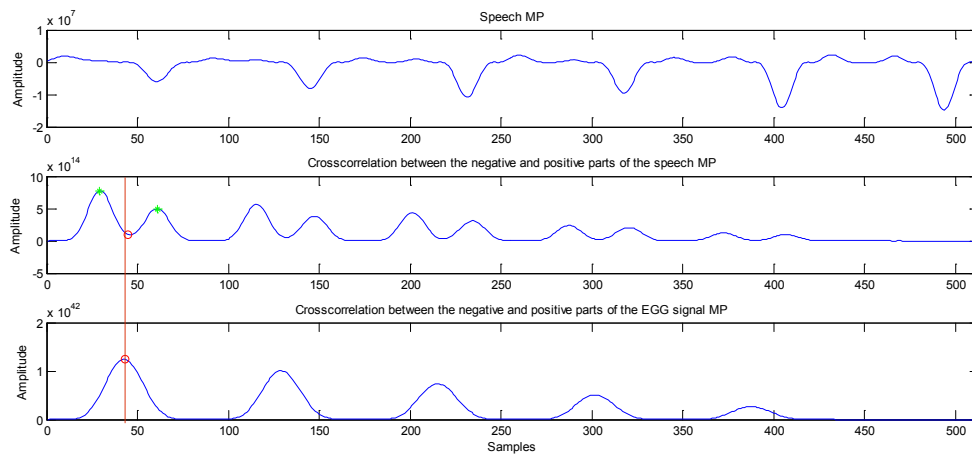


Figure 10 Speech MP, crosscorrelation of the negative and positive parts of the speech MP and the crosscorrelation of the negative and positive parts of the EGG MP.

Observing the OQ accuracy representation in Figures 12 and 14, we conclude that the OQ estimation is more precise when considering the mean OQ value over the voiced segments.

Gross deviation of the OQ estimation is caused by the errors of the open phase estimation happening when the opening peaks are doubled or imprecise.

The OQ estimation is unbiased in all cases. The error is much larger in Figures 13 and 14 than in Figures 11 and 12, showing that the GOI localisation from the speech signal is less accurate than from the EGG signal in the second case.

4. Experiments and results

4.1. Data

To evaluate the performance of our algorithm for OQ estimation, we use the Keele University database. This database includes the acoustic speech signals and laryngograph signals (single speaker recording). Five adult female speakers (f_i) and five adult male speakers (m_i) with $i \in \{1, \dots, 5\}$ are recorded in low ambient noise conditions using a sound-proof room. Each utterance consists of the same phonetically balanced English text: "The North Wind Story." In each case, the acoustic and laryngograph signals are time-synchronised and share

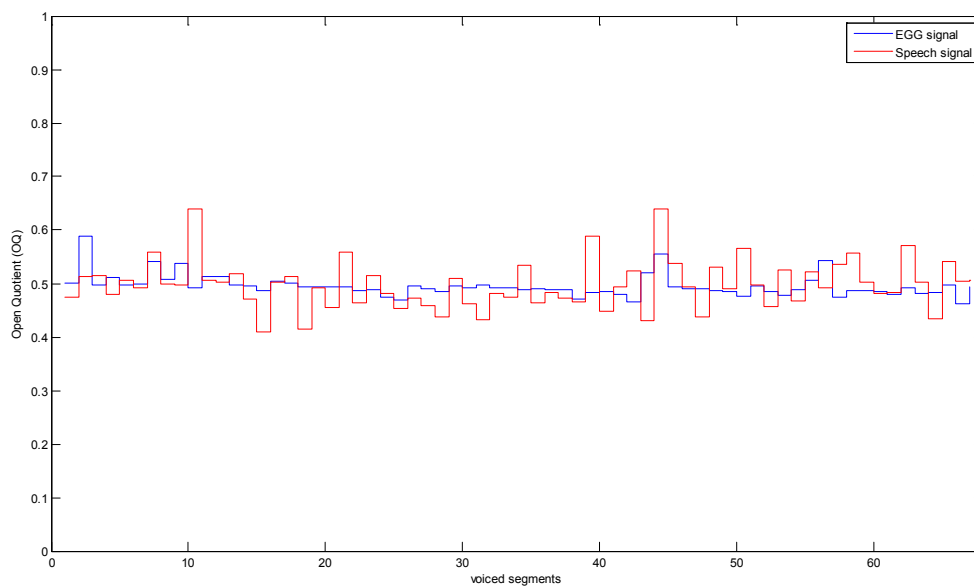


Figure 11 OQ estimated from the speech signal and OQ estimated from the EGG signal for each voiced segments of speaker f4.

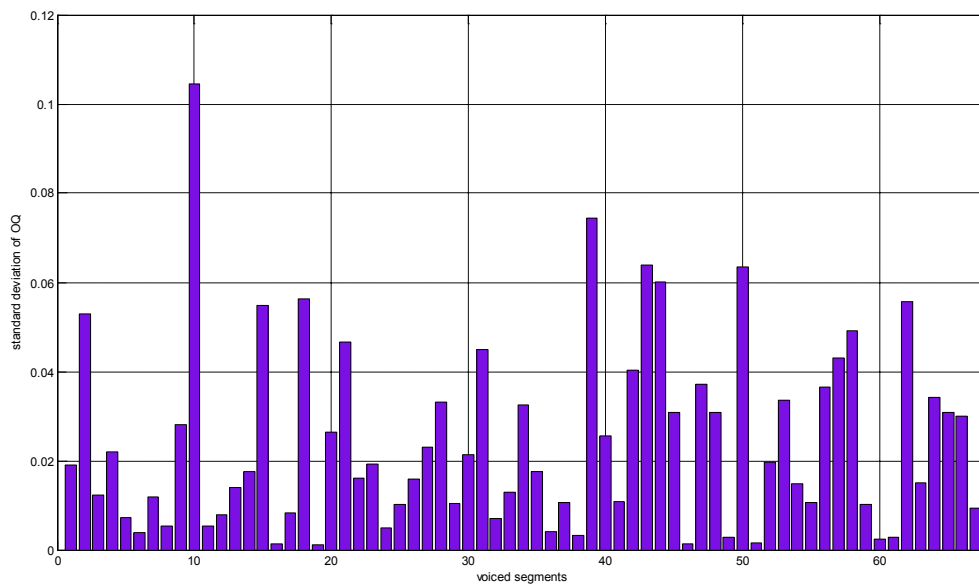


Figure 12 OQ estimation accuracy over voiced segments for speaker f4.

the same sampling rate value of 20 kHz [32]. The Keele database includes reference files containing a voiced/unvoiced segmentation and a pitch estimation of 25.6 ms segments with 10 ms overlapping. The reference files also mark uncertain pitch and voicing decisions. The database is open source and it available on [33].

4.2. Results

The Keele University database consists of running speech containing voiced, unvoiced and silence parts.

Only voiced segments extracted from the database are handled by our algorithm.

To evaluate the performance of our approach for OQ estimation, we calculate absolute and relative errors between OQ estimated from the speech signal and the reference OQ estimated from the EGG signal.

We consider the indexes $\{1, \dots, 10\}$ corresponding to speakers $\{f_1, f_2, f_3, f_4, f_5, m_1, m_2, m_3, m_4, m_5\}$. Each speaker k is characterised by N_k the number of voiced

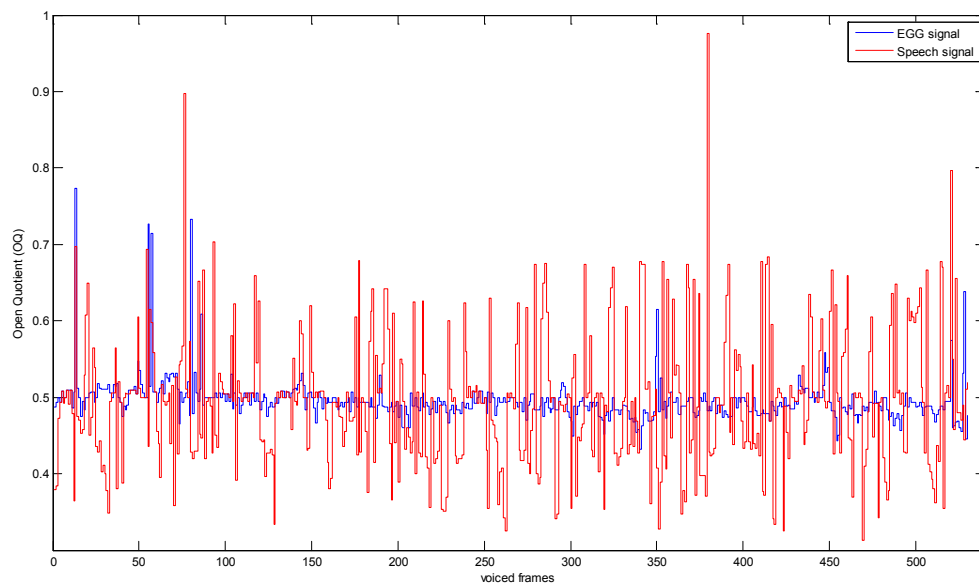


Figure 13 OQ estimated from the speech signal and OQ estimated from the EGG signal over voiced frames.

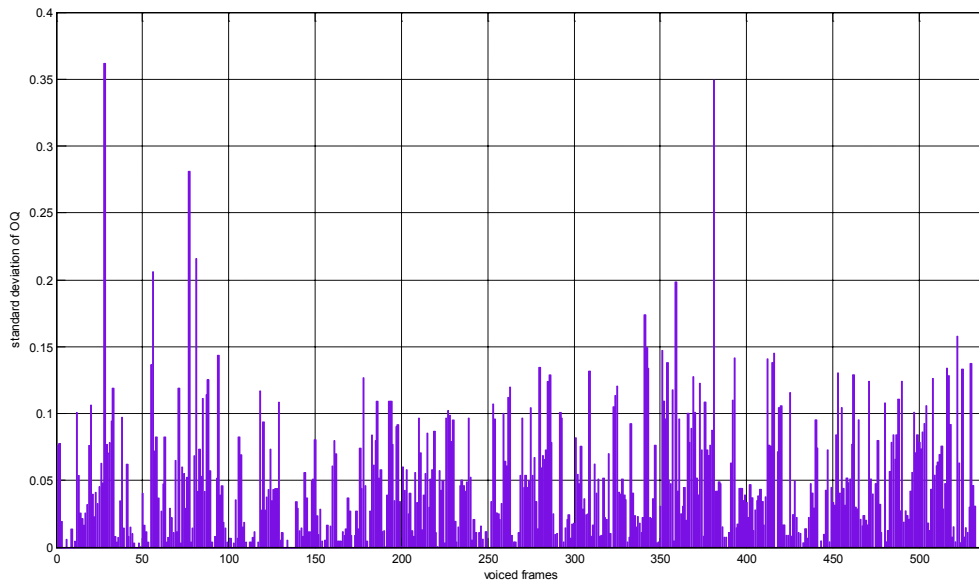


Figure 14 OQ estimation accuracy over voiced frames for speaker f4.

segments. Each segment is divided into n_{ki} frames where $k \in \{1, \dots, 10\}$ and $i \in \{1, \dots, N_k\}$.

In the first evaluation case, absolute or relative errors over the whole frames for each speaker k are defined as follow

$$e_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \frac{1}{n_{ki}} \sum_{j=1}^{n_{ki}} |\text{oq}_{nki}(j) - \text{oqegg}_{nki}(j)| \quad (5)$$

$$er_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \frac{1}{n_{ki}} \sum_{j=1}^{n_{ki}} \left| \frac{\text{oq}_{nki}(j) - \text{oqegg}_{nki}(j)}{\text{oqegg}_{nki}(j)} \right| \quad (6)$$

where $\text{oq}_{nki}(j)$ is the estimated OQ over a frame j that belongs to a voiced segment i uttered by a speaker k . $\text{oqegg}_{nki}(j)$ is the reference OQ value for the same frame calculated from the EGG signal.

For the second case, absolute and relative errors are defined by the mean values of the OQ estimated over the frames constituting the voiced segment:

For a given speaker k , the absolute and the relative errors are given by

$$\varepsilon_k = \frac{1}{N_k} \sum_{i=1}^{N_k} |\text{OQ}_{ki} - \text{OQegg}_{ki}| \quad (7)$$

$$\varepsilon r_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \frac{|\text{OQ}_{ki} - \text{OQegg}_{ki}|}{\text{OQegg}_{ki}} \quad (8)$$

where OQ_{ki} is the mean value calculated over a segment referring to the frames constituting this voiced segment.

Tables 1 and 2 depict the absolute and relative errors of the OQ estimation, from the speech signal compared to the EGG signal, for all the speakers of the Keele University database.

Table 1 gives errors referring to voiced frames. However, Table 2 gives errors referring to voiced segments.

Overall results show that the estimation of the OQ with the proposed method is competitive especially when considering the errors calculated over voiced segments of the database. In this case, absolute errors are at most 0.1 for speakers M1 and M5 and 0.07 for speakers f1 and f3. Relative errors do not exceed 13% for female speakers and 21% for male speakers.

Besides, the proposed approach for the OQ estimation can be considered as interesting and efficient regarding the error values and the lack of developed works in this field.

This research is a first step considered in our global project to give an accurate estimation of instantaneous OQ from the speech signal. That's why, the proposed measure is of great importance as it permits to give an

Table 1 Performance of the MP for the OQ estimation over voiced frames of the Keele University database

| Speakers | Absolute error | Relative error (%) | speakers | absolute error | Relative error (%) |
|----------------|----------------|--------------------|----------------|----------------|--------------------|
| F ₁ | 0.08 | 18 | M ₁ | 0.10 | 21 |
| F ₂ | 0.07 | 16 | M ₂ | 0.09 | 28 |
| F ₃ | 0.08 | 18 | M ₃ | 0.12 | 30 |
| F ₄ | 0.05 | 10 | M ₄ | 0.08 | 21 |
| F ₅ | 0.07 | 16 | M ₅ | 0.11 | 30 |

Table 2 Performance of the MP for the OQ estimation over voiced segments of the Keele University database

| Speakers | Absolute error | Relative error (%) | speakers | absolute error | Relative error (%) |
|----------------|----------------|--------------------|----------------|----------------|--------------------|
| F ₁ | 0.07 | 13 | M ₁ | 0.10 | 19 |
| F ₂ | 0.04 | 9 | M ₂ | 0.07 | 17 |
| F ₃ | 0.07 | 13 | M ₃ | 0.07 | 16 |
| F ₄ | 0.04 | 8 | M ₄ | 0.06 | 15 |
| F ₅ | 0.05 | 10 | M ₅ | 0.10 | 21 |

approximate interval more little than the period to localise the GOI. Once the GOIs are accurately located, we can turn back to estimate once again the OQ with more precision and for each period.

5. Conclusion

In this article, an approach for the OQ estimation from the speech signal is presented. It is based upon the correlation of the speech MP.

The MP is used to provide a simplified transformed speech signal that reminds the derivative of the EGG signal shape representing the global source activity.

The OQ estimation is obtained by calculating the ratio of the open phase over the pitch period. The open phase is referred as the index non-null of the first maximum localised on the inter-correlation function between the positive and the negative parts of the speech MP. As the same way, the pitch period is indexed by the first maximum of the speech MP correlation function.

Evaluation computes the absolute and relative errors between the OQ values determined from the speech signal and the OQ measured on the EGG signal considered as a reference. The evaluation is done on the Keele University database. The proposed approach reveals interesting performance.

Competing interests

The authors declare that they have no competing interests.

Received: 21 January 2011 Accepted: 10 November 2011

Published: 10 November 2011

References

1. Fant, G. *Acoustic Theory of Speech Production* (Mouton, La Hague, 1960)
2. N Gaubitch, P Naylor, Spatio-temporal averaging method for enhancement of reverberant speech, in *5th International Conference on Digital Signal Processing*, 607–610 (2007)
3. P Jinachitra, Glottal closure and opening detection for flexible parametric voice coding. *INTERSPEECH* (2006). paper 1359-Thu2BuP.2
4. D Guerchi, P Mermelstein, Low-rate quantization of spectral information in a 4 kb/s pitch-synchronous CELP coder, in *IEEE Workshop on speech coding*, 111–113 (2000)
5. J Gudnason, M Brookes, Voice source cepstrum coefficients for speaker identification, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 4821–4824 (2008)
6. P Alku, E Vilkmán, A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and

7. male speakers. *Folia Phoniatr (Basekl)* **48**, 240–254 (1996). doi:10.1159/000266415
7. D Klatt, L Klatt, Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am.* **87**, 820–857 (1990). doi:10.1121/1.398894
8. PA Keating, C Esposito, Linguistic voice quality, in *11th Australasian International Conference on Speech Science and Technology* (Auckland, NZ, December 2006)
9. M Echternach, S Dippold, J Sundberg, MF Zander, B Richter, High-speed imaging and electroglottography measurements of the open quotient in untrained male voices' register transitions. *J Voices* **24**(6), 644–650 (2010). doi:10.1016/j.jvoice.2009.05.003
10. R Winkler, W Sendmeier, Open quotient (EGG) measurements of young and elderly voices: results of production and perception study. *ZAS Papers Linguistics* **40**, 213–225 (2005)
11. DG Hanson, BR Gerratt, GS Berke, Frequency, intensity and target matching effects on photoglottographic measures of open quotient and speed quotient. *J Speech Hear Res.* **33**, 45–50 (1990)
12. P Kitzing, B Sonesson, A photoglottographical study of the female vocal folds during phonation. *Folia Phoniatr (Basekl)* **26**, 138–149 (1974). doi:10.1159/000263776
13. N Henrich, C d'Allessandro, M Castellengo, B Doval, Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *J Acoust Soc Am.* **117**(3), 1417–1430 (2005). doi:10.1121/1.1850031
14. N Henrich, C d'Allessandro, M Castellengo, B Doval, On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am.* **115**(3), 1321–1332 (2004). doi:10.1121/1.1646401
15. N Henrich, B Doval, C d'Allessandro, M Castellengo, Open quotient measurements on EGG, speech and singing signals, in *Proceedings of the 4th International Workshop on Advances in Quantitative Laryngoscopy, Voice and Speech Research*, Jena (April 2000)
16. A Bouzid, N Ellouze, Voice source measurement based on multiscale analysis of electroglottographic signal. *Speech Commun*
17. YL Shue, J Kreiman, A Alwan, a novel codebook search technique for estimating the open quotient, in *Interspeech*, 2895–2898 (2009)
18. N Sturmél, C d'Allessandro, B Doval, A spectral method for estimation of the voice speed quotient and evaluation using electroglottography, in *7th Conference on Advances in Quantitative Laryngology* (Groningen, The Netherlands, October 6-7, 2006), p. 6
19. P Jinachitra, JO Smith, Joint estimation of glottal source and vocal tract for vocal synthesis using Kalman smoothing and EM algorithm, in *WASPAA'2005*, New Paltz, NY
20. N Sturmél, C d'Allessandro, B Doval, Glottal parameters estimation on speech using the zeros of the z-transform. in *INTERSPEECH 2010*, 665–668 (2010)
21. S Mallat, S Zhong, Characterization of signals from multiscale edges. *IEEE Trans Pattern Anal Mach Intell.* **14**(7), 710–732 (1992). doi:10.1109/34.142909
22. C Wendt, AP Petropulu, Pitch determination and speech segmentation using the discrete wavelet transform, in *Proceedings of ISCAS 96*, Atlanta **2**, 45–48 (1996)
23. VN Tuan, C d'Allessandro, Robust glottal closure detection using the wavelet transform, in *Proceedings of the European Conference on Speech Technology*, 2805–2808 (1999)
24. JF Wang, SH Shen, Wavelet transforms for speech signal processing. *J Chin Inst Eng.* **22**(5), 549–560 (1999). doi:10.1080/02533839.1999.9670493
25. A Rosenfeld, A nonlinear edge detection. *Proc IEEE.* **58**, 814–816 (1970)
26. Y Xu, JB Weaver, DM Healy, J Lu, Wavelet transform domain filters: a spatially selective noise filtration technique. *IEEE Trans Image Process.* **3**(6), 747–758 (1994). doi:10.1109/83.336245
27. A Bouzid, N Ellouze, Electroglottographic measures based on GCI and GOI detection using MP. *Int J Comput Commun Control.* **III**(1), 21–32 (2008)
28. W Saidi, A Bouzid, N Ellouze, Evaluation of multi-scale product method and DYPSA algorithm for glottal closure instant detection, in *3rd International Conference on Information and Communication Technologies: From Theory to Applications*, 2008. *ICTTA 2008*, 1–5 (April 7-11, 2008)
29. W Saidi, A Bouzid, N Ellouze, MPM method and DYPSA algorithm evaluation for GCI detection in noisy speech signal. *Int J Comput Inf Technol and Comp.* **1**(1), 93–105 (2010)

30. LR Rabiner, On the use of autocorrelation analysis for pitch detection. *IEEE Trans Acoust Speech Signal Process.* **25**(1), 24–33 (1977). doi:10.1109/TASSP.1977.1162905
31. S Anastalpo, MP Karnell, Synchronized videoscopic and electroglottographic examination of glottal opening. *J Acoust Soc Am.* **83**, 1883–1890 (1988). doi:10.1121/1.396472
32. F Plante, G Meyer, WA Ainsworth, A pitch extraction reference database, in *Proc of EUROSPEECH 1995*, 837–840 (1995)
33. Keele Pitch Database, Psychology Home page–Human Machine Perception, (University of Liverpool, 1995) <http://www.liv.ac.uk/Psychology/hmp/projects/pitch.html>

doi:10.1186/1687-4722-2011-8

Cite this article as: Saidi et al.: Correlation analysis of the speech multiscale product for the open quotient estimation. *EURASIP Journal on Audio, Speech, and Music Processing* 2011 **2011**:8.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
