EVALUATION OF A COMPLEMENTARY HEARING AID FOR SPATIAL SOUND SEGREGATION

L. Giuliani, L.Brayda *

Fondazione Istituto Italiano di Tecnologia
Genova, Italy

S. Sansalone, S. Repetto, M. Ricchetti *

Linear s.r.l.
Genova, Italy

ABSTRACT

Spatial segregation of sounds is a common and simple task for healthy hearing people. Unfortunately, people who suffer from partial hearing loss have great troubles in separating sound sources in crowded and noisy environments. Social isolation is the most common consequence and hearing aids are not a solution, especially in severe noisy conditions, because of their limited directionality.

In this work we present the Glassense, a wearable device designed to be a complement for traditional acoustic prostheses. It exploits microphone arrays to spatially filter sounds, so that frontal speech sources are preserved, while competing noise from the sides and the back is attenuated.

This behaviour, like an "acoustical lens", changes the ratio between interesting frontal sounds and other sources, improving speech reception performance of subjects. The use of Glassense in severe noisy environments can promote social inclusion for people who suffer from partial hearing loss.

Index Terms— Hearing aid, microphone arrays, beamforming, real time, binaural, disability

1. INTRODUCTION

Partial hearing loss is a very diffuse pathology across the world. 360 million people suffer from disabling loss greater than 40dB on the better hearing ear (30dB for children) and it affects approximately one-third of people over 65 years of age [1]. Modern prostheses are able to give back hearing to patients in most occasions of daily life. These devices perform noise reduction and dereverberation tasks, but their ability to increase the ratio between interesting sounds and noise is not sufficient to ensure clear speech comprehension in noise.

The brain usually processes binaural cues to locate and isolate possible disturbances [2], solving the so-called "cocktail party problem". The presence of hearing loss and the use of acoustical prostheses modify the cited cues in a way that forces the patient to make a great mental effort to tell apart different sound sources and that limits the performance in this task. A simple acoustic amplification is not an answer to this problem since what hearing impaired people really need is to amplify the sounds they are interested in more than competing noise.

Previous works like [3] show that the use of one microphone per ear is better than having one microphone only in the most damaged ear. Moreover, it has been shown that binaural multi-microphone hearing aids can decrease the effect of competing spatial sound sources [4]. The use of arrays of microphones allows to apply beamforming: a spatial signal processing technique that exploits constructive and destructive interferences to filter sounds coming from specific directions.

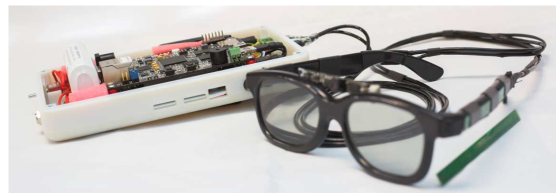


Fig. 1. The Glassense device

A lot of different microphone arrays devices have been designed in the recent years in order to apply spatial sound segregation techniques. Some of these solutions directly exploit prostheses [5] or cochlear implant [6] microphones as two elements arrays.

An interesting and effective device has been presented by Widrow in 2001 [7] and contained a microphone array within a necklace. The solution appeared to be really effective in increasing speech comprehension capabilities of the users, but it does not seem to be available on the market any more.

A recent device which has some characteristics in common with the one we are illustrating in this work has been proposed by Boone in 2006 [8], as a follow up of Merks work [9]. It is a pair of glasses equipped with two array of microphones, one for each temple. The device is currently available on the market with the name of *hearing glasses* and is produced by Varibel company.

In 2011 Mens [10] analysed the performance of the hearing glasses using several beamforming modalities on some hearing impaired subjects. The results of this last work witness a solid reinforcement in speech comprehension capabilities of hearing impaired subjects using the hearing glasses in cocktail party-like noise.

It is worth noting that the hearing glasses act as a substitute for a hearing prosthesis and that it is necessary to configure the device on the subjects hearing loss profile. The device we propose in this work, on the other side, is a generic system that can potentially be a complement, and not a substitute, to any existing - already personalized - hearing prosthesis.

In our last work [11] we tested the efficacy of the beamforming algorithm of Glassense on healthy hearing subjects providing the acoustical feedback by common earplugs. In this work we repeated the last experiment, with a mixed sample of healthy hearing and hearing impaired subjects, connecting the Glassense to their hearing prostheses and observing the improvement in speech under-

*This work is partly supported by the Ligurian PAR-FAS grant Glassense (CUP G35C13001360001) and partly by the EU FP7 grant BLINDPAD (grant number 611621).

standing capabilities.

2. THE GLASSENSE SYSTEM

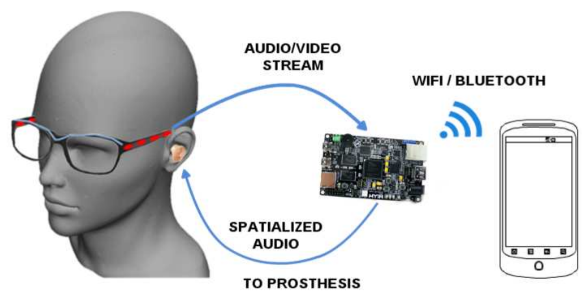


Fig. 2. The Glassense system. The audio captured by the microphones is sent to the elaboration board, which performs filtering and sends the resulting signal to hearing aids. It is possible to control the device with a smartphone.

In this work we illustrate and validate a wearable device designed to complement common hearing aids in order to improve speech comprehension capabilities of the patient in noisy conditions. We deliver spatial selectivity to binaural audio input by means of two microphone arrays. Using a beamforming algorithm the Glassense can spatially filter acoustical sources and attenuate sounds coming from the sides and the back of the listener, while preserving those coming from the frontal direction. In principle, the listener can use the head motion as a spatial selector, deciding the direction of interest in which the sounds are preserved.

The Glassense system [12], depicted in Fig. 2, is composed by two superdirective microphone arrays positioned on the temples of a pair of glasses. An elaboration board is used to acquire and process audio signals and to send them to the ears by hearing aids or common earplugs, granting that binaural audio output is processed in quasi real-time.

Each microphone array contains 4 digital MEMS microphones. The elaboration board consists in a MIYR Z-turn, a low-cost Linux-based development board based on ARM processing system, paired with a custom daughterboard. A data-independent filter-and-sum beamforming has been used to design two specular end-fire linear microphone arrays [13]. The array has 4 equally spaced elements over 0.10 m and the beamformer structure is composed by 4 FIR filters of order 127.

The working frequency band ranges from 400 Hz to 4 kHz and the sampling frequency is 16 kHz. The designed array is superdirective for frequencies up to 2300 Hz (i.e. more than 2 octaves). The directivity index (DI) has a mean value of 8.6 dB over the array working frequency band. The same array design and beamformer structure has been used for the microphone array pair.

In Fig. 4 we compared the theoretical representation of the Beam Power Pattern (BPP) with some measures made using the actual device. Image (a) shows the theoretical BPP of the filter at various frequencies in a Cartesian plot. As one notes, the array has the highest response at 90°and a sidelobe at the opposite end-fire, i.e. 270°, between 5 and 15 dB lower.

Image (b) is a measure of half of the symmetrical polar pattern in free field (i.e. without other objects near the device), performed with a pure tone at 1 KHz frequency generated step by step around

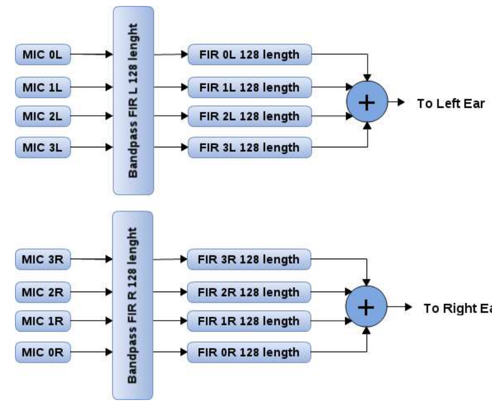


Fig. 3. A block diagram illustrating the implementation of the filtering logic of the Glassense

the device. The 15° steps applied collecting the data, together with imperfections in the room, caused some differences from the theoretical one. Image (c) represents the stereo beam pattern (left signal in the first graph, right signal in the second) perceived by the device mounted on a KEMAR-like dummy-head. It is possible to notice the masking effect of the head, which partially modifies the shape of the pattern.

The purpose of this work is to verify if the use of an acoustic beamforming algorithm can increase the speech intelligibility for hearing impaired people in noisy environments. The methodology we followed in order to verify the research question is illustrated in the next section.

3. EXPERIMENTAL SETUP

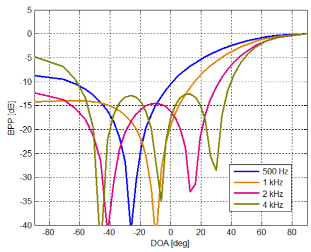
In the next paragraphs we will describe the methodology adopted during the execution of the experiment.

3.1. Participants

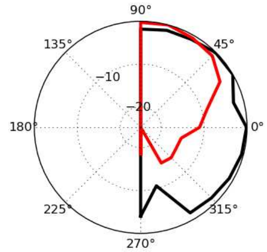
Eight subjects (mean age 72, range 59-78 yr) were recruited. Six of them suffers from light to severe sensorineural or mixed (sensorineural and transmissive) hearing loss. The two remaining subjects hearing was healthy. Three of the subjects also had visual impairment (fragile or low vision).

3.2. Setup

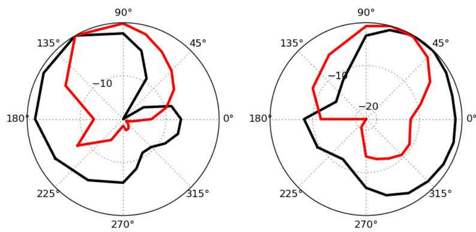
The experiment has been performed in a sound attenuated audiometric booth (size = 3.2 x 4.8 x 2.73 m). The room was lined with 50 mm acoustic foam to create a semianechoic environment (T60 = 0.2 s). Subjects were seated in the middle of Four Behringer active loudspeakers, 1 m away from them at 0°(directly ahead), 180°(back) and ±90°on either side of the midline. All subjects were facing towards the loudspeaker at 0°, although their heads were not constrained. They had to wear the Glassense device and a couple of in-the-ear hearing aids realized and calibrated for them by Linear s.r.l. according to their hearing loss. Healthy hearing subjects also worn prostheses configured without hearing loss compensation. Each prosthesis contained a microphone, a speaker and an Onsemi SOC SA3291 module (Wireless DSP for hearing prostheses).



(a) Theoretical Cartesian pattern



(b) Free field polar pattern at 1 KHz



(c) Dummy-head stereo polar pattern at 1 KHz

Fig. 4. Comparison between the theoretical and measured Beam Power Patterns of the Glassense. Image (a) shows the theoretical Cartesian shape of the pattern. Image (b) shows half of the symmetrical polar pattern of the Glassense in *unfiltered* (black line) and *beamforming* (red line) condition in free field at 1 KHz. Image (c) shows the 1 KHz stereo beam patterns of the Glassense worn by a KEMAR-like dummy-head.

The Glassense had been connected to an Entratech relay-device able to send audio signals by Near-Field Magnetic Induction (NFMI) technology to Onsemi SOC SA3291 modules. The relay device acted as a bridge, sending audio signal from Glassense to the hearing aids of the subjects. The volume of the signal sent to the prostheses had been calibrated in order to homogenize acoustical intensities of frontal sources in each acoustical modality of the Glassense.

3.3. Stimuli

Two kinds of stimuli, part of a wide-band corpus between 20 Hz and 20 kHz, have been proposed to the subjects according to ISO 8253-3:2012 [14] which specifies the requirements for the composition, validation and evaluation of speech test materials.

- The *target speech* stimuli consisted in 20 lists of 10 bisyllabic italian words extracted from the Bocca Pellegrini elaborate corpus [15] and was played from the speaker located in front of the user, at fixed volume.
- The *competing noise* stimuli was a four channel registration of a cocktail party speech, reproduced by all the four speakers.

Since the *target speech* volume was fixed for each subject, we modified the volume of the *competing noise* across trials, in order to obtain differences in the ratios of the two stimuli, i.e. the Signal to Noise Ratio (SNR).

Fig. 5. The subject head was always directed towards the frontal speaker, which reproduced the target speech signal

3.4. Procedure

Before starting the experiment, all the subjects carried out a free ear diffuse-field speech audiometry test. The purpose of this was to find the correct volume of the *target speech* at which they were able to correctly listen and repeat the words. Once found, the volume of the speech was fixed for the subject across all the trials.

After this procedure the experiment started: the subjects had to wear the Glassense device and perform a sequence of listen and repeat tasks. During each trial they had to listen to the *target speech* and repeat the words they could understand, while the *competing noise* was reproduced by all the speakers.

Two conditions affected the trials:

- The SNR value, which spanned between -15 and +15 dB. An SNR equal to 0 means the same level of signal and noise. The various ratios were obtained modifying the volume of the *competing noise* while the *target speech* was constant.
- The acoustical modality, which had three possible values (randomized across subjects):
    1. **Hearing Aids** - The subject listens directly by the hearing aid.
    2. **Unfiltered** - The Glassense provides the acoustical signal to the hearing aid. Only the microphone positioned above the pinna is exploited for each temple and no filter is applied. This is a simulation of the behaviour of an omnidirectional behind-the-ear prosthesis.
    3. **Beamforming** - The Glassense provides the acoustical signal to the hearing aid. The device exploits all the microphones and applies the beamforming algorithm to suppress part of the noise coming from the sides and the back.

For each task the registered performance consisted in the percentage of correctly repeated words.

4. DATA ANALYSIS AND RESULTS

Each subject performed listen and repeat tasks in presence of cocktail party noise at several SNRs in each of the acoustical modalities described in the procedure. The performance profile across SNR values is common for each of them: with louder noise (i.e. low SNR) the performance decreases until it goes to 0%. On the contrary, low noise brings the performance to increase, up to 100%. We interpolated the performance over SNR values obtaining three logistic curves for each subject, related to the various acoustical modalities.

The standard threshold used to analyse speech comprehension capabilities is 50% and is called Speech Reception Threshold or SRT. In our data analysis we compared the SRT values obtained from the curves related to each acoustical modality in all subjects. For the analysis of the collected data we adopted non-parametric tests, since the SRT populations were made by only 8 values each and it is difficult to assert normality of such populations.

We performed a Friedman Test and we found a significant difference between at least one of the SRT populations from the others (p-value = 0.007635). Since we were interested in beamforming capabilities of increasing speech recognition, we tested the corresponding SRTs against the other two populations by Wilcoxon-Mann-Whitney test, applying Bonferroni post-hoc correction. The beamforming modality mean SRT resulted 2.4 dB lower respect to the hearing aid (p-value = 0.007813) and 3.3 dB lower respect to the unfiltered modality (p-value = 0.02344).

A lower SRT value corresponds to a lower SNR related to 50% guess rate performance. This means that using the beamforming filtering algorithm the subjects were able to correctly understand half of the target speech words with an higher level of competing noise with respect to the hearing aids and unfiltered modality.

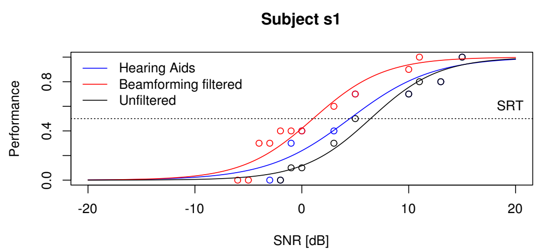


Fig. 6. Example of how the performance of subjects across different SNR values has been interpolated with logistic curves. The dotted horizontal line indicates the 50% performance, corresponding to Speech Reception Threshold.

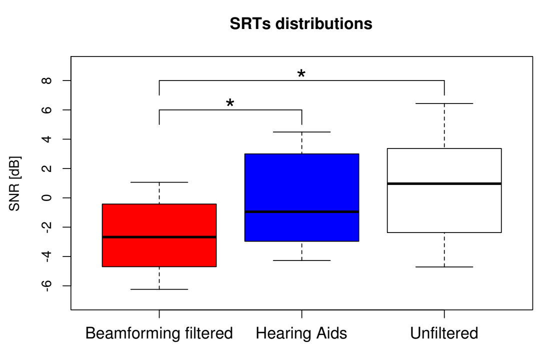


Fig. 7. Distribution of the Speech Reception Threshold of the subjects across the different acoustic modalities. The beamforming condition gives the best results, as it enables the subjects to reach the 50% guess rate with higher levels of competing noise.

Compared to our results, the SRT improvement asserted by Mens in [10] using the hearing glasses is higher: 6.3 dB between the "highly directional 4-microphone mode" and the "Omni BTE", which are similar to *beamforming* and *unfiltered* conditions we presented in our work. Nonetheless, there are some differences in the way his experiment was performed. Mens tested the Varibel hearing glasses using phrases as target speech, while we used lists of unrelated words. This is quite an important difference, since we excluded that a better recognition of isolated words could lead to better understanding of contextual informations. It is entirely possible that an improvement in SRT assessed using list of words may correspond to higher SRTs when testing phrases.

There are other differences between our experiment and the one proposed by Mens with Varibel hearing glasses, including the kind of cocktail party noise and the target speech used. Future sessions will attempt to match the testing scenarios.

The illustrated tests have been done in a semi-anechoic environment, but higher reverberation levels seem to decrease speech intelligibility [16] and previous findings indicate that microphone array devices can reduce reverberation effects on speech comprehension [17]. In the future we will consider to repeat the speech comprehension tests exposed in this work in places with higher reverberation which is, by the way, more typical of everyday life contexts.

5. CONCLUSION

In this study we investigated a possible solution to increase speech comprehension capabilities of hearing impaired people in noisy conditions. After a brief description of the context, we analysed the state of the art of hearing devices exploiting beamforming spatial filtering techniques. We also provided some technical details about our device, Glassense, and described the experimental session performed.

In our last work [11] we claimed an improvement of 3.8 dB in SRT mean values between the *beamforming* and *unfiltered* conditions on normal hearing subjects wearing the Glassense with earplugs. The main differences in this experimental session were that most of the subjects had hearing impairment and that we provided the acoustical feedback of the device by hearing aids instead of earplugs. The data analysis showed a result very similar to the previous one: in the *beamforming* condition the SRT mean values improved by 3.3 dB respect to the *unfiltered*. Nonetheless, the most important result showed by the collected data is a difference of 2.4 dB between the *hearing aids* and the *beamforming* condition. This is a real improvement: beamforming spatial segregation performance is higher than the inner-ear hearing aid. It is important to notice that this also means that Glassense is able to improve speech comprehension capabilities of the subjects even without the contribution of the pinna, which is exploited in the *hearing aids* condition, but not in the *beamforming*, since the microphones are out of the ear.

Our results confirm that the proposed solution can represent an advantage for hearing impaired people in noisy conditions. Future works will include studies with an increased number of subjects, comparison with similar devices, like the Varibel *hearing glasses*, and tests in higher reverberation environments.

6. ACKNOWLEDGEMENTS

We would like to thank Federico Traverso and Andrea Trucco from the University of Genoa for the development of the beamforming algorithm and Francesco Diotalevi and the IIT Electronic Design Laboratory for the hardware development of the Glassense.

7. REFERENCES

[1] "World Health Organization — Deafness and hearing loss," http://www.who.int/mediacentre/factsheets/fs300/en/, 2016.

[2] S. Haykin and Z. Chen, "The cocktail party problem," *Neural computation*, 2005.

[3] James F. Feuerstein, "Ear and hearing," 1992.

[4] Simon Doclo, Marc Moonen, Tim Van Den Bogaert, and Jan Wouters, "Reduced-bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 1, pp. 38–51, 2009.

[5] Matthias Froehlich, Katja Freels, and Thomas A. Powers, "Speech Recognition Benefit Obtained from Binaural Beamforming Hearing Aids: Comparison to Omnidirectional and Individuals with Normal Hearing," *AudiologyOnline*, , no. 14338, pp. 1–8, 2015.

[6] R. J. Van Hoesel and G. M. Clark, "Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients," *The Journal of the Acoustical Society of America*, vol. 97, no. 4, pp. 2498–2503, 1995.

[7] B. Widrow, "A microphone array for hearing aids," *IEEE Transactions on Circuits and Systems*, vol. 1, no. 2, pp. 26–32, 2001.

[8] M. M. Boone, "Directivity Measurements on a highly directive hearing aid: the hearing glasses," 2006.

[9] I. L. D. M. Merks, *Binaural application of microphone arrays for improved speech intelligibility in a noisy environment*, Ph.D. thesis, 2000.

[10] Lucas H. M. Mens, "Speech understanding in noise with an eyeglass hearing aid: asymmetric fitting and the head shadow benefit of anterior microphones," *International journal of audiology*, vol. 50, no. 1, pp. 27–33, 2011.

[11] Luca Giuliani, Sara Sansalone, Stefania Repetto, Federico Traverso, and Luca Brayda, "Compensating Cocktail Party Noise with Binaural Spatial Segregation on a Novel Device Targeting Partial Hearing Loss," pp. 384–391. Springer International Publishing, 2016.

[12] Luca Brayda, Federico Traverso, Luca Giuliani, Francesco Diotalevi, Stefania Repetto, Sara Sansalone, Andrea Trucco, and Giulio Sandini, "Spatially selective binaural hearing aids," *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers - UbiComp '15*, pp. 957–962, 2015.

[13] Andrea Trucco, Federico Traverso, and Marco Crocco, "Maximum Constrained Directivity of Oversteered End-Fire Sensor Arrays," *Sensors (Basel, Switzerland)*, vol. 15, no. 6, pp. 13477–502, jan 2015.

[14] *ISO 8253-3:2012, Acoustics – Audiometric test methods – Part 3: Speech audiometry*.

[15] E. Bocca and A. Pellegrini, "Studio statistico sulla composizione fonetica della lingua italiana e sua applicazione pratica all'audiometria con la parola," *Archivio Italiano di Otologia, Rinologia e Laringologia*, vol. 56, no. 5, pp. 116–141, 1950.

[16] J. P. Moncur and D. D. Dirks, "Binaural and monaural speech intelligibility in reverberation.," *Journal of speech and hearing research*, vol. 10, no. 2, pp. 186–95, 1967.

[17] James M. Kates and Mark R. Weiss, "A comparison of hearing-aid array-processing techniques," vol. 99, no. January 1995, pp. 3138–3148, 1996.