Applied Acoustics 127 (2017) 215–222

Contents lists available at ScienceDirect

Applied Acoustics

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

Improved transient evoked otoacoustic emission screening test using simple regression model and window optimization

Zoran Šarić [a,*], Miško Subotić [a], Ružica Bilibajkić [a], Slobodan Jovičić [b], Marko Barjaktarović [b]

[a] Laboratory of Acoustics, Life Activities Advancement Center, Gospodar Jovanova 35, 11000 Belgrade, Serbia
[b] Faculty of Electrical Engineering, University of Belgrade, Serbia

ARTICLE INFO

Article history:
Received 27 October 2016
Received in revised form 17 May 2017
Accepted 14 June 2017

Keywords:
Transient-evoked otoacoustic emission
Hearing screening
Wavelet
Simple regression model
Minimum mean square error criterion

ABSTRACT

The underlying problem in detection of transient evoked otoacoustic emission (TEOAE) is its extremely low level which is much under the level of noise that appears in the auditory canal. The algorithms based on wavelet transformation (WT) and frequency (e.g. scale) dependent windows are proved to have higher specificity and sensitivity of TEOAE test in comparison to the FFT-based algorithms using single analysis window. In this paper, a new algorithm for TEOAE screening test with improved performances is proposed. It is based on three improvements. The first one is based on simple linear regression model applied to series of signal-to-noise ratio (SNR) estimates and correction of the reproducibility estimate by mean square error criterion. The second one exploits inter-subjects variability of TEOAE latency by selecting the best window position. The third is based on the use of specific window shape developed from approximate of minimum mean square error criterion. The performance of the proposed TEOAE algorithm was tested on real TEOAE measurements embedded in artificial noise as well as on pure noise extracted from TEOAE measurements. The results provided evidence for the superiority of its performance in comparison with two referential algorithms.

1. Introduction

Analysis of otoacoustic emission (OAE) gives reliable and objective information about the quality of the cochlear function. Transient evoked OAEs (TEOAEs) are especially interesting because measurement of TOAEs elicited by high-level click stimuli can distinguish between normal hearing and impaired hearing and is widely used in the screening programs [1].

A serious problem in OAE signal detection is its extremely low level, between 10 and 20 dB SPL [2]. This is significantly under the level of noise that appears in the auditory canal. In order to provide a favorable signal-to-noise ratio (SNR) and to improve the reliability of OAE detection, different procedures are applied based on improvement of recording conditions, optimization of stimulus signal characteristics, as well as the improvement of the algorithms for processing of OAE signal [3]. Some of the processing algorithms use specific time windowing to overcome the low-frequency noise effect for newborn hearing screening [1]. Contrary to this approach, some of the researchers proposed the use of number of analysis windows in order to increase signal to noise ratio in each of TEOAE sub-bands [4]. This idea is based on dispersion of TEOAE latency for different frequency components [5].

In order to improve signal to noise ratio (SNR), the time-coherent ensemble averaging is commonly used. The omission of the noisy parts from the analysis interval further improves the SNR. This method of "time windowing" has been studied in order to find the exact duration of the analysis interval that provides the best separation between normal hearing and impaired hearing [4].

The wavelet technique, by which the signal is decomposed into a set of frequency components (scales), is well-suited for TEOAE analysis [5]. By computing the reproducibility from various scales, it is possible to improve the pass/fail separation during TEOAE hearing screening [6], as well as to reveal subtle differences that may exist between normal and pathological ears [7].

An advanced technique for time-frequency decomposition of TEOAE signal relies on assumption that TEOAE signal can be considered as a superposition of the resonant modes of characteristic frequencies and latencies [8]. By this approach, decomposition of TEOAE signal is performed by matching pursuit (MP) algorithm which finds a sub-optimal solution in a redundant set (dictionary) of functions [9]. The method has proved to be a reliable tool for

* Corresponding author.
E-mail addresses: sariczoran@yahoo.com (Z. Šarić), ifp2@ikomline.net (M. Subotić), r.bilibajkic@add-for-life.com (R. Bilibajkić), jovicic@etf.rs (S. Jovičić), marko.barjaktarovic@gmail.com (M. Barjaktarović).

http://dx.doi.org/10.1016/j.apacoust.2017.06.011
0003-682X/

extraction of time-frequency properties of TEOAE components such as amplitude, latency and time span [10].

It was shown that TEOAE latency is inversely proportional to the central frequency of the scale [7]. A considerable increase in reliability of TEOAE detection can be achieved by adjusting the window position and its width in accordance to the latency and click response duration of each scale. The method proposed in [6] combines wavelet signal decomposition, nonlinear denoising, and scale-dependent time windowing. This method is shown to have a much higher specificity in TEOAE signal detection, compared to the previously used procedures. An additional effort has been done in designing the subject specific time window [11]. The windows were designed with reference to a minimum mean square error criterion involving the correlation properties of the ensemble of responses.

The method proposed in [12] was focused on the further reduction of the variance of the calculated cross-correlation caused by noise in TEOAE measurements. Instead of use one pair of buffers, usually denoted as buffers A and B, it uses N paired buffers. The variance of the cross-correlation calculated between buffers A and B is reduced by averaging of the N cross-correlations calculated on N paired buffers.

In this paper, we proposed three improvements of the detection algorithm. The first one is based on a linear regression model applied to series of signal to noise (SNR) estimates. The second exploits inter-subject variability of TEOAE latency by using different windows positions, while the third optimizes the shape of the window using the least squares criterion in signal estimation. The integral algorithm which includes these three proposed improvements is compared with the referent algorithms [6,12] on two test cases. The first test case includes test examples of real TEOAE records artificially noised by method [12]. The aim of this test case is to assess the specificity of tested algorithms. The second test case includes test examples of pure noise extracted from real TEOAE measurements [12]. The aim of the second test case is to assess sensitivity of the tested algorithms. The results are discussed in Section 4.

2. Materials and methods

2.1. TEOAE acquisition

Sixty volunteers participated in this study. Written consent was obtained from all participants prior to testing. The test protocol was reviewed and approved by ethical committee of the Institute for experimental phonetics and speech pathology (Belgrade, Serbia) where all measurements were conducted. Experiments were performed in a quiet room in which SPL of the ambient noise was SPL = 30 dBA. TEOAE recording was performed by our own acquisition system designed for research purpose, also used in [13]. The level of click stimuli was 85 dB SPL. Signal analysis was performed off-line in MATLAB. From sixty normally hearing subjects we selected TEOAE measurements of 46 ears with reproducibility greater than 70%.

According to the derived non-linear response (DNLR) method [1], four evoked responses associated with each set of the four-click stimuli were averaged. The result is referred to as the subaveraged response. Set of 256 packets of stimuli, each consisting of 4 clicks generates 256 subaveraged responses. They were stored in 256 buffers for further processing.

In order to test TEOAE algorithm on TEOAE measurements contaminated with a controlled amount of noise, we used noise generation similar to the method proposed in [12]. The noise was extracted from real TEOAE measurements by subtracting two neighboring subaveraged responses, i.e. buffers 1 and 2, 3 and 4, and so on. To obtain 256 noise buffers, we extended TEOAE measurement until 512 subaveraged responses were recorded. Set of 256 noise buffers was used to generate 20 noise distributions using random permutation generator.

TEOAE responses of 46 ears with artificially generated noise were used to generate two test groups [12]: (*a*) Test group that consisted of 920 TEOAE measurements mixed with artificially generated noise in order to evaluate specificity of tested algorithms; (*b*) The test group that consisted of 920 artificially generated noise measurements in order to evaluate sensitivity of tested algorithms.

2.2. Basic structure of TEOAE detection algorithm

Fig. 1 briefly depicts TEOAE signal processing based on discrete wavelet transformation (DWT), including improvement based on N paired buffers proposed in [12], and a module for reproducibility calculation based on the method proposed in this paper.

During TEOAE measuring session, 256 subaveraged responses of four click stimuli were recorded in 256 buffers for further processing. In commonly used method [1], odd buffers form a set A, while even buffers form a set B. According to the approach [12], any combination of 128 responses of the set of 256 available responses can be used to form buffers A and B. The method [12] proposes the use of N different buffer distributions $P_i = \{A_i, B_i\}$, $i = 1,..,N$, denoted as paired buffers, for which sets A_i and B_i may be defined as $A_1 = \{1,3,5,\ldots,255\}$, $B_1=\{2,4,6,\ldots,256\}$, $A_2=\{1,2,5,6\ldots,254\}$, $B_2=\{3,4,7,8\ldots,256\}$, and so on. A minimal number of paired buffers is one (A_1 and B_1), but for the larger N, the variance of reproducibility estimate becomes lower [12]. In the next processing step, discrete wavelet transformation (DWT) is performed using Coiflet5 [14]. Only scales 5, 6, and 7 with central frequencies 1150, 2200 and 4400 Hz were used [11,12]. These scales were processed by scale dependent windows [6] before reproducibility calculation. The high-frequency scale 7 was windowed from 2.5 to 7.5 ms, middle-frequency scale 6 was windowed from 2.5 to 9 ms, and low-frequency scale 5 from 3.5 to 14 ms.

Cross-correlation coefficient of jth scale of lth paired buffer, usually referred to as "wave reproducibility" is calculated by [15,16] and [17]

$$r(j,l) = \frac{\sum_{t=1}^{T} \bar{x}_{A,l}(t,j)\bar{x}_{B,l}(t,j)}{\sqrt{\sum_{t=1}^{T} \bar{x}_{A,l}^2(t,j)}\sqrt{\sum_{t=1}^{T} \bar{x}_{B,l}^2(t,j)}}, \quad j=5,6,7, \quad l=1,\ldots,N, \tag{1}$$

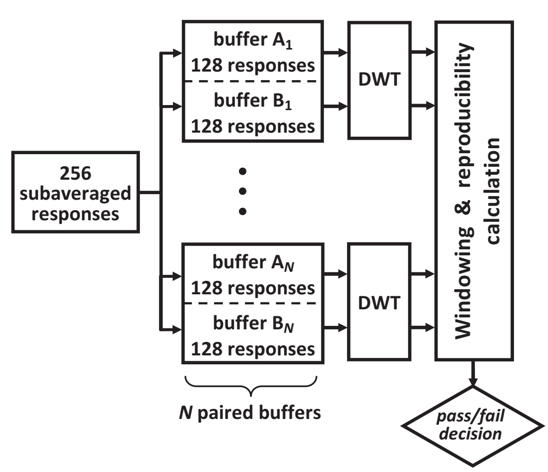

Fig. 1. Block diagram of the proposed wavelet-based algorithm.

where t is time index. $\bar{x}_{A,l}(t,j)$ and $\bar{x}_{B,l}(t,j)$ are weighted and ensemble averaged jth scale signals

$$\bar{x}_{A,l}(t,j) = \frac{1}{128}\sum_{i=1}^{128} w(t,j)x_{A,l}(t,i,j),\ \bar{x}_{B,l}(t,j) = \frac{1}{128}\sum_{i=1}^{128} w(t,j)x_{B,l}(t,i,j), \tag{2}$$

where $w(t,j)$ is scale dependent window function. Averaged cross-correlations of N paired buffers are

$$r(j) = \frac{1}{N}\sum_{l=1}^{N} r(j,l). \tag{3}$$

Finally, three scale cross-correlations are averaged to obtain single cross-correlation measure r, by

$$r = \frac{1}{3}\sum_{j=5}^{7} r(j), \tag{4}$$

which is used for pass/fail decision.

Algorithm [6] may be presumed as a special case of the algorithm [12], for $N = 1$, in which odd subaverages are distributed in the buffer A_1, while the even subaverages are distributed in the buffer B_1. The advantage of the method [12] is a lower variance of the calculated cross-correlations due to the averaging of the N cross-correlations of N paired buffers.

In the next sections a method based on a number of improvements in TEOAE detection, which can be used independently as well as in combination with the methods [6] and [12], will be described. The first improvement is based on simple linear regression model applied to series of SNR estimates in subaveraged responses. The second is based on the use of inter-subject variability of the TEOAE latency. And the third is based on suboptimal optimization of the windowing function in minimum mean square error sense.

2.3. Improvement by linear regression model of SNR estimates

In this section, we will present a new approach for reproducibility estimation. This approach is based on linear regression modeling of the series of SNR estimates and relationship between reproducibility and SNR. Reproducibility estimated by the linear regression model will be used for correction of the reproducibility obtained by (1) using the approximate solution of minimum mean square error criterion.

Proposed processing is the same in all paired buffers. For the sake of simplicity, the paired buffers index l will be omitted in the following text. We model discrete wavelet transformation (DWT) coefficients $x_q(t,i,j)$ of each subaveraged response i as random variable by

$$x_q(t,i,j) = s(t,j) + \xi_q(t,i,j), q \in \{A, B\} \tag{5}$$

where q denotes buffer A or B, $s(t,j)$ is noiseless response of the scale j, $\xi_q(t,i,j)$ is zero mean random process originated from the noise in ear canal, t is time index relative to the click instant, $i = 1,\ldots,M$ is index of the subaveraged response of buffer A or B, where $2M$ is number of available subaverages ($M = 128$). We assume that $\xi_A(t,i,j)$ and $\xi_B(t,i,j)$ are uncorrelated, and they are also uncorrelated with $s(t,j)$. We also assume that variance of the noise

$$E\left(\xi_q^2(t,i,j)\right) = \sigma_n^2(j) \tag{6}$$

is constant in time. Average signal-to-noise ratio (SNR) based on subaveraged responses from 1 to m is

$$SNR(m) = \frac{1}{3}\sum_{j=5}^{7} SNR_j(m), SNR_j(m) = \frac{\left(\sum_{t=1}^{T}\left(\frac{w(t,j)}{m}\sum_{i=1}^{m}\frac{x_A(t,i,j)+x_B(t,i,j)}{2}\right)\right)^2}{\left(\sum_{t=1}^{T}\left(\frac{w(t,j)}{m}\sum_{i=1}^{m}\frac{x_A(t,i,j)-x_B(t,i,j)}{2}\right)\right)^2}, \tag{7}$$

Mathematical expectation $E(SNR(m))$, $m = 1,\ldots, M$ is linearly dependent on number of responses m

$$E(SNR(m)) = \bar{\rho}m + 1, \tag{8}$$

where $\bar{\rho} = E(SNR(1))$ is mathematical expectation of the original SNR for a single subaveraged response. According to linear dependence $E(SNR(m))$ on m (8), we apply simple linear regression model [18].

$$SNR(m) = \alpha + \beta m + \varepsilon(m), \tag{9}$$

where α and β are intercept and slope constants, while $\varepsilon(m)$ is series of residuals of the model. An example of linear regression lines for the series $SNR_j(m)$, of scales $j = 5$, 6 and 7 is displayed in Fig. 2. The last element $y(M)=\alpha + \beta M$ is used as an approximation of unknown mathematical expectation $E(SNR(M)) \approx \hat{\rho}(M) = y(M)$. By applying the approximate formula that connects SNR and cross-correlation [16] the cross-correlation estimated by a series $SNR(M)$ is

$$\hat{r}_{by\_snr} = \frac{\hat{\rho}(M) - 1}{\hat{\rho}(M) + 1} \tag{10}$$

We are using the estimate $\hat{r}_{by\_snr}$ in order to improve the estimate r, obtained by (4). We will assume that the r can be modeled as a stationary random process $r = \bar{r} + \varsigma$, where $\bar{r}$ is mathematical expectation of the averaged cross-correlation r, and ς is error uncorrelated with $\bar{r}$. The goal is to calculate a coefficient $\bar{\alpha}$ such that estimate

$$\hat{r} = \bar{\alpha}r \tag{11}$$

has the minimum mean square error $E\{e^2\} = E\{(\bar{\alpha}r - \bar{r})^2\}$. The weighting coefficient $\bar{\alpha}$ that fulfills this requirement is

$$\bar{\alpha} = \frac{E(\bar{r}^2)}{E(\bar{r}^2) + E(\varsigma^2)}. \tag{12}$$

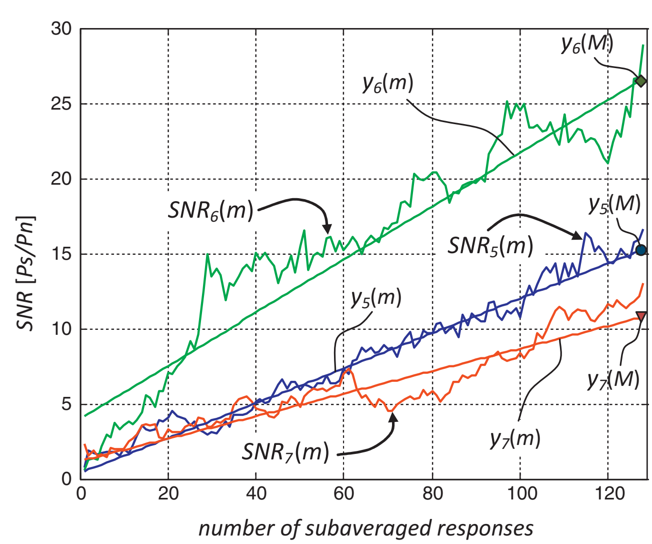

Fig. 2. Series $SNR_l(m)$, $m = 1,\ldots,128$, and estimated linear regression lines $y_l(m)$ for scales $l = 5$, 6 and 7.

Both variances $E(\bar{r}^2)$ and $E(\varsigma^2)$ are unknown. The first variance we approximate by $\hat{r}^2_{by\_snr}$, (10), while the second $E(\varsigma^2)$ is adopted as constant $E(\varsigma^2) = 0.01$ experimentally determined on available samples of TEOAE measurements. Substituting $E(\bar{r}^2) = \hat{r}^2_{by\_snr}$ and $E(\varsigma^2) = 0.01$ into (12), and taking into account (11), the improved estimate of the averaged r is

$$\hat{r} = \frac{\hat{r}^2_{by\_snr}}{\hat{r}^2_{by\_snr} + 0.01} r \tag{13}$$

Estimate $\hat{r}$ will be used as reproducibility measure for pass/fall decision.

2.4. The use of inter-subject variability of TEOAE latency

In [19], it is shown that inter-subject variability of TEOAE latency is approximately ±10%. The correct position of the analysis window adjusted in accordance to the subject latency increases reproducibility and therefore the specificity of TEOAE test. Measurements and detailed analysis of TEOAE latency on the wider population were carried out by [5], using the wavelet transformation. However, accurate estimation of TEOAE latency during screening test is quite difficult in the real conditions with ambient noise especially for patients with low TEOAE. The errors of window position estimate additionally increase the variance of the reproducibility. Therefore, this paper proposes a coarse and robust procedure for selection of the windows positions, which increases reproducibility with less influence to its variance. The central position of the window is taken from [6,12], while the other two windows are moved to ±10 % relative to the central window position. The rule for the selection of the windows is simple: Select the window with maximal reproducibility.

2.5. The optimization of the window shape

In TEOAE analysis the commonly used window is Tukey [20,6,12] with 2.5 ms onset and offset. The idea is to apply more accurate window shape according to assumed energy contour of TEOAE response and to calculate the near-optimal estimate of TEOAE response in the minimum mean square error sense. Assume that the energy contour of TEOAE response $P_{anv}(t,j)$ can be modeled by

$$P_{anv}(t,j) = P_{\max}(j)h(t,j), \tag{14}$$

where $h(t,j)$ is normalized energy contour of scale j, such that $\max_t h(t) = 1$, $P_{\max}(j)$ is maximal (short time) power of the scale j.

By using observation model (5), our aim is to define weight function $w(t,j)$ that is to be applied to ensemble average

$$\hat{s}(t,j) = w(t,j)\,\bar{x}_q(t,j) \tag{15}$$

such that the mean square error between the original signal $s(t,j)$ and its estimate $\hat{s}(t,j)$ is minimal. The solution for $w(t,j)$ is

$$w(t,j) = \frac{\rho_w(t,j)}{\rho_w(t,j)+1}, \tag{16}$$

where $\rho_w(t,j)$ is instantaneous signal-to-noise ratio in ensemble averaged signal. Taking into account model (14) and assumption that the noise power is constant in time (6), signal to noise ratio is

$$\rho_w(t,j) = \rho_{\max}(t,j)h(t,j) \tag{17}$$

By substituting (17) into (16) optimal weight function is

$$w(t,j) = \frac{\rho_{\max}(j)h(t,j)}{\rho_{\max}(j)h(t)+1}. \tag{18}$$

Regarding the selection of $h(t)$, we tested several functional dependencies. Despite the fact that some authors used gammatone functions for modeling the TEOAE response [5,21], the envelope of this function did not provide the best results. The reason for this may be that gammaton function has an infinite response, while our window is time-limited. The best experimental results were obtained when the envelope of TEOAE signal energy is modeled with the shape that has a rapidly cosine raised type of growth at the beginning, on ¼ width of the window, which is then followed by a slow decline on the ¾ width of the window. Fig. 3a displays envelopes $h(t,j)$, for j = 5,6,7 scale, while Fig. 3b displays the corresponding window functions $w(t,j)$ for $\rho_{\max} = 4$ that remains fixed for all three scales. The position and the width of the windows are maintained the same as in [6,12].As $\rho_{\max}(j)$ is unknown, it is possible to estimate it from available measurements. We did not do that because estimation error increases the final variance of the calculated reproducibility. Instead, we used fixed value which is optimal for the case when SNR is around the threshold. In some of the screening test protocols [1,4], the threshold for the decision based on SNR is 3 dB. Taking into account that maximal SNR is 3 dB over its averaged value, we set $\rho_{\max}(j) = \rho_{\max} = 4$, $(10\log_{10}4 = 6\,dB)$.

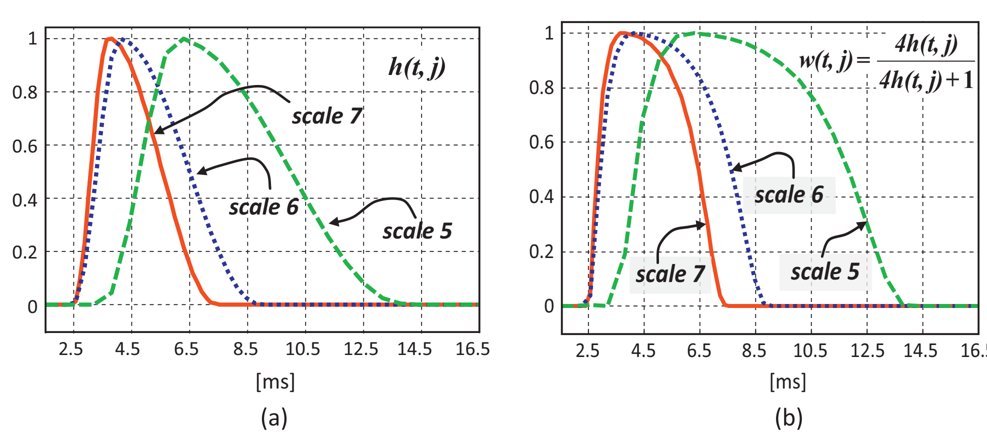

Fig. 3. (a) The shape of the energy contours $h(t,j)$ for scale 5 (green dashed line), scale 6 (blue dotted line) and scale 7 (red solid line). (b) Optimal window function $w(t,j)$, for scale 5 (green dashed line), scale 6 (blue dotted line) and scale 7 (red solid line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Performance of the algorithms assessed by their reproducibility.

	Case 1				Case 2			Number of members in overlapped area
	Average [%]	Std [%]	Min [%]	<50%	Average [%]	Std [%]	Max [%]	
Alg 1.	78.18	14.79	23.79	40	0.20	16.15	44.44	30 + 75 = 105
Alg 2.	78.93	13.30	30.86	37	0.42	13.07	40.13	18 + 12 = 30
Alg 3a	79.35	12.07	41.72	27	−1.65	9.56	33.57	*No overlap*
Alg 3b	80.08	11.84	41.72	27	1.42	8.91	37.44	*No overlap*
Alg 3c.	81.69	10.69	50.02	–	0.72	8.27	33.87	*No overlap*

3. Results

In order to evaluate performance of the proposed method for TEOAE detection, we tested following algorithms under the same conditions:

Alg 1: Januškausas 2001 [6] - DWT with specified window positions
Alg 2: Yang 2002 [12] - improved Alg 1 by averaging of $N = 4$ paired buffers
Alg 3: Šarić 2016
  Alg 3a - improved Alg 2 by simple linear regression model of SNR estimates
  Alg 3b - improved Alg 3a by subject dependent windows positions
  Alg 3c - improved Alg 3b by optimized shape of the window function, i.e. complete proposed algorithm Alg 3.

The procedure of wavelet based denoising [6] was omitted in [12], so that its uncertainty effect on the variance could then be ignored. In order to compare both algorithms [6] and [12] with our algorithm we also omitted denoising. The algorithms are tested in two test cases. In the first test case, we used 46 real TEOAE measurements embedded in the set of 920 realizations of noise synthesized by the method proposed in [12]. The aim of this test case is to compare performances of the tested algorithms related to their specificities. Results are displayed in Table 1, columns 1–4. Averaged reproducibility and its standard deviation are displayed in columns one and two for each algorithm. For non-Gaussian distributions, the mean and standard deviation don't provide sufficient information about the performance of detection algorithm. So, we examined two additional features. In column three we presented minimal reproducibility of 920 tested measurements. Finally, in column four there is the number of test examples with reproducibility under 50%, e.g. the number of miss-detected TEOAEs when the decision threshold is 50%. Fig. 4f–j display histograms of reproducibility measure for algorithms Alg 1, Alg 2 and Alg 3a–c respectively.

The aim of the second test case is to assess sensitivities of tested algorithms. Algorithms were tested on a set of 920 realizations of the noise synthesized by the method proposed in [12]. In our experiments, all test examples had reproducibility less than 50% so the sensitivity was 100% for all algorithms. To compare its sensitivity under this circumstance, we investigated several features: (a) averaged reproducibility, (b) its standard deviation and (c) maximal reproducibility of the examined test examples. The results are displayed in Table 1 in columns 5–7. Fig. 4a–e display histograms of reproducibility for algorithms Alg 1, Alg 2 and Alg 3a–c respectively.

To compare separation of the classes "there is TEOAE" (case 1) and "there is no TEOAE" (case 2), we examined the overlapping of the histograms obtained by tested algorithms. Column 8 of Table 1 shows a number of test examples with reproducibility between the maximal value obtained in the case 2 and the minimal value obtained in the case 1. Overlapping parts for Alg 1 and Alg 2 are graphically displayed in Fig. 4a, f, and Fig. 4b, g respectively. Histograms for Alg 3a–c displayed in Fig. 4c,f show that the classes are completely separated.

Statistically significant difference of the results obtained in test case 1 by algorithms Alg 1, Alg 2 and Alg 3c was evaluated by ANOVA test using SPSS software package. Single factor ANOVA test revealed that there is statistically significant difference ($F(2,2757) = 18.438$; $p < 0.001$) in the reproducibility obtained by the compared algorithms. In order to evaluate pairwise statistical difference of the tested algorithms, the post hoc test with Bonferroni correction of α, ($\alpha = 0.05$) was applied. Fig. 5 shows the mean values and standard deviations of the three compared algorithms with statistically significant differences between them. Post hoc test found significantly higher reproducibility for the Alg 3c when compared to both Alg 1 and Alg 2.

4. Discussion

In the test case 1, with normal hearing subjects, we examined averaged reproducibility as the first indicator of the quality of TEOAE detection. Column 1 of Table 1 shows that proposed algorithm Alg 3c provides larger averaged reproducibility in comparison with algorithms Alg 1 and Alg 2 (81.69 % compared with 78.18 and 78.93 %).

To evaluate the individual contributions of the each method we examined their incremental contributions by successively introducing one method after another. The advantage of this approach is that it takes into account possible interconnections among the proposed methods. From the data displayed in column 1 of Table 1 we see that the proposed methods included in algorithms Alg 3a, Alg 3b and Alg 3c incrementally increase averaged reproducibility by 0.42% (79.35–78.93), 0.73% (80.08–79.35) and 1.61 % (81.69–80.08) respectively. We conclude that each of the proposed methods increases averaged reproducibility.

As the second indicator of algorithms accuracy, we examined standard deviation of the estimated reproducibility. The smaller standard deviation indicates the better accuracy of the estimated reproducibility [22]. The standard deviation of reproducibility calculated by Alg 3c in test case 1 is less than those calculated by algorithms Alg 1 and Alg 2 (10.69 % in comparison with 14.79 % and 13.30 %, column 2, Table 1). From the data displayed in column 2 of Table 1, the differential decrease of reproducibility variance obtained by algorithms Alg 3a, Alg 3b and Alg 3c are 1.23% (13.30–12.07), 0.23 % (12.07–11.84) and 1.15% (11.84–10.69) respectively. Hence, we conclude that all three proposed methods contribute to the variance reduction. The larger averaged reproducibility and its smaller standard deviation indicate the large specificity of TEOAE test i.e. the smaller probability of misdetection of TEOAE.

As the distribution of reproducibility is not strictly Gaussian we examined two additional features: (*i*) *minimal* reproducibility (Table 1, column 3) and (*ii*) the number of test examples with reproducibility under threshold less than 50 % (Table 1, column 3). The larger minimal reproducibility of algorithm Alg 3c indicates its larger specificity. The column four of Table 1 also indicates

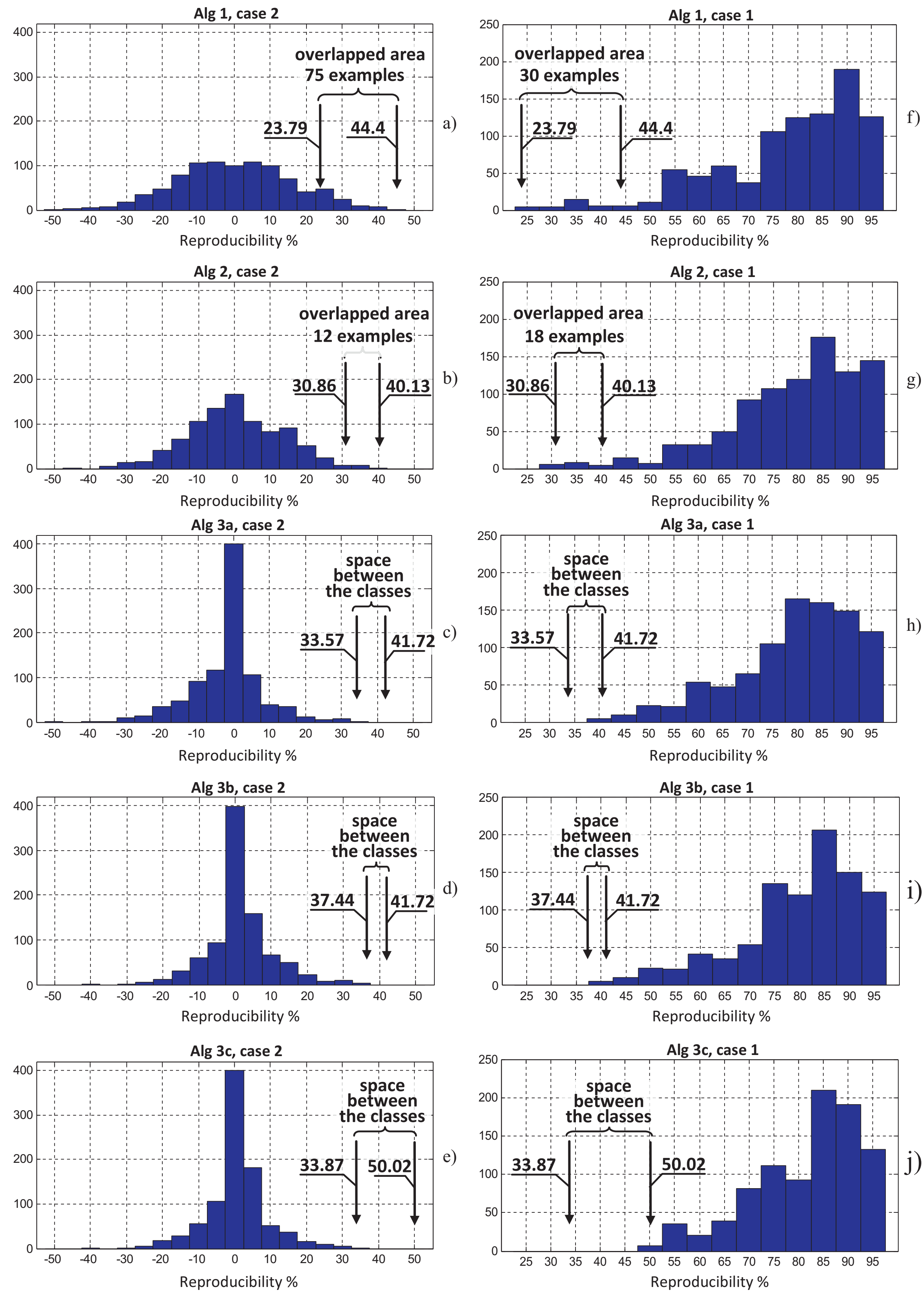

Fig. 4. Reproducibility histograms: (a) and (d) for Alg 1; (b) and (e) for Alg 2; (c) and (f) for Alg3.

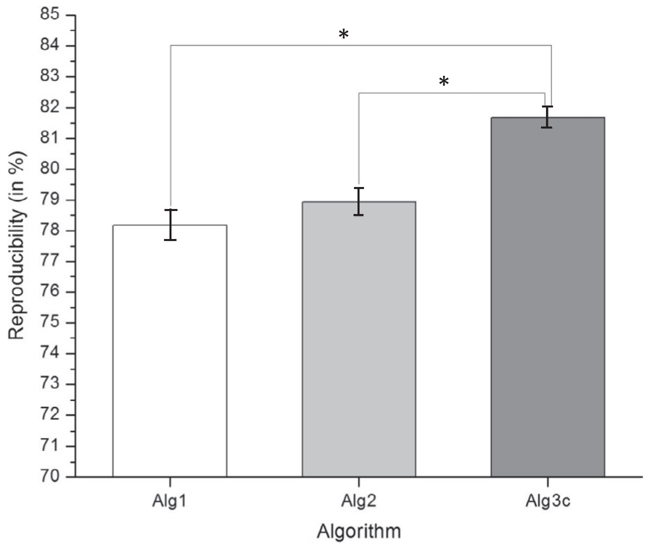

Fig. 5. Mean values and standard deviations of reproducibility for algorithms Alg 1, Alg 2 and Alg 3c. Asterisk denotes that the mean difference is significant at the 0.05 level.

larger specificity for algorithm Alg 3c, as there is no example with reproducibility under 50%. For the sake of comparison, algorithms Alg 1 and Alg 2 had respectively 40 and 37 potentially misclassified examples under threshold 50%. Taking into account larger averaged reproducibility, its smaller standard deviation, larger minimal reproducibility and absence of the examples under 50% we conclude considerable improvement of specificity for algorithm Alg 3c in comparison with algorithms Alg 1 and Alg 2.

In the test case 2, tested algorithms have no false TEOAE detection. So, we couldn't directly compare its sensitivities. To overcome this problem, we compared the features that are related to algorithms' sensitivities as (*i*) averaged reproducibility, (*ii*) its standard deviation and (*iii*) maximal reproducibility in a population of 920 pure noise examples. Algorithm Alg 3c had a much lower standard deviation in comparison to the algorithms Alg 1 and Alg 2 which is displayed in column six of Table 1, as well as in histograms Fig. 4a–e. The smaller standard deviation indicates larger sensitivity, e.g. the smaller probability of false TEOAE detection. As in the case 1, we checked whether the each method contributes to the reduction of reproducibility standard deviation. From the data displayed in column 6 of Table 1, we calculated differential decrease of the standard deviation. The results obtained by algorithms Alg 3a, 3b and 3c are 3.41 % (13.07–9.56), 0.65% (9.56–8.91) and 0.64% (8.91–8.27) respectively. Having in mind the results for the test cases 1 and 2 it seems reasonable to include all three methods in TEOAE detection algorithm, i.e. to apply algorithm Alg 3c.

A little increase in averaged reproducibility for Alg 3c (0.72%) in comparison to 0.20% and 0.42% for algorithms Alg 1 and Alg 2, column five, has negligible influence on the sensitivity of the test having in mind influence of the reduction of its standard deviation. Algorithm Alg 3c also had the lower maximal reproducibility. For the comparison sake, maximal reproducibility for algorithms Alg 1 and Alg 2 were 44.44 and 40.13 % respectively in comparison to 33.87 % for Alg 3c, column seven, Table 1. This is the consequence of using the method described in subSection 2.3. The smaller maximal reproducibility indicates the larger sensitivity of the test. So, we can conclude that results obtained in test case 2 confirm the better sensitivity of algorithm Alg 3c.

Based on the analyses of experimental results, the performance of the proposed method of TEOAE detection, Alg 3c, surpasses the performance of algorithms Alg 1 and Alg 2 in respect of all analyzed quality indicators, except for the slight increase of average value in the case of pure noise. As an explicit indication of increased separability of classes "there is TEOAE" and "there is no TEOAE", column eight, Table 1, displays the number of class members in the overlapping parts of histograms, for each algorithm. While the populations of the classes are overlapped in the case of algorithms Alg 1 and Alg 2, by 105 and 30 test examples respectively, the algorithm Alg 3c provides complete separation of the classes (Fig. 4).

Statistical analysis displayed in Fig. 5 showed that the differences in reproducibility obtained by algorithm Alg 3c and the other two algorithms Alg 1 and Alg 2 are statistically significant by ($p = 0.000$).

5. Conclusions

In this paper, three methods for improvement of TEOAE detection algorithm were proposed. The first one is the correction of calculated reproducibility based on the linear fitting of the sequence of calculated SNRs in term of a number of processed subaveraged responses. The second proposed method exploits inter-subject variability of TEOAE latency. The third proposed method modifies weight function in accordance with the assumed energy contour of TEOAE response and optimizes it for specific SNR threshold. Conducted experiments proved that each of the proposed methods and their combination increase reproducibility in the case 1 and reduce its variance in both cases 1 and 2, and hence, increase both sensitivity and specificity of the TEOAE test.

In the future work the proposed algorithm for TEOAE screening test may be optimized for reduction of the measuring time while keeping the high specificity and sensitivity of the test.

6. Ethical approval

"All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institute for experimental phonetics and speech pathology (Belgrade, Serbia), and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

Acknowledgements

This research was supported by grants 178027 and TR32032 from the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] Kemp DT, Ryan S. The use of transient evoked otoacoustic emissions in neonatal hearing screening programs. InSeminars in Hearing 1993 Feb (Vol. 14, No. 01, pp. 30–44). Copyright© 1993 by Thieme Medical Publishers Inc.
[2] Robinette MS, Glattke TJ, editors. Otoacoustic emissions: clinical applications. New York: Thieme; 1997 Mar.
[3] Zhang VW, McPherson B. A review of otoacoustic emission hearing screening technology. Audiolog Med 2008;6(2):100–14.
[4] Whitehead ML, Jimenez AM, Stagner BB, McCoy MJ, Lonsbury-Martin BL, Martin GK. Time-windowing of click-evoked otoacoustic emissions to increase signal-to-noise ratio. Ear Hear 1995;16(6):599–611.
[5] Tognola G, Grandori F, Ravazzani P. Time-frequency distributions of click-evoked otoacoustic emissions. Hear Res 1997;106(1):112–22.
[6] Janušauskas A, Marozas V, Engdahl B, Hoffman HJ, Svensson O, Sörnmo L. Otoacoustic emissions and improved pass/fail separation using wavelet analysis and time windowing. Med Biol Eng Compu 2001;39(1):134–9.
[7] Tognola G, Grandori F, Ravazzani P. Wavelet analysis of click-evoked otoacoustic emissions. IEEE Trans Biomed Eng 1998;45(6):686–97.
[8] Jedrzejczak WW, Blinowska KJ, Konopka W, Grzanka A, Durka PJ. Identification of otoacoustic emission components by means of adaptive approximations. J Acoust Soc Am 2004;115(5):2148–58.
[9] Mallat SG, Zhang Z. Matching pursuit with time–frequency dictionaries. IEEE Trans Signal Process 1993;41(12):3397–415.

[10] Jedrzejczak WW, Hatzopoulos S, Martini A, Blinowska KJ. Otoacoustic emissions latency difference between full-term and preterm neonates. Hear Res 2007;231(1):54–62.
[11] Janusauskas A, Sornmo L, Svensson O, Engdahl B. Detection of transient-evoked otoacoustic emissions and the design of time windows. IEEE Trans Biomed Eng 2002;49(2):132–9.
[12] Yang LP, Young ST, Kuo TS. Modification of the wavelet method used in transiently evoked otoacoustic emission pass/fail criterion to increase its accuracy. Med Biol Eng Compu 2002;40(1):34–40.
[13] Subotić M, Šarić Z, Jovičić ST. Adaptive filter based two-probe noise suppression system for transient evoked otoacoustic emission detection. Ann Biomed Eng 2012;40(3):637–47.
[14] Daubechies I. Ten Lectures on Wavelets (Philadelphia, PA: Society for Industrial and Applied Mathematics). DaubechiesTen Lectures on Wavelets 1992. 1992.
[15] Bray PJ. Click evoked otoacoustic emissions and the development of a clinical otoacoustic hearing test instrument. London University (Doctoral dissertation, Ph. D. thesis). 1989.
[16] Prieve BA, Gorga MP, Schmidt A, Neely S, Peters J, Schultes L, et al. Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears. J Acoust Soc Am 1993;93(6):3308–19.
[17] Welzl-Müller K, Stephan K. Confirmation of transiently evoked otoacoustic emissions based on user-independent criteria. Audiology 1994;33(1):28–36.
[18] Soong TT. Fundamentals of probability and statistics for engineers. John Wiley & Sons; 2004.
[19] Harte JM, Pigasse G, Dau T. Comparison of cochlear delay estimates using otoacoustic emissions and auditory brainstem responses. J Acoust Soc Am 2009;126(3):1291–301.
[20] Whitehead ML, Stagner BB, Lonsbury-Martin BL, Martin GK. Measurement of otoacoustic emissions for hearing assessment. IEEE Eng Med Biol Mag 1994;13 (2):210–26.
[21] Zhang ZG, Zhang VW, Chan SC, McPherson B, Hu Y. Time–frequency analysis of click-evoked otoacoustic emissions by means of a minimum variance spectral estimation-based method. Hear Res 2008;243(1):18–27.
[22] Tognola G, Grandori F, Ravazzani P. Data processing options and response scoring for OAE-based newborn hearing screening. J Acoust Soc Am 2001 Jan 1;109(1):283–90.