BeBeC-2018-D19

SIMULATION, VISUALIZATION AND LOCALIZATION OF SOUND IN A REAL AND A VIRTUAL ROOM ACOUSTIC ENVIRONMENT USING BEAMFORMING

M. NAVVAB [1], G. HEILMANN [2], A. MEYER [2], S. BARRÉ[2], M. BÖCK[2,]

[1]University of Michigan, Taubman College of Architecture and Urban Planning,
Arch Bldg., 2000 Bonisteel Blvd., Ann Arbor MI 48109, E-Mail: moji@umich.edu

[2] GFai Tech GmbH, Volmerstrasse 3, 12489 Berlin-Adlershof, Germany

ABSTRACT

This paper describes a method to simulate, visualize and localize the sound in both real and virtual interior room acoustic environments to test human sound perception. Real world measured results using an acoustic camera along with Noise Image software combined with beamforming techniques are applied to examine the real as well as simulated room acoustic conditions. Furthermore, the transition from the three-dimensional sound recording inside a real architectural space to the three-dimensional virtual acoustic mapping is described. The measured results using sound mapping, binaural dummy head and application of the head related transfer function are used to present the findings that demonstrate and contribute toward sound source localization within real and simulated environments. The results show that the users are able to navigate and locate a real and virtual sound source in a dynamic virtual acoustic environment. The findings from these simulations, auditory navigation experiments via noise mapping, demonstrate that the beamforming method could be used to calibrate the room acoustic characteristic and examine sound perception in a virtual environment, and provide opportunities to study sound localization and fine tune the current HRTF for various interior architectural acoustic designs.

Key words: Interior Sound Mapping, Visualization, Sound Localization, Auditory Navigation, Virtual Acoustics, Spatial Hearing, Perception of Sound, HRTF.

1 INTRODUCTION

Human ability to hear is based on sensing by the ear and the neural system to process the signal in this case sound pressure. This activity within the brain has conscious as well as subconscious effects. Given the living and working environments; we are exposed to a variety of familiar sound sources, and due to prolong exposure to them there are obvious subconscious effects such as ear damage that includes loss of spectral sensitivity.

The cognitive part of the brain evaluates the known or unknown sounds such as alarms, thunderstorms, music or speech, and provides a conscious response such as immediate attention, pain or pleasure. With todayֹs growth in audio technologies and use of virtual environment, humans are exposed to many real or virtual visual and audio cues; as such, the utilization of conscious response to comprehend and react and perceive this large set of information is a very complicated process.

Psychoacoustics field of study provides a path to understand the usersô perception of sound by examining its conscious effects. Physical characteristics of sound that produce hearing sensations in the listener and their corresponding hearing sensations as relate to their perception are Physical Characteristic (e.g.; sound pressure level, frequency, duration) and Hearing Sensation (loudness, pitch, subjective duration).

These sensations and physical characteristic do correlate directly, but not to one physical characteristic. It is known that loudness, frequency and sound pressure level do impact each other physical characteristic. Human ear size, age and genders do contribute to this complex system toward sound quality evaluations that are usually specific to room acoustic environment.

One of the objectives is to conduct an experiment in such a way that the perceived sound changes according to the movements of the subject (e.g. distance & orientation) in a given simulated room for its acoustic characteristics to be evaluated. The results could be used as part of an auditory navigation experiment. Many auditory navigation tests have been conducted in the past, and an overview of these different techniques needed in auditory navigation shown in research work by **[1-3]** are directly related examples of such experimental work.

In our experiment we applied the capability of Noise Image simulation software, binaural recording and FMOD auralization systems within a VR Laboratory to obtain close approximation of the simulated room architectural characteristics of an actual 3D-virtual reality laboratory space called the "CAVE". The locations of the loud speakers and their directionality were modeled based on the actual measurements using the acoustic camera recording system. **[4-7,19].** These sound localizations are achieved by the application of the Head Related Transfer Function (HRTF) through the inclusion of the cues exclusive of differences between the signals reaching each ear as Inter-aural Time Delay (ITD) and Inter-Aural Level Differences (ILD). "Relying on a variety of cues, including intensity, timing, and spectrum, our brains recreate a three-dimensional image of the acoustic landscape from the sounds we hear" **[8, 9].**

2 METHODOLOGY

We have used simulation techniques under real and virtual settings to create a room acoustic environment that provides the opportunity to examine the human sound perception. Physical characteristics of sound sources are measured, mapped and subjectively evaluated. Head Related Transfer Function (HRTF) is used as a sound quality metric to correlate better to human perception. An acoustic camera sphere array with 120 Mics, utilizing Beamforming techniques within Noise Image software and a Sennheiser MK2002 binaural microphone placed at the entrances to the middle ear of the dummy head are used to measure sounds and recorded data used for subjective evaluation. The sound quality matrices offered by **ISO 3382 standards** along with sound loudness and power are calculated based on the measured data at selected points and reported.

2.1 Binaural Audio Measurements

Binaural audio recording is a method of capturing audio by creating a 3D-audio experience. This concept includes the acoustic shadow created by the subject head, while incorporating audio directionality in the scene along all-three axis, X, Y, ***and*** Z, using two microphones. Each microphone is positioned into the ear canal of a dummy head. We used a Sennheiser MK2002, within the outer ear of the dummy head and human subjectsô head.

An advantage of the MK2002 system over other dummy head systems is that the MK2002 can be used with either the provided dummy head or a real head. This allows users to wear the microphone system connected to a portable audio recorder to gather realistic, 3D sound within real or virtual reality audio room acoustic environment. Simulated sound using FMOD system within VR (see section 2.4) were recorded and later on played back for the subjects and the data from different observations were used for comparison of the real time localization in the VR space and through listening to audio recordings with headphones.

2.2 Head Related Transfer Function (HRTF)

The HRTF for distance sources is a complicated function of azimuth, the elevation as presented in reference by Algazi and Duda **[20]**. The HRTF database from CIPIC is a set of anthropometric measurements that can be used for scaling studies. The CIPIC HRTF Database is a public-domain database of high-spatial-resolution HRTF measurements for 45 different subjects, including the KEMAR mannequin with both small and large pinna. The database includes 2,500 measurements of head-related impulse responses for each subject. These ñstandardò measurements were recorded at 25 different interaural-polar azimuths and 50 different interaural-polar elevations.

2.3 Implementation of the HRTF in Noise Image Software

The relative sensitivity of the ear in decibels was fit as a function of frequency and angular location with a mix of polynomial trigonometric functions. The data is not smooth and the fit needed 214 independent parameters to obtain a 94% correlation. The fit range covers frequencies from 500 to 16,000 hertz, polar angles from 0 to 170 degrees, and all azimuthal angles.

Ear sensitivity for four subjects from the CIPIC HRTF database was calculated at 71 frequencies over the range from 500 to 16,000 hertz for each of two sets of angles: 25 polar angles with the zenith oriented along the line connecting the left and right ears, and 50 azimuthal angles with 0° in front of the face. The average data set thus consists of 88,750 data points over angles and frequencies. The input sound level was arbitrary, so the data was normalized to have an average value of zero. The functional form used for the frequency variable are not bounded, so the fit should not be used for frequencies significantly outside the measured data range of 500 to 16,000 hertz. Similarly, the data for polar angles covers the range from 10 to 170 degrees from zenith, so the fit may be inaccurate at the limiting values of 0 and 180 degrees. The fit using trigonometric functions for the azimuthal angles, and therefore should be accurate over the entire 0 to 360 degree range. This function is implemented in Noise Image software and provides users the ability to view the sound localization with and without the use of HRTF at different viewing direction or elevation.

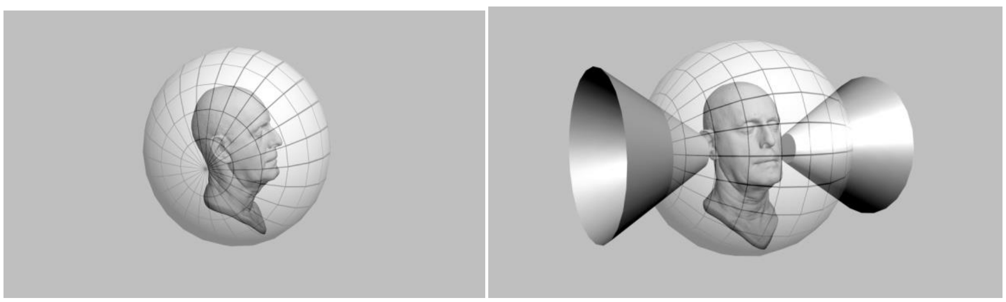

Figure 1: Azimuth and elevation grids used for fitted HRTF with respect the head position

2.4 Audio and Visual Cueing for Spatial Perception

Application of immersive Virtual Environment (VE) technology for sound perception is achieved through auditory stimuli based on the results of simulated, auralized, and reproduced sounds within computer-simulated spaces of the existing conditions. This immersion capability allows stimulation of all human sensory subsystems in a natural way within this immersive environment. The subject uses special viewing glass and headphones (for best realistic sensation utilizing Head Related Transfer Function (HRTF) with a head-tracking device to listen to the auditory event of a simulated space in real-time while viewing the interior space. In our experiment we applied the capability of simulation software, and FMOD auralization systems within a VR Laboratory. This FMOD sound system allows playing a very close approximation of the room acoustic characteristics within 3D-virtual reality laboratory space settings. The locations of the loud speakers and their directionality were modeled based on the actual measurements using the acoustic camera recording system **[11-12 and 15-16].** FMOD is a programming library and toolkit for the creation and playback of interactive audio. It supports all leading operating systems and game platforms. FMOD is a proprietary audio library made by Firelight Technologies that plays music files of diverse formats on many different operating system platforms, used in games and software applications to provide audio functionality within the 3D virtual Laboratory. The FMOD sound system has an advanced plug-in architecture, that can be used to extend the support of audio formats or to develop new output types, e.g. for streaming. FMOD sound system now contains three main parts: 1) FMOD Ex, the low-level sound engine, 2) FMOD Event System, more abstract, higher level application layer to simplify playback of content created with FMOD Designer, 3) FMOD Designer, the sound designer tool used for authoring complex sound events and music for playback **[4, 11, and 19]. Figure 2** shows subjectœs positions while experiencing the sound and real views of the virtual laboratory.

3 HUMAN SUBJECT EXPERIMENTAL SET UP

The subjects were free to move and turn in a virtual space to analyze the effect of various noise factors (e.g., computers, projectors and HVAC system within the space) which impact the sound stimulus and hence their directional cues. This takes place within a simulated acoustics environment while the recorded sound of a Mozart string quartet piece is played through 4 real speakers. The FMOD system provided or simulated the additional 4 virtual speakers between the real speakers. The time spent from starting to end position, and trajectory of the subjectœs motion, were recorded digitally using a digital camera equipped with an equal distance projection fisheye lens and a video recorder. Specific observations were made on each subjectœs ability to locate or localize the real or virtual source by looking and aiming the laser pointer closest to each speakerœs location. In the experiment, instructions were given aurally. In the beginning of the experiment there were intentionally no rehearsal tests to help the experiment so that the subjects would not notice or recognize the locations of the real loudspeakers within the virtual simulated environment. A complete test set included 8 speakers (4 real and 4 virtual) and each source played an identical WAVE file within the virtual space. These 8 positions for all speakers are shown in **Figure 3** as virtual speakers (VS1, VS2, VS3, and VS4) and the real loudspeakers as (RS1, RS2, RS3, RS4) in the real space.

Moving in the virtual space was controlled with the use of Jugular software and a motion tracking system. The subject was able to move forward, backward, and to turn left and right at normal human speed. When subject assumed that he/she has found the sound source, he/she indicated that by stopping at the final location. This experiment was done in the horizontal and vertical plane given the 45 degree downward tilt of angle by the speakers aimed toward the center of the space at 5 feet height. The sound source was a point source. The target area was a one meter (~3.3 feet) sphere around the source. Starting positions were in order from speaker 1 to 8 and the approximate or exact locations were identified by all subjects. The entire experiment did not exceed 5 minutes per subject. WAVE files were displayed or simulated within the virtual environment using the FMOD sound system **[10]**.

Figure 2: Views within VR Lab showing subjectœs positions while experiencing the sound

3.1 The UM3D Lab

The UM 3D virtual reality laboratory includes an immersive virtual-reality like environment, measuring 10 ft (3.048 m) in width, depth, and height. It runs on a cluster of six workstations, with one control computer, one motion-tracking computer, and four rendering computers. The software is an ongoing in-house development, named Jugular that integrates several open-source, proprietary, and custom-developed subsystems for graphics, sound, animation, physics, motion-tracking, data management, and networking. [9, 10].

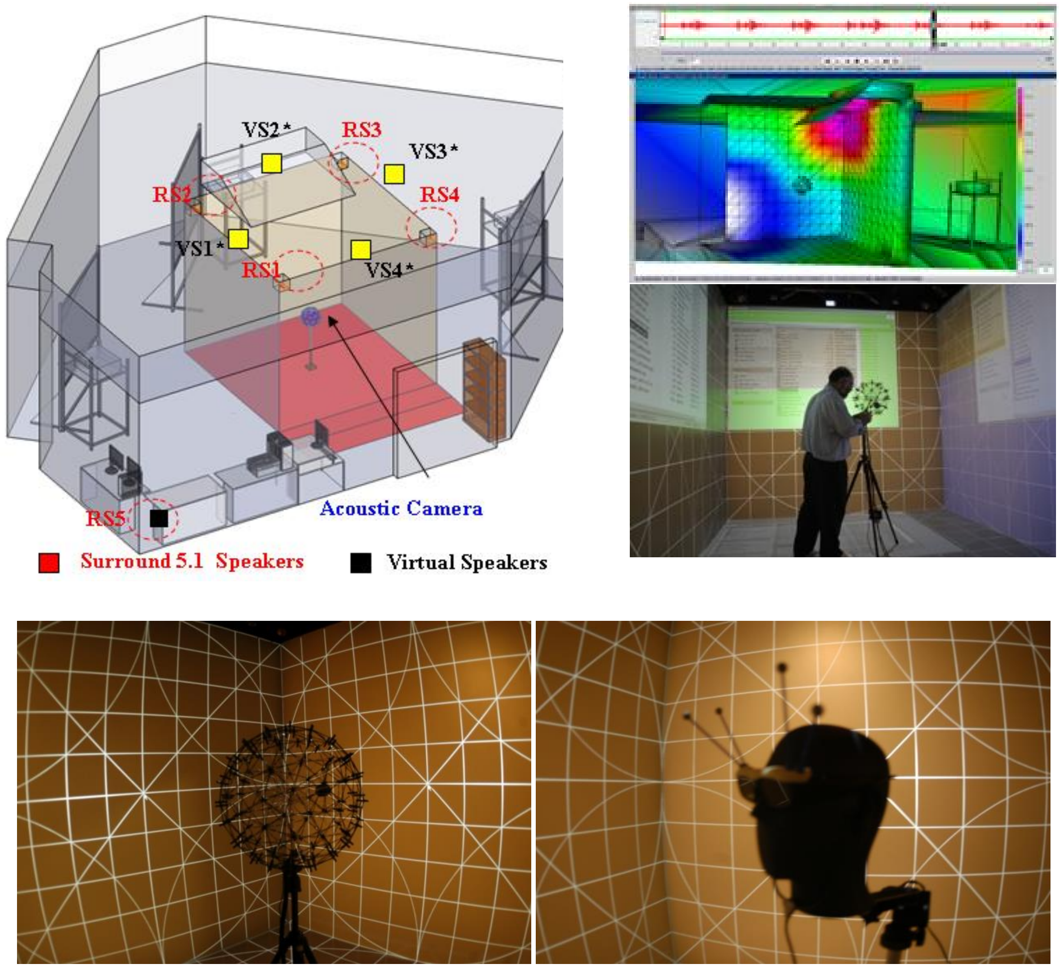

Figure 3: Schematic computer model and real views of the VR speakers (real/Virtual) and their locations and Acoustic Camera sphere array with 120 mics and dummy head.

3.2 The Virtual Reality Modeling Language (VRML) & Interactive Audio

The VRML is representative of simulation capabilities in the Virtual Environment. VRML version 2, also known as VRML97, was adopted as an International Standard ISO/IEC 14772 in 1997. The standards specify a file format, a content model, and algorithms for its interpretation. The model is a directed acyclic graph that includes nodes for geometry, color, texture, and light, as well as sound. However, it provides for only one texture per shape and one pair of texture coordinates per vertex. VRML version 2 was amended in 2002 to add geospatial and NURBS support, but the shape, material, lighting, and sound specifications remain unchanged [11].

3.3 Principle of the Delay-and-Sum Beamformer

The time domain calculation of a delay-and-sum beamformer within the Noise Image software is well documented and validated [7]. Despite its extreme simplicity, the delay-and-sum method in the time domain is quite robust and powerful and has shown its practical usability in an extraordinary wide range of acoustic localization and troubleshooting applications for several years [15, 16]. The output devices such as speakers at a distance from the listener, or headphones impact the user experience with "3D Sound". Passive "3D Sound" where the results always sound the same, the sound specialization information can be used within a sound playback file, and heard through a stereo amplifier with speakers or headphones. Interactive "3D Sound" situations where the listener's apparent position within the audio scene is required, the "3D Sound" rendering algorithms should be done on listener equipment in real time. The CAVE's sound system utilizes the FMOD and it is known that the system in its speaker mode will never be as good as the headphone mode for "3D sound" rendering. Since the use of HRTF rendering algorithm must interpolate for locations in between the given HRTF locations, it requires high real time computing [4, 10]. **Figure 4** shows sound mapping representation for selected speakers and application of HRTF as a filter.

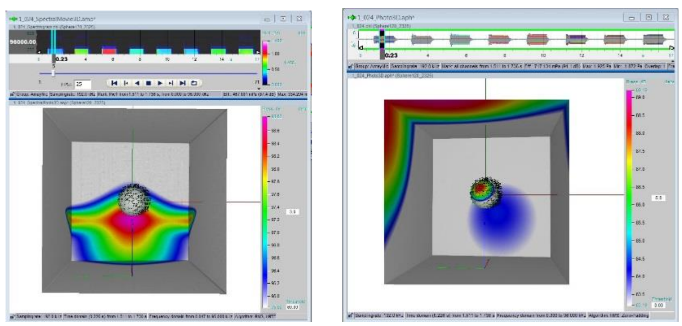

Figure 4: Sound mapping representation for 8 speakers with and without HRTF filter at Up and down position in elevation viewing angles.

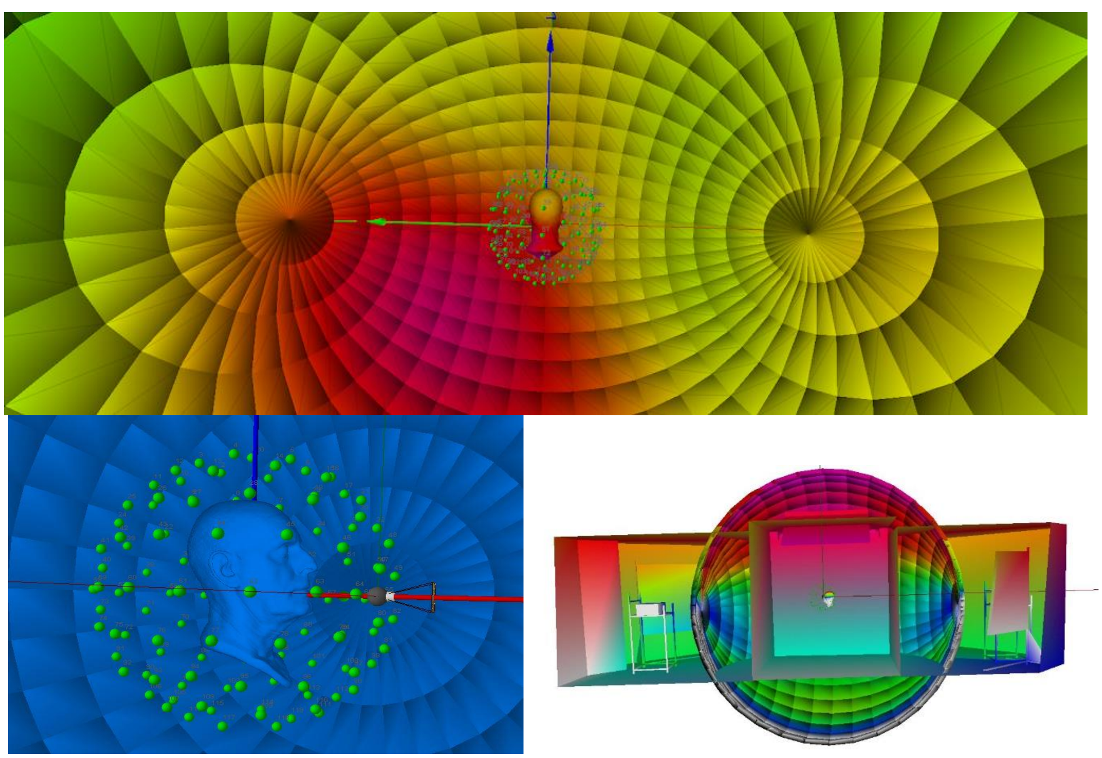

Figure 5: Sound pressure level readings taken inside VR lab and projected on a scene with HRTF gridded surface. The circle on each sides of the figure head shows the field of view for Left and Right ears.

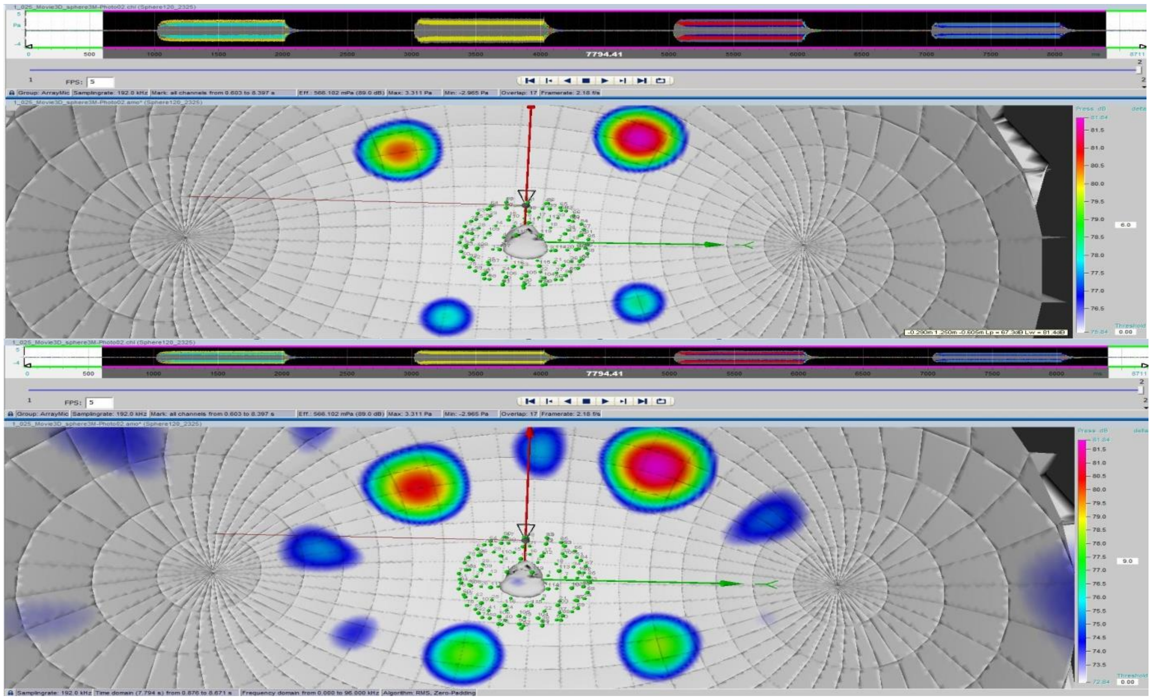

Figure 6: Sound output from real speaker 1-4 Source 1.5KHZ pure tone, 8 Sec.6dB, 9d Delta.

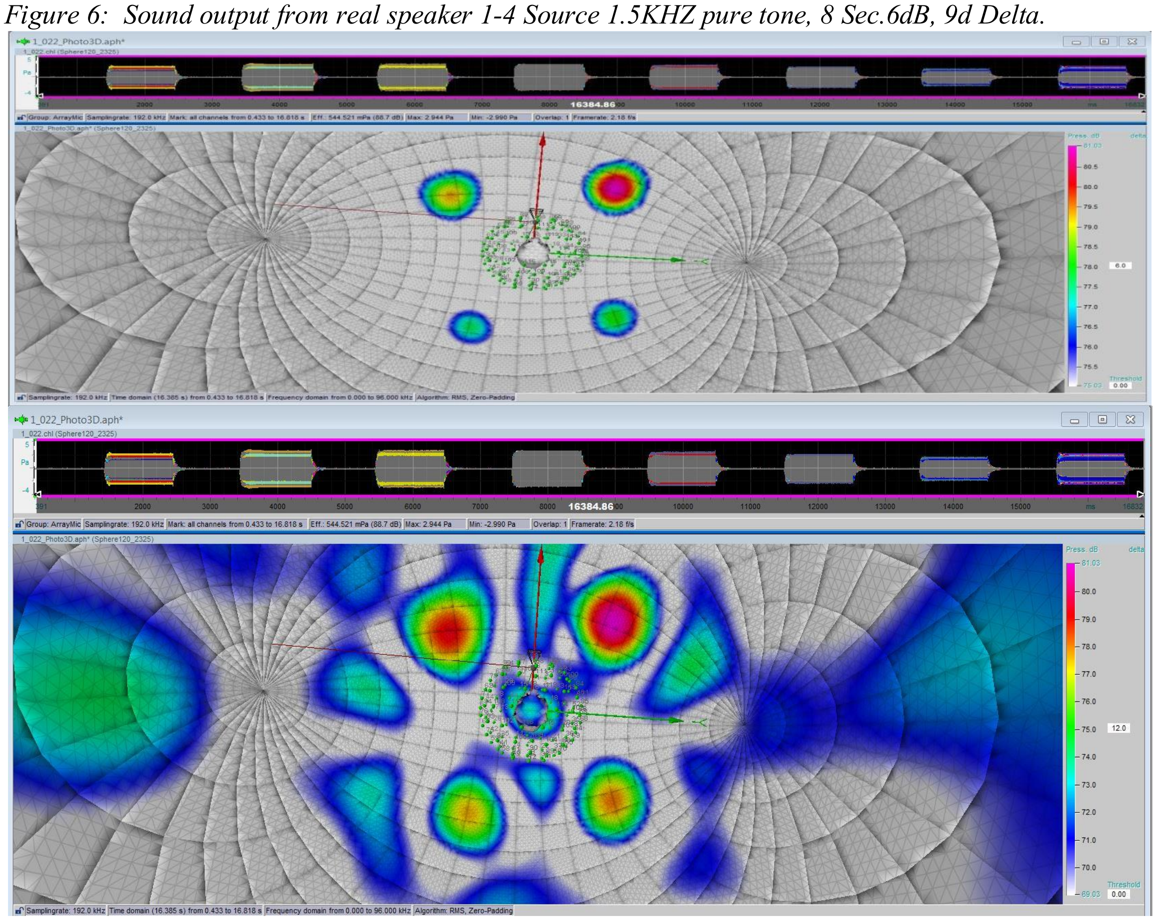

Figure 7: Sound output from real+virtual speaker Source 1.5KHZ pure tone,16 Sec.6dB,12dB Delta.

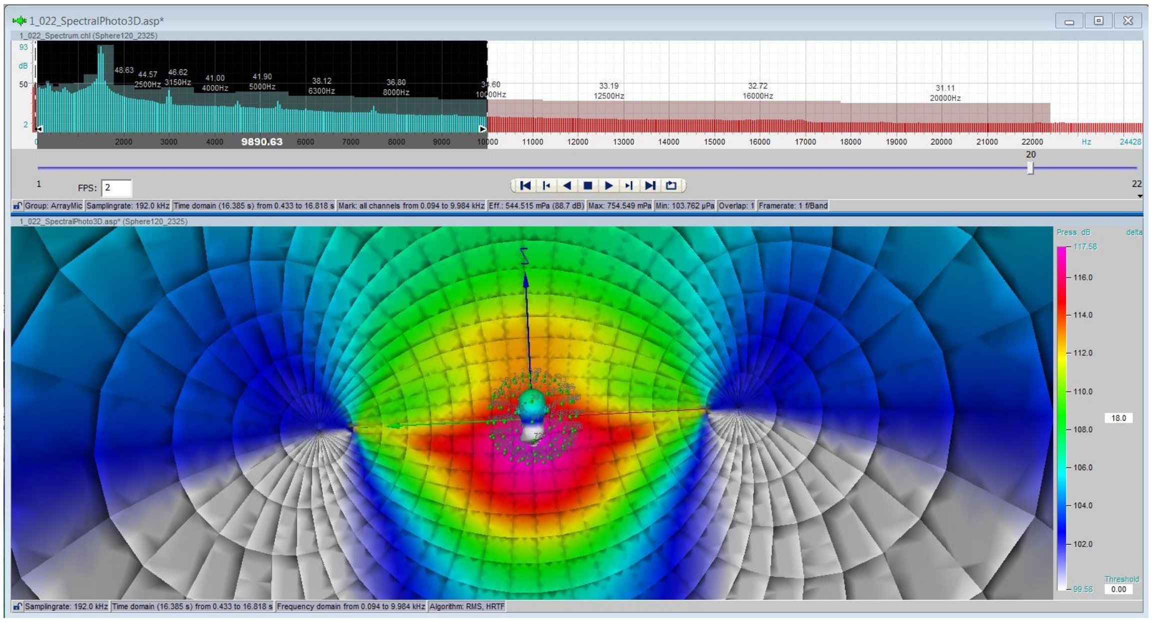

Figure 8: Sound output from 8 speakers, Filtered with HRTF at 0 Elevation and +10dB sensitivity

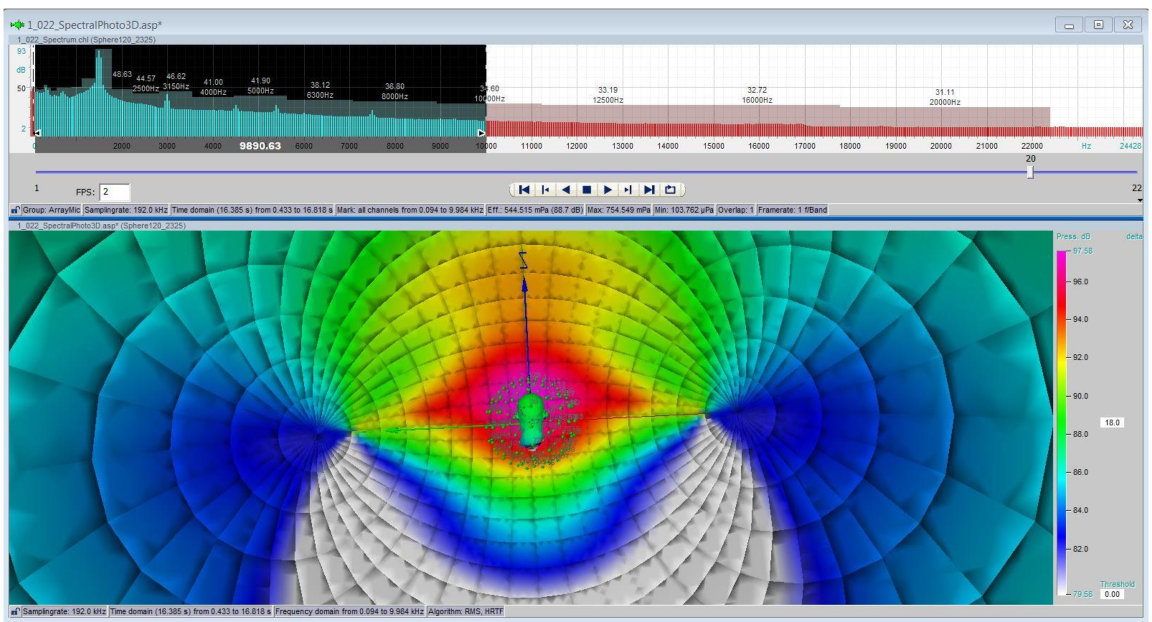

Figure 9: Sound output from 8 speakers, Filtered with HRTF at 60 Elevation and 0dB sensitivity

As it was stated earlier, the sound quality matrices offered by ISO 3382 standards are implemented within Noise image software and below are listed all acoustic measures recommended by ISO standard 3382; Noise Image software calculates the most important ones: Reverberation Time, Strength Measure, Center Time, Echo Criterion, Definition Measure C50 for Speech, Speech Transmission Index, Articulation Loss, Clarity Measure C80 for Music, Lateral Efficiency. The real time recording, monitoring and simulation procedures used can be applied to evaluate real or virtual room acoustic while designing for acoustic comfort or pleasure. This immersion capability allows stimulation of all human sensory subsystems in a natural way within this virtual environment. Sample output by Noise Image using ISO indexes allows evaluating design alternatives shown in **Table-1**. *Sample output by Noise Image using selected (ISO/DIS 3382) indexes is shown in* **Table-2.**

Table -1: Parameters for description of the acoustical quality in rooms (ISO/DIS 3382)

T30/s = Reverberation time, computed from -5 to -35 dB of the decay curve
EDT/s = Early Decay time, computed from 0 to -10 dB of the decay curve
D/% = Definition, Early (0-50ms) to total energy ratio
Ts/ms = Center Time, time of 1.0 moment of the energy impulse response
G/dB = Sound level related to omnidirectional free filed radiation over 10m distance
LF/% = Early lateral (5-80ms) energy ratio, $\cos^2$ (lateral angle)
LFC/% = Early lateral (5-80ms) energy ratio, cos (lateral angle)

Table-2: Sample output by Noise Image using selected (ISO/DIS 3382) indexes.

	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz	4000 Hz
Early Decay Time	0.72	0.31	0.44	0.4	0.28	0.27	0.33	0.35	0.27	0.24	0.22	0.21	0.24
T20	0.68	0.65	0.48	0.47	0.41	0.33	0.36	0.34	0.33	0.33	0.31	0.33	0.4
Definition	81.48	90.04	86.76	88.69	92.64	93.86	91.42	83.36	92.36	92.58	94.96	94.19	93.35
C80	8.97	12.09	10.41	11.32	14.98	17.48	16.41	13.66	17.06	16.67	16.93	16.32	16.75
C50	6.44	9.56	8.16	8.94	11	11.85	10.28	7	10.83	10.96	12.75	12.1	11.47
C7	3.15	-9.92	-7.8	-2.71	-0.62	1.81	-4.09	-20.09	-11.86	-9.51	-8.4	-8.5	-6.78
Center Time	23.63	25.82	30.82	23.13	16.51	12.75	21.52	27.7	19.35	22.84	20.04	20.21	19.85

4.0 RESULTS

Measured data within VR ï CAVE lab space provided the data base for these parametric audio perception studies. The goal is to test and seek results in terms of observations or obtain clear evidence of the inter-aural level and time difference (ITD, LTD). Through use of AC and NI, it is possible to calculate the difference in loudness, frequency and sound distribution. The difference between the two ears as simulated or measured within the real or simulated environment are the key index. For example when the loudness of a sound is blocked due to the acoustic shadowed area on the side of the head, with a large difference in ITD, we are able to locate the source. Human perception is also confused when the ITD or ILD cues are too small to be differentiated. The direction of the source in three planes, in x, y, z, path differences in polar coordinates contributes to the human perception/sound localization. ITD provides the cue for path-length difference due to azimuthal variation or direction and elevation (Figure 1) and ILD along with the associated frequencies cues contribute to the sense of distance. The results from subjective studies demonstrate the ability or lack of subject ability to localize the sound under constant azimuthal variation.

Stimulus, panning method for localizing the sound within the acoustic environment was used to conduct this experiment. Each WAVE file had equal loudness. Each speaker cycle of time play was about 8 seconds long (4 speakers) including one second dead band as played in a sequential loop for all 8 speaker positions it was 16 seconds long. The sound source had an Omni-directional pattern. The synthesized or aurulized sections of the sound were produced by a physical-based model **[17, 18].** Panning or seeking for the sound source methods had no limitation on subject movements. The interaural time difference (ITD), was included as an auditory cue to all test conditions. The ITD was calculated from a spherical head model within EASE program **[19]** and implemented with a short delay line. When the subject asked to locate the sound, they started at the center of the space facing the center wall with full degree of freedom to move within the virtual space.

Once the sound (e.g.; music, clapper, 256HZ, 1024HZ and 1500 HZ pure tone) started to play, subjects were asked to seek and locate the sound and its direction, Selection of the sources are based on the past studies by other researchers. **[21-23]** Subject movement and recognition of the source location were recorded using a camera and a video recorder. The results and sound distribution within the space were also measured using an acoustic camera in absence of the subject. The summery of the findings using VR space and headphone are shown in **Table 3.**

Table -3: Subjective evaluation results using VR space and headphone.

Subject Performance in VR Lab	1SPK-RS1	2SPK-VS1	3SPK-RS2	4SPK-VS2	5SPK-RS3	6SPK-VS3	7SPK-RS4	8SPK-VS4
# of Subjects from 32 (total) located the Spk	22	28	29	28	29	25	22	20
% Number of subjects located SPK	0.69	**0.88**	0.91	**0.88**	0.91	**0.78**	0.69	0.63
Subject Performance using Headphone	**1SPK-RS1**	**2SPK-VS1**	**3SPK-RS2**	**4SPK-VS2**	**5SPK-RS3**	**6SPK-VS3**	**7SPK-RS4**	**8SPK-VS4**
# of Subjects from 32 (total) located the Spk	20	28	23	28	23	28	20	28
% Number of subjects located SPK	0.63	**0.88**	0.72	**0.88**	0.72	**0.88**	0.63	**0.88**
Speakers Positions								
Speaker within two wall	NO	YES	YES	YES	YES	YES	NO	NO
SPK Field of projection (Horizontal/Vertical)	90/90	90/90	90/90	90/90	90/90	90/90	90/90	90/90
Aiming direction downward center	45/45	45/45	45/45	45/45	45/45	45/45	45/45	45/45

The best results in virtual sound source localization are achieved if and only if the binaural difference-based cues, and the monaural and binaural cues that arise from the scattering process of the Head Related Transfer Function (HRTF) for the particular individual are included within the virtual scene. If the primary output device is a pair of headphones or speakers, the "3D Sound" rendering process will result in a stereo signal, e.g. one signal for the "Left ear" and one signal for the "Right ear", but the signal components tend to cancel cross talk effects between Left and Right channels. The effectiveness to which this can be done varies both with distance from the listener, as well as distance (or angle) between the speakers. The effect is best with speakers with the angle range of 27-45 degrees between the listener and speaker on either side of the listener when facing forward. **[8]**. Measured results for different viewing conditions of the surfaces within the virtual laboratory provide some clue to the incoming sound and its directions. **Figure 6** shows the sequence of 4 real speaker as source with 1.5KHZ pure tone for 8 Sec. at 6dB and 9dB delta. **Figure 7** shows sound output from 8 speaker as source with 1.5KHZ pure tone for 16 Sec. at 6dB and 12dB delta in sound pressure levels. The sequence of the speakers and the sound intensity distribution in terms of calculated Delta dB within each viewing scene for each real and virtual speaker provide an insight into the sound projected from each speaker. These results show the impact of the VR screen on sound distribution within the space and as to how it is perceived by the human subjects with respect to each speaker (real and or virtual) as indicated. The measured data were viewed and examined using the Noise Image software with and without the HRTF as a filter for the sound source directionality as it was perceived by the subject at the center of the space. See **Figures 8 and 9** for HRTF at 10 and 60th elevations for 8 Sources at 1.5KHZ. The measurement and simulation path within real and virtual environment for all experimental procedures are shown in a flow chart diagram including the steps for HRTF filtering Noise Image software within **Figure 10.**

5. DISCUSSIONS

The results indicate that the inter-reflection and or the missing 4th wall (back wall) within the virtual environment has impacted the subject sound localization of the sound source within VR due to the loss of reflections. The locations and field of projection for each of the speakers have impacted the subjectsô ability to identify each speakerõs location within VR. Survey results based on subject listening to recorded binaural files using dummy head show the impact of such condition on source localization. See the noticeable differences within 3rd and 6th rows in **Table 3.** These results validates the hypothesis that the wall surfaces do contribute to the interaural time difference (ITD). However, parametric tests for examining the interactions between large numbers variables require a larger number of subjects. This experiment includes only 32 subjects.

6. FINAL REMARK AND FUTURE WORK

We have demonstrated in both real and simulated room acoustic environments that when the sound sources are positioned with an equal azimuth and with a small difference in the ITD index; will create most confusing cues toward subjectõs ability in sound localization. We are able to recreate or simulate audio conditions within the VR environments to stimulate human perception positively, and provide supportive evidence that it is possible to create an audio environment that could be used for room acoustic studies and architecture design evaluations.

This paper has shown the implementation of an integrated method utilizing new scalable technique for approximating indirect sound in fully dynamic scenes for real-time simulation in a fully immersive virtual environment. The application of such methods allows various scenes in combination with wide-spread real-time rendering techniques for sound to be utilized for experiencing virtually designed spaces under investigation. Future wireless physiological/neurological monitoring in the VR offers a great opportunity for unobtrusively quantifying of human response (conscious or subconsciously) to a precisely controlled and readily modulated VR representation of an environment.

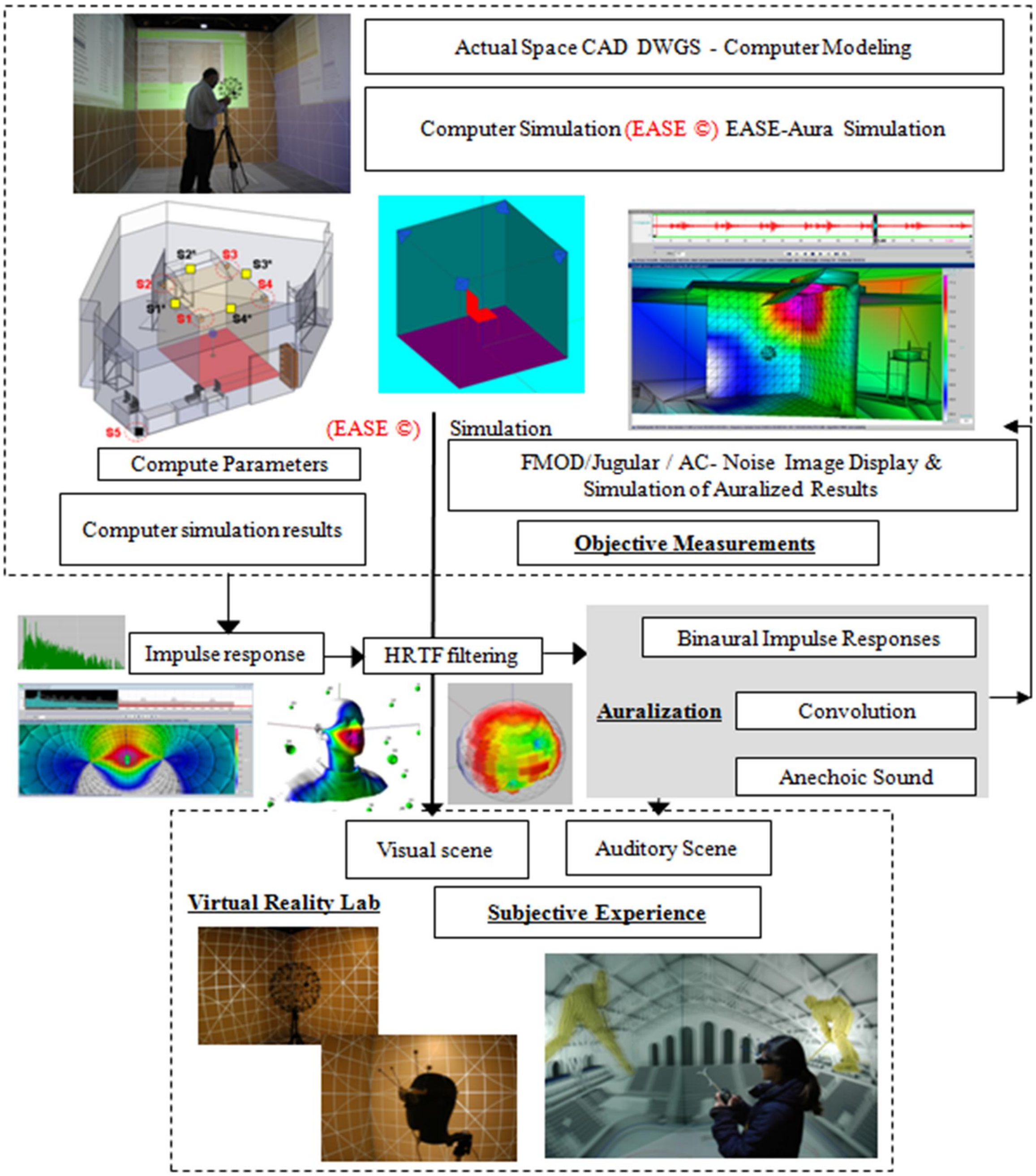

Figure 10 Measurements and simulation path within real and virtual environment.

REFERENCES

[1] Loomis, J., Golledge, R., and Klatzky, R. Navigation system for the blind: Auditory display modes and guidance. Presence: Teleoperators and Virtual Environs 7, 2 (Apr 1998), 193ï 203.

[2] Rutherford, P. Virtual acoustic technology: Its role in the development of an auditory navigation beacon for building evacuation. In Proc. 4th UK Virtual Reality SIG Conference (London, UK, 1997), R. Bowden, Ed., Brunel University.

[3] Begault, D. 3-D Sound for Virtual Reality. Academic Press, Cambridge, MA, 1994.

[4] Hall Ted W., Navvab, M., 2011, Ch. 11, PART III: "Virtual Reality as a Surrogate Sensory Environment. Book by Tauseef GulrezñAdvances in Robotics and Virtual Reality", Springer's book/978-3-642-23362-3, http://www.springer.com/engineering/computational + intelligence+and+complexity/book/978-3-642-23362-3

[5] G. Heilmann, A. Meyer, D. Döbler: Time-domain beamforming using 3Dmicrophone arrays, Proceedings of the BeBeC 2008, Berlin, Germany, 2008

[6] Don H. Johnson, Dan E. Dudgeon, ñArray Signal Processing,ò1993 by PTR Prentice- Hall, Inc.

[7] Dirk Döbler and Gunnar Heilmann, ñPerspectives of the Acoustic Camera,ò The 2005 Congress and Exposition on Noise Control, August 2005.

[8] Hartmann, W. M. (1999). "How We Localize Sound," in Physics Today, pp. 24-29.

[9] Beier KP (2000) Web-Based Virtual Reality in Design and Manufacturing Apps. In: Hansa International Maritime Jour 137/5:42ï 47

[10] http://www.fmod.org/index.php/fmod

[11] ISO/IEC (1997) the Virtual Reality Modeling Language (ISO/IEC 14772¬1:1997). ISO, Geneva http://www.web3d.org/x3d/specifications/vrml/ISO-IEC-14772-VRML97/

[12] Howard, D. M., and Angus, J. A. S. (2006). Acoustics and psychoacoustics (Focal Press, Amsterdam).

[13] Foley, J. D, van Dam, A. Feiner, S., Comp Graphics, Principles Practice, 2nded. (Addison-Wesley, 1992).

[14] O. Jaeckel: Strengths and Weaknesses of calculating Beamforming in the time domain, Proceedings of the BeBeC 2006, Berlin, Germany, 2006

[15] Any. Meyer, D. Döbler, "Noise source localization within a car interior using 3D microphone arrays, Proceedings of the BeBeC 2006, Berlin, Germany, 2006

[16] Andy Meyer, Dirk Döbler, Jan Hambrecht, Manuel Matern, "Acoustic Mapping on three-dimensional models", International Conference on Computer System & Technologies, 2011.

[17] Valmiki, V. Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters. Doctoral thesis, Helsinki University of Technology, Lab. of Acoustics and Audio Signal Processing, Report 37, 1995. Available at http://www.acoustics.hut./ -vpv/publications/vesa phd.html.

[18] Tapio Lokki, Matti Gr¨ohn, Lauri Savioja, and Tapio Takala, A Case Study of Auditory Navigation in Virtual Acoustic Environments, IEEE, 2000, **ISSN:** 1070-986X http://ieeexplore.ieee.org/servlet/opac?punumber=93

[19] Enhanced Acoustic Simulator for Engineers - E.A.S.E. Software- Acoustic modeling, http://afmg.eu/index.php/company.html.

[20] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, ñThe CIPIC HRTF Database,ò Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electro acoustics, pp. 99-102, Mohonk Mountain House, New Paltz, NY, Oct. 21-24, 2001. [http://www.ece.ucdavis.edu/cipic/spatial-sound/hrtf-data/]

[21] E.J. Macaulay, W.H. Hartman, B. Rakerd, Physics today, March 2010

[22] G.W. Stwart, ñThe function of intensity and phase in binaural location of pure tone.ò Physics, Rev.15, May1920, pp.425.

[23] Hartley, Fry, ñThe binaural location of pure toneò, Physics, Rev.18, Dec1921, pp.431.

[24] F.A. Firstone, ñThe phase differences and amplitude ratio at the ears due to a source of pure toneò, Journal of Acoustic Society, Oct, 1930, pp. 260-280.