Speech enhancement in a smartphone-based hearing aid

Maksim I. Vashkevich,
Candidate of technical Sciences, associate Professor of the Belarusian state University of Informatics and Radioelectronics (BSUIR)

Iliy S. Azarov,
Doctor of technical Sciences, associate Professor, BSUIR

Aleksandr A. Petrovsky,
Doctor of technical Sciences, Professor of the chair of electronic computing BSUIR

Abstract

The paper presents speech enhancement techniques for advanced smartphone-based hearing aid which originates from our free smartphone application "Petralex" recently released for iOS and Android devices. In the present contribution we develop a new solution which overcomes limitations of full-band processing and introduces extended functionality. The new processing scheme decomposes the signal into perceptually matched sliding bands and implements spectral gain shaping for hearing loss compensation, dynamic range compression, noise reduction and acoustic feedback suppression. We propose an acoustic feedback suppression algorithm that is based on spectral subtraction rule. The algorithm is robust to rapid changes in acoustic feedback path and according to experiments allows to achieve added stable gain up to 24 dB. The paper contains theoretical background, description of the implemented techniques and some experimental results.

Keywords: hearing aid, noise reduction, acoustic feedback suppression

INTRODUCTION

Qualitative improvement of hearing aids in the last decade occurred due to increase of computational power of portable devices, their power resources and improvement of analog-to-digital/digital-to-analog converters. There is a miniaturization tendency in hearing aid de-sign which can be noticed in retrospection [1]: pocket hearing aids were superseded by aids inserted in spectacle frame then appeared devices placed behind the ears and now they be-come small enough to be hidden inside ear's channel. Recently a wide spread of mobile mul-timedia platforms (especially smartphones) gave new life to pocket hearing aids. A smartphone is capable of functioning as a hearing aid under special software which takes con-trol over audio subsystem of the device. Recently a number of hearing aid applications have been introduced for portable multimedia devices. Although a smartphone cannot be consid-ered as an adequate substitute for a small-sized hearing aid it still might be advantageous for the following reasons [2]:

- functionality of the device can be very flexible regarding both signal processing algorithms and user interfaces;
- large power and computing resource of a smartphone allows implementing sophisticated real-time processing algorithms;

- hearing loss compensation algorithm can be applied to various multimedia content such as music, audio books, movies etc.;
- personal fitting of the hearing aid can be carried out without assistance of audiologist using in situ audiometry;
- it is possible to use different external headsets for different life situations;
- using a smarthpone is psychologically comfortable since it is not recognized as a hearing aid by surrounding people;
- for hearing impaired smartphone users there is no need to buy and wear an additional device.

As a pocket hearing aid smartphone has additional advantages:

- a large distance between microphone and speaker prevents occurring of acoustical feedback on considerably high gain levels;
- large physical dimensions of the device can be convenient for persons with con-strained motor function;
- using speakers with bone conduction does not lead to mechanical feedback.

Some time ago we released "Petralex" — a free application for hearing loss compensa-tion with in situ audiometry [3]. The application proved to be helpful and for now it is consid-ered as one of the useful hearing assistive technology[1] [4]. Recently we completed a survey with more than 1500 participants among "Petralex" users that clearly indicated applicability of a smartphone as a self-fitting hearing aid. Compared to conventional hearing aids the ap-plication provided the same average change in the hearing ability and turned out to be even more effective in noisy situations.

Considering significant social impact of smartphone-based hearing aids we redesigned "Petralex" in accordance with accumulated user experience. Designing of an original signal processing algorithm is rather difficult considering requirements of the target platform. One of the main problems is processing delay. In has been shown that long processing delays are undesirable due to the comb filter effect, which occurs when the processed sound and the un-processed sound are mixed at the eardrum [5]. It is known that even very short delays (4–8ms) can noticeably reduce sound quality [6].

Although smartphone-based hearing aids are not capable of reaching such short delays due platform limitations. However it is still important to minimize inherent delay of the algorithm. In "Petralex" this problem was solved by using full-band processing scheme [2]. However the scheme has a problem in achieving required sound pressure level which is limited by available dynamic range of the device. Full-band digital amplification leads to clipping effect while applying full-band limiters restrict maximum gain of perceptually important components. Considering that in the preposed solution we implemented a subband processing scheme that processes the sound in narrow frequency bands and controls amplitude each of them individually.

---

[1] "Petralex" downloaded more than 300 000 users; according to iTunes Connect App Analytics iOS users have more than 1'000 active sessions per day.

It is known that 52 % of hearing impaired people use hearing aids in noisy disturbing situations [7]. Many studies have shown that noise reduction increases hearing comfort and significantly reduces harmful impact on user's hearing [1]. Along with background noise a hearing aid user suffers from acoustic feedback, which occurs when the processed signal leaks from the speaker back to the microphone. In context of smartphone-based solution this be-comes a serious problem because the user normally applies standard headphones with heavy sound leakage. Despite that the microphone and speaker are separated by each other acoustic feedback often arises at the desired amplification level. Adaptive feedback cancellation pre-sented in diversity of least mean squares (LMS) techniques is currently the mainstream of acoustic feedback cancellation in hearing aids [8–12]. However practical experience shows that this approach is ineffective for smartphone implementation. When using smartphone as a hearing aid the feedback path is very unstable because of changing distance between micro-phone and speakers. In such conditions adaptive filtering cannot noticeably improve maximum stable gain: when using low adaptation rates the reaction to changes becomes unpredictable, when using sufficiently high adaptation rates the speech signal drastically degrades. It was shown that the room acoustic also makes a considerable contribution to feedback path [13], however robust modeling of room acoustic by means of adaptive filtering can hardly be done in real-life environment. Another problem is robustness: adaptive filtering can be applied only when the system is stable and once stability is lost it cannot be recovered. A known approach for robust feedback control is notch-filtering howling suppression [14–17] which is able to stabilize a system without reducing the broadband gain. The approach is suitable for a smartphone; however its weak side compared to adaptive feedback cancellation is a low max-imum stable gain increase and signal distortion [18]. Considering that we propose an original algorithm based on spectral subtraction instead of notch-filtering. The algorithm applies a weighting rule derived specially for feedback and can be combined with noise reduction which attenuates both background noise and feedback residual. According to experimental results the proposed solution provides close performance to adaptive feedback cancellation in terms of maximum stable gain increase and speech quality, however and at the same time is very robust against changes in feedback path. Combination of noise reduction implies that both algorithms share the same analysis/synthesis framework which is advantageous regarding computational efficiency.

1. IMPLEMENTED PROCESSING SCHEME

In modern hearing aids, signal processing is usually performed in frequency subbands introducing analysis-synthesis delays in the forward path. Many research efforts have been focused on this problem [19–20], however the delays of these solutions are still high (6-8 ms). Good low-delay filtering schemes based on peaking filters [6], cochlear filters [21] and side-branch processing [22] has been recently proposed. Some common frequency-dependent am-plification schemes are shortly described below. Existing mobile platforms can process the signal in real-time by separate frames of 6 ms or longer, that requires block by block pro-cessing. It is impossible to eliminate delays introduced by analog-to-digital and digital-to-analog converters which can reach 0,4 to 2 ms depending on implementation [23]. Inherent hardware delay of a smartphone is much longer due to implementation of audio processing pipeline (10–20 ms for iPhone and 50–300 ms for Android).
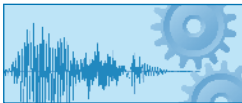
1.1 Full-band processing

Considering constrains of the mobile platform it is possible to use full-band processing scheme that uses finite impulse response (FIR) filtering and dynamic range compression (DRC) for hearing loss compensation. The scheme is shown in Figure 1.
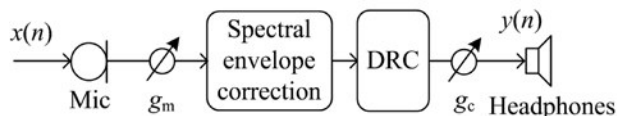

Figure 1. Full-band processing scheme

Spectral envelope correction is done using FIR filter which is designed using prescription gain formulas. There are two loudness controls: microphone sensitivity g_m and output level g_c, which the user can adjust according to the current acoustic conditions. The block of dynamic range compression applies time-varying gain for recruitment correction. Compression ratio is chosen according to the degree of hearing loss. Considering that smartphone uses a stereo headset it is possible to apply binaural hearing compensation processing left and right channels separately. In the previous version of "Petralex" we applied linear phase filter with group delay ≈ 3 ms, which is synthesized using the windowing method.

The full-band processing scheme has the following advantages: low processing delay (which consists of the group delay of the equalizer filter and platform delay), low computa-tional cost and simplicity in design. However the scheme is not capable of controlling loud-ness of separate spectral components which requires time-frequency transform of the signal.

1.2 Sub-band processing

Functionality of the hearing aid can be significantly extended using sub-band decom-position of the signal into separate frequency components. Processing in this case can be car-ried out using individual time-varying amplification of each subband channel [6].

There are sub-band processing systems with reduced processing delay. In [22, 24] a scheme of sub-band amplification is proposed that does not require synthesis filter. Processing in the forward path is carried out using FIR filter, coefficients of which are updated for each processing frame according to amplifications gains derived from subband side branch. It is also possible to use parametric band-pass filters [6] or cochlear filters [21] summing outputs of the filters after amplification. Both cochlear filter bank and peaking filters decompose the signal into perceptually matched subband components and has a very low group delay. However these approaches are computationally more consuming compared to general sub-band processing scheme.

1.3 Proposed processing scheme

We assume that the input signal can be represented in the frequency domain as a sum of clean speech signal $X(w)$, acoustic feedback $A(w)$ and background noise $N(w)$:

$$\bar{\bar{X}}(w) = X(w) + A(w) + N(w) = \bar{X}(w) + N(w), \qquad (1)$$

where $\bar{X}(w) = X(w) + A(w)$ is the speech signal with acoustic feedback. Let $R_{\bar{\bar{x}}}(w)$, $R_{\bar{x}}(w)$, $R_n(w)$ are power spectral densities (PSD) of $X(w)$, $A(w)$ and $N(w)$ respectively, then $\bar{X}(w)$ can be estimated from $\bar{\bar{X}}(w)$ by using noise reduction factor

$$G_{NR}(\omega) = \sqrt{1 - \frac{R_n(\omega)}{R_{\bar{\bar{x}}}(\omega)}} . \qquad (2)$$

Feedback suppression factor can be estimated in the same way:

$$G_{AF}(\omega) = \sqrt{1 - \frac{R_a(\omega)}{R_{\bar{x}}(\omega)}} . \qquad (3)$$

On the basis of described approach the following processing scheme is proposed (Fig-ure 2). The signal processing includes three consecutive stages: 1) noise reduction; 2) acoustic feedback suppression and 3) hearing loss compensation.
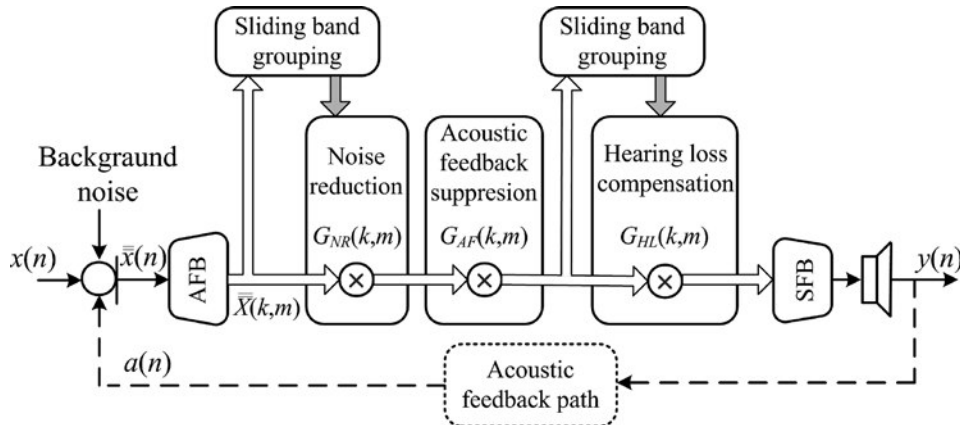

Figure 2. Implemented processing scheme

The input signal $\bar{\bar{x}}(n)$ is decomposed into complex subbands $\bar{\bar{X}}(k,m)$ by the analysis filter bank (AFB), where k and m are frequency and time indices respectively, and the processed full-band signal $y(n)$ is reconstructed by synthesis filter bank (SFB). For reasons of computational efficiency we use an oversampled DFT-modulated filter bank. Calculation of noise reduction coefficients $G_{NR}(k,m)$ requires estimation of noise PSD. In order to make noise statistics more reliable subband signals are combined in a wide sliding bands. Acoustic feedback suppression coefficients $G_{AF}(k,m)$ is calculated based on estimation of acoustic feedback signal PSD. At the last stage subband signals multiplied by the $G_{NR}(k,m)$ and $G_{AF}(k,m)$ are combined into sliding bands for determining required hearing compensations gains $G_{HL}(k,m)$ which are calculated using to a desired prescription formula and DRC pro-file.

2. ANALYSIS-SYNTHESIS BASED ON DFT-MODULATED FILTER BANK

2.1. Analysis-synthesis framework

Filter banks are commonly used tool to organize subband signal processing in modern hearing instruments [6, 19–22, 24]. A DFT (or complex) — modulated filter bank with poly-phase implementation of FIR prototype filter is one of the most efficient and popular [19, 24]. For example in [25] for hearing aid system was used an oversampled, polyphase DFT filter bank with 16 frequency bands. Decimation of the subband signals reduces computational cost, however decimation/interpolation in this solution inevitably distorts the output signal. A well-known techniques such as aliasing compensation that used in perfect (near perfect) recon-struction filter banks are not suitable for hearing aid since gains applied to subbands are sig-nificantly different. For this reason a special procedure for FIR prototype design should be used [19].

Another different form for implementing DFT filter bank is weighted overlap-add (WOLA) structure [20, 26]. WOLA structure is more general than poly-phase structure in which number of channels and decimation factor have the following restriction

$$K = MI. \tag{4}$$

where I — a positive integer (I = 1,2,3 ...) called oversampling ratio. In WOLA is unrelated to K.

The output signal for k-th channel of analysis DFT filter bank is expressed as follows

$$X(k,m) = \sum_{n=-\infty}^{\infty} h(mM - n)x(n)W_K^{-kn}, \qquad k = 0, 1, \dots K - 1. \tag{5}$$

where $W_K = e^{j2\pi/K}$, $x(n)$ — input signal, $h(n)$ — filter prototype (of length L) that is a sliding analysis window that selects and weights a frame of the input signal. Output signals $X_k(m)$ are referred to as the short-time spectrum of the signal at time $n = mM$. Expression (5) can be rewritten as DFT

$$X(k,m) = \sum_{n=-\infty}^{\infty} y(n)W_K^{-kn} \tag{6}$$

where $y(n) = h(mM - n)\, x(n)$ — windowed input sequence.

The WOLA synthesis structure can be expressed in the form

$$\hat{x}(n) = \sum_{m=-\infty}^{\infty} f(mM - n)\frac{1}{K} \sum_{k=0}^{K-1} \hat{X}(k,m)W_K^{kn} \tag{7}$$

where $f(n)$ — synthesis filter (or synthesis window). Simplified structure of WOLA filter bank (when length of analysis and synthesis windows L equals to K) is given in Figure 3 where the following notation is used

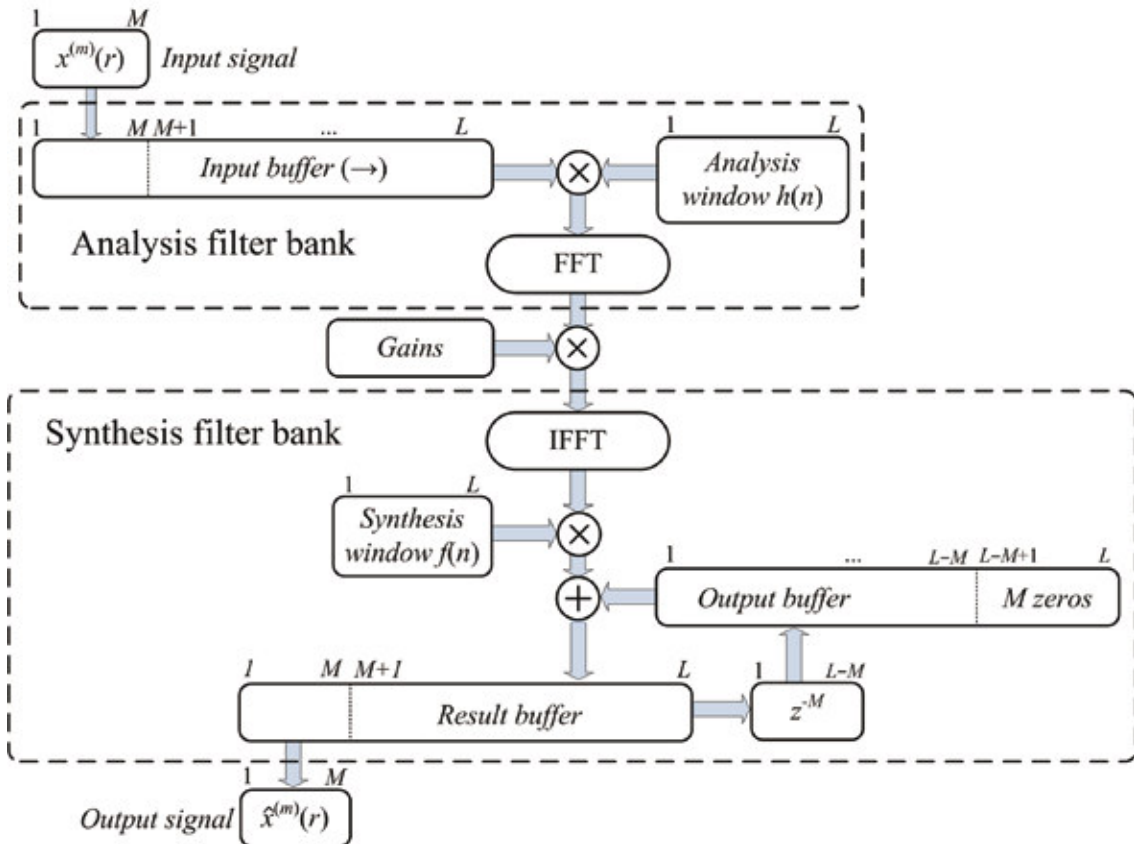$$x^{(m)}(r) = x(mM + r), r = 0, 1, \dots M - 1. \tag{8}$$

Figure 3. WOLA structure of DFT-modulated filter bank

We use a simple method of calculation $h(n)$ and $f(n)$ that allows to obtain good fre-quency resolution and low aliasing of the reconstructed signal. For analysis we use the Hamming window, which provides good trade-off between main-lobe width and side-lobes attenuation in short-time spectrum:

$$h(n) = h_{hamm}(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{L-1}\right). \tag{9}$$

where $n = 0 \dots L - 1$. Assuming that L and M are odd the synthesis window is defined as

$$f(n) = \begin{cases} \dfrac{h_{hann}(n)}{h_{hamm}\left(n + \frac{L-1}{2} - M\right)} = \dfrac{0.5 - 0.5\cos\left(\frac{2\pi n}{2M}\right)}{0.54 - 0.46\cos\left(\frac{\pi(2n+L-1-2M)}{L-1}\right)}, & \text{for } n \in [0; \, 2M] \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

where numerator is the Hanning window of length $2M + 1$. The synthesis window attenuates phase breaking effect between adjacent frames. According to (10) applying both windows $h(n)$ and $f(n)$ is equivalent to applying Hanning window that ensures perfect concatenation of the reconstructed signal since summation of two shifted version of Hanning windows gives one

$$h_{hann}(n) + h_{hann}(n + M) = 1. \tag{11}$$

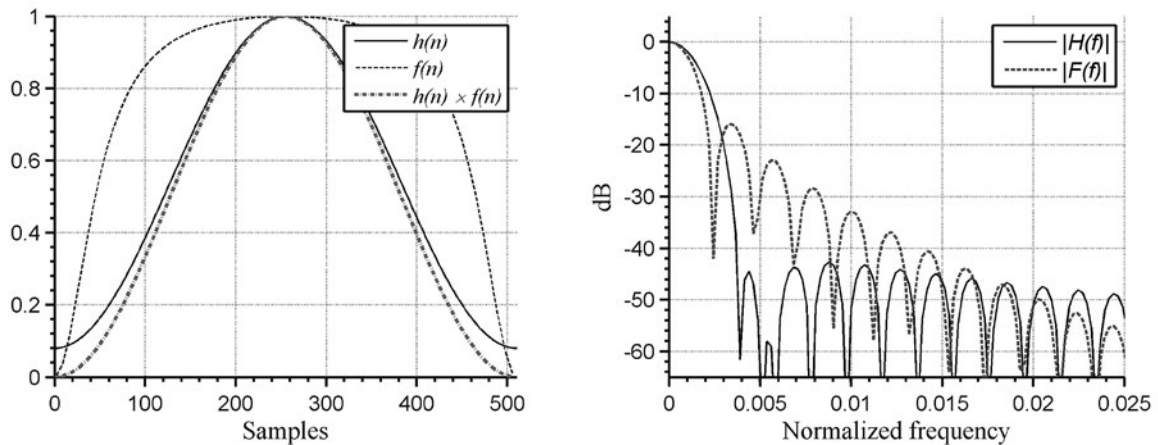Figure 4 shows windows calculated according to (9)–(10) that guarantee perfect signal recon-struction.

Figure 4. Impulse and magnitude frequency responses of $h(n)$ and $f(n)$ ($L = 511$ and $M = 255$)

2.2 Sliding band grouping

Using overlapping frequency bands is advantageous for estimating level of the sub-band signals [27] since it prevents from decrease of the spectral contrasts and modulation depths in speech signal. Subbands signals $|X(k,m)|^2$ are grouped in sliding bands both in time and frequency domains

$$S(k,m) = \frac{1}{N_t} \sqrt{\sum_{i=-bw(k)}^{bw(k)} \sum_{j=0}^{N_t-1} |X(k+i, m-j)|^2} . \qquad (12)$$

where $bw(k)$ determines frequency bandwidth, N_t — number of summed frames. Following the psychoacoustic principle that the bandwidth is proportional to the central frequency $bw(k)$ is calculated as
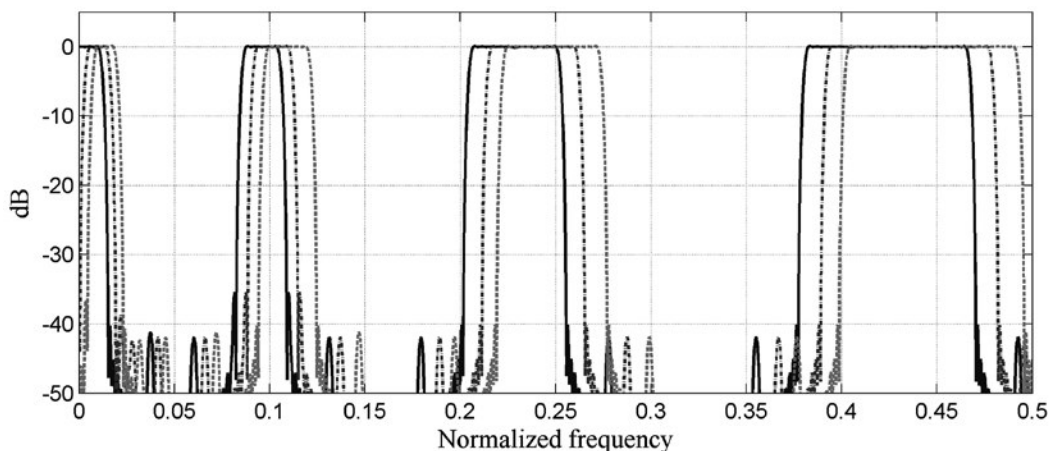

Figure 5. Overlapping frequency bands for spectral energy estimation (filter bank parameters: $L = 511$, $M = 40$)

$$bw(k) = \max\left\{\text{round}\left(\frac{k}{20}\right), bw_{\min}\right\}. \tag{13}$$

where is minimum bandwidth. Using such grouping reduces effect of musical artifacts in noise reduction algorithm since fragments of residual are not perceived as tones.

An illustration of using overlapping frequency bands is given in Figure 5, where parti-tion of the frequency range into sliding bands is shown.

3. NOISE REDUCTION

Implemented noise estimation algorithm is generally based on the minima controlled recursive averaging (MCRA) [28], where noise spectrum is obtained by averaging past spectral values $S(k,m)$:

$$\tilde{D}(k,m) = \tilde{a}_d(k,m)\tilde{D}(k,m-1) + \left(1 - \tilde{a}_d(k,m)\right)S(k,m). \tag{14}$$

where

$$\tilde{a}_d(k,m) = \alpha_d + (1 - \alpha_d)p(k,m) \tag{15}$$

is a time-varying smoothing parameters that depends on estimation of conditional signal presence probability $p(k,m)$. In (15) $a_d (0 < a_d < 1$) is a smoothing parameter determines aver-aging time in the absence of speech. $p(k,m)$ is obtained by recursive averaging

$$p(k,m) = \alpha_p p(k,m-1) + (1 - \alpha_p)I(k,m)$$

where

$$I(k,m) = \begin{cases} 1 & \text{if } S_r(k,m) > \delta, \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

denotes function that indicates the presence of speech component. In (16) the hard decision is done based on the ratio between the local noisy speech spectrum estimation and its derived minimum $S_{min}(k,m)$:

$$S_r(k,m) \triangleq \frac{S(k,m)}{S_{\min}(k,m)}. \tag{17}$$

$S_{min}(k,m)$ is updated using temporary minimum $S_{tmp}(k,m)$. Initially $S_{min}(k,0)$ and $S_{tmp}(k,0)$ are set to $S(k,0)$ and for each frame they are updated in following way
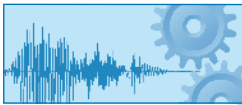
$$S_{\min}(k,m) = \min\{S_{\min}(k,m-1), S(k,m)\}. \tag{18}$$

$$S_{\text{tmp}}(k,m) = \min\{S_{\text{tmp}}(k,m-1), S(k,m)\}. \tag{19}$$

Every *L* frames the following update rule is used

$$S_{\min}(k,m) = \min\{S_{\text{tmp}}(k,m-1), S(k,m)\}. \tag{20}$$

$$S_{\text{tmp}}(k,m) = S(k,m) \tag{21}$$

Parameter *L* determines the resolution of the local minima search. The local minimum is searched on a window of at least *L* frames, but not more than 2*L* frames. A good result is ob-tained for window duration of 0.5–1.5 s [28].

Spectral gain $G_{NR}(k,m)$ is determined as

$$G_{NR}(k,m) = \max\left\{\sqrt{\frac{S^2(k,m) - v\tilde{D}^2(k,m)}{S^2(k,m)}}, 10^{-RL/20}\right\}. \tag{22}$$

where *v* — subtraction factor (1 < *v* < 6), *RL* — adjustable parameter that de-fines the desired residual noise level in dB.

4. ACOUSTIC FEEDBACK SUPPRESSION

When the system is stable the feedback loop can be considered as linear and the feed-back signal typically occurs as a single sine wave. However when the system is unstable the loop becomes non-linear and feedback signal appears as a number of harmonics with unstable parameters as shown in Figure 6.

In both cases acoustic feedback occurs as a quasi periodical signal which is generated by re-cursive summation of the output with a time offset t_o. Periodicity of acoustic feedback en-sures that its spectral components
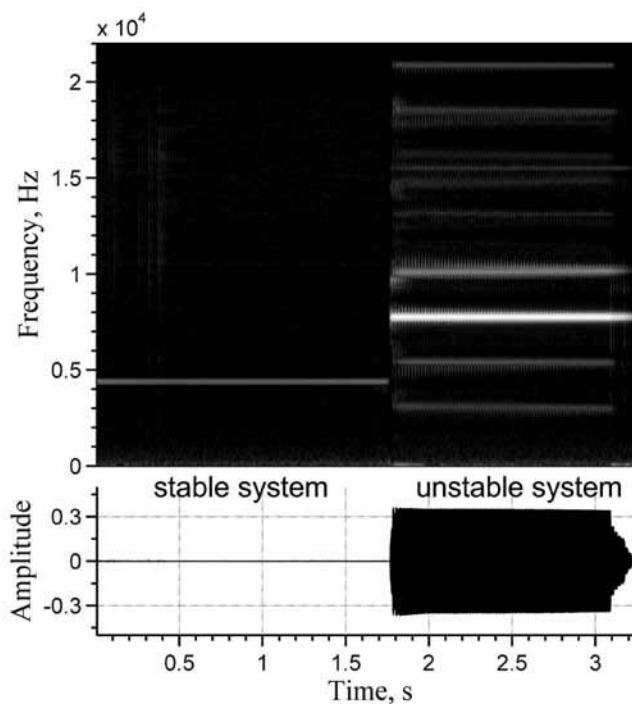

Figure 6. Acoustic feedback in quiet conditions (recorded on iPhone using built-in microphone and standard headphones)

are spaced in frequency domain by fundamental frequency $f_0 = 1 / t_0$ and therefore feedback affects only a subset of $X(k,m)$:

$$\bar{X}(k,m) = \begin{cases} X(k,m) + A(k,m), & \min_{v}\left|k - \dfrac{f_0 v}{f_s/K}\right| \leq d \\ X(k,m), & \min_{v}\left|k - \dfrac{f_0 v}{f_s/K}\right| > d \end{cases}. \tag{23}$$

where v — number of feedback harmonics, f_s — sampling frequency and d — frequency offset, specified by the main lobe of frequency response of the analysis window.

For clean speech expected value of spectral amplitude can be roughly estimated from adjacent frequency components as $E[|X(k,m)|] \approx E[|X(k \pm d, m)|]$ for any sufficiently small frequency offset d. Using (23) and assuming that acoustic feedback increases mean spectral amplitude i.e. $E[|\bar{X}(k,m)|] \geqslant E[|X(k,m)|]$ we get

$$E[|X(k,m)|] \approx \min_{-d<i<d}\left[E[|\bar{X}(k \pm i, m)|]\right]. \tag{24}$$

$$E[|A(k,m)|] \approx E[|\bar{X}(k,m)|] - \min_{-d<i<d}\left[E[|\bar{X}_{k+i}(m)|]\right] \tag{25}$$

According to (23) and (24) expected feedback gain $E[|\bar{X}(k,m)|] / E[|X(k,m)|]$ can be esti-mated from short-time spectral amplitudes close to the corresponding sample $\bar{X}(k,m)$. We introduce the following measure of feedback gain $\chi(k,m)$ based on l previous frames and $2d$ neighboring frequency bins:

$$\chi(k,m) \triangleq \frac{\min_{-l+1 \leq j \leq 0}|\bar{X}(k,m+j)|}{\min_{-d \leq i \leq d}\left[\max_{-l+1 \leq j \leq 0}|\bar{X}(k+i,m+j)|\right]}. \tag{26}$$

In order to avoid overrating of feedback level we use local minima over previous time samples for estimating $E[\bar{X}(k,m)]$ and local maxima for estimating $E[X(k,m)]$. Figure 7 shows probability density function $\rho(\chi)$ obtained experimentally in quiet conditions and during loud speaking for different signal-to-feedback ratios. According to experimental data $\chi > 1$ indi-cates that acoustic feedback is present with probability around 95%. Feedback is not detected when the speech signal is very loud compared to feedback.

Feedback is smoothly controlled, using fixed smoothing value $\alpha_{AF}(0 < \alpha_{AF} < 1)$ that speci-fies averaging time and a time-varying smoothing parameter $\tilde{\alpha}_{AF}$ that depends on feedback intensity

$$\tilde{\alpha}_{AF} = \alpha_{AF} + (1 - \alpha_{AF})\left(\frac{\chi(k,m)}{\chi_{th}}\right)^{\beta}. \tag{27}$$

where β is equalizing parameter that balances reaction to quiet and loud feedback and χ_{th} is threshold value for hard decision. Suppression gain $G_{AF}(k,m)$ is updated using the following expression:

$$G_{AF}(k,m) = \begin{cases} \tilde{\alpha}_{AF} G_{AF}(k,m-1) + \dfrac{(1-\tilde{\alpha}_{AF})}{\max(\chi(k,m),1)}, & \chi(k,m) < \chi_{th} \\ 1/\chi(k,m), & \chi(k,m) \geq \chi_{th} \end{cases}. \tag{28}$$
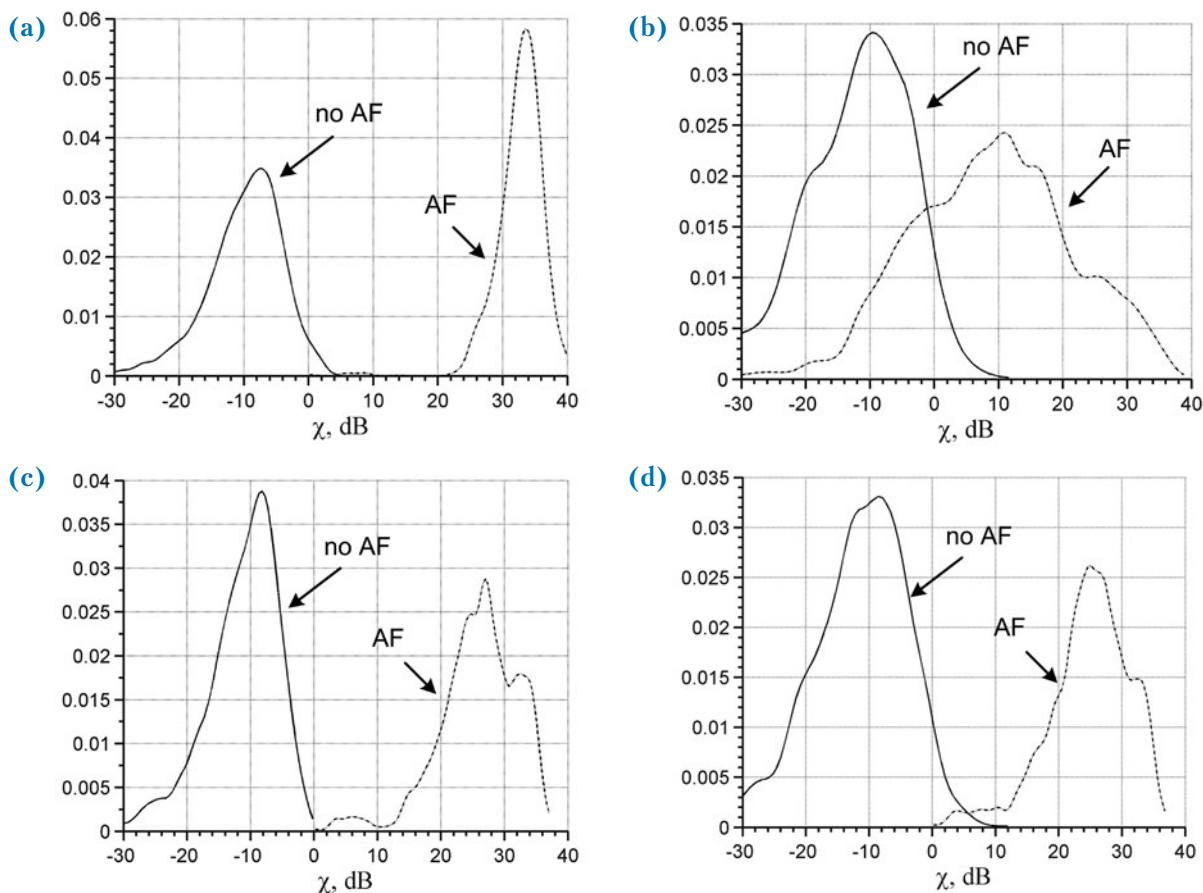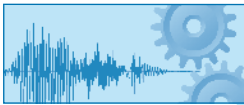
73

Figure 7. Probability density functions $\rho(\chi)$ for acoustic feedback presence and absence in different conditions: **(a)** stable system in quiet (signal-to-feedback ratio −7dB), **(b)** stable system, loud speaking (signal-to-feedback ratio 35dB), **(c)** unstable system in quiet (signal-to-feedback ratio −38dB), **(d)** unstable system, loud speaking (signal-to-feedback ratio −7dB)

When $\chi(k, m)$ exceeds χ_{th} an intense acoustic feedback is detected. In this case suppression gain is updated instantaneously and then slowly released.

5. HEARING LOSS COMPENSATION

In order to determine personal target amplification gains we use in situ audiometry. Hearing threshold levels in quiet are measured using increasing tonal sounds with frequencies 125, 250, 500, 1000, 2000, 4000 and 8000 Hz. We calculate required amplification using conventional formulas: Berger [29], POGO (Prescription of gain and output) [30] and NAL–R (National Acoustic Laboratories, Australia) [31]. Correspondent calculation rules were im-plemented as described in [1]. Recruitment correction is implemented using subband com-pressors with compression ratio derived from hearing loss profile. For each channel the hearing loss compensation gain is calculated according to a given prescription formula and current compression gain.

6. EXPERIMENTAL RESULTS

6.1 Design aspects and implementation

According to Figure 2 the processing of incoming block of samples is carried out using the following steps:

1. New M samples block moved to input buffer, multiplied by analysis window $h(n)$ followed by FFT (see Figure 3 and eq. (5)) to obtain $\bar{\bar{X}}(k, m)$;

2. In block "Sliding band grouping" (Figure 5) $S(k, m)$ is calculated using eq. (12);

3. The estimated spectrum $S(k, m)$ passed to "Noise reduction" block, where coefficients $G_{NR}(k, m)$ calculated using eq. (14)–(22);

4. Modified filter bank outputs $\bar{\bar{X}}(k, m)G_{NR}(k, m)$ are passes to "acoustic feedback suppression" block, where $G_{AF}(k, m)$ are calculated using eq. (26)–(28);

5. Modified filter bank outputs $\bar{\bar{X}}(k, m)G_{NR}(k, m)G_{AF}(k, m)$ are passes to "Sliding band grouping" block, where smoothed estimation of clean speech spectrum is obtained;

6. Based on estimation of clean speech spectrum obtained in step 5, prescription gains and DRC settings hearing loss compensation coefficients $G_{HL}(k, m)$ are calculated in hearing loss compensation block. Filter bank outputs modified as

$$\hat{X}(k, m) = X(k, m)G_{NR}(k, m)G_{AF}(k, m)G_{HL}(k, m);$$

7. Subband signals $\hat{X}(k, m)$ are sent to synthesis filter bank, where output block of sample of length M is obtained (see Figure 3).

The proposed signal processing system was implemented and tested using iPhone-5s and personal computer. The sampling frequency is 44.1 kHz and frame size $L = 511$, signal is captured by blocks of $M = 255$ (50% overlap) that corresponds to 5.8 ms processing delay. In order to apply efficient radix-2 FFT we used zero padding. The program was written using combination of C++ and objective C languages, with Apple's IDE "Xcode, ver.8.2". The processing algorithm insignificantly reduces discharge time of a smartphone (on iPhone-5s the algorithm can continuously work more than 24 hours).

6.2 Noise reduction

Three types of different noises were added to create noisy signals with segmental SNRs in range [–5, 10] dB. The segmental signal-to-noise ratio (SSNR) is defined by [32]

$$SSNR = \frac{10}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \log \frac{\sum_{k=0}^{K/2} |X_k(m)|^2}{\sum_{k=0}^{K/2} |D_k(m)|^2}, \tag{29}$$

where M represents the set of frames that contain speech and $|M|$ its cardinality.

Five male and five female speech samples of duration over 40 s were used in the experiment. The following parameters of noise reduction algorithm were used: $bw_{min} = 3$,

α_d = 0,95, α_p = 0,2, L = 172 (minimum search window is about 1 s), $\delta \sqrt{5}$, v = 2 , RL = 9, N_t = 4. Table 2 shows the average SSNR improvement obtained for different noises.

Segmental SNR improvement for various noise types and levels

Input SegSNR, dB	White noise	Cafeteria noise	Traffic noise
−5	9,81	6,04	6,64
0	7,87	4,03	4,43
5	6,13	2,36	2,68
10	4,59	0,91	1,31

Figure 8 shows the response of the algorithm to rapid change of noise intensity. Figure 8 (b) shows a speech signal corrupted by the traffic noise which starts at 2 s. The algorithm reacts in less than 2 s after appearance of noise.
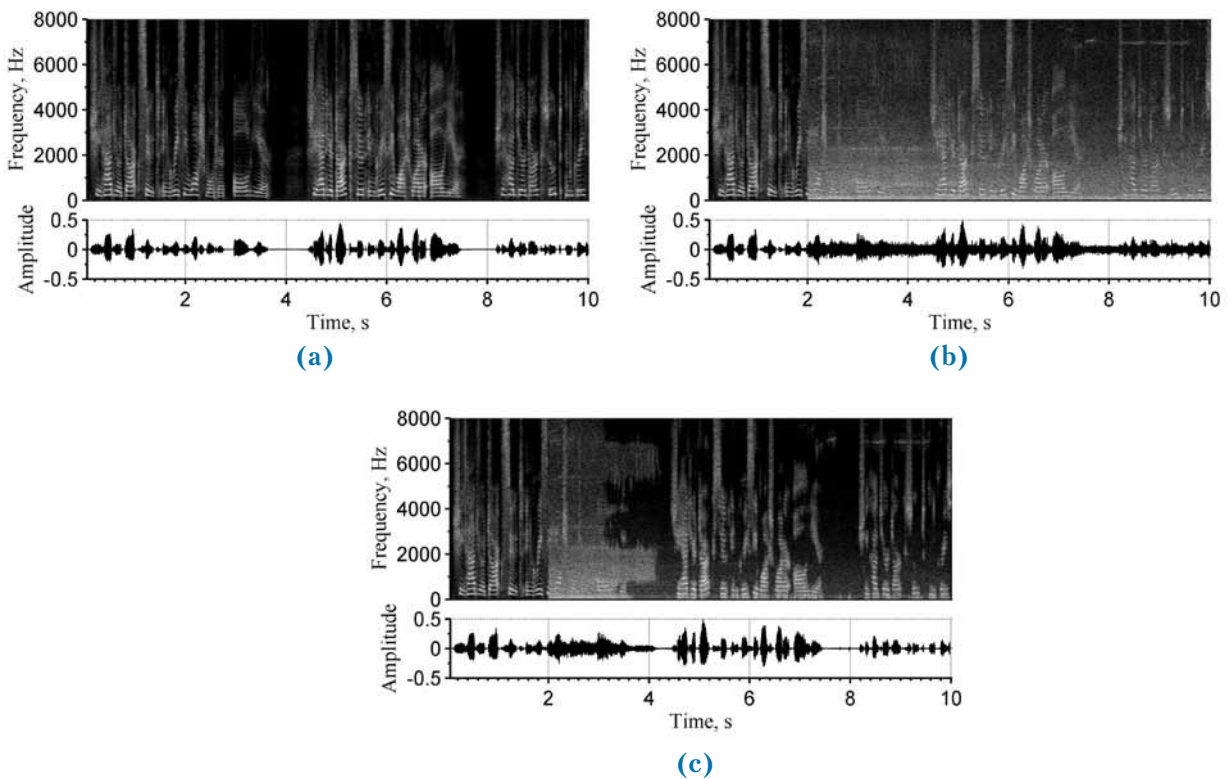

(a)

(b)

(c)

Figure 8. Performance of the noise reduction algorithm: **(a)** clean speech; **(b)** noisy speech signal (noise appears at 2 s); **(c)** processed noisy speech

6.3. Acoustic feedback suppression

The proposed combined noise and acoustic feedback reduction algorithm was evaluat-ed using a feedback path model similar to [10]. Feedback path was modeled as a FIR filter with 279 coefficients, frequency response of the filter is shown in Figure 9. The hearing loss compensation gain was constant for all subbands. The following parameters of the feed-back suppression algorithm were used: $\alpha_{AF} = 0{,}997$, $\beta = 0{,}15$ and $\chi_{th} = 10$.
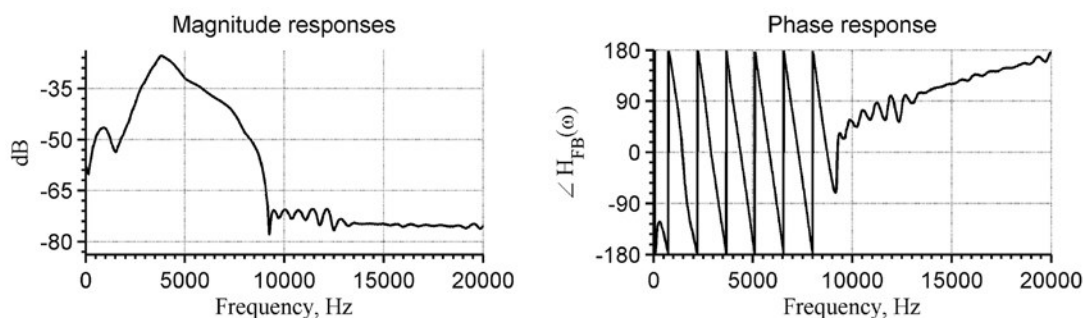

Figure 9. Frequency response of the acoustic feedback path

At first the maximum stable gain of the system was determined [9] that can be applied to signal without feedback control. Then we evaluated performance of the system at different added stable gains ΔG using the proposed feedback suppression algorithm and the LMS adap-tive filtering algorithm (279 coefficients) [8]. Table 3 shows the obtained SSNR values ob-tained in the experiment ('US' means unstable system). The noise signal was obtained as the difference between output signal (with feedback loop and suppression) and the output signal in ideal conditions (without feedback loop and without suppression).

Table 3

SSNR for different added stable gains ΔG

ΔG,dB	No AFR, dB	LMS, dB	Proposed AFR, dB
0	8,12	17,66	12,27
4	US	5,57	11,94
8	US	5,35	11,10
12	US	3,18	10,22
16	US	1,72	9,25
20	US	US	7,59
24	US	US	4,56

The proposed suppression algorithm provided much higher SSNR than LMS in all cases and keeps the system stable even at high added stable gain of 24 dB.

We also evaluated performance of combined feedback and noise suppression in noisy conditions. A speech signal mixed with pink noise at different SSNR levels was used as input to the system with $\Delta G = 12$dB. Table 4 presents obtained SSNR measures.

Table 4

SSNR in noisy condition, 12dB

Input , dB	Proposed feedback suppression, dB	Proposed feedback suppression and noise reduction, dB	LMS, dB	LMS and noise reduction, dB
15	10,54	11,29	5,83	7,78
10	7,07	9,34	4,53	5,62
5	3,51	6,57	2,31	5,00
0	−0,73	3,12	−1,17	2,42
−5	−5,41	−1,24	−5,01	−0,93

The implemented noise reduction algorithm considerably improves SSNR, suppressing both feedback residual and background noise. Suppression of feedback residual significantly improves subjective perception of the processed speech, removing audible tonal components.

In order to evaluate performance of the algorithm in real-life environment we used a PC-based real-time mockup and standard multimedia headset with large headphones. The mockup was placed in a big reverberant room. During the test we changed orientation and location of the headset in order to model time-varying feedback path. The proposed algorithm showed similar performance to previous modeling experiments and never became unstable. An example of performance of the proposed algorithm is given in Figure 10. When feedback suppression is off the system quickly becomes unstable and feedback emerges as multiple to-nal components,
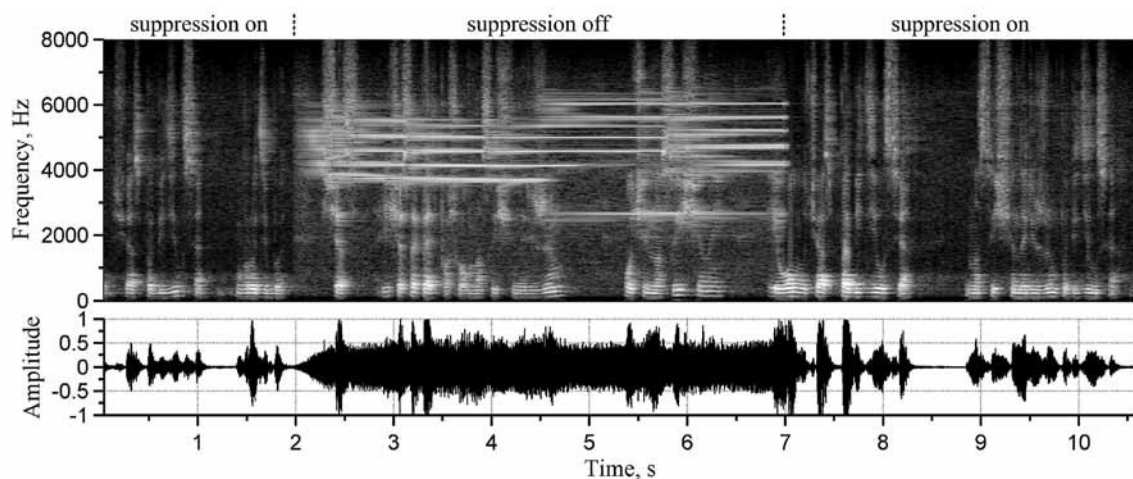

Figure 10. Output signal of real-time mockup: feedback suppression is turned-on at 2 s (distance between microphone and headphones is approximately 30 cm, is approximately 12 dB)

turning on feedback suppression stabilizes the system and eliminates feed-back components completely. The response of the algorithm is very short due to derived short-time weighting rule. In the same conditions the LMS algorithm was unable noticeably increase the maximum stable gain.
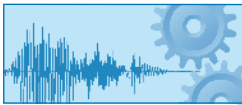
CONCLUSION

The paper presents speech enhancement techniques for a smartphone-based hearing aid. The processing of the signal is performed using DFT-modulated filter bank and include noise reduction and acoustic feedback suppression. The paper introduces an acoustic feedback suppression algorithm based on spectral subtraction that is robust to rapid changes in feedback path. According to experimental results the technique provides high additional gain and high quality of processed speech.

ACKNOWLEDGEMENT

REFERENCES

1. A. Vonlanthen, H. Arndt. *Hearing instrument technology for the hearing health care professional 3rd Edition,* New York: Thomson Delmar Learning, Clifton Park, 2006.
2. E.S. Azarov, M.I. Vashkevich, S.V. Kozlova, A.A. Petrovsky. "Hearing correction system based on mobile computing platform," Informatics, vol.42, no. 2, pp. 7–25, April 2014. (in Russian).
3. IT ForYou. (2014). "Petralex hearing aid v1.4.3." [online] Available: play.google.com/store/apps/details?id=com.it4you.petralex, itunes.apple.com/us/app/petralex-hearing-aid/id816133779?mt=8
4. J. Ismaili, El H.O. Ibrahimi. "Mobile learning as alternative to assistive technology devices for special needs students," *Educ. and Inform. Technol.*, no.1, pp. 1–17, Jan. 2016.
5. R. W. Bäuml and W. Sörgel, "Uniform polyphase filter banks for use in hearing aids: design and constraints," in Proc. of *Proc. European Signal Process. Conf.*, Lausanne, Switzerland, Aug. 2008.
6. A. Pandey and V.J. Mathews "Low-delay signal processing for digital hearing aids," *IEEE Trans.s on Audio, Speech, and Lang. Process.*, vol. 19, no. 4, pp. 699–710, May 2011.
7. S. Bertoli, K. Staehelin, E. Zemp, C. Schindler, D. Bodmer, R. Probst "Survey on hearing aid use and satisfaction in Switzerland and their determinants," *Intern. journal of audiology,* vol. 48, no.4, pp. 183-195, 2009.
8. J.A. Maxwell and P.M. Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Transactions on Speech and Audio Processing,* vol. 3, no. 4, pp. 304–313, 1995.
9. R. Vicen-Bueno, A. Martinez-Leira, R. Gil-Pita, and M. Rosa-Zurera, "Modified LMS-based feed-back-reduction subsystems in digital hearing aids based on WOLA filter bank," *IEEE Trans. on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3177–3190, May 2009.
10. H. Schepker, and S. Doclo, "A semidefinite programming approach to min-max estimation of the common part of acoustic feedback paths in hearing aids," *IEEE/ACM Trans. on Audio, Speech and Lang. Process.*, vol. 24, issue. 2, pp. 366–377, Feb. 2016.
11. H. Schepker, L.Tran, S. Nordholm, and S. Doclo, "Improving adaptive feedback cancellation in hearing aids using an affine combination," in Proc. *IEEE Int. Conf. on Acoust., Speech, and Signal Process.*, Shanghai, China. March 2016, pp. 231–235.

12. F. Strasser and H. Puder, "Correlation detection for adaptive feedback cancellation in hearing aid," *IEEE Signal Processing Letters*, vol. 23, no. 7, pp. 979–983, June 2016.

13. J.M. Kates, "Room reverberation effects in hearing aid feedback cancellation," *Journal Acoust. Soc. Am.*, vol. 109, no. 1, pp. 367–378, Jan. 2001.

14. A. F. Rocha and A. J. S. Ferreira "An accurate method of detection and cancellation of multiple acoustic feedbacks," in *Preprints AES 118th Conv.*, Barcelona, Spain, May 2005, AES Preprint 6335.

15. T. van Waterschoot and M. Moonen "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," *J. Audio Eng. Soc.*, vol. 58, no. 11, pp. 923-940, Nov. 2010.

16. S. M. Kuo and J. Chen "New adaptive IIR notch filter and its application to howling control in speakerphone system," I*EE Electron. Lett.*, vol. 28, no. 8, pp. 764–766, Aug. 2002.

17. K. Ngo, T. van Waterschoot, M. G. Christensen, M. Moonen, S. H. Jensen and J. Wouters "Prediction-error-method-based adaptive feedback cancellation in hearing aids using pitch estimation," in *Proc. European Signal Process. Conf.*, Aalborg, Denmark, Aug. 2010, pp. 40–44.

18. T. Van Waterschoot and M. Moonen "Fifty years of acoustic feedback control: state of the art and future challenges," *Proc. of the IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

19. D. Alfsmann, H. G. Göckler and T. Kurbiel "Filter bank for hearing aids applying subband amplification: a comparison of different specification and design approaches," in *Proc. European Signal Process. Conf.*, Glasgow, UK, Aug. 2009, pp. 2663 — 2667.

20. M. Rosa-Zurera, R. Gil-Pita, E. Alexandre Cortizo, M. Utrilla Manso, L. Cuadra-Rodriguez, "WOLA filter bank design requirements in hearing aids," in *Proc. Intern. Conf. on Pattern Recogn. and Inform. Process.*, Minsk, Belarus, May 2009, pp. 215–218.

21. M. Vashkevich, E. Azarov and A. Petrovsky "Low-delay hearing aid based on cochlear model with nonuniform subband acoustic feedback cancellation," in *Proc. European Signal Process. Conf.*, Bucharest, Romania, Aug., 2012, pp. 514–518.

22. K. M. Kates and K. H. Arehart "Multichannel dynamic-range compression using digital frequency warping," *EURASIP J. Appl. Signal Process.*, vol. 18, no. 1, pp. 3003–3014, Dec. 2005.

23. J. Ryan, S. Tewari "A digital signal processor for musicians and audiophiles," *Hearing Reveiw,* vol. 16, no. 2, pp. 38–41, Feb. 2009.

24. H. W. Lollmann and P. Vary "Generalized filter-bank equalizer for noise reduction with reduced signal delay," in *Proc. Interspeech,* Lisbon, Portugal, Sept. 2005, pp. 2105-2108.

25. T. Schneider and R. Brennan "A multichannel compression strategy for a digital hearing aid," *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Process.*, Munich, Germa-ny, Apr. 1997, pp. 411–414.

26. R. E. Crochiere and L. R. Rabiner *Multirate digital signal processing*, New Jersey: Prentice-Hall Inc., 1983.

27. N. Tiwari, P. C. Pandey and P. N. Kulkarni "Real-time implementation of multiband frequency compression for listeners with moderate sensorineural impairment," in *Proc. Interspeech*, Portland, USA, Sept. 2012, pp. 1860-1863.

28. I. Cohen and B. Berdugo Noise "Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement," *IEEE Signal Processing Letters*, vol 9, no. 1, pp.12–15, Jan. 2002.

29. K.W. Berger, E.N. Hagberg and R.L. Rane "Determining hearing aid gain," *Hearing Instruments.* vol. 30, no.4, pp. 26–44, 1980.
30. G.A. McCandless and P.E. Lyregaard "Prescription of gain/output (POGO) for hearing aids," *Hearing Instruments*, vol. 35, no. 1, — pp. 16–21, 1983.
31. D. Byrne and H. Dillon "The national acoustic laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear and Hearing*, vol. 7, no.7, pp. 257–265, 1986.
32. S. R. Quackenbush, T. P. Barnwell, and M. A. Clements *Objective Measures of Speech Quality*, Englewood Cliffs, New Jersey: Prentice Hall, 1988.

ОБРАБОТКА РЕЧЕВОГО СИГНАЛА В СЛУХОВЫХ АППАРАТАХ НА ОСНОВЕ СМАРТФОНА

Максим Иосифович Вашкевич,
кандидат технических наук, доцент Белорусского государственного университета информатики и радиоэлектроники (БГУИР)

Илья Сергеевич Азаров,
доктор технических наук, доцент БГУИР

Александр Александрович Петровский,
доктор технических наук, профессор кафедры электронных вычислительных средств БГУИР

Аннотация

В статье предложены методы обработки речевого сигнала для усовершенствованного слухового аппарата на основе смартфона, в основу которого положено наше бесплатное, недавно представленное, приложение «Petralex» для устройств iOS и Android. В данной работе показано новое решение, в котором преодолеваются ограничения обработки в широкополосном частотном диапазоне и расширяется функциональность аппарата. В новой схема обработки осуществляется декомпозиция речевого сигнала на перцептуально согласованные частотные полосы и осуществляется спектральное усиление для компенсации потери слуха, сжатие динамического диапазона, снижение шума и подавление акустической обратной связи. Мы предлагаем алгоритм подавления акустической обратной связи, основанный на правиле спектрального вычитания. Алгоритм устойчив к быстрым изменениям пути акустической обратной связи и, согласно экспериментам, позволяет добиться стабильного усиления до 24 дБ. Статья состоит из теоретического обзора, описания реализованных методов и некоторых экспериментальных результатов.

Ключевые слова: слуховой аппарат, редактирование шума, подавление обратной акустической связи