The Acoustically Transparent Hearing Device: Towards Integration of Individualized Sound Equalization, Electro-Acoustic Modeling and Feedback Cancellation

Florian Denk\textsuperscript{1,4}, Steffen Vogl\textsuperscript{3}, Henning Schepker\textsuperscript{2,4}, Birger Kollmeier\textsuperscript{1,4}, Matthias Blau\textsuperscript{3,4}, Simon Doclo\textsuperscript{2,4}

\textsuperscript{1} Medizinische Physik, University of Oldenburg, Germany
\textsuperscript{2} Signal Processing Group, University of Oldenburg, Germany
\textsuperscript{3} Institut für Hörtechnik und Audiologie, Jade Hochschule Oldenburg, Germany
\textsuperscript{4} Cluster of Excellence Hearing4all

\{florian.denk, simon.doclo\} @uni-oldenburg.de

Abstract

Assistive hearing devices often suffer from a low acceptance by the end user due to poor sound quality. Recently, a novel acoustically transparent hearing device was developed that aims at increasing the acceptance and benefit, also for (near-to) normal hearing people, by providing better sound quality. The hearing device integrates three microphones and two receivers and can be calibrated in-situ in an attempt to conserve the open-ear sound transmission characteristics of an individual person.

To further improve the quality of acoustic transparency and extend the functionality of the hearing device, we outline the integration of further models and algorithms. Electro-acoustic models of the device can improve adjustment to transparency by providing a better estimate of the pressure at the eardrum with an in-ear microphone. In addition, the multi-microphone device layout allows the development of custom feedback cancellation algorithms by means of a beamformer in order to robustly steer a spatial null towards the hearing device receiver.

1. Introduction

Despite a great improvement in hearing technology in the past decades, the acceptance of assistive hearing devices is still limited, partially due to poor sound quality \cite{1, 2, 3}. This is particularly true for potential first-time users with a mild-to-moderate hearing loss or even (near-to) normal hearing. While they would benefit from features like speech enhancement or amplification in acoustically challenging situations, they are usually not willing to accept a general degradation of the sound quality. Therefore, an important challenge is to develop a device that is acoustically transparent, i.e., that allows hearing comparable to that of the open ear while being capable of providing a desired sound enhancement at the eardrum. These principles can be applied not only to hearing aids, but also to consumer products, e.g., hearables \cite{4, 5}.

We recently developed a prototype of an acoustically transparent hearing device that can be individually calibrated aiming to preserve the open-ear sound transmission characteristics of the particular user, even if the ear canal is partially occluded \cite{6}. The used sound equalization approach exploits the microphone positions of a novel vented multi-microphone earpiece, including an in-ear microphone for monitoring the pressure in the ear canal. Acoustical transparency on the perceptual level was verified in a subjective listening experiment \cite{6}, and convincing sound quality with the device was observed for normal hearing subjects \cite{7}. Nevertheless, the need for improving transparency in a physical sense was revealed in a recent technical evaluation \cite{8}. Furthermore, other processing stages might interact with the desired goal of acoustic transparency.

After presenting the hardware of the device in Section 2, in this paper we first review the sound equalization approach to achieve acoustic transparency in Section 3, and then present two approaches that aim at improving and completing its functionality towards a full acoustically transparent hearing device. To improve the acoustic transparency feature, a promising approach is to include electro-acoustic models of the device \cite{9}. These models provide an accurate estimate of the sound pressure at the eardrum with an in-ear microphone, which is key to precise sound equalization in a non-occluding fit. Principles and first results comparing the estimated and measured pressure at the eardrum are outlined in Section 4. In addition, the multi-microphone hardware layout facilitates feedback cancellation using a beamformer with a spatial null steered towards the receiver \cite{10, 11, 12} in addition to state-of-the-art adaptive feedback cancellation methods \cite{13}. The principle is briefly introduced and potential interactions of the null-steering approach with the aim of providing acoustic transparency are evaluated in Section 5. Challenges resulting from integrating all approaches are discussed in Section 6.

2. Hardware

The custom in-the-ear type earpiece with relatively open acoustical properties is depicted in Figure 1. A schematic drawing is shown in Figure 2, together with the filter stages in sound equalization (see Section 3) and feedback cancellation (see Section 5), as well as references to the electro-acoustic model (see Section 4).

All electronic components are removably fitted into an individual silicone earmould that fills the concha bottom. In total, the device contains 3 microphones and 2 receivers. Two microphones (Type Knowles GA-38) and two balanced armature receivers are located in an acrylic tube referred to as the core, which is inserted into a bore through to the ear canal. The first microphone is located at the inner face of the core and points towards the eardrum (“in-ear microphone” with output voltage $y_1$ and pressure $p_1$) and serves to monitor the sound pressure in the ear canal. The second microphone is located at the outer face of the core and points outwards (“entrance microphone”, with output voltage $y_2$ and pressure $p_2$). The third microphone

89
Assembled earpiece, from [6]. 3 microphones and 2 receivers are fitted into an individual silicone earmould.

("concha microphone", Type Knowles FG-23329, with output voltage $y_3$ and pressure $p_3$) is placed in the back of the concha by flush insertion into a hole. The two independent receivers are positioned next to the microphones at both ends of the core, but both pointing towards the eardrum. The inner one is a tweeter (Knowles WBFLK-30019, with input voltage $u_1$) and the outer one a woofer (Knowles FK-26768, with input voltage $u_2$). Note that although included in the electro-acoustic model, the woofer is not currently used in operation, i.e., it is not considered in sound equalization and feedback cancellation, which is indicated by a dashed line in Figure 2. The hearing device is connected to a PC for real-time signal processing via a soundcard and a custom supply and amplifier box.

The residual space in the core between the microphones and receivers forms a vent, to increase wearing comfort by ventilation. Lower part: Corresponding elements and circuit of the electro-acoustic model. Approximatively contains the direction-dependent portion of the transfer function to the eardrum, and the optimal gain function is the relative transfer function between the concha microphone location and the eardrum in the individual ear. In [6], a flat gain function was used, with the extension that the direct sound leaking through the individual core is considered.

To achieve sound equalization to the target pressure, the filter $G$ of the hearing device is adjusted in a calibration routine conducted in-situ, i.e., when the device is inserted into the ear. The concha microphone is used to pick up external sound. Assuming that the pressure at the eardrum and the in-ear microphone are similar, the pressure at the eardrum generated by the external sound source and the active device is estimated using the in-ear microphone. Based on the observed deviation from the target pressure, the filter $G$ is adapted until convergence is achieved.

3. Achieving Acoustic Transparency by Individualized Sound Equalization

3.1. Principles

Acoustic transparency is achieved, when the superposition of direct sound leaking through the core and the electro-acoustically reproduced sound at the eardrum is physically or perceptually equal to the pressure that would be present with an open ear. Achieving acoustic transparency can be separated into two problems: First, the pressure at the eardrum with an open ear has to be estimated based on the available microphone signals to compute the so-called target pressure. Second, the device has to be adjusted such that the target pressure is generated at the eardrum of the individual subject when the device is in the ear, i.e., sound equalization is performed.

In [6], the target pressure was defined as the pressure at the concha microphone, multiplied with an appropriate frequency-dependent gain function. This strategy is justified by observations from spatial audio technology showing that the relative transfer function between a recording point near the (blocked) ear canal entrance and the eardrum of an open ear is not direction-dependent [16]. Thus, the concha microphone approximately contains the direction-dependent portion of the transfer function to the eardrum, and the optimal gain function is the relative transfer function between the concha microphone and the eardrum in the individual ear. In [6], a flat gain function was used, with the extension that the direct sound leaking through the individual core is considered.

To achieve sound equalization to the target pressure, the filter $G$ of the hearing device is adjusted in a calibration routine conducted in-situ, i.e., when the device is inserted into the ear. The concha microphone is used to pick up external sound. Assuming that the pressure at the eardrum and the in-ear microphone are similar, the pressure at the eardrum generated by the external sound source and the active device is estimated using the in-ear microphone. Based on the observed deviation from the target pressure, the filter $G$ is adapted until convergence is achieved.

3.2. Current limits and possible extensions

In psychoacoustic experiments with normal-hearing subjects, satisfactory results in terms of acoustic transparency on the perceptual level were observed [6, 7]. However, physical evaluations still reveal some deficits with the current sound equalization approach. Figure 3 shows measurements of the Real-Ear Insertion Gain (REIG) of the transparent hearing device prototype as presented in [6]. The REIG is the difference between the sound pressure at the eardrum measured when the device is inserted and with an open ear. Acoustic transparency on a physical level is achieved if the REIG is 0 dB across all frequencies. The measurements were conducted in a free-field environment in both ears of 12 subjects, and include 3 incident directions in the horizontal plane (azimuth $\theta = 0^\circ, 90^\circ, -135^\circ$).

The measured REIGs deviate from 0 dB, particularly for
frequencies above 2 kHz. The error is notably different between subjects and incidence directions, and the variation increases with frequency. This result shows that there is room for improvement in acoustic transparency, which may be tackled with various approaches.

Most of the observed error in the REIG can be explained by two factors: errors in the estimation of the target pressure, and inaccuracies in the sound equalization due to incorrect estimation of the pressure at the eardrum. To estimate the pressure at the eardrum, in [6] the pressure at the in-ear microphone was used, which in most cases introduces an individual estimation error of up to ±20 dB, which is highly variable across frequencies [8]. Thus, the sound equalization error could be reduced if a better estimate of the pressure at the eardrum were available. Electro-acoustic modeling approaches can be used for this purpose, which are treated in Section 4.

In addition, the occurrence of acoustic feedback due to the acoustic coupling between the receiver and the concha microphone has been neglected so far. While this is possible when only the concha microphone is utilized for sound pickup and the applicable gain is limited, appropriate feedback management is a prerequisite when larger amplification than for acoustic transparency is required, or when both external microphones are used for sound pickup, e.g., when implementing a directional microphone. Feedback cancellation techniques tailored to the custom hardware layout are reviewed in Section 5, where possible interactions with acoustic transparency are examined.

4. Electro-Acoustic Model

In our previous work [9], we proposed an electro-acoustic model, which serves to better understand the underlying physical principles of sound transmission in the hearing device, and to estimate quantities at locations where they cannot be directly measured, e.g., the sound pressure at the eardrum. The current focus is to predict the sound pressure at the eardrum $p_d$ in vivo, based on measurements using the microphones of the hearing device only.

The model is made up of lumped elements and two-port networks, as depicted in Figure 2. The middle part is the core, which can be regarded as fixed over individual subjects. On the other hand, both terminations, i.e., the external sound field and the residual ear canal, are individual to every ear. The complete model cannot be determined in one step, but is built up in a series of measurements and calculations that are described in the following.

4.1. Model of the Core

First, the model of the core is obtained. It consists of:

- two microphones, characterized by their sensitivity measured prior to assembling the core, each converting its output voltage signal $v_{m}$ to the corresponding pressure $p_m$.
- two receivers, which are modeled as ideal volume velocity sources, delivering the flux $q_m$. This source parameter was also measured prior to assembling the core, according to the technique described by [17].
- the vent, represented by three acoustic transmission lines modeled as two-port networks $A_{1,2,3}$ according to [18]. The three parameters of each transmission line (length, radius, and a loss factor) need to be fitted by referring to acoustic measurements. The microphones and receivers are coupled into the vent at locations depicted in Figure 2.

To fit the free parameters of the transmission lines, the assembled core was coupled to a training setup with known termination impedances, and all four transfer functions between the two receivers and microphones 1 and 2 were measured. The medial termination was an IEC711 coupler, while at the lateral end the core was mounted in a baffle. The optimal parameters of the two-port networks were found by minimizing the differences between the measured and modeled transfer functions. Good agreement and computational effectiveness could be achieved with the Nelder-Mead-Simplex [19] algorithm, where the parameter values were constrained to realistic boundaries.

4.2. Model of the Individual Ear

In a second step, a model of the individual ear is estimated. It contains both terminations of the core, as shown in Figure 2. The external sound field (outer termination of the core) is characterized by the radiation impedance $Z_{rad}$, which can be further split into the transfer impedance $Z_{2,3}$ between the outer core end and the concha microphone, and a remaining impedance $Z_{rad} - Z_{2,3}$. $Z_{rad}$ is approximated by the physical model of a piston in baffle. The model of the individual ear canal $E$ (medial termination of the core) is individualized based on measurements in the ear of a subject. It is composed of four cascaded acoustic transmission lines with four radii and one total length as parameters, and two parallel load impedances $Z_l$ and $Z_{l,\text{residual}}$, located medially (across them $p_d$ is produced), $Z_l$ is a purely resistive frequency-independent load to represent losses, $Z_{l,\text{residual}}$ is complex valued and frequency dependent.

Assuming the core model and the outer termination are known and using the transfer function measurements from any of the two receivers to the in-ear microphone, the acoustic impedance $Z_{ee}$ at the point of $p_1$ (i.e., the in-ear microphone) in the direction towards the ear canal can be calculated. Then, the parameters of the individual ear canal model $E$ are fitted by minimizing both level and phase differences between measured and modeled impedances $Z_{ee}$, summed across the frequencies from approximately 1 to 15 kHz. Again, the Nelder-Mead-Simplex algorithm was applied with realistic boundaries. Since it was observed that results depended on initial values, 500 random initial values were used and the lowest cost result taken.

Several studies (e.g., [20, 21]) have shown that $Z_{ee}$ - or the reflectance derived from it - can be used to estimate an ear canal model $E$ and ultimately predict the sound pressure at the eardrum $p_d$.  

Figure 3: Real-Ear Insertion Gain (REIG) after 1/6 octave smoothing measured with the hearing device prototype as presented in [6]. The data includes measurements in both ears of 12 subjects, with 3 incident directions in the horizontal plane.
of the microphones to steer a spatial null towards the position of the hearing device receiver. Note that only the inner receiver of the device is considered here.

Several optimization approaches for calculating the null-steering beamformer coefficients have been proposed, including a robust least-squares design [10, 11] and a robust min-max design [12] aiming at directly maximizing the maximum stable gain of the hearing device, i.e., the gain before the closed-loop system becomes unstable. Furthermore, the benefit of combining a fixed null-steering beamformer and an AFC algorithm based on the prediction-error-method to cancel residual feedback has recently been shown [13]. However, in none of the presented null-steering beamformer optimization approaches [10, 11, 12], the preservation of the pickup microphone directional response that is required for achieving acoustic transparency has been taken into account. This implies that the null-steering beamformer may alter spectral directional cues and bias spatial perception, e.g., sound localization. Therefore, after briefly introducing the optimization procedure, in the following we analyze the directional response of the fixed null-steering beamformer.

We assume time-invariance of the acoustic feedback paths \( H_m(n) = H_m \), \( m = 1, \ldots, M \) between the receiver and the \( m \)th microphone. Assuming the availability of \( I \) measurements of the acoustic feedback paths (e.g., obtained by prior measurement), the coefficients of the null-steering beamformer \( \mathbf{B} \) are obtained by minimizing the following least-squares cost-function [11]

\[
J_{LS}(\mathbf{b}) = \sum_{i=1}^{I} \| (\mathbf{H}^{(i)})^T \mathbf{b} \|_2^2.
\]

where \( \mathbf{b} \) is the \( ML_B \)-dimensional vector of the beamformer coefficients and \( \mathbf{H}^{(i)} \) is the \( ML_B \times (L_d + L_H - 1) \)-dimensional matrix of concatenated convolution matrices of the acoustic feedback paths from the \( i \)th measurement, \( i = 1, \ldots, I \), with \( L_d \) the number of beamformer coefficients for each microphone and \( L_H \) the length of the acoustic feedback path. To prevent the trivial solution of \( \mathbf{b} = 0 \), the beamformer coefficients in a reference microphone \( m_0 \) are constrained to correspond to a delay of \( L_d \) samples, i.e.,

\[
\mathbf{b}_{m_0} = \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^T_{L_d}.
\]

5. Feedback Cancellation

Acoustic feedback occurs when a signal is picked up by a microphone, amplified, played back by a receiver and picked up again by the microphone, creating a closed-loop system. In hearing devices, adaptive feedback cancellation (AFC) is typically used to reduce the detrimental effect of acoustic feedback, which is most often perceived as howling or whistling. In AFC, an adaptive filter is used to estimate the acoustic feedback path between the hearing device receiver and the microphone, theoretically allowing for perfect feedback cancellation [22]. However, due to the closed-loop electro-acoustic system, the estimate of the acoustic feedback path is generally biased [23, 24]. Several algorithms have been proposed with the aim of reducing this bias, where the so-called prediction-error-method [24] seems most promising. While an AFC algorithm can be applied for any hardware layout, the considered multi-microphone setup (cf. Figure 2) additionally allows for the use of multi-microphone feedback cancellation approaches. This includes a fixed null-steering beamformer that exploits the spatial diversity of the feedback paths (e.g., obtained by prior measurement) to steer a spatial null towards the position of the hearing device receiver. Note that only the inner receiver of the device is considered here.

Several optimization approaches for calculating the null-steering beamformer coefficients have been proposed, including a robust least-squares design [10, 11] and a robust min-max design [12] aiming at directly maximizing the maximum stable gain of the hearing device, i.e., the gain before the closed-loop system becomes unstable. Furthermore, the benefit of combining a fixed null-steering beamformer and an AFC algorithm based on the prediction-error-method to cancel residual feedback has recently been shown [13]. However, in none of the presented null-steering beamformer optimization approaches [10, 11, 12], the preservation of the pickup microphone directional response that is required for achieving acoustic transparency has been taken into account. This implies that the null-steering beamformer may alter spectral directional cues and bias spatial perception, e.g., sound localization. Therefore, after briefly introducing the optimization procedure, in the following we analyze the directional response of the fixed null-steering beamformer.

We assume time-invariance of the acoustic feedback paths \( H_m(n) = H_m \), \( m = 1, \ldots, M \) between the receiver and the \( m \)th microphone. Assuming the availability of \( I \) measurements of the acoustic feedback paths (e.g., obtained by prior measurement), the coefficients of the null-steering beamformer \( \mathbf{B} \) are obtained by minimizing the following least-squares cost-function [11]

\[
J_{LS}(\mathbf{b}) = \sum_{i=1}^{I} \| (\mathbf{H}^{(i)})^T \mathbf{b} \|_2^2.
\]

where \( \mathbf{b} \) is the \( ML_B \)-dimensional vector of the beamformer coefficients and \( \mathbf{H}^{(i)} \) is the \( ML_B \times (L_d + L_H - 1) \)-dimensional matrix of concatenated convolution matrices of the acoustic feedback paths from the \( i \)th measurement, \( i = 1, \ldots, I \), with \( L_d \) the number of beamformer coefficients for each microphone and \( L_H \) the length of the acoustic feedback path. To prevent the trivial solution of \( \mathbf{b} = 0 \), the beamformer coefficients in a reference microphone \( m_0 \) are constrained to correspond to a delay of \( L_d \) samples, i.e.,

\[
\mathbf{b}_{m_0} = \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^T_{L_d}.
\]
To obtain the beamformer coefficients, we first measured the acoustic feedback paths of the hearing device in the left ear of a dummy head with adjustable ear canals [25], both in free-field and with a hand very close to the ear, using a sampling rate of 32 kHz. The beamformer coefficients were then computed by minimizing (1) subject to the constraint in (2) for $M = 3$ microphones (in-ear, entrance and concha microphone), $L_B = 32$, $m_0 = 2$ (entrance microphone), $L_d = 16$ and $I = 2$. The resulting added stable gain, i.e., the increase in gain margin compared to using only the entrance microphone ($m = 2$), was 18.3 dB and 22.6 dB for the free-field condition and the hand condition, respectively.

To compute the directional response of the null-steering beamformer for an incoming signal, the acoustic transfer functions to the microphones $D(\theta_j)$, $j = 1, \ldots, J$ were measured for $J = 24$ equidistantly spaced angles $\theta_j$ surrounding the dummy head at a distance of approximately 2.5 m in the horizontal plane. Figure 5 shows the directional response of the beamformer $\tilde{D}(\theta_j) = B^T D(\theta_j)$ for multiple frequencies relative to the directivity $D_2(\theta_j)$ of the entrance microphone. Ideally, the relative directional response would be equal to 0 dB for all frequencies and incidence angles. However, the response is different from 0 dB for most of the considered frequencies and directions. Nevertheless, for most frequencies and incident angles the null-steering beamformer alters the directivity only by approximately ±4 dB.

6. Discussion and Summary

The principles of an acoustically transparent hearing device presented in [6] and physical evaluation results have been reviewed in Section 3, and possible extensions towards improving and extending its functionality have been presented in Sections 4 and 5. While the good performance of electro-acoustic modeling and customized feedback cancellation for the hearing device has been demonstrated for these approaches individually, a next challenge is the integration of the two approaches with the transparency feature of the hearing device in real-time operation.

Unbiased estimation of the pressure at the eardrum can improve acoustic transparency by improving sound equalization to a target pressure at the eardrum. The electro-acoustic model presented in Section 4 is able to predict the pressure at the eardrum that is generated by the hearing device receiver accurately up to approximately 6-7 kHz in magnitude and phase. However, when estimating the sound pressure at the eardrum in normal operation, the superposition with the direct sound leaking through the vent needs to be considered. Since the model is in principle also able to predict the pressure generated by the receiver at the in-ear microphone, this predicted pressure can be subtracted from the observed pressure at the in-ear microphone to obtain an estimate of the direct sound only. The pressure at the eardrum generated by the direct sound alone can then also be predicted. It should be noted that for integrating the electro-acoustic model with the acoustically transparent hearing device, it is sufficient to extract all relevant transfer functions from the model after calibration measurements.

Although the null-steering beamformer presented and evaluated in Section 5 yields impressive results in terms of feedback cancellation, it was also noted that it introduces a direction-dependent bias compared to the reference microphone. Spectral directional cues contained in the reference microphone signal are thus altered, which may introduce perceptual errors regarding spatial hearing or other undesired artifacts. However, the deviations are in the range of about ±4 dB, and their perceptual relevance is not yet clear. Another issue is the delay of $L_d$ samples introduced by the beamformer, which should be considered when designing the equalization filter $G$. In principle, this can be achieved by performing the in-situ calibration [6] with the beamformer output as hearing device input signal.

The electro-acoustic models of the individual ear, as well as the null-steering beamformer require knowledge of the transfer functions between the hearing device receivers and microphones. They can be measured in-situ in the individual ear using only the device as part of calibration measurements, which are also necessary to achieve transparency [6]. Gathering these data is therefore no practical obstacle to integrating the electro-acoustic model and the null-steering beamformer into a future version of the prototype.

In conclusion, there seem to be no principal problems hindering the integration of electro-acoustic models and customized feedback cancellation methods into our prototype hearing device. Both are promising approaches to improving the sound equalization to achieve acoustic transparency in our prototype hearing device, as well as increasing its functionality in more realistic application scenarios with higher gain settings and more than one pickup microphone in each side. Future work will hence focus on the implementation of the presented approaches to construct an improved version of our acoustically transparent hearing device.

7. Acknowledgement

This work was supported by the Research Unit FOR 1732 Individualized Hearing Acoustics and the the Cluster of Excellence Hearing4all, both funded by the German Research Council DFG.

8. References


