

SINGULARITY DETECTION OF ELECTROGLOTTOGRAM SIGNAL BY MULTISCALE PRODUCT METHOD

Aïcha Bouzid, and Nouredine Ellouze

Signal, Image and Pattern Recognition, National School of Engineers of Tunis

ENIT B. P. 37, le Belvédère 1002, Tunis, Tunisia

phone: + (216) 98211085, fax: + (216) 71874722, email: bouzidacha@yahoo.fr, N.Ellouze@enit.rnu.tn

ABSTRACT

This paper deals with singularity detection in electroglottogram (EGG) signal using multiscale product method. Wavelet transform of EGG signal is operated by a windowed first derivative of a Gaussian function. This wavelet transform acts as a derivative of a smoothed signal by the Gaussian function. The wavelet coefficients of EGG calculated for different scales, show modulus maxima at discontinuities. The detected singularities correspond to glottal opening and closure instants called GOIs and GCIs. Multiscale product is the multiplication of wavelet coefficients of the signal at three successive scales. This multiscale analysis enhances edge detection, and gives better estimation of the maxima. Geometric mean of the three scale wavelet coefficients is calculated by applying cubic root amplitude function on product. This method gives a good representation of GCI and a best detection of GOI, so as the product is a nonlinear combination of different scales which reduces noise and spurious peaks. The presented method is effective and robust in all cases even for particular signal showing undetermined GOIs and multiple closure peaks.

1. INTRODUCTION

Electroglottography is a non invasive exploration of glottal activity, the resulting EGG signal is a data reference for pitch measurement and best estimation of glottal closure instant (GCI) and glottal opening instant (GOI) [1]. This reference is commonly used in some speech processing techniques like voiced/unvoiced classification and accurate source parameter estimation. Robust instant detection method can help in synchronous speech analysis and synthesis.

Referring to Childers [1], EGG signal has singularities at GCIs corresponding to signal discontinuities, these events indicate rapid change behaviour of the source (glottis), the derivative of the EGG shows a strong peak at glottal closure instant.

Referring to Mallat [2], interesting information lies in sharp transitions on the signal or its derivative, singularities and edges are detected by following the wavelet transform local maxima at fine scales.

As it was shown in previous work [3], wavelet transform is efficient in most cases for detecting singularities in EGG signal corresponding to closure instants, and is not

efficient for glottal opening instant. In fact, abrupt variation of amplitude on EGG signal can be observed at GCI, however, at GOI, the discontinuity is less obvious and can't be detected in most cases from this signal [4].

Unlike many multiscale techniques, that first form the edge maps at several scales and then synthesize them together, the multiscale product consists of determining the edges as the local maxima directly in the scale products. Multiscale product method (MPM) was introduced by Sadler et al. [5], [6], for signal edge detection. It was shown that the scale multiplication achieves better results than any scales, especially on the localization performance [7], [8], [9].

The aim of this paper is to use multiscale product method to improve the accuracy of measures operated on EGG signal for glottal opening and closure instant detection.

The present paper is organised as follows. Section 2 presents a brief description of singularities in EGG signals and methods for their detection. Besides, it is shown that EGG signal can present in many cases undetermined GOI and in less cases ambiguous GCI. Section 3 concerns wavelet transform as a useful tool for singularity detection. Section 4 aims with multiscale product for peak detection and singularity time estimation. Section 5 deals with application of multiscale product method to strengthen peaks of singularities in EGG signal.

2. EGG SIGNAL SINGULARITIES

Glottal parameter estimation can be operated on Electroglottogram using cross level method, Rottenberg et al. [10] use 50% level in the case of modal and tense voice. The major drawback of such methods is lack of accuracy. Howard [11] uses 3/7 threshold on EGG signal for GOI detection, and has more precise GCI detection using derivative of the EGG called DEGG [12]. Childers characterises the EGG inflexion points by the EGG derivative [13], [14]. Experimental investigation shows that GOI and GCI correspond respectively to strong and weak peaks observed on DEGG [15], [16], [17]. Derivative of EGG signal shows in each period two opposite peaks. The apparent peak corresponds to GCI and the weak one corresponds to GOI. In several cases, the GOI can't be determined. This situation is frequently present in different EGG databases. Many cases of multiple peaks appear on DEGG at glottal closure instants [4]. Different examples are presented in figures 1, 2 and 3, it is obviously noted that de-

derivative of EGG signal and threshold methods are not sufficient in all cases for GOI and GCI detection.

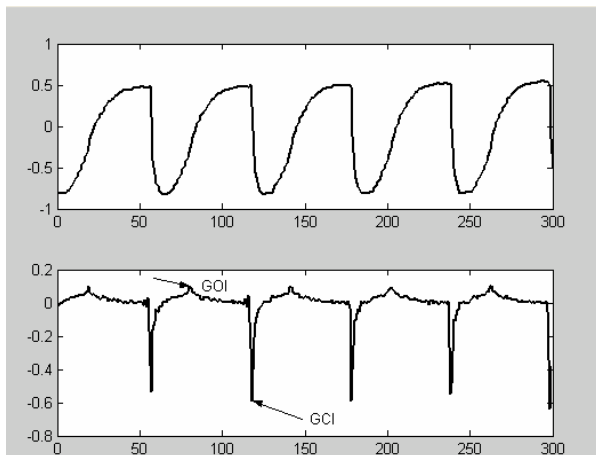


Figure 1 - EGG signal of female voice f4 /o/, and its derivative.

Figure 1 depicts the case of EGG signal of female voice pronouncing the vowel /o/, whose derivative shows clear peaks at GOI and GCI. Positive peaks correspond to GOI, and negative ones to GCI.

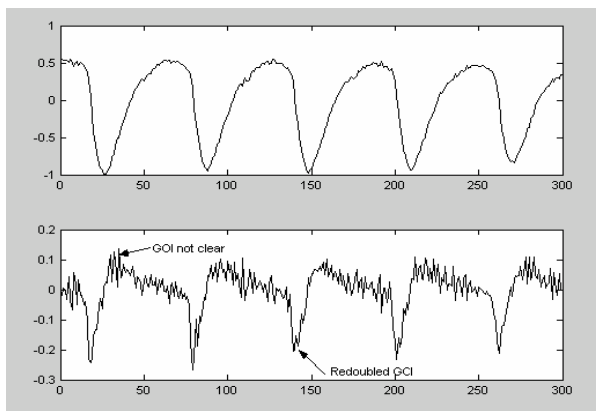


Figure 2 - EGG signal of female voice f2 /i/, and its derivative.

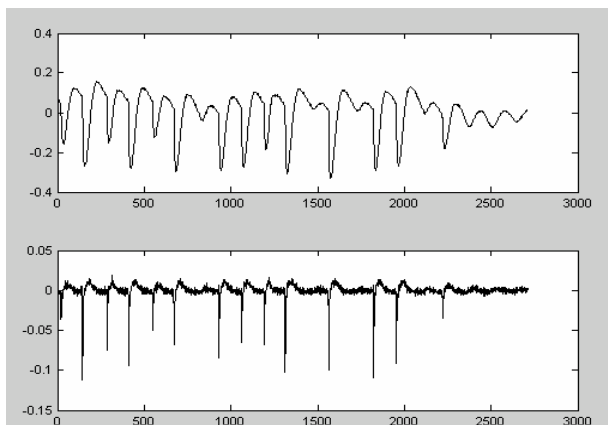


Figure 3 - EGG signal of male voice m2 (voiced sound), DEGG.

Figure 2 shows the case of undetermined GOI and multiple peaks of glottal closure, on signal of female voice pronouncing the vowel /i/.

Figure 3 illustrates other problematical cases where threshold methods on EGG and DEGG fail to detect GCIs and GOIs. In fact, Some GCIs and GOIs are indiscernible on the DEGG. This glottal behaviour was observed by Anastaplo and Karnell [21]. Pérez et al. underline the difficulty of GOI detection from the derivative of EGG signal [18].

3. WAVELET MODULUS MAXIMA

It is well known [2] that wavelet transform may be used for detecting and characterizing signal singularities. Singularities are detected by finding abscissa, where the wavelet modulus maxima converge at fine scales, and wavelet vanishing moments characterise the type of the singularity [19]. Modulus maxima of wavelet transform coefficients are a localized maxima appearing at specific singularity points, depending on wavelet vanishing moments. This is explained by the fact that wavelet transform with n vanishing moments can be interpreted as a multiscale differential operator of n th order of the smoothed signal. This provides a relationship between the differentiability of the signal and wavelet modulus maxima decay at fine scales. So if wavelet has one vanishing moment, modulus maxima appear at signal discontinuities, and represent the maxima of the first derivative of the signal smoothed by a given function. For wavelet having two vanishing moments, the modulus maxima correspond to discontinuities on derivative of the smoothed signal. If the wavelet transform has no modulus maxima at fine scales, then the function is locally regular. Hence singularities can be determined by locating abscissa where the wavelet modulus maxima converge at fine scales [2].

In previous work [3], we have shown that local regularity of the EGG signal can be characterised by wavelet transform modulus maxima and localized singularities indicate glottal closure and opening instants. Wavelet coefficients of EGG signal across scales, present local maxima and minima when using a wavelet that is a first derivative of a smoothing function.

Figure 4 shows an EGG signal (same utterance than figure1) at the top followed by its wavelet transforms at the first three scales. The wavelet transforms of EGG signal show 2 kinds of peaks across scales. A greater one corresponding to glottal closure instant; it's about sharp variations in the signal. The second kind of peak is linked to the glottal opening instant, it's about slow variations. Hence the GCI and GOI are indicated by the location of the local modulus maxima obtained by one vanishing moment wavelet. Analysis of EGG wavelet transform modulus maxima behaviour across scales shows that singularities detected at glottal closure and opening instants are characterized as signal discontinuities and EGG signal is more regular at glottal opening instant than at glottal closure instant [3]. That's why observed peaks at opening instants are weak comparing to those at closure instants.

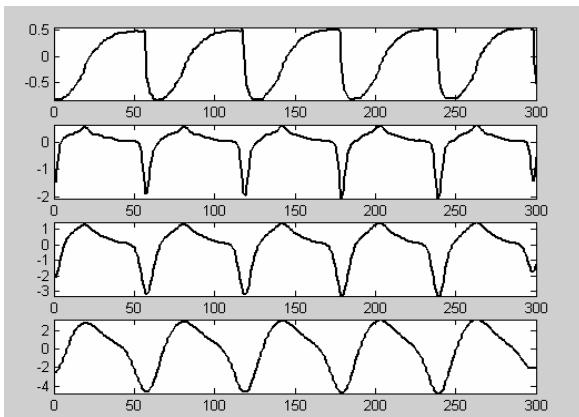


Figure 4 – Speaker f4 EGG /o/, wavelet transforms at 3 adjacent scales.

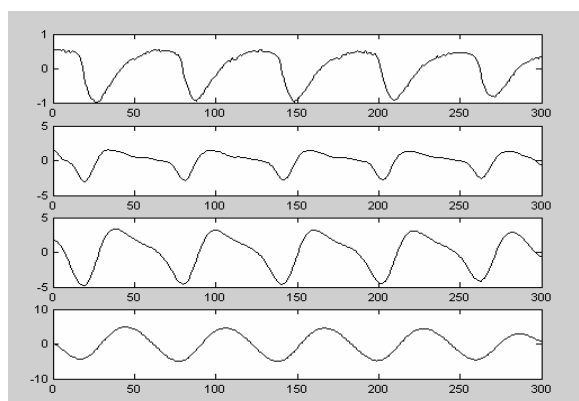


Figure 5 – Speaker f2 EGG /i/, wavelet transforms at 3 adjacent scales.

Figure 5 illustrates the fact that no scale can give a best and accurate detection of GCI and GOI. Effectively, singularities are too smoothed to be well localized.

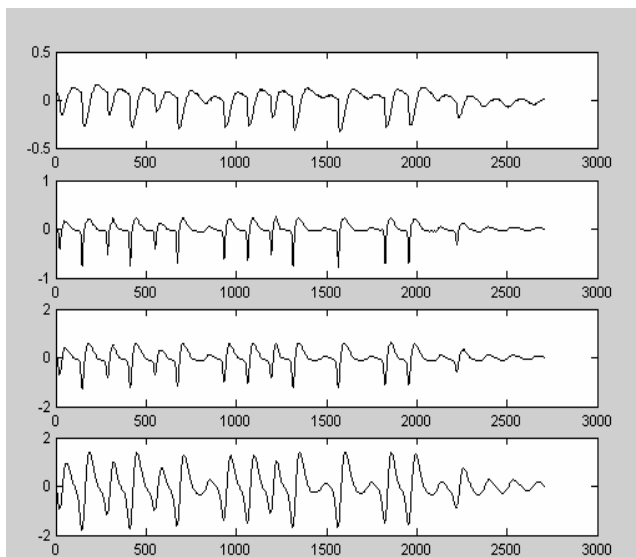


Figure 6 – Speaker m2 EGG /o/, wavelet transforms at 3 adjacent scales.

Figure 6 depicts an example where the wavelet transform as well as DEGG signal, can't detect a number of singularities at any scales.

4. MULTISCALE PRODUCTS

The products of coefficients across scales are frequently used for image analysis. Witkin [21] provided the foundation for scale space theory by generalizing Rosenfeld's work [20], in which smoothing filters at dyadic scales were used. Based essentially on forming multiscale products of smoothed gradient estimates, this approach attempts to enhance the peaks of the gradients caused by true edges, while suppressing false peaks due to noise. The wavelet transform acts as an edge detector, and the detail coefficients should be equivalent to the estimated gradients.

To distinguish edge maxima from noise and inappropriate maxima, Mallat and Zhong [22] analyze the singularity properties of wavelet transform domain maxima across various scales. First derivative of Gaussian and quadratic spline are used to play this role. Xu et. al. [9] relies on the variations in scale of the wavelet transform, direct multiplication of wavelet transform data at adjacent scales is used to distinguish important edges from noise. Sadler and swami [6] studied multiscale product method of signal in presence of noise. In wavelet domain, it is well known that edge structures are present at each subband while noise decreases rapidly along the scales. It has been observed that multiplying the adjacent scales could sharpen edges while diluting noise [6], [9].

Consider a multiscale analysis by forming the product of the wavelet transforms of a function $f(n)$ at some dyadic adjacent scales

$$p(n) = \prod_j w_{2^j} f(n) \quad (1)$$

Where $w_{2^j} f(n)$ is the wavelet transform of the function $f(n)$ at scale 2^j .

This expression is distinctly a non linear function. The product $p(n)$ reveals peaks at signal edges, and has relatively small values elsewhere. Singularities produce peaks along scale in wavelet transform, these peaks are reinforced by the product $p(n)$. Although particular smoothing levels may not be optimal, the non linear combination tends to reinforce the peaks while suppressing spurious peaks.

The signal peaks will align across scales for the first few scales, but not for all scales because increasing the amount of smoothing will spread the response and cause singularities separated in time to interact. Thus, choosing too large scales will result in misaligned peaks in $p(n)$. An odd number of terms in $p(n)$ preserves the sign of the edge. Choosing the first three dyadic scales is an optimal solution in multiscale product for detecting small peaks.

Motivated by the efficiency of the multiscale product in improving the edge detection, this method is applied on EGG signal and then can outperform the wavelet transform precision in weak singularity detection.

5. EGG MULTISCALE PRODUCTS

EGG signal used in this work is obtained from the Keele University database. Five adult female and five adult male speakers were recorded in low ambient noise using a sound proof room. Each utterance consisted of the same phonetically balanced English text. EGG signals are sampled at the frequency of 20 kHz.

The mother wavelet used is a derivative of a Gaussian function. Derivatives of Gaussian are often used to guarantee that all maxima lines propagate up to the finest scale.

Wavelet coefficient products at three adjacent levels are used to enhance the maxima. Cubic root is applied on the product to make zooming effect. The resulting processing, corresponding to modulus cubic root of coefficient product, is called the three geometric mean of the signal.

The method is tested on the all EGG signals of Keele database and efficiency is proved by all detected extrema. Specific examples are presented in figures 7, 8 and 9; they show derivative of EGG signal, multiscale product and geometric modulus mean results.

The proposed method is first applied on EGG where derivative signal shows clearly GOI and GCI. The first example taken is a frame of vowel /o/ uttered by a female speaker f4.

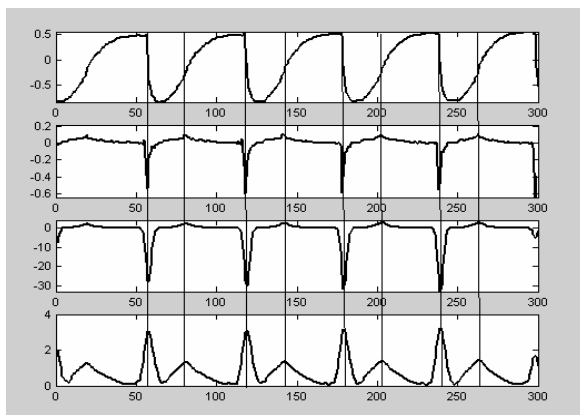


Figure 7 – f4 speaker EGG /o/, DEGG, MP and its modulus cubic root.

Figure 7 shows respectively the EGG signal, derivative of the signal, multiscale product of three adjacent scales (MP) and the cubic root of the products. The cubic root gives a zooming structure of the signal particularly for weak amplitudes.

The multiscale product, depicts the resulting cross-scale product $p(n)$ for three scales, and shows clean peaks aligned with the DEGG signal edges. First, we note in the cross scale product two types of peaks; those corresponding to GCI are more distinguishable than those related to GOI.

The geometric means show maxima at GOI and GCI, the weak peaks at GOIs given by the modulus cubic root of the product are better represented and effectively reinforced than obtained by multiscale product.

Figure 8 illustrates the example of a female utterance of vowel /i/ where the GOIs are undetermined and the peaks of GCI present irregular structures. We note double peaks of

glottal closure that engender aberrant measures. Besides, we can see clearly the effect of the product in suppressing the additional peaks and consequently the best detection of GOI. The modulus cubic root permits to reinforce peaks corresponding to GOI.

Figure 9 underlines the importance of the geometric means consisting of computing the modulus cubic root of wavelet coefficient product. In fact the cubic root takes out peaks at GOI and GCI that don't exist not only at DEGG signal but also at the product. It's about complex cases. In these special cases, small peaks denoting GCIs are indicated by solid lines however the GOIs corresponding to smaller peaks are visualised with dotted lines. Thus, we can detect the inexistent GCI and GOI at DEGG signal by the modulus cubic root of multiscale product.

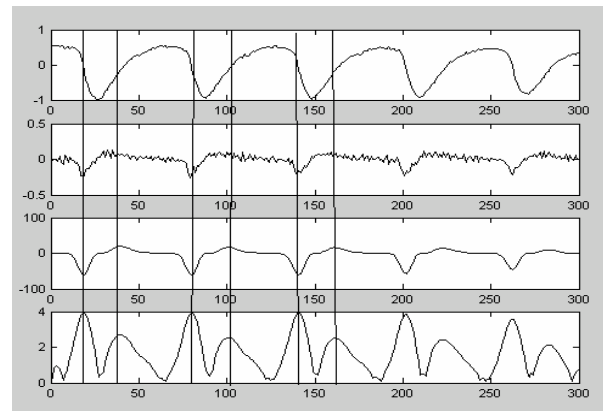


Figure 8 - Speaker f2 EGG /i/, DEGG, MP and its modulus cubic root.

6. CONCLUSION

Detection of GCI and GOI on EGG signal using multiscale product is considered. The proposed method consists of computing the wavelet transform of EGG signal at three adjacent scales. The wavelet used is the first derivative of a Gaussian function, which has one vanishing moment; wavelet transform has been calculated at first three adjacent scales. Then we multiply the wavelet transform coefficients of the corresponding scales, calculate the cubic root amplitude of the product, and finally, we look for maxima of determined cubic root.

This method gives better localization of GCI and GOI even in typical cases of multiple peaks and undetermined GCI and GOI on DEGG. The non linear products reinforce the cross-scale peaks produced at GCI and especially at GOI and reduce spurious noisy peaks. This method was tested on all EGG signals of Keele database; efficiency was proved by detected extrema examination.

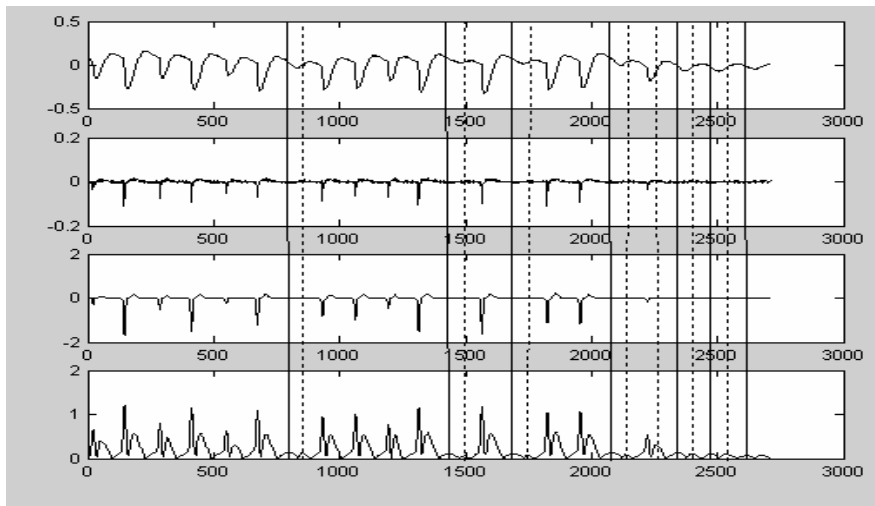


Figure 9 - Speaker m2 EGG /o/, DEGG, MP, and its modulus cubic root.

REFERENCES

- [1] D. G. Childers, A. M. Smith and G. P. Moore, "Relationships between electroglottograph, speech, and vocal cord contact," *Folia Phoniatr.*, vol. 36, pp. 105-118, 1984.
- [2] S. Mallat, *A Wavelet Tour of Signal Processing*. Second Edition, Academic Press, San Diego 1999.
- [3] A. Bouzid, and N. Ellouze, Local regularity analysis at glottal opening and closure instants in electroglottogram signal using wavelet transform modulus maxima, in *Proc. Eurospeech 2003*, Geneva, September 2003, pp. 2837-2840.
- [4] N. Henrich, *Etude de la Source Glottique en Voix Parlée et Chantée: Modélisation et Estimation, Mesures Acoustiques et Electroglottographiques*. Phd Thesis, Paris 6 University, 2001.
- [5] B. M. Sadler, T. Pham, and L. C. Sadler, "Optimal and wavelet-based shock wave detection and estimation," *Journal of Acoustical Society of America*, vol. 104, no. 2, pp. 955-963, 1998.
- [6] B. M. Sadler, and A. Swami, "Analysis of multiscale products for step detection and estimation," *IEEE Trans. Inform. Theory*, vol. 45, no. 3, pp. 1043-1051, 1999.
- [7] L. Zhang, and P. Bao, "Edge detection by scale multiplication in wavelet domain," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1771-1784, 2002.
- [8] P. Bao, L. Zhang, and X. Wu, "Canny edge detection enhancement by scale multiplication," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1485-1490, 2005.
- [9] Y. Xu, J. B. Weaver, D. M. Healy, and J. Lu, "Wavelet transform domain filters: a spatially selective noise filtration technique," *IEEE Trans. Image Processing*, vol. 3, no. 6, pp. 747-758, 1994.
- [10] M. Rothenberg, and J. J. Mahshie, "Monitoring Vocal Fold Abduction through Vocal Fold Contact Area," *Journal of Speech and Hearing Research*, vol. 31, pp. 338-351, 1988.
- [11] D. M. Howard, "Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers," *Journal of Voice*, vol. 9, no. 2, pp. 1212-1223, 1995.
- [12] D. M. Howard, G. A. Lindsey, and B. Allen, "Toward the quantification of vocal efficiency," *Journal of Voice*, vol. 4, no. 3, pp. 205-212, 1990.
- [13] D. G. Childers, and A. K. Krishnamurthy, "A Critical review of electroglottography," *CRC Critical Reviews in Biomedical Engineering*, vol. 12, pp. 131-161, 1985.
- [14] D. G. Childers, D. M. Hooks, G. P. Moore, L. Eskenazi, and A. L. Lalwani, "Electroglottography and vocal fold physiology," *Journal of Speech Hearing Research*, vol. 33, pp. 245-254, 1990.
- [15] D. G. Childers, and J. N. Lara, "Electroglottography for laryngeal function assessment and speech analysis," *IEEE Trans. on Biomedical Engineering BME*, vol. 31, No. 12, pp. 807-817, 1985.
- [16] S. Anastaplo, and M. P. Karnell, "Synchronized videoscopic and electroglottographic examination of glottal opening," *Journal Acoustical Society of America*, vol. 83, no. 5, pp. 1883-1890, 1988.
- [17] M. H. Hess, and M. Ludwigs, "Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns," *Journal of Voice*, vol. 14, no. 2, pp. 255-271, 2000.
- [18] J. Pérez, and A. Bonafonte, "Automatic voice-source parametrization of natural speech," in *Proc. ICSLP 2005*, Lisboa, Portugal, September 2005.
- [19] S. Mallat, and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. On Information Theory*, vol. 38, no. 2, pp. 617-643, 1992.
- [20] A. Rosenfeld, "A non linear edge detection," *Proc. IEEE*, Vol. 58, pp. 814-816, May 1970.
- [21] A. Witkin, "Scale-space filtering," *Proc. Int. Joint Conf. Artif. Intell.*, pp. 1019-1021, 1983.
- [22] S. Mallat, and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710-732, 1992.
- [23] F. Plante, G. F. Meyer, and W. A. Ainsworth, "A Pitch Extraction Reference Database," in *Proc. Eurospeech 1995*, September 1995, pp. 837-840.