Compression Fitting of Hearing Aids and Implementation

Kashyap Patel[1], **Issa M. S. Panahi**[1]

[1]Department of Electrical and Computer Engineering, University of Texas at Dallas, TX 75080, USA.

Abstract

A compressor in hearing aid devices (HADs) is responsible for mapping the dynamic range of input signals to the residual dynamic range of hearing-impaired (HI) patients. Gains and parameters of the compressor are set according to the HI patient's preferences. In different surroundings depending upon noise level, the patient may seek to tune the parameters to improve performance. Traditionally, fitting of the hearing aids is done by an audiologist using hearing aid software and the HI patient's opinion at a clinic. In this paper, we propose a frequency-based multi-band compressor implemented as a smartphone application, which can be used as an alternative to that of the traditional HADs. The proposed solution allows the user to tune the compression parameters for each band along with a choice of compression speed and fitting strategy. Exploiting smartphone processing and hardware capabilities, the application can be used for bilateral hearing loss. The performance of this easy-to-use smartphone-based application is compared with traditional HADs using a hearing aid test system. Objective and subjective evaluations are also carried out to quantify the performance.

I. Introduction

The normal human ear is capable of sensing a broad range of sound levels ranging from extremely soft sounds to extremely loud sounds that occur in an individual's surroundings. The range of sound a person can hear describes the dynamic range of hearing. Sensorineural hearing loss reduces the sensitivity of the ear to sounds that are presented at low-intensity levels while sensitivity to high-intensity sounds remains mostly unaffected [1], [2]. Hearing-impaired (HI) individuals have reduced dynamic auditory range in one or both ears. The hearing threshold is defined as the softest sound a person can hear. In general, HI listeners have a considerably high threshold of hearing compared to the normal listener. At present, hearing aid devices (HADs) are the most common first step in hearing improvement. Out of many complex signal processing algorithms performed on HADs, compression is responsible for utilizing the residual dynamic range of hearing [3].

Compression works on the principle of amplifying the soft sounds and preserving the comfortable listening levels for loud sounds. Prescribed gain is applied if the input level is less than the compression threshold; else, compression is applied. Comfort level can be addressed by never allowing the output of the HADs to exceed the uncomfortable loudness level. In a real environment, the input signal is unconstrained and may fluctuate from soft to moderate or loud sounds often. In literature, compression schemes use an approach of wide dynamic range compression (WDRC) to map the input signal range into multiple regions

[4]. Speed of compression is also accountable for the comfort level and audibility of the incoming signal or speech [5], [6].

The amount of amplification is decided based on the hearing thresholds. For a hearing-impaired ear, hearing thresholds are not the same across the audible frequency range. Practically, human hearing is measured by an audiologist via the pure tone audiometry method. The hearing thresholds are plotted on a graph of intensity versus frequency, an audiogram for 0.25, 0.5, 1, 2, 4 and 8 kHz frequencies [7]. The amount of amplification or insertion gain for HI listeners for each band is calculated based on prescriptive threshold fitting strategies like NAL-NL2, DSL-v5 (Desired Sensation Level version 5), half gain rule, etc [8], [9]. These fitting strategies calculate the gain required for soft, moderate, and loud sound levels to be heard by a listener. Multi-band compression (MBC) architecture is used to provide the required insertion gain for each frequency band according to input sound level [10]. Based on the audiogram, the audiologist will follow a fitting strategy to determine optimal amplification and compression settings. [11].

The useful features of a smartphone make it a suitable stand-alone device for hearing improvement. In this paper, we present a frequency-based multi-band compression strategy implemented on a smartphone working as an assistive hearing device. Related work has been concerned with updating the compressor parameters via a webpage-based method [12]. However, our motivation behind implementing compression on the smartphone is to provide the ease of fitting the HADs according to one's audiogram and listening preferences. Smartphones carrying all the essential components of traditional HADs, enormous processing capability, and ease of accessibility makes them an attractive alternative to explore for hearing aid applications. The implemented application delivers a bilateral compression experience, which runs on the iPhone in real time with user-friendly graphical user interface (GUI).

Section 2 presents a design of a proposed compressor based on hearing loss, calculation of insertion gain, and effect of compression speed on listening of a HI patient. Section 3 presents the model of compression used and the method for determining insertion gain. Smartphone implementation and performance comparison are carried out in Section 4. Finally, Section 5 summarizes the overall work.

II. Compression Design

Frequency-based nine-band dynamic range compression using a smartphone is proposed in this study. Compression is applied in side-chain configuration with feed-forward topology [13]. Nine bands are considered with center frequency as 0.25, 0.5, 1, 1.5, 2, 3, 4, 6, and 8 kHz. The goal of the proposed scheme is to ensure that the signal is audible while preserving the quality and intelligibility. Audibility can be addressed by mapping the dynamic range of input signal to the residual dynamic range of a HI person. Employed compression can be explained in three stages as per Fig.2.

1. Insertion gain stage: It calculates the amount of gain each frequency band needs to meet the hearing threshold requirement of a HI patient and amplifies the signal accordingly.

2. Compression function stage: It attenuates the high-intensity signals while leaving the low-intensity signals untreated according to the input/output characteristic curve of each band. It determines the gain reduction to be applied after the first stage to prevent the uncomfortable loudness.

3. Gain smoothing stage: It ensures that gain reduction is applied smoothly and not instantaneously to preserve the signal characteristics.

A. Insertion Gain

The audiogram of a person gives the hearing thresholds in dB hearing level (dB HL) for octave band frequencies. These frequencies are selected as they are most important for understanding speech signals. Popular prescriptive fitting strategies like NAL-NL2, DSL-v5, half-gain rule are employed by an audiologist to convert hearing thresholds into actual gain values in dB [8], [9], [14]. These fitting strategies provide the insertion gain for each band, depending on the input level: soft (55 dB SPL), moderate (65 dB SPL) or loud (75 dB SPL).

DSL v5 method has an insertion gain table available for three input levels and different hearing thresholds and bands [15]. Fig. 1 shows the insertion gains computed by the DSL-v5 method for a person with a moderate hearing loss. The half-gain rule method simply calculates the gain by dividing each threshold by two. As an exception to the rule, gain at 250 Hz and 500 Hz is further reduced by 10 and 5 dB respectively. An example of calculating the insertion gain by half-gain rule is explained in Table I. NAL-NL2 being the licensed product, is not described in this paper.

B. Compression Parameters

Compression Threshold (T) defines a level above which compression is applied. T is typically kept at 65 dB SPL. For wide dynamic range compression, audio range is divided into multiple thresholds which are often around 55,65,75, and 85 dB SPL [4].

Compression Ratio (CR) defines the amount of compression to be applied. Each band can have a different CR. Value of CR is preferred between 1.5 to 5 by HI patients [16]. In simple mathematical form, it can be written as

$$CR = \frac{input - T}{output - T}, \text{ for } input > T \tag{1}$$

where T is compression threshold.

Knee width (W) option controls the transition of the input-output ratio from unity to set ratio. The soft knee is preferred because of less obvious effects.

These variables can be adjusted with designated bands across the frequency range. In general, the signal is divided into nine-band range with center frequency as 0.25, 0.5, 1, 1.5, 2, 3, 4, 6, and 8 kHz as shown in Fig. 3.

Makeup Gain (M) control is usually provided at the compressor output. Generally, it works as a volume control. Given a positive M, it will increase the level of the whole signal.

C. Compression Speed

Attack Time (τ_A) is defined as how quickly a compression will get into action once the threshold is crossed. In compression, τ_A is generally kept around 5 ms to 20 ms. One reason for attack time to be short is that if the input signal gets louder, we want to compress the output as soon as possible.

Release Time (τ_R) is defined as the time it takes for gain to return to the normal once the signal falls below the threshold. Based on the value of τ_R, the speed of the compression is decided. Fast-acting compression has τ_R around 50 ms to 200 ms, whereas slow-acting compression has τ_R around 500 ms to 1 s.

The speed of the compression affects the quality and intelligibility of sound [17]. Fast-acting compression preserves the audible cues in quiet speech, but in the presence of surrounding noise, it can degrade the sound quality by introducing pumping and breathing effects [18]. Alternatively, slow-acting compression maintains the temporal cues and the listening comfort, but may provide inadequate gain for soft inputs that come right after loud inputs [19]. Therefore, a fixed speed in compression might affect the performance of the hearing aids. Thus we employ an adaptive compression strategy based on spectral flux as a measure of the transient of the input signal. The more transient the signal, the higher its spectral flux will be and shorter the release time needed for compression.

III.   Compression Model

Fig. 2 shows the block diagram of the proposed compression model. Frame based input signal, $\mathbf{x}(n) = [x(0)x(1)\ldots x(N-1)]^T$ is transformed into $\mathbf{X}(n) = [X(0)X(1)\ldots X(K-1)]^T$ time-frequency domain via STFT filter bank, where n is the time indicator of the frame, N represents the number of samples in one frame and K is the number of points in single sided FFT.

$$X_L(n,k) = 10log_{10}|X(n,k)|^2 \tag{2}$$

where $X_L(n,k)$ is a log scale of input signal. Based on the energy of the signal in the frame, insertion gain values are calculated for each of the nine bands. Giving each band a required gain will assure that all the frequency ranges are audible to the patient. But sharp changes in gain level from one band to another may introduce temporal artifacts. In our proposed method, these artifacts are minimized by interpolation as shown in Fig. 3, which gives a smooth transition of gain between bands. Gain can be inserted as;

$$X_G(n,k) = X_L(n,k) + A(k). \tag{3}$$

$X_G(n,k)$ will be now passed to compression function stage. Output $V_G(n,k)$ can be given by Eq. 4. Indices k and n have been omitted for readability for the next two equations.

$$V_G = \begin{cases} X_G, & 2(X_G - T) \leq -W \\ X_G + \dfrac{\left(\frac{1}{CR} - 1\right)\left(X_G - T + \frac{W}{2}\right)^2}{2W}, & 2|X_G - T| \\ T + \dfrac{X_G - T}{CR}, & 2(X_G - T) > -W, \end{cases} \tag{4}$$

where, T, CR and W are the compression threshold, ratio, and knee width parameters respectively. Now the gain smoothing stage employs peak detector to compute $Z_G(n,k)$, which can be formulated as

$$Z_G = \begin{cases} \alpha_A V_G^{Z-1} + (1 - \alpha_A) W_G, & W_G > V_G^{Z-1} \\ \alpha_R V_G^{Z-1} + (1 - \alpha_R) W_G, & W_G \leq V_G^{Z-1} \end{cases} \tag{5}$$

Where, W_G can be calculated as $W_G = V_G - X_L$; attack time and release time constants are given as $\alpha_A = \exp\left(-\dfrac{K}{\tau_A f_s}\right)$ and $\alpha_A = \exp\left(-\dfrac{K}{\tau_A f_s}\right)$ respectively. Here, f_s is the sampling rate and τ_A and τ_R are given in milliseconds. Additionally, to make compression adaptive, we adapt τ_R with respect to spectral flux as follows;

$$\tau_R(n) = max\left(\tau_{Rmin}, \frac{\tau_{Rmin}}{SF(n)}\right) \tag{6}$$

where τ_{Rmin} is set to be a value of release time for fast-acting compression (approx. 100 ms). The maximum value of τ_R can be limited to the value of release time for slow-acting compression (approx. 2 s). $SF(n)$ is the spectral flux at the nth frame, given as

$$SF(n) = \frac{1}{N} \sum_k \left(|X(n,k)| - |X(n-1,k)|\right)^2 \tag{7}$$

$SF(n)$ is near to zero when the signal is not transient and has a higher value when the signal is transient. After the smoothing stage of the compression, the frequency domain output signal frame $Y(n)$ is generated by shaping the frequency domain input signal based on a gain factor calculated from processing the side chain. Thus,

$$Y(n,k) = X(n,k) \times 10^{\left(\frac{M + Z_G(n,k)}{20}\right)} \tag{8}$$

where M is the make-up gain which controls the overall volume. The time-domain output signal frame $y(n)$ can be obtained via inverse short-time Fourier transform.

IV. Implementation and Results

1) Implementation:

To investigate the behavior and performance of the proposed design, the algorithm presented is implemented on an iOS-based smartphone (iPhone). The algorithm is written in C/C++ on Xcode IDE. Low-level Core Audio framework was implemented on iOS to achieve real-time recording and playback for minimum input/output latency. Input/output signal runs at a 16 kHz sampling rate and a frame size of $N = 256$ samples, optimized for low latency. One reason to operate at 16 kHz and not 48 kHz is to allow Bluetooth Low Energy (LE) connection to existing HAD's with wireless connectivity and wireless earbuds. Windowed input frame is transformed into the frequency domain with 1024 point FFT and a HOP size of 512 samples. Each frame is processed within 1.8 ms which is less than the 16 ms frame size ensuring that the application operates without any audio glitches. The input signal is further divided into two channels, right and left for bilateral hearing loss and processed separately to have 2-channel output. Fig. 4 shows the self explanatory graphical user interface (GUI) of the application. The setting page has options to choose the prescriptive fitting strategy "DSL - v5" or "Half - gain rule." User can also select the compression speed from the "fast-acting compression", "slow-acting compression" and the "adaptive compression."

2) Objective Evaluation:

A comparison of the performance of a hearing aid with wireless connectivity and the proposed application is done based on Fonix 8000 hearing aid analyzer software [20]. ANSI weighted Digital Speech is used as an input signal at 60 dB SPL. A "Starkey Halo 2" hearing aid is fitted with the software for hypothetical moderate hearing loss as per Table I, fast-acting compression, and DSL-v5 fitting rule. Similar parameters and fitting choices were made on the smartphone application. The output of the smartphone application when the volume was set to maximum and the Starkey hearing aid are fed into the hearing aid analyzer microphone. In Fig. 5, we can observe the performance over different frequencies and overall RMS output. The maximum difference between output is around ±10 dB SPL but overall RMS output level is nearly identical (90.1 and 93.5 dB SPL) with both devices.

Objective measure of the quality and intelligibility were also carried out to quantify the performance of the proposed compression. For quality measurement, Perceptual Evaluation of Speech Quality (PESQ) and Hearing Aid Speech Quality Index (HASQI) are considered. For intelligibility measurement, Coherence and Speech Intelligibility Index (CSII) and Hearing Aid Speech Perception Index (HASPI) are considered. The Table II shows the objective scores of the proposed method with other compression parameter set as $CR = 2$, $CT = 65dB$, $W = 20dB$, $M = 0dB$, $\tau_A = 10ms$, and adaptive compression mode.

3) Subjective Evaluation:

Perception based subjective evaluation was performed on 10 normal hearing person. Subjects were told to rate the sound quality of the smartphone application when compression is "ON" and "OFF" between 1–10, 1 being poor and 10 being excellent quality and comfortability. Subjective tests were conducted in a room with one speaker playing 6 short

speech sentences from HINT database with variable intensity and another speaker constantly playing machinery noise in the background (SNR 10–15 dB). Intensity variation of speech determines the performance of the proposed compression. Noise in the background is played with SNR 10 15 dB to emulate the typical real life noisy case. The parameter when compression was ON were following: all hearing thresholds set to 0 dB, $CR = 2$, $CT = 65dB$, $W = 20dB$, $M = 5dB$, $\tau_A = 10ms$ and adaptive compression mode. Mean opinion scores when compression was "ON" and "OFF" were 9.5 and 6.25 respectively. The Institution's Ethical Review Board approved all experimental procedures involving human subjects.

V. Conclusion

To optimally utilize the auditory range of a hearing-impaired patient, smartphone-based hearing assisting application have been developed and proposed in this paper. The setting of each parameter is explained briefly. Hearing thresholds are inserted into the application by user via the developed GUI. The application has a set of default parameters, speed of compression, and fitting strategy. These parameters and listening preferences can be easily set or changed by the user based on his/her preference. Electroacoustic characteristics were compared with that of a hearing aids and the developed/proposed application using smartphone on a hearing aid analyzer. Further, objective and subjective evaluations suggested favorable agreement between the hearing aid and the application.

Acknowledgments

References

[1]. Kuk Francis K, "Theoretical and practical considerations in compression hearing aids," Trends in Amplification, vol. 1, no. 1, pp. 5–39, 1996. [PubMed: 25425854]

[2]. Banerjee Shilpi, "The compression handbook," 2011.

[3]. Dillon Harvey, Hearing aids, Hodder Arnold, 2008.

[4]. Jenstad Lorienne M, Seewald Richard C, Cornelisse Leonard E, and Shantz Juliane, "Comparison of linear gain and wide dynamic range compression hearing aid circuits: Aided speech perception measures," Ear and Hearing, vol. 20, no. 2, pp. 117–126, 1999. [PubMed: 10229513]

[5]. Rosengard Peninah S, Payton Karen L, and Braida Louis D, "Effect of slow-acting wide dynamic range compression on measures of intelligibility and ratings of speech quality in simulated-loss listeners," Journal of Speech, Language, and Hearing Research, 2005.

[6]. May Tobias, Kowalewski Borys, and Dau Torsten, "Signal-to-noise-ratio-aware dynamic range compression in hearing aids," Trends in hearing, vol. 22, pp. 2331216518790903, 2018. [PubMed: 30117366]

[7]. Yund E William and Buckles Krista M, "Multichannel compression hearing aids: Effect of number of channels on speech discrimination in noise," The Journal of the Acoustical Society of America, vol. 97, no. 2, pp. 1206–1223, 1995. [PubMed: 7876443]

[8]. Polonenko Melissa J, Scollie Susan D, Moodie Sheila, Seewald Richard C, Laurnagaray Diana, Shantz Juliane, and Richards Andrea, "Fit to targets, preferred listening levels, and self-reported outcomes for the dsl v5. 0a hearing aid prescription for adults," International journal of audiology, vol. 49, no. 8, pp. 550–560, 2010. [PubMed: 20438300]

[9]. Keidser Gitte, Dillon Harvey, Flax Matthew, Ching Teresa, and Brewer Scott, "The nal-nl2 prescription procedure," Audiology research, vol. 1, no. 1, 2011.

[10]. McCormack Leo, Välimäki Vesa, et al., "FFT-based dynamic range compression," in Proceedings of the 14th Sound and Music Computing Conference, July 5–8, Espoo, Finland, At Espoo, Finland, 2017.

[11]. Moore BCJ, Alcántara JI, and Marriage J, "Comparison of three procedures for initial fitting of compression hearing aids. i. experienced users, fitted bilaterally," British Journal of Audiology, vol. 35, no. 6, pp. 339–353, 2001. [PubMed: 11848176]

[12]. Alamdari Nasim, Lobarinas Edward, and Kehtarnavaz Nasser, "An educational tool for hearing aid compression fitting via a web-based adjusted smartphone app," in ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019, pp. 7650–7654.

[13]. Giannoulis Dimitrios, Massberg Michael, and Reiss Joshua D, "Digital dynamic range compressor design—a tutorial and analysis," Journal of the Audio Engineering Society, vol. 60, no. 6, pp. 399–408, 2012.

[14]. Berger Kenneth W, Hagberg Eric N, and Rane Robert L, "A reexamination of the one-half gain rule.," Ear and hearing, vol. 1, no. 4, pp. 223–225, 1980. [PubMed: 7409361]

[15]. Western University, "Desired sensation level by hand v5," https://www.dslio.com/wp-content/uploads/2014/06/DSL-5-by-Hand.pdf.

[16]. Neuman Arlene C, Bakke Matthew H, Hellman Sharon, and Levitt Harry, "Effect of compression ratio in a slow-acting compression hearing aid: Paired-comparison judgments of quality," The Journal of the Acoustical Society of America, vol. 96, no. 3, pp. 1471–1478, 1994. [PubMed: 7963011]

[17]. Kuk Francis, Slugocki Chris, Korhonen Petri, Seper Eric, and Hau Ole, "Evaluation of the efficacy of a dual variable speed compressor over a single fixed speed compressor," Journal of the American Academy of Audiology, 2019.

[18]. Schneider Todd and Brennan Robert, "A multichannel compression strategy for a digital hearing aid," in 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE, 1997, vol. 1, pp. 411–414.

[19]. Stone Michael A and Moore Brian CJ, "Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task," The Journal of the Acoustical Society of America, vol. 116, no. 4, pp. 2311–2323, 2004. [PubMed: 15532662]

[20]. Frye Kristina, "The effect of analysis methods and input signal characteristics on hearing aid measurements," http://www.frye.com/wp/wp-content/uploads/2013/08/article.pdf, 2019.
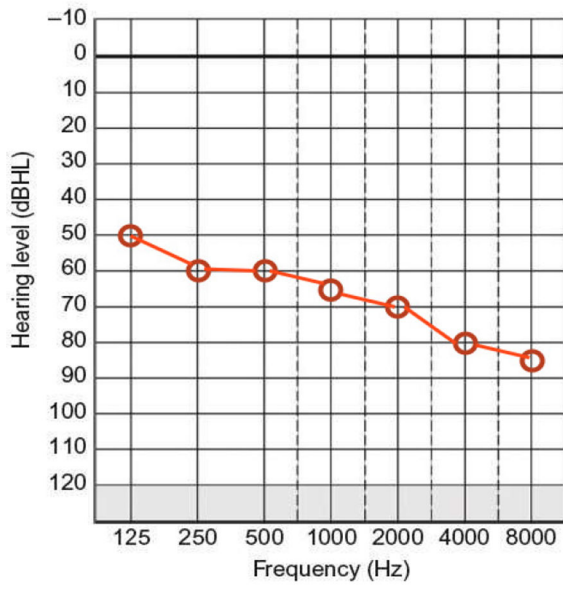
(a) Hypothetical Audiogram
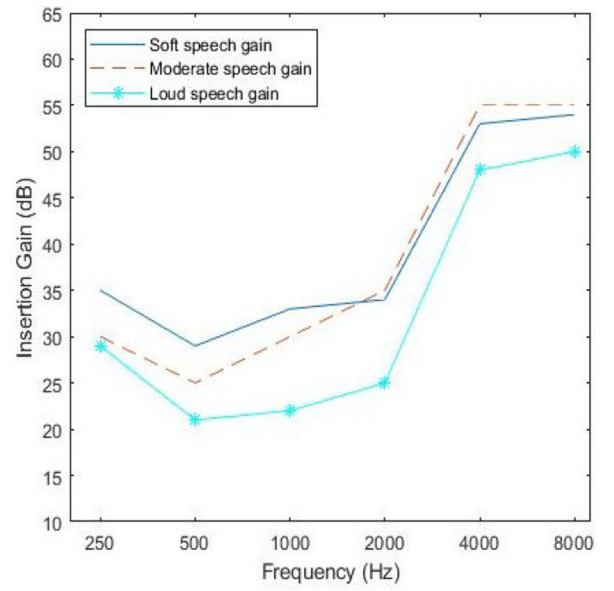
(b) Insertion gain

Fig. 1:
Hypothetical audiogram and corresponding insertion gain based on DSL-v5 prescriptive fitting
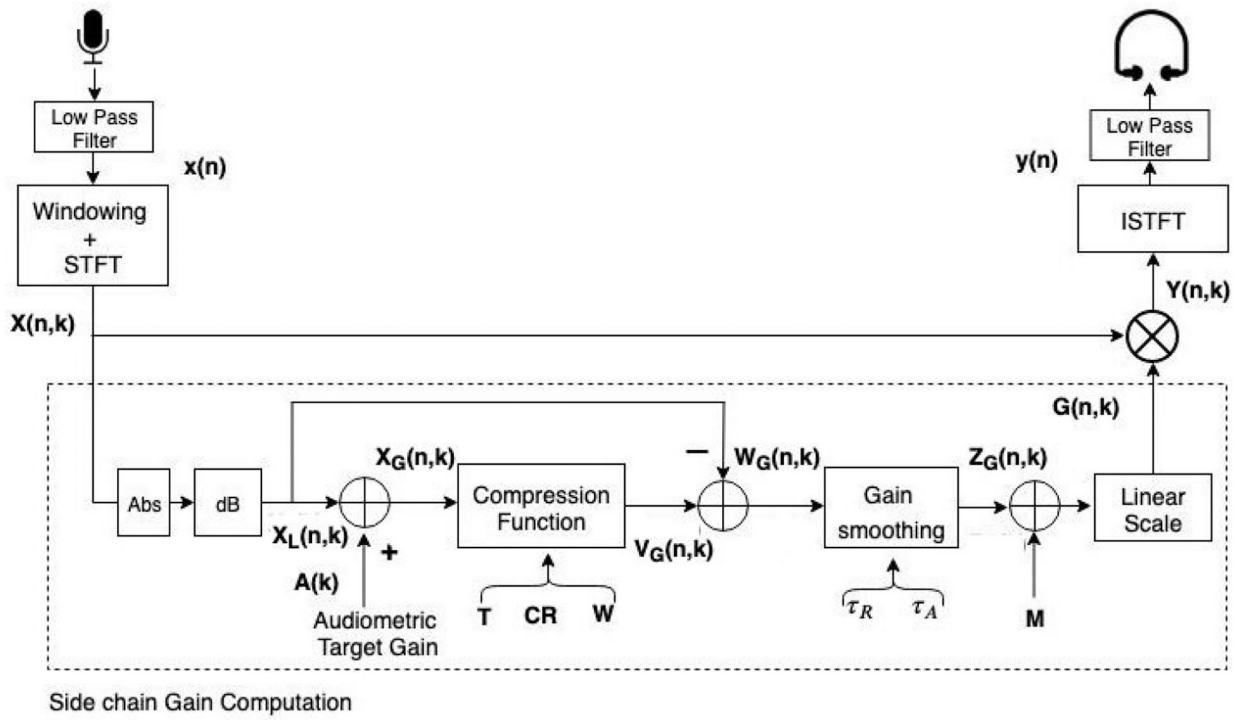
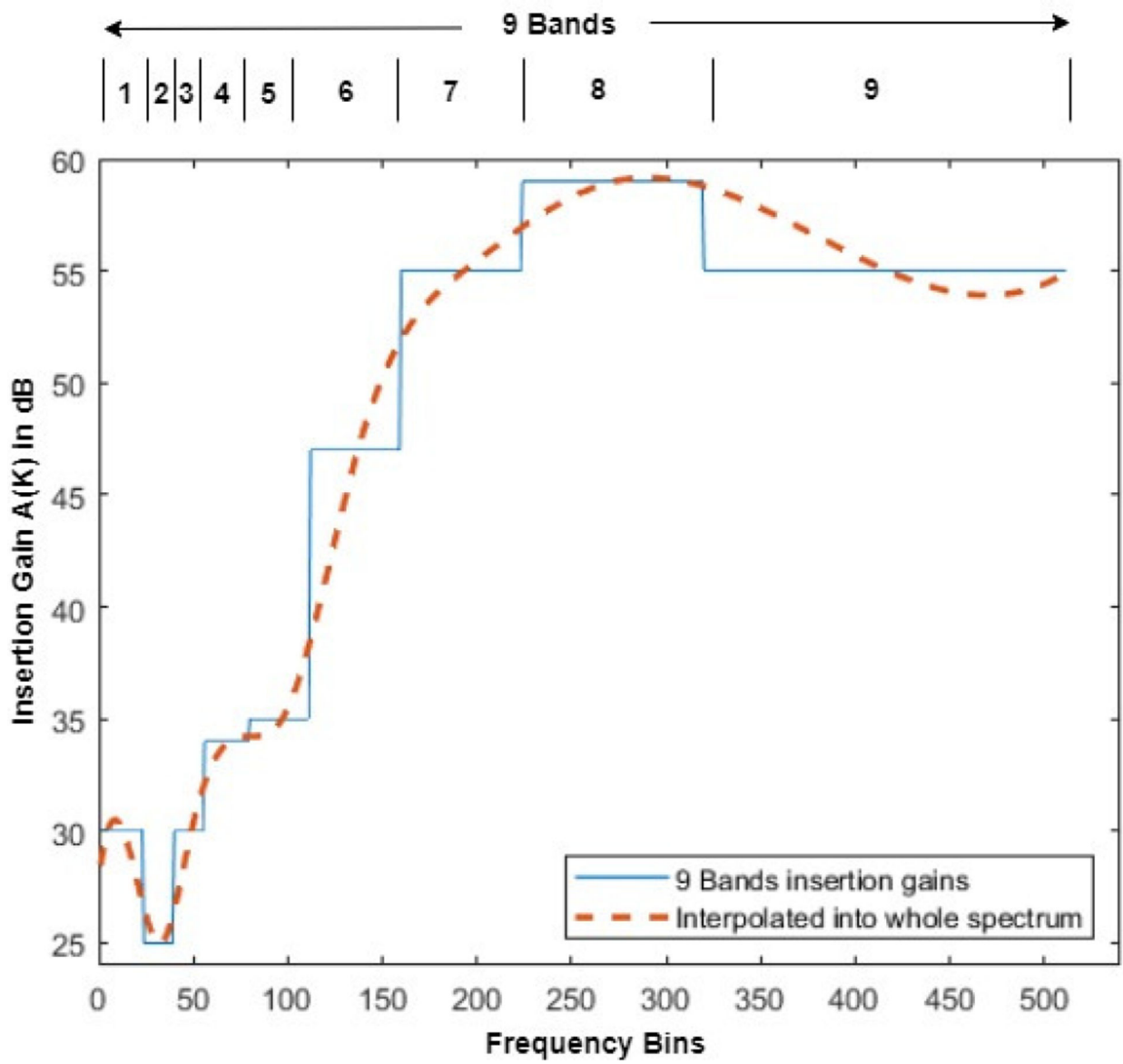**Fig. 2:**
Block diagram of compression pipeline

Fig. 3:
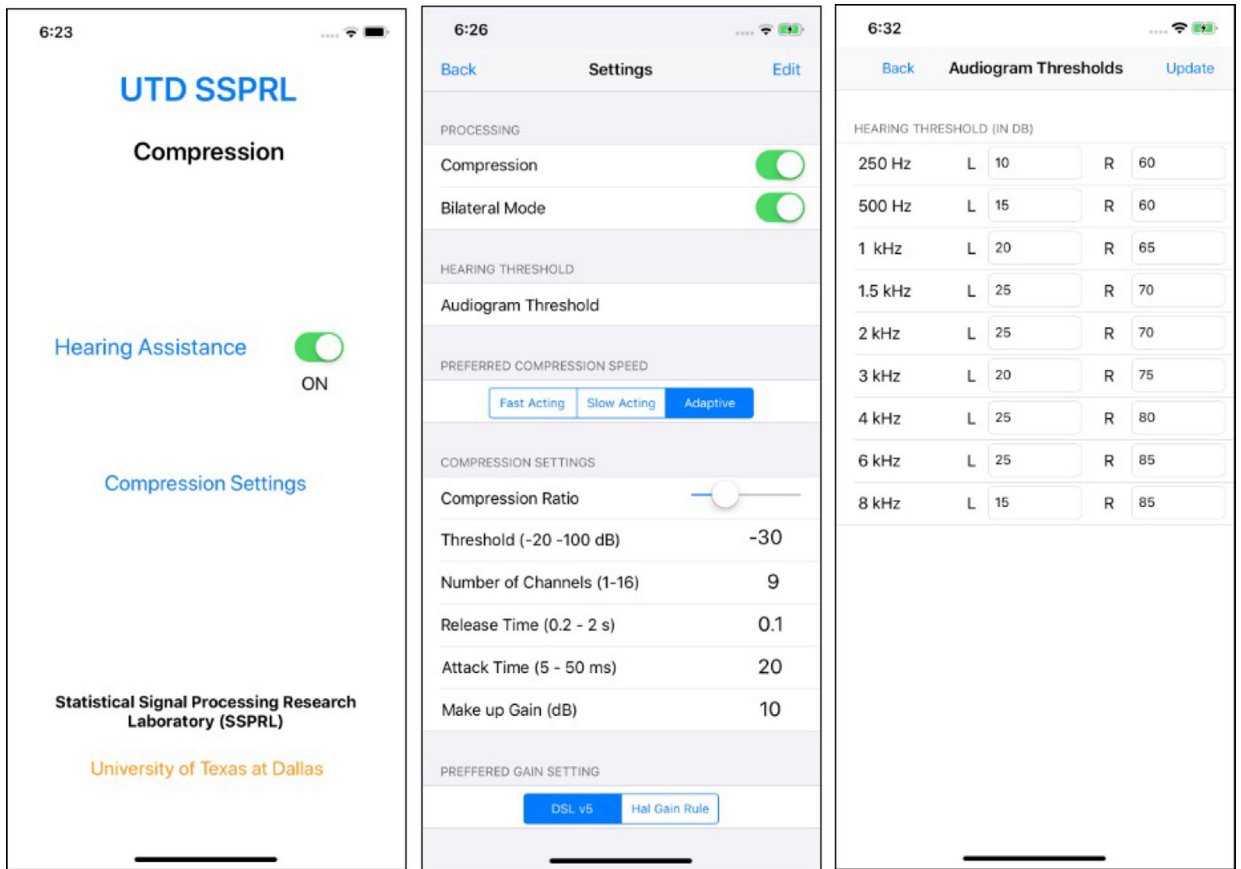Interpolation of hypothetical insertion gain for nine bands, $K = 1024$, $f_s = 16000$

Fig. 4:
Screenshot of the developed smartphone application

Fig. 5:
Comparison of hearing aid output with developed smartphone application on Fonix 8000 hearing aid analyzer. Curve 1 (green) represents the output of the smartphone-based hearing assistance, and Curve 2 (blue) represents the output of a Starkey Halo Hearing aid

TABLE I:

Insertion gain values based on half-gain rule

Center Frequencies of 9 Bands (kHz)	0.25	0.5	1	1.5	2	3	4	6	8
Hearing Threshold (dB HL)	60	60	65	70	70	75	80	85	85
Insertion gain (dB)	20	25	32.5	35	35	37.5	40	42.5	42.5

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table II:

Objective performance of the proposed compression and respective scores

HASPI (0–1)	HASQI (0–1)	CSII (0–1)	PESQ (0–4.5)
0.99	0.89	0.92	4.2