Archive of SID

Aud Vest Res (2016);25(4):207-214.

RESEARCH ARTICLE

Acoustic feedback suppression in binaural hearing aids

Fatmah Tarrab[1*], Zouheir Marmar[1], Isam Alamine[2]

[1]- Department of Biomedical Engineering, Faculty of Electrical and Mechanical Engineering, Damascus University, Damascus, Syria

[2]- Department of Otorhinolaryngology, Faculty of Medicine, Damascus University, Damascus, Syria

Received: 1 Sep 2016, Revised: 23 Sep 2016, Accepted: 27 Sep 2016, Published: 29 Nov 2016

Abstract

Background and Aim: The hearing loss and acoustic feedback signal in the binaural hearing aids may annoy users, make speech reorganization difficult, and affect the sound quality. This paper presents a new method for acoustic feedback suppression in binaural hearing aids, based on the combination of binaural information and blind source separation. This method can enhance sound from the specific direction of the user and cancels acoustic feedback signal so that the user can hear clearly.

Methods: In the proposed method, the binaural information (interaural time differences (ITDs) and interaural level differences (ILDs)) was generated using computational auditory scene analysis (CASA). In addition, we used underdetermined blind source separation (BSS) for the automatic classification of the time-frequency (T-F) units of the speech mixture spectrogram. The system performance was evaluated using 32 acoustic English speech mixtures. The sound quality was assessed using 19 normal-hearing listeners of both genders.

Results: The system achieved a good acoustic feedback suppression performance by giving a higher signal to noise ratio and 18.07 dB on average for the signal-to-feedback ratio. In addition, the sound after processing had a high quality according to the subjective assessment.

Conclusion: Our system allowed the user to increase the gain of the hearing aid without affecting the sound quality.

Keywords: Source separation; computational auditory scene analysis; acoustic feedback; hearing aids

[*] **Corresponding author:** Department of Biomedical Engineering, Faculty of Electrical and Mechanical Engineering, Damascus University, Damascus, Syria. Tel: 00963-938616542,
E-mail: eng.fatmah.tarrab@gmail.com

Introduction

Many people with hearing loss use binaural hearing aids to keep the directional selectivity of the auditory system and to compensate the binaural imbalance. The advantages of binaural hearing aids are its effectiveness on increasing the binaural sound pressure and improving the binaural speech-to-noise ratio [1,2]. The performance of binaural hearing aids decreases when acoustic feedback from the receiver (earphone) to sensor (microphone on hearing aid) occurs due to the violation of Nyquist stability criterion (a graphical technique for determining the stability of a dynamical system) at one or more frequencies. This phenomenon is known as howling or squealing of aids. It limits the maximum gain, affects speech understanding and acoustic comfort, and may cause inability to hear especially when the gain of the hearing aid is increased [2-5]. Therefore, it is important to separate acoustic feedback from a speech signal prior to further processing.

Researchers have used the adaptive feedback canceller (AFC) to solve the howling problem [4,6,7]. Aforementioned method cancels the

www.SID.ir

Archive of SID

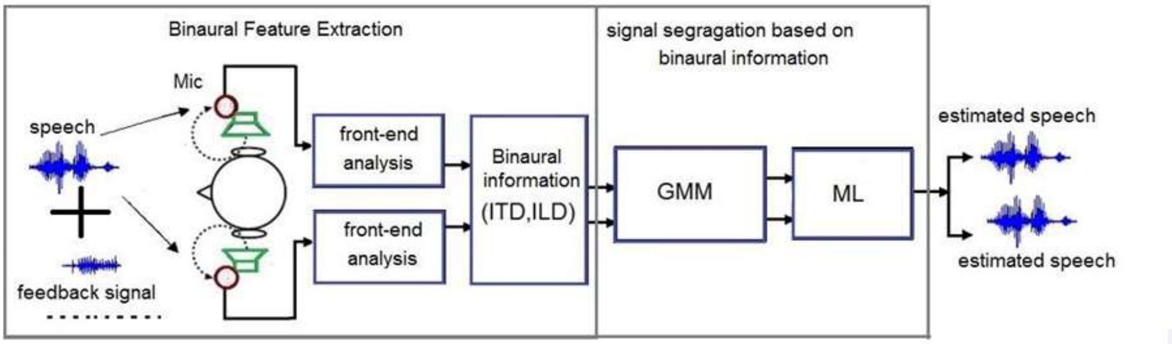

Fig. 1. The block diagram of proposed system for feedback suppression in binaural hearing aids.

feedback components by the use of an adaptive procedure which estimates the feedback path. The coefficients of the adaptive filter are updated at the same time when the speech signal is delivered to the user so the feedback loop causes some correlations between the input speech and the feedback signal. As a result, the adaptive filter has a bias problem; this bias severely limits the performance of the feedback cancellation system and reduces the speech quality. However, the bias problem is still an issue in spite of different proposed solutions [6,8].

Because the howling on hearing aids occurs at the specific frequency or frequencies for each of channels separately, the binaural information can be used to segregate the speech signal which must be emphasized from the feedback components causing howling on the assumption that the frequencies to be howled for each channel are not the same at a time. Based on this idea, this paper presents a new method for acoustic feedback suppression in binaural hearing aids by using the combination of binaural information and blind source separation. In the proposed system, at first, we generate the binaural information (interaural time differences (ITDs) and interaural level differences (ILDs)) using computational auditory scene analysis (CASA) that is inspired by the human auditory system [9,10]. Then, for each source, we compute the probability density functions (PDFs) of time-frequency (T-F) units by modeling the binaural feature space using Gaussian mixture model. After that, we use underdetermined blind source separation (BSS), which is a promising approach to deal with multiple overlapping sources, by only considering T-F units that are believed to be dominated by the target [11,12]. An important advantage of this approach is that the number of sources can be equal to or exceed the number of microphones, which are usually four in binaural hearing aids.

Methods

We can estimate the direction-of-arrival (DOA) of sound based on ITD and ILD. ITD is defined for the frequency range around 1500 Hz or less and ILD for the frequency range around 1000 Hz or above. It should be note that estimation on overlapped frequency range depends on both ITD and ILD.

Based on ITD and ILD information, we can separate the concurrent speeches with difference DOAs and concentrate the specific speech even if the surrounding noise is loud.

In this paper, we proposed a new method for acoustic feedback suppression in binaural hearing aids, based on the combination of binaural information (ITD, ILD) and blind source separation. The proposed method performed speech separation without using pre-trained speaker models; instead, it used the binaural information that was independent of the signal structure and available from a mixture signal. In addition, it did not need a voice-activity detector; thus, it was capable to restore original sources through the information available from microphones only. The proposed system is shown in Fig. 1

www.SID.ir

Archive of SID

where the acoustical and possibly mechanical feedback is illustrated by rounded arrows from loudspeakers to microphones. The proposed system for acoustic feedback suppression is constructed with two major parts as follows: 1) binaural feature extraction (ITD and ILD), and 2) signal segregation based on the binaural information.

Binaural feature extraction (ITD and ILD)

Assumption based on an assumption, the binaural signals of each microphone in the binaural hearing aids consist of speech and feedback signals.

We generated the binaural information based on CASA, which used binaural front-end analysis to simulate the processing of the human auditory system.

Binaural front-end analysis

The both input binaural signals, left and right mixture, observed by microphones attached to hearing aids were sampled at a rate of 16 kHz, then split into auditory channels using a bank of Q=128 gammatone filters that cover the range of 50 to 8000 Hz because the sampling rate is set to 16 kHz. The processing of the inner hair cells is simulated by half-wave rectification and square-root compression [9]. The resulting binaural auditory signals of the left and the right ith gammatone channel are represented by (l_i) and (r_i), respectively.

Interaural time and level differences were independently estimated for each auditory channel using overlapping frames of 20 ms with a 10 ms shift to capture rapid changes of the feedback signal.

Estimating of interaural time differences (ITD)

The time difference between the binaural auditory signals in the ith gammatone channel is estimated using the normalized cross-correlation function ($R_i(t,\tau)$),which is defined in equation (1) [13].

$$(1)\ R_i(t,\tau) = \frac{\sum_{k=0}^{k-1} \bar{l_i}(t.k\ /2-k)\ \bar{r_i}(t.k\ /2-k-\tau)}{\sqrt{\sum_{k=0}^{k-1} \bar{l_i}^2(t.k\ /2-k)}\sqrt{\sum_{k=0}^{k-1} \bar{r_i}^2(t.k\ /2-k-\tau)}}$$

Where:

i=1, …, 128 channels.

τ: time lag.

t: frame number.

$(\bar{l_i})$, $(\bar{r_i})$: denote the mean values of the left and right auditory signals.

The mean values of the left and right auditory signals were estimated over the frame number by finding the peak value of $R_i(t,\tau)$ through equation (2) to obtain the time lag $C_i(t)$ corresponding to the estimation of ITD.

$$(2)\ C_i(t) = \arg\max R_i(t,\tau)$$

In order to increase the ITD accuracy while keeping the computational complexity moderate, we applied exponential interpolation, as defined in equation (3).

$$(3)\ \delta_i(t) = \frac{\log R_i(t,C_i(t)+1) - \log R_i(t,C_i(t)-1)}{4\log R_i(t,C_i(t)) - 2\log R_i(t,C_i(t)+1) - 2\log R_i(t,C_i(t)-1)}$$

The maximum ith corresponds to the estimated ITD (as in samples) that was obtained by equation (4):

$$(4)\ ITD_i(t) = (C_i(t)+\ \delta_i(t))/Fs$$

Where:

$ITD_i(t)$: The ITD information.

$C_i(t)$: The real part of estimated ITD.

$\delta_i(t)$: The imaginary part of estimated ITD.

Fs: sampling frequency.

Estimating of interaural level differences (ILD)

The ILD was estimated by comparing the energy integrated across the time interval W between the left and right binaural signals. The ILD information in the ith gammatone channel expressed in dB was obtained by equation (5).

$$(5)\ ILD_i(t) = 20\log_{10}\left\{\frac{\sum_{k=0}^{k-1} r_i^2(t.k\ /2-k)}{\sum_{k=0}^{k-1} l_i^2(t.k\ /2-k)}\right\}$$

A sound source positioned at the left hand side would result in a negative ILD, whereas a positive ILD would be caused by a source lateralized to the right hand side.

Binaural information for sound source

According to the probabilistic nature of the binaural information, Gaussian mixture models (GMMs) modeled a combination of ITDs and ILDs as the binaural feature space by equation (6):

www.SID.ir

Archive of SID

Table 1. Descriptions of TIMIT database [15]

Speaker ID	Utterance
FS1	"What movies have you seen recently"
FS4	"Only the best players enjoy popularity"
FS6	"Development requires a long-term approach"
FS8	"Change involves the displacement of form"
MS2	"Singapore is a fine place"
MS5	"Our aim must be to learn as much as to teach"
MS10	"Most assuredly ideas are invaluable"
MS11	"False ideas surfeit another sector of our life"

MS; male speaker, FS; female speaker

(6) $P(x\backslash\lambda_i) = \sum_{q=1}^{Q} Л_{i,q}\, N\left(X; \mu_{i,q}, \Sigma_{i,q}\right)$

Where:
(X): denotes the observed signal.
(λ): denotes the specific direction of speech source.
(μ, Σ): denotes the mean and covariance matrix of the qth Gaussian component of the mixture model. (Л): denotes the mixture weight (probability) of the qth distribution.
The qth component density is represented by:

(7) $N(X; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^D\, |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (X-\mu)\right)$

Where: $(-\infty < x < +\infty,\ \Sigma_{i,q} > 0,\ Л_{i,q} > 0,\ \sum_{q=1}^{Q} Л_{i,q} = 1$

Signal segregation based on binaural information

The Gaussian mixture model parameters and the assigned regions of the time-frequency (T-F) units were refined iteratively using the Expectation-Maximization (EM) algorithm [12,14]. A maximum likelihood estimator (MLE) was applied to determine the localization of T-F units and estimate posterior probabilities of the dominant Gaussian to detect the target speech from the binaural signal. Then, the mask was set to the posterior probabilities of the dominant Gaussian components as follow:

(8) $M_i(t,f) = p(x\backslash\lambda)$

The spatial image estimate of the nth signal received at the mth sensor was then obtained through the application of mask Mn to the mth observation as shown in equation (9):

(9) $\widehat{s_{mn}}(t,f) = M_i(t,f) \,.\, X_m(t,f)$

Finally, the estimated source images were reconstructed in the time-domain by Inverse Short Time Fourier Transformation (ISTFT) to obtain the estimates and individual frequency components of the recovered signal.

Evaluation of the performance of the proposed system

To evaluate the performance of the proposed system, we generated 32 binaural mixtures consisting of real speech and feedback signals under four different gain conditions. In all cases, a real speech source was randomly selected from TIMIT database as given in Table 1 with the length of approximately 2 seconds and from both genders (4 male and 4 female) [15].
The acoustic feedback signal is a real speech filtered by the impulse response of the measured feedback path, shown in Fig. 2, and then amplified by different gains from 20 dB to 80 dB to get 32 binaural mixtures of target speech and feedback.
Our system was implemented in MATLAB and the experiments were run on an Intel Core 2 Duo (2.00 GHz) processor with 2 GB RAM. The quality of the separated signals was

www.SID.ir

Archive of SID

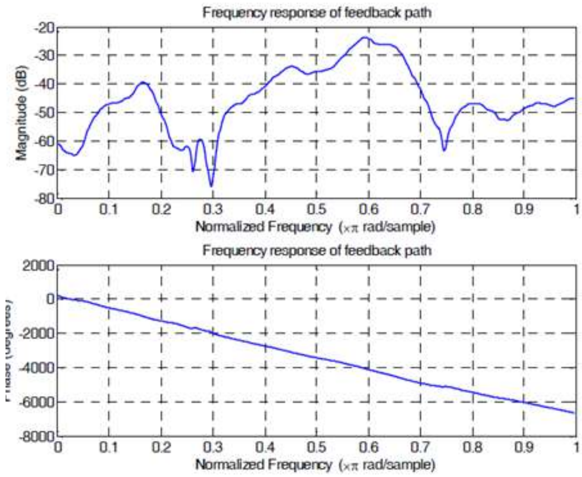

Fig. 2. The impulse response of the measured feedback path.

measured by calculating the signal to noise ratio (SNR) using equation (10) and also the improvement of the signal-to-feedback ratio (SFR) [16,17].

(10) $\mathrm{SNR} = 10\,\log10 \frac{\|s_{target}\|^2}{\|s_{target} - s_{estimate}\|^2}$

To measure the improvements in the SFR, we calculated measured SFR before applying our method using equation (11) and after applying the method using equation (12) and then, SFR improvements were quantified through equation (13).

(11) $\text{SFR-before} = 10\,\log10 \frac{E[S]^2}{E[\mathrm{F}]^2}$

(12) $\text{SFR-after} = 10\,\log10 \frac{E[S]^2}{E[\mathrm{F} - \widehat{F}]^2}$

(13) SFR improvement = SFR (after) - SFR (before)

A subjective test was conducted using headphone on 19 normal hearing listeners of both genders (11 male and 8 female subjects) with a mean age of 5.8±0.6 years for children and 44.8±13 years for adults. The listener's normal Peripheral hearing was ensured through screening. Our assumption was that if the sound distortion is acceptable for normal hearing people, then it will also be acceptable for hearing impaired people. At the beginning of each test, the laptop and headphone sound volume was controlled to ensure a comfortable hearing.

Table 2. Descriptions of mean opinion scores (MOSs)

MOS	Quality	Description of Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

The scores are adopted from Viswanathan and Wiswanathan [18]

Then, test subjects rated each test signal on a sound quality scale [18], called the mean opinion scores (MOSs) in the range 1−5, as given in Table 2.

The paired samples t-test was then performed to compare the signal-to-feedback ratio before and after applying our proposed system. In addition, the ANOVA was performed to compare sound quality by different gains after applying our method. The obtained data were analyzed using SPSS 19.

Results

The proposed system was implemented in MATLAB with four different gain conditions (20, 40, 60, and 80 dB) using 32 binaural mixtures consisting of speech and feedback sources to measure the quality of the separated signals. The results are shown as follows:

Quality of the separated signals

In Fig. 3, we showed the quality of the separated target speech in terms of the SNR for eight target speech signals under each of the four different gain conditions.

As one would expect, our method was able to achieve much larger SNR for mixtures with speech and feedback sources for all four different gain conditions because it was able to discriminate between speech and feedback signal by binaural information, which was independent of signal structure.

In Fig. 4, we also showed the quality of the

www.SID.ir

Archive of SID

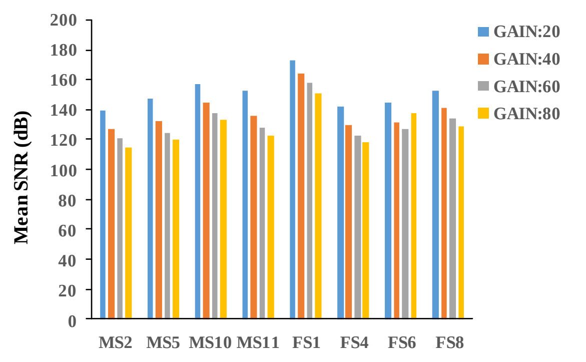

Fig. 3. Mean signal-to-noise ratio (SNR) for binaural mixtures of four gain conditions. MS; male speaker, FS; female speaker

separated feedback signal in terms of SFR improvement for 32 binaural mixtures consisting of speech and feedback sources for each of the four different gain conditions.

The proposed system also achieved higher SFR for mixtures with speech and feedback sources under all four different gain conditions.

The results of the paired samples t-test in Table 3 showed statistically significant differences between the quality of sound in the change of the SFR before and after applying our method

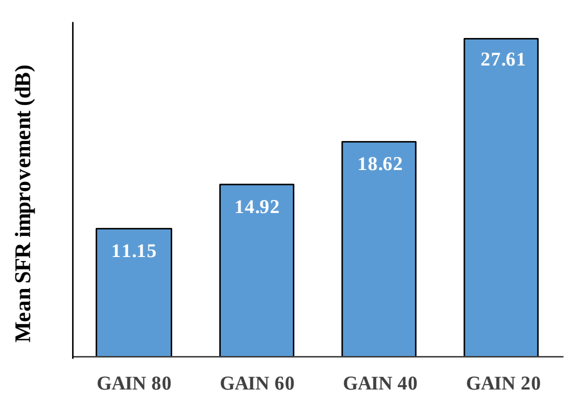

Fig. 4. Mean signal-to-feedback ratio (SFR) improvement for four gain conditions.

for each of the four different gain conditions ($p<0.05$).

Subjective test for the separated sound quality

The separated target speech quality was also assessed by subjective test on (19) normal listeners of both genders using MOSs. We generated 20 binaural mixtures consisting of speech and feedback sources for four different gain conditions.

The ANOVA showed no significant difference between MOSs obtained in any of the four subtests ($p>0.05$) and our proposed method achieved a high sound quality in the subjective test results. This finding illustrated that increasing gain of the hearing aid does not have any effect on sound quality because ILD contribution reduces at higher gain and thus the mixing vectors provide more distinct information without any effect on the results.

Discussion

According to Figures 3 and 4, the speech in noise (SNR) and signal-to-feedback ratio (SFR), in sequence, were the largest at the lower gain values where the feedback signal was not high in the mixture. As gain increases, the feedback signal of the mixture increases; as a result, the SNR and SFR improvements decrease although they are still high and acceptable.

The adaptive algorithm proposed in [4] was capable of achieving improvements on average in SNR and SFR due to the combination of a beamformer and feedback canceller with dual microphones, but the system has not been assessed well under different gain conditions and the feedback and noise canceller updated their coefficients only when instability was detected.

The adaptive system proposed in Lombard et al. [6] achieved good signal enhancement and a new stability condition was adapted to the considered binaural scenario by combining adaptive feedback cancellation (AFC) and adaptive binaural filtering (BAF), but a steady-state analysis showed that the AFC suffered from a bias in its optimum (Wiener) solution. This bias, similar to the unilateral case, was due to the correlation between feedback and external source signals.

www.SID.ir

Archive of SID

Table 3. Mean (variance) signal-to-feedback ratio (SFR) of four gain condition before and after application of the present method

	Mean (variance) SFR (dB)		
Gain (dB)	**Before**	**After**	**p**
20	-20.63 (6.22)	6.98 (60.44)	0.000
40	-26.65 (6.22)	-8.03 (45.86)	0.000
60	-30.17 (6.22)	-15.25 (53.86)	0.000
80	-32.67 (6.22)	-21.52. (40.85)	0.000

SFR; signal to feedback ratio, p; probability

The correlation between feedback and external source signals caused a bias of the BAF solution, but contrary to the bias encountered by the AFC, the BAF bias increased with increasing hearing aid amplification levels.

We can see that the performance of the proposed method was consistently better than the adaptive algorithm on the basis of SNR and SFR. It achieved higher SNR and higher SFR (18.07 dB) on average compared to the SNR and SFR (14.43 dB) of the feedback canceller in the adaptive algorithm [4]. Therefore, it achieved advances in terms of SFR on average 3.64 dB in comparison with the adaptive algorithm. In addition, the proposed method performed feedback suppression without using training mode or voice-activity detector; thus, it restored original speech through the information available from microphones only.

Our system achieved a good acoustic feedback suppression performance in comparison with other studies in which AFC was used [4,6] and allowed the user to increase the gain of the hearing aid without affecting the sound quality. The theoretical findings and the validity of the proposed system suggest the visibility of practical testing on binaural hearing aid users. It provided significant improvements to existing feedback canceller system in binaural hearing aids.

Conclusion

In binaural hearing aids, the adaptive feedback canceller causes signal correlations between the input speech signal and the feedback signal due to a closed loop. These correlations result in a biased estimate of the feedback path; thus, this paper proposed a new method for improving acoustic feedback suppression using binaural information and blind source separation. This is achieved by combining ITDs and ILDs for sound source segregation and unsupervised learning classification to select source positions from a set of candidate positions that are most likely speech sources. Evaluation of the proposed system indicated that our system improved acoustic feedback suppression using binaural information and unsupervised learning and it achieved high separated sound quality to support binaural hearing aid users and allowed them to increase the gain of the hearing aid without affecting the sound quality.

Acknowledgments

We would like to thank the team of Damascus University, electrical engineer Ammar Akhras and Dar –aloloom & Alttarbia School for their collaboration and facilities provided during this work.

REFERENCES

1. Avan P, Giraudet F, Büki B. Importance of binaural hearing. Audiol Neurootol. 2015;20 Suppl 1:3-6.
2. Dillon H. Hearing aids. 2^{nd} ed. London: Thieme Medical Pub; 2001.
3. Guo M, Jensen SH, Jesper J. Evaluation of state-of-the-art acoustic feedback cancellation systems for hearing aids. J Audio Eng Soc. 2013;61(3):125-37.
4. Lee HW, Jeon MY. A combined feedback and noise cancellation algorithm for binaural hearing aids. Advances in Electrical and Computer Engineering. 2011;11(3):35-40.
5. Chisaki Y, Matsuo K, Usagawa T. Howling canceler using interaural level difference for binaural hearing assistant system. Acoust. Sci. & Tech. 2007;28(2):90-7.
6. Lombard A, Reindl K, Kellermann W. Combination of adaptive feedback cancellation and binaural adaptive filtering in hearing aids. EURASIP J Adv Signal Process. 2009;2009:1-15.
7. Spriet A, Rombouts G, Moonen M, Wouters J. Adaptive feedback cancellation in hearing aids. Journal of the Franklin Institute. 2006;343(6):545-73.
8. Boukis C, Mandic DP, Constantinides AG. Toward bias minimization in acoustic feedback cancellation systems. J Acoust Soc Am. 2007;121(3):1529-37.
9. Shao Y, Srinivasan S, Jin Z, Wang D. A computational auditory scene analysis system for speech segregation and robust speech recognition. Comput Speech Lang. 2010;24(1):77-93.

www.SID.ir

Archive of SID

10. May T, van de Par S, Kohlrausch A. A probabilistic model for robust localization based on a binaural auditory front-end. IEEE Trans Audio Speech Lang Process. 2011;19(1):1-13.
11. Sawada H, Araki S, Makino S. Underdetermined convolutive blind source separation via frequency binwise clustering and permutation alignment. IEEE Trans Audio Speech Lang Process. 2011;9(3):516-27.
12. In: Ganesh RN, Wang W, editors. Blind source separation: advances in theory, algorithms and applications. Berlin: Springer; 2014.
13. Trahiotis C, Bernstein LR, Stern RM, Buell TN. Interaural correlation as the basis of a working model of binaural processing: an introduction. In: Popper AN, Fay RR, editors. Sound Source Localization.1st ed. New York: Springer; 2005. p. 238-71.
14. Makino S, Sawada H, Mukai R, Arki S. Blind source separation of convolutive mixtures of speech in frequency domain. IEICE Trans Fundamentals. 2005;E88-A(7):1640-55.
15. Garofolo JS, Lamel LF, Fisher WM, Fiscus JG, Pallett DS, Dahlgren NL, et al. TIMIT acoustic-phonetic continuous speech corpus LDC93S1. Web Download. Philadelphia: Linguistic Data Consortium, 1993.
16. Jafari I, Haque S, Togneri R, Nordholm S. Evaluations on underdetermined blind source separation in adverse environments using time-frequency masking. EURASIP J Adv Signal Process. 2013;2013(1):162.
17. Vincent E, Sawada H, Bofill P, Makino S, Rosca JP. First stereo audio source separation evaluation campaign: data, algorithms and results. Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation; 2007 Sep 9; London, UK. Berlin: Springer.
18. Viswanathan M, Viswanathan M. Measuring speech quality for text-to-speech systems: development and assessment of a modified mean opinion score (MOS) scale. Comput Speech Lang. 2005;19(1):55-83.

www.SID.ir