19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

HIGH FIDELITY MODELING AND EXPERIMENTAL EVALUATION OF BINAURAL BONE CONDUCTION COMMUNICATION DEVICES

PACS: 43.55.Ka Computer simulation of acoustics, modeling

Walker, B. N.[1]; Stanley, R. M.; Przekwas, A.[2]; Tan, X.G.[2]; Chen, Z.J.[2]; Yang, H.W.[2]; Wilkerson, P.[2]; Harrand, V.[2]; Chancey, C.[3]; Houtsma, A.J.M.[3]

[1] Sonification Lab, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA, 30332 USA. Contact: bruce.walker@psych.gatech.edu
[2] CFD Research Corp., Huntsville, AL, USA; Contact: ajp@cfdrc.com
[3] U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL, USA; Contact: valeta.carol.chancey@us.army.mil

ABSTRACT

Bone conduction (BC) headsets can present audio when ambient sounds must be either blocked by hearing protection, or preserved to maintain situational awareness. Binaural BC is especially desirable to enable simultaneous talkers using spatialized audio. Development of advanced binaural BC devices has been limited by lack of empirical data, inadequate modeling and design tools, and difficult experimental procedures. We discuss our research with binaural BC, and present advanced modeling/simulation tools that use novel coupled Acoustics-Structures Interaction modeling to analyze sound propagation from a BC speaker, through the elastic cranial bone, brain tissue, and cerebral spinal fluid to the cochlea, and further on to the hair cells on the basilar membrane. This high-fidelity anatomy- and biomechanics-based modeling is used to analyze BC device performance at various head contact points, and to obtain parameters for compact (mass-spring-damper) models and “bone related transfer functions”. Those functions are used to predict speech intelligibility, which is then measured using psychophysical studies involving human listeners, and fed back into the models. The resulting listener-validated models can be used to further study BC processes in humans, and to develop more advanced binaural BC devices for use in the most challenging acoustical environments.

INTRODUCTION

Auditory displays are often delivered to a listener through headphones. Headphones allow private presentation of high-fidelity dichotic (stereo) sounds to a listener, without the perception changing as a person moves and turns, all in a portable package. On the other hand, there are problems with headphones. Covering the ears with headphones deteriorates detection and localization of ambient sounds in the environment. These external sounds may be of particular interest in augmented reality and tactical situations that require high situational awareness, as well as for visually impaired users who rely on environmental audio cues as their primary sense of orientation. Furthermore, headphones do not allow auditory display to occur simultaneously with earplugs. Bone-conduction headsets may lead to a useful alternative to headphones, since they avoid covering the ears of the listener. This may facilitate improvement in the detection and localization of environmental sounds, and allow the display of auditory information with hearing protection inserted into the ear canal. Bone-conduction devices also cater to the preferences of users who would rather not have their ears occluded for social, aesthetic or other reasons (e.g., visually impaired users [1].

With binaural bone conduction devices now becoming available, it is important to investigate whether they can serve as headphone replacements in situations where spatialized audio (stereo, 2D, or full 3D) has been beneficial. For example, one of the important uses of virtual 3D audio displays is the ability to increase detectability of signals amidst distracters and noise [2]. There has been considerable development in new bone-conduction hearing apparatus designed for use in auditory displays (see www.oiido.com; www.temcoj.co.jp/english/boneconduction/index.html; www.dowumi.com/eng_index.php). However, despite these developments, there has been relatively little published research into the psychophysical properties of using these devices in binaural applications. Processing sounds for spatial audio displays is a complex process, based on considerable psychophysical research investigating how to manipulate acoustic cues to produce a reliable percept of sounds originating from different locations. These concepts must be better understood before one can implement spatial audio displays with bonephones.

Spatialized audio has been suggested as a way to increase the effectiveness of radio (speech) systems, by separating the speech channels in the audio space [2]. Thus, in addition to, and perhaps in advance of, studying spatial audio with bonephones there is a need to understand how effective bonephones can be at transmitting speech signals. Thus, the present project is concerned with assessing and modeling the psychophysics of bone conduction audio, for ultimate use in spatialized speech applications.

In the past, auditory researchers have often predicted that spatial audio through bone conduction would not be feasible at all, due to insufficient differences in the signal arriving at the two ears [3-5]. This was primarily due to the presumption that there would be too little interaural attenuation to support interaural level differences (ILDs). In addition, personal correspondence and discussions at conferences has made it clear to us that many presume the speed of sound in the skull to be too fast to allow for sufficient interaural time differences (ITDs). ILDs and ITDs are the major cues for spatial audio (at least in terms of azimuth). There has been recent empirical evidence, however, demonstrating sufficient interaural attenuation to support ILDs, and slower-than-expected speed of sound in living tissue, and thus sufficient ITDs to support spatial audio cues [see 6, 7-14]}. In addition, in our own work with binaural bone-conduction headsets we have clearly experienced good stereo separation [11, 15]. Further, other recent work has studied how effectively speech can be understood through bone-conduction headsets [16, 17]. Thus it is evident that dichotic listening is possible, which means that virtual spatialization should also be possible. We are currently involved in more studies in that area.

In addition to behavioral studies, it is useful to build computer models and software-based simulations to help understand the behavioral results and underlying mechanisms, as well as predict future outcomes. Unfortunately, there has not been much research on establishing detailed computer models or simulations that incorporate the air, bone, and fluids involved in transmission of bone conduction signals through the human head and into the cochlea (but see [6] for a model of skull motion). To address this gap, the present project also includes the development of sophisticated computer simulations of the human head and related structures (neck, etc.). We use empirical testing with human listeners to validate and refine the models, which can make for more detailed predictions, which are then tested with human listeners. Ultimately, both the models and the knowledge gained from testing lead to more effective bone conduction transducers, and a thorough understanding of the effects of placing a given transducer at any location on the head. To begin with, the present study is focusing on predicting and then improving intelligibility of speech signals that are presented via bone conduction headsets.

COLLABORATIVE TEAM

Recent collaboration between the Georgia Tech Sonification Lab, USAARL at Fort Rucker, and CFD Research Corporation has began to address issues relevant to the use of bone-conduction headsets for spatialized speech communications delivery. Each of the labs has complimentary strengths that allow for a strong inter-disciplinary team. USAARL has done some of the pioneering work on speech intelligibility through bone-conduction microphones and speakers [16], as well as other uses of bone-conduction sound transmission [18]. Work in the Sonification Lab has demonstrated the feasibility of using binaural bone-conduction headsets in spatial audio displays. This includes evaluations of bonephones for spatial audio tasks [11, 17], assessment of lateralization through bone-conduction headsets [9], and empirically-based bone-to-air shift functions [15]. CFDRC has developed advanced models of waveforms traveling through the head and cochlea.

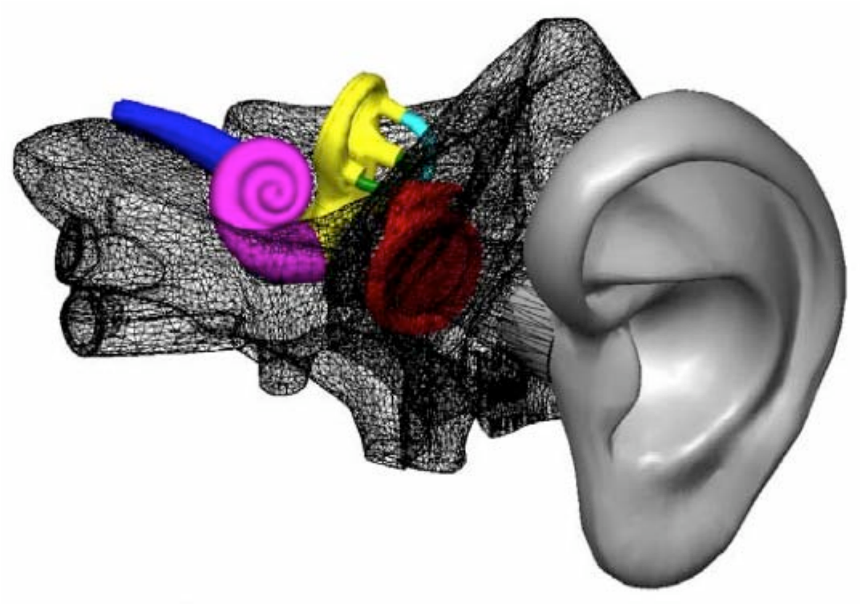

Figure 1. Portion of the finite elements model surrounding the cochlea.

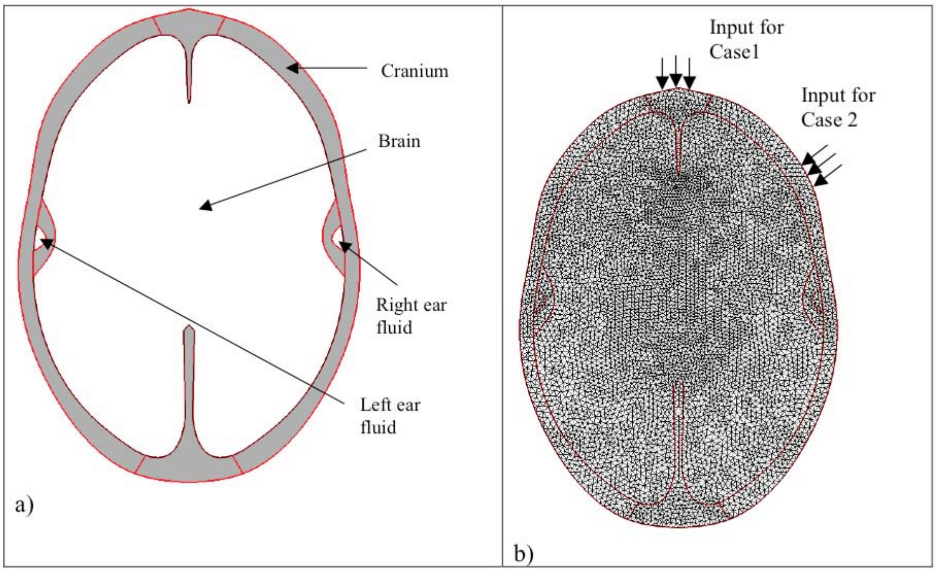

Figure 2. Slice of the fluid portion of the model.

COMPUTER MODELS

Computational Fluid Dynamics Research Corp. (CFDRC) has developed a highly sophisticated computer model of the human head that (1) models the skull through finite element models, (2) models the soft tissue (e.g., brain) and fluids (e.g., cerebrospinal fluid), (3) models the cochlea and basilar membrane, and (4) combines the sub-models for a complete analysis tool. See Figure 1 for an example of the finite elements sub-model of the skull, highlighting the portion surrounding the cochlea, and Figure 2 for a slice through a fluid sub-model. The overall model can be perturbed with a pulse or signal simulating the vibration of a bone conduction transducer, and the resulting vibration at the cochlea or on the basilar membrane can be determined. See Figure 3 for an example of a pulse applied to the forehead.

There is a certain transit time from the point of vibration to the cochlea, which depends on the actual structures through which the wave travels. Further, there can be spectral changes, as certain frequency components can be attenuated or even enhanced (e.g., there are resonant

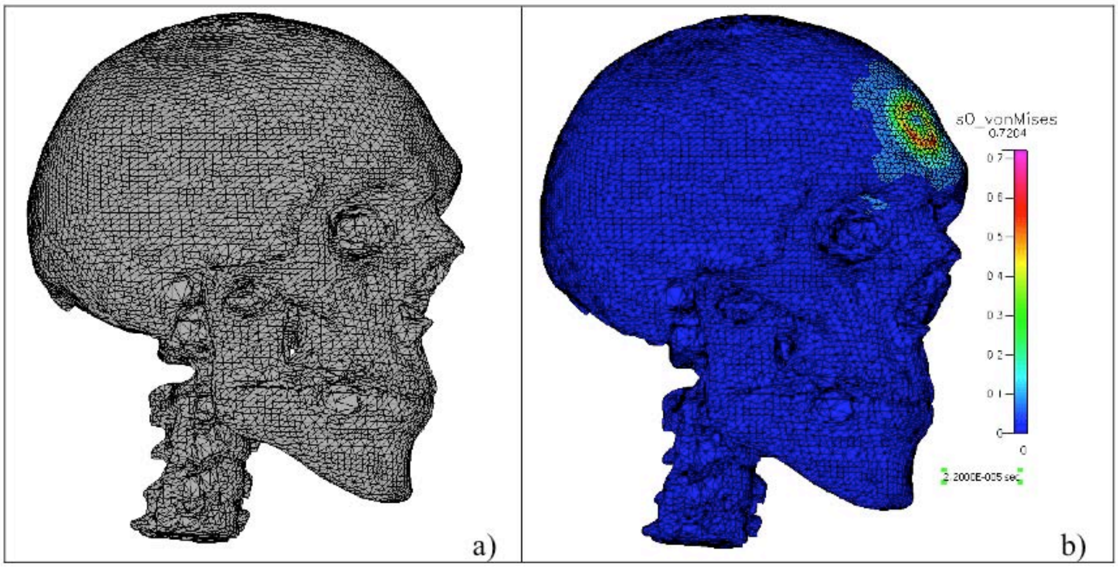

Figure 3. (a) Finite elements mesh for entire skull. (b) Perturbation of model by Gaussian pulse applied to forehead.

frequencies for the head). The overall effect on a speech sound can affect intelligibility. For example, if only the high frequencies are attenuated, the intelligibility of fricatives will be reduced, compared to vowels.

Examining the overall filtering effects of the head from the various transducer locations allows us to develop a “bone related transfer function,” or BRTF. The BRTF can help us to predict effects on overall detection and also speech intelligibility. These predictions, and in turn the computer model itself, are then validated using empirical listening studies. If human listeners do not experience changes in detection and intelligibility in ways that are consistent with the model’s predictions, then the model is refined and improved.

BEHAVIORAL VALIDATION STUDIES

At Georgia Tech we are currently conducting validation studies, as described above. BRTFs are determined by perturbing the computer model at various skull locations, and with various signals (pure tones, noise, speech signals). The speech intelligibility of the resulting stimuli is then evaluated. Speech intelligibility is measured empirically through human listeners with the Diagnostic Rhyme Test (DRT). The DRT is a standardized speech intelligibility test that contrasts initial consonants of words to reveal how well a system has preserved voicing features [19]. In addition to testing with the filters applied to air-conducted stimuli, speech intelligibility can also be tested with unfiltered stimuli delivered through bone conduction transducers placed at the same locations simulated by the filters. Figure 4 shows standard bone conduction transducers applied to three different locations on the head (mastoid, condyle, and vertex). By comparing speech intelligibility between simulations delivered through air conduction and actual signals delivered through bone-conduction transducers, these results will demonstrate the degree to which the models are accurate in the prediction of speech intelligibility. This data will also provide some insight into optimal location of a transducer for speech communication. In addition, the empirical psychophysical performance data will be used to adjust the models to improve their fidelity.

With a validated model, more complex predictions can be made. Specifically, with a validated model of waves traveling through the skull we can begin to consider binaural aspects of bone-conducted stimuli. How the phase, amplitude, and frequency of waveforms are altered as they travel through the head will allow an understanding of the abilities and limitations of binaural bone-conduction, and together with previous monaural models of speech intelligibility, will allow predictions about spatializing communication signals to improve detectability. It may also describe some adjustment filters that can be applied to bone-conducted stimuli to improve the resulting perception of spatialized sounds. Future work will also consider the effect of other

Figure 4. Headband bracket for securing the RadioEar B-71 BC transducer (shown), for use in mastoid (left image), condyle (center), and vertex (right) locations.

variables on speech intelligibility, such as simultaneous headgear with bone-conduction transducers, the effect of hearing protection on the waveforms, and transducer design.

All of this will then be taken to the binaural case, where multiple stimulation sites can be used in the model, to predict the best ways to maximize both speech comprehension and general spatialized audio, using binaural bonephones.

CONCLUSIONS

There is a clear need for headphone alternatives, and modern binaural bone conduction devices may be acceptable, or even preferred in many cases. The project discussed here involves the development of highly sophisticated computer models of the head, to predict changes to a vibration/sound signal as it passes from a spot on the surface of the head to the cochlea. This, in turn, is used to predict speech intelligibility, among other things. The utility of the model is carefully validated using behavioral research, in which listeners complete the Diagnostic Rhyme Test (DRT, and the results are used to evaluate the effectiveness of the model, and subsequently to refine it. This research is being extended to binaural conditions, which then allow for a complete model of spatialized audio presented through bone conduction transducers, rather than the typical headphone presentation. Ultimately the generation of bone related transfer functions should allow us to tune general audio spatialization systems for bone conduction, by adjusting or replacing the HRTFs with BRTFs.

References:

[1] B. N. Walker, R. M. Stanley, and J. Lindsay, "Task, user characteristics, and environment interact to affect mobile audio design," in *PERVASIVE*, Munich, Germany, 2005.

[2] D. S. Brungart and B. D. Simpson, "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal," *Journal of the Acoustical Society of America,* vol. 112, pp. 664-676, 2002.

[3] B. A. Goldstein and C. W. Newman, "Clinical masking: A decision-making process.," in *Handbook of Clinical Audiology*, 4 ed, J. Katz, Ed. Baltimore: Williams & Wilkins, 1994, pp. 109-131.

[4] J. Katz and J. Lezynski, "Clinical Masking," in *Handbook of Clinical Audiology (5th ed.)*, 5 ed, J. Katz, Ed. Philadelphia: Williams & Wilkins, 2002, pp. 124-141.

[5] J. Blauert, *Spatial hearing: The psychophysics of human sound localization*. Cambridge, MA: The MIT Press, 1983.

[6] P. Y. S. Stenfelt and R. L. Goode, "Transmission properties of bone conducted sound: Measurements in cadaver heads," *Journal of the Acoustical Society of America,* vol. 118, p. 2373, 2005.

[7] K. Kaga, M. Setou, and M. Nakmura, "Bone-conducted sound lateralization of interaural time difference and interaural intensity difference in children and a young adult with bilateral microtia and atresia of the ears.," *Acta Oto-Laryngologica,* vol. 121, pp. 274-277, January 2001.

[8] A. F. M. Snik, A. J. Beyon, C. T. M. van der Pouw, E. A. M. Mylanus, and C. W. R. J. Cremers, "Binaural application of the bone-anchored hearing aid," *Annuals of otology, rhinology, and laryngology,* vol. 107, pp. 187-193, 1998.

[9] R. M. Stanley and B. N. Walker, "Lateralization of sounds using bone-conduction headsets," in *Human Factors and Ergonomics Socieety 50th Annual Meeting*, San Francisco, CA, 2006, pp. 1571-1575.

[10] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," *Human Factors,* vol. 48, pp. 265-278, 2006.

[11] B. N. Walker and J. Lindsay, "Navigation performance in a virtual environment with bonephones," in *International Conference on Auditory Display*, Limerick, Ireland, 2005, pp. 260-263.

[12] E. Franke, "Response of the human skull to mechanical vibrations," *The Journal of the Acoustical Society of America,* vol. 28, pp. 1277-1284, 1956.

[13] J. Tonndorf and A. F. Jahn, "Velocity of propagation of bone-conducted sound in a human head," *Journal of the Acoustical Society of America,* vol. 70, pp. 1294-1297, November 1981.

[14] J. Zwislocki, "Acoustic attenuation between the ears," *Journal of the Acoustical Society of America,* vol. 25, pp. 752-759, July 1953.

[15] R. M. Stanley, "Toward adapting spatial audio displays for use with bone conduction: The cancellation of bone-conducted and air-conducted sound waves.." vol. Masters in Science Atlanta: Georgia Institute of Technology, 2006, p. 97.

[16] B. Acker-Mills, A. J. M. Houtsma, and W. Ahroon, "Speech inteligibility with acoustic and contact microphones," in *New Directions for Improving Audio Effectiveness*, Amersfoort, The Netherlands: NATO RTO-MP-HFM-123/RSY, Paper 07, 2005, pp. 7-1 – 7-14.

[17] B. N. Walker, R. M. Stanley, N. Iyer, B. D. Simpson, and D. S. Brungart, "Evaluation of bone-conduction headsets for use in multitalker communication environments," in *Annual Meeting of the Human Factors and Ergonomics Society*, 2005.

[18] A. J. M. Houtsma, I. P. Curry, J. M. Sewell, and W. N. Bernhard, "Auscultation in high-noise environments using hybrid elctromechanical and ultrasound-Doppler techniques," *The Journal of the Acoustical Society of America,* vol. 120, p. 3361, 2006.

[19] W. D. Voiers, "Evaluating processed speech using the diagnoistic rhyme test," *Speech Technology,* pp. 30-39, Jan/Feb 1983.