Evaluation of a New Stereophonic Reproduction Method with Moving "Sweet Spot" Using a Binaural Localization Model

SEBASTIAN MERCHEL AND STEPHAN GROTH

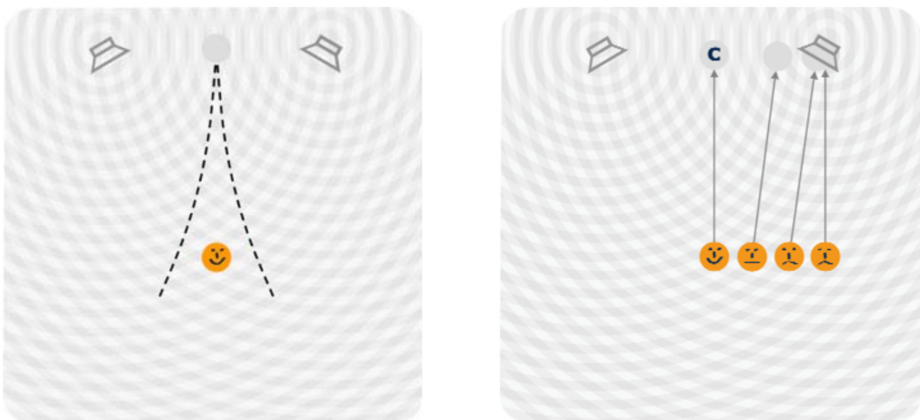
Chair of Communication Acoustics, Dresden University of Technology, 01062 Dresden, Germany

ABSTRACT

This paper describes the use of a binaural localization model to evaluate the utility of a new stereophonic play back system. The system is designed to adjust the "sweet spot" to a moving listener position in real time. This is done by adaptively manipulating time delay and level differences between the loudspeaker signals in reference to the center of the listeners head. The localization in such a system is one of the most important quality parameters. It is investigated on the basis of a binaural model after Braasch [1]. The modeled localization angle is compared to the target angle for different phantom source and listener positions. The absolute value of the difference between modeled localization angle and target angle to phantom source is plotted and discussed using "vector maps" and "quality maps". The results indicate that an adjustment of signals corresponding to the center of the listener's head improves the localization over the whole listening area. Some localization error remains, which can be estimated and compensated for using the binaural model. A real time implementation of the system confirms the modeled results.

INTRODUCTION

Stereophonic multichannel reproduction is widely used today. One major disadvantage of stereophonic play back systems is the narrow "sweet spot" in which correct audio localization is possible (Figure 1). Due to this limitation, the listener's freedom of movement is drastically restricted. To improve stereophonic reproduction by releasing the listener from its static hearing position was an issue since the beginning of stereophony.



(a) **(b)**

Fig. 1: "Sweet Spot" in a conventional stereophonic reproduction system. **(a)** Schematic figure. **(b)** Localization shift of a center phantom source C depending on the listener's position.k

LITERATURE

A number of methods exist which deal with broadening the area of stereophonic perception by *adjusting loudspeaker characteristics*. Those methods lead to contradicting localization cues from interaural level differences (ILDs) and interaural time differences (ITDs). This results in localization blur for excentric listening positions [2].

Another method is the *adjustment of the delay and amplitude difference between the loudspeaker signals* to compensate for different distances to the listener's position. This adjustment can be done manually or by using a measuring microphone. In all cases, existing systems are designed for static positions and do not respond to movement of the listener. For the first time Kyriakakis [3] described a system which uses head tracking and time delay between the loudspeaker signals to move the "sweet spot" automatically. Unfortunately, there are no publications about the usefulness and emerging artifacts of time delay and amplitude adjustment in off-center listening.

A comprehensive literature survey can be found in [4].

ADAPTIVE SIGNAL ADJUSTMENT

The present work presents a play back system which manipulates the loudspeaker signals depending on the listener's position in real-time. Therefore, the x-y position of the listener is tracked. The delay and amplitude ratio of the loudspeaker signals is calculated in such a manner that a phantom source ideally stays at its position independent of the listener position [4].

BINAURAL MODEL

To evaluate the utility of the system, a binaural model after Braasch [1] is used (Figure 2). Two sources generate the left and right channel of a stereophonic system. The outer and inner ear is modeled. Binaural localization cues (ITDs and ILDs) are estimated using a cross-correlation analysis in several frequency bands. The corresponding localization angle was found with a remapping algorithm using measured HRTFs. The interaction of ITDs and ILDs in the localization process is still object of research. Thus resulting localization angles (φ_{ITD}, φ_{ILD}) are examined separately.

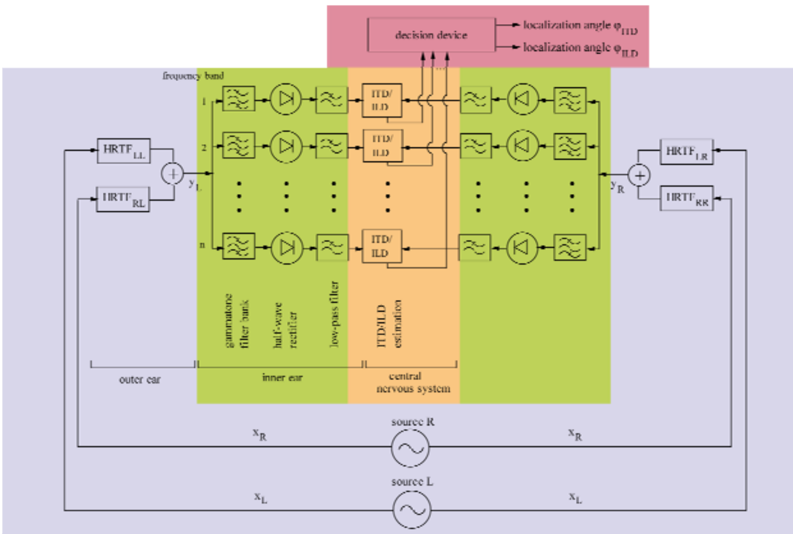


Fig. 2: Adapted binaural model for phantom source localization in a stereophonic setup after Braasch [1].

MODELED LOCALIZATION IN A STEREO SETUP

By means of the described model, different play back and hearing situations with or without signal adjustment are simulated in a stereo reference setup with a base width of 2 meters. A band-limited noise stimulus is used from 300 Hz to 4500 Hz.

Modeling of a center source

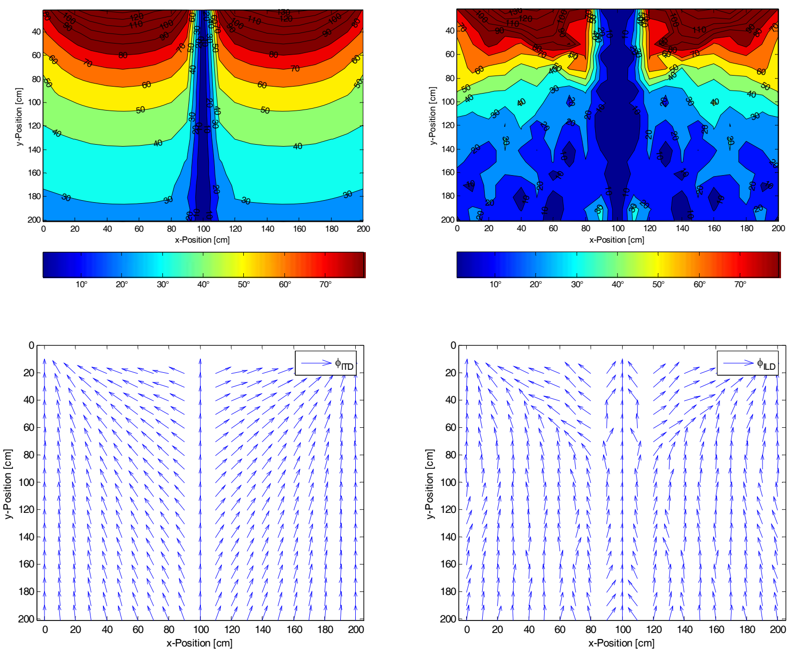
First, both loudspeakers are driven with identical signals to produce a center source. The improvement of the system with signal adjustment can be illustrated by using “quality maps”. They show the localization error (absolute value of the difference between target angle to phantom source and modeled localization angle) over the whole listening area. The upper row of Figure 3 shows the resulting error for a center source (a) based on ITD cues and (b) based on ILD cues *without* signal adjustment. Without signal adjustment only a narrow area with a minor localization error emerges – the “sweet spot”. When leaving this area, the error increases rapidly until the precedence effect appears. The upper row of Figure 4 shows the resulting error for a center source *with* signal adjustment. It can be seen that the area with a minor localization error can be considerably broadened, except for extreme hearing positions near the loudspeakers. This can be shown even more intuitively by using “vector maps”. In the lower row of Figure 3 and 4 the modeled localization angle is plotted as a direction vector over the whole listening area. Again a center source is modeled without (Figure 3) and with (Figure 4) signal adjustment. As can be seen, without signal adjustment the localization vector rapidly points to the nearer loudspeaker while leaving the “sweet spot”. With signal adjustment, the position of the perceived phantom source almost remains stable at the center. Apart from a center source, different stereo recording techniques are simulated to investigate the stability of a shifted source.

Modeling of a 10° left phantom source using intensity stereo

In intensity stereo a shifted source is produced by applying different levels to the right and the left channel. Figure 5 and 6 (top) show the “quality maps” that are based on (a) ITD and (b) ILD cues for a phantom source which is shifted about 10° to the left (relative to the “sweet spot”). This relates to a level difference of $\Delta L = 6$ dB between the left and the right loudspeaker. Note that in this case the “sweet spot” becomes asymmetric. A comparison of the “quality maps” shows a broadening of the listening area if signal adjustment is applied. This is true for both ITD and ILD cues. The “vector maps” in Figure 6 (bottom) show an almost stable localization angle to the shifted source; ITD and ILD cues lead to the same direction. Therefore, a sharp image is produced.

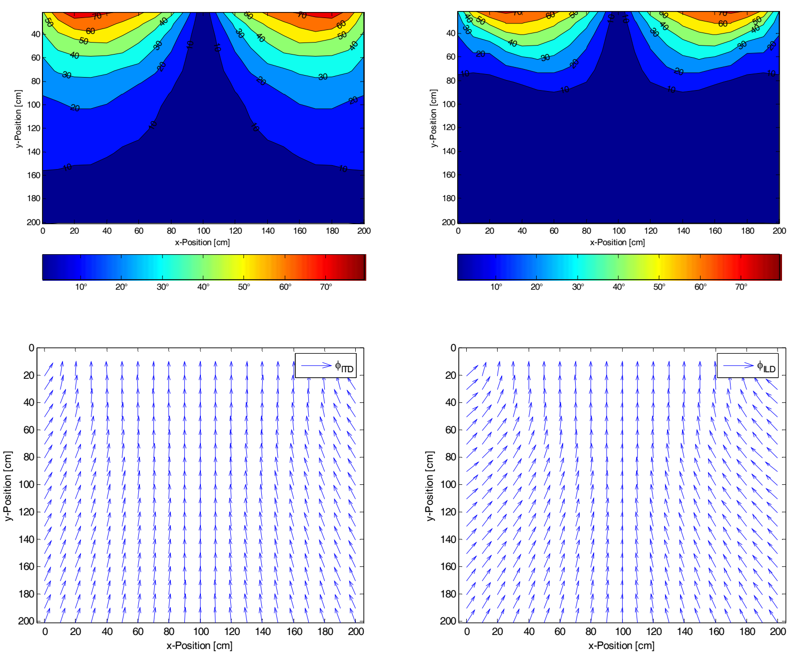
Modeling of a 10° left phantom source using time stereo

However, there is a different situation when using time stereo. In time stereo, the shifted source is produced via time delay between the channels. It can be shown that time delay between the loudspeaker signals lead to unnatural level differences at the listener’s ears (see also Lipshitz [5]). Localization cues from ITDs and ILDs contradict. This can be seen in Figure 7 where a shifted source is simulated throughout a time difference of $\Delta t = 0.14$ ms. Figure 8 shows the same situation but with signal adjustment. For ITD cues the listening area is broadened similar to “intensity” stereo, but for ILD cues the error becomes larger. In the “vector maps” in Figure 8, it can be seen that this leads to contradicting localization cues and thus to localization blur and image broadening. Still, if compared to no signal adjustment, the “sweet spot” enlarges.



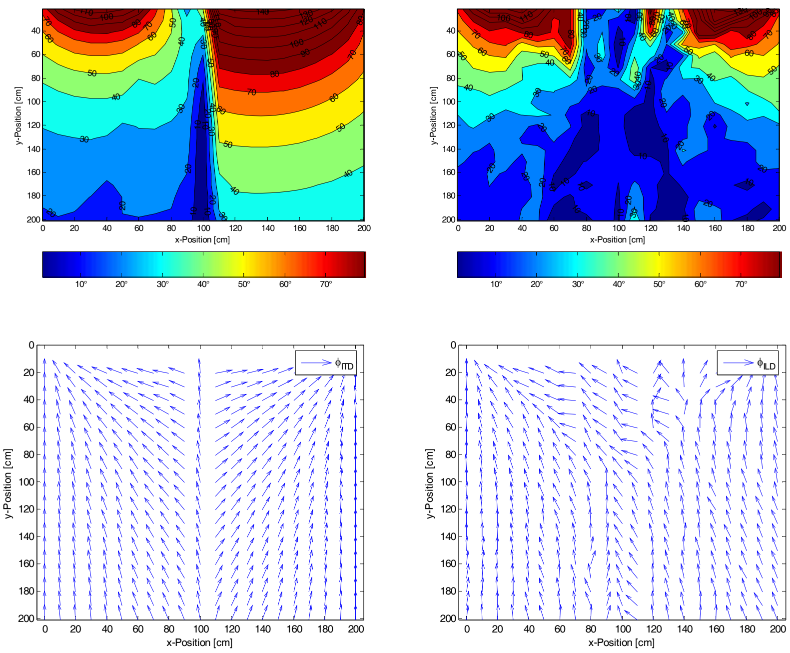
(a) **(b)**

Fig. 3: “Quality maps” (top) and “vector maps” (bottom) for a simulated **center** phantom source **without signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.



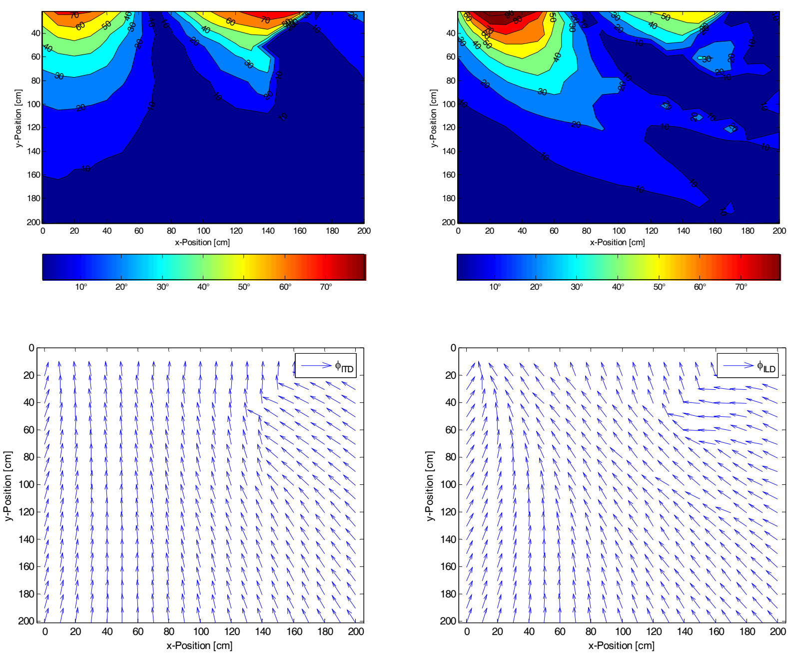
(a) **(b)**

Fig. 4: “Quality maps” (top) and “vector maps” (bottom) for a simulated **center** phantom source **with signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.



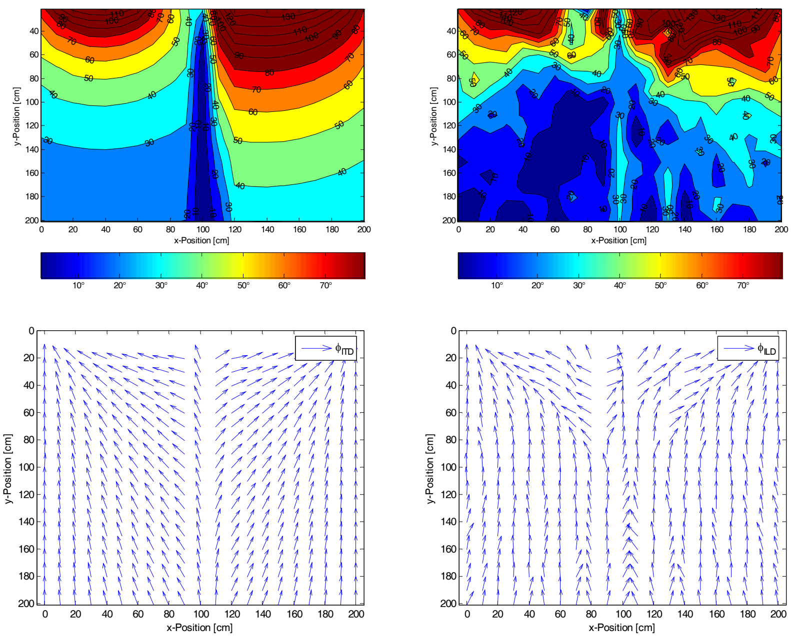
(a) **(b)**

Fig. 5: "Quality maps" (top) and "vector maps" (bottom) for **intensity stereo** (ΔL = 6 dB) **without signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.



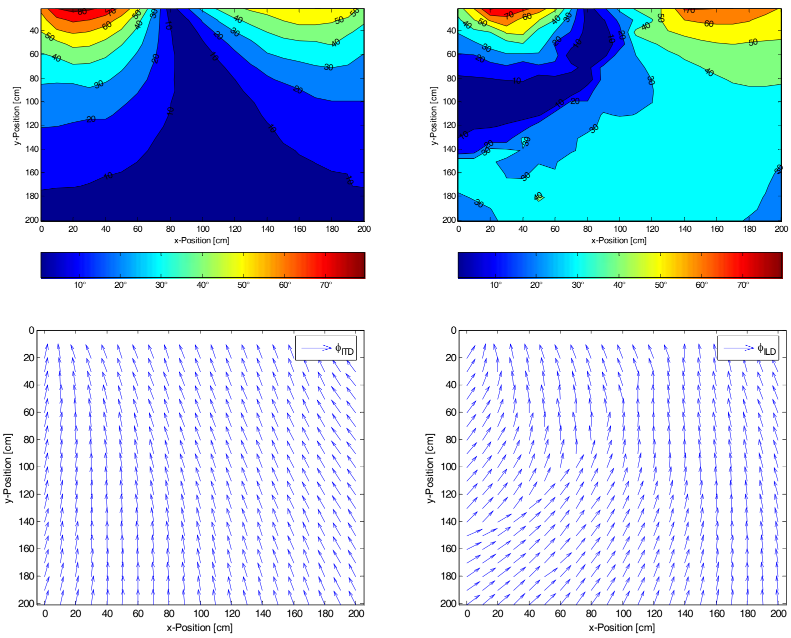
(a) **(b)**

Fig. 6: "Quality maps" (top) and "vector maps" (bottom) for **intensity stereo** (ΔL = 6 dB) **with signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.



(a) **(b)**

Fig. 7: “Quality maps” (top) and “vector maps” (bottom) for **time stereo** (Δt = 0.14 ms) **without signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.



(a) **(b)**

Fig. 8: “Quality maps” (top) and “vector maps” (bottom) for **time stereo** (Δt = 0.14 ms) **with signal adjustment**; **(a)** based on ITD cues; **(b)** based on ILD cues.

IMPLEMENTATION

The system was implemented on a PC using C++. The response time of the system depends mainly on the length of the audio buffer (which is necessary for adaptively delaying the loudspeaker signals) and the reaction time of the head tracker. An update interval of 100 ms was found to be suitable. The influence of inaccurate head position detection can be seen in Figure 9 and 10. Figure 9 shows the absolute error of the delay between the loudspeaker signals with signal adjustment if the listener position is erroneously detected 10 cm to the left (x_{ERR}) or to the fore (y_{ERR}). The error is relatively large and results in an additional phantom source shift (compare 0.14 ms delay corresponds to 10° phantom source shift for the standard listening position). It can be seen that an inaccurate detected x-position is more crucial than an inaccurate y-position. Figure 10 shows the absolute error of the amplitude difference between loudspeaker signals with signal adjustment if the listener position is detected wrong as above. Again inaccurate detected x-position is more crucial. The resulting error of the amplitude difference is relatively small.

In summary, the spatial precision of the tracking system needs to be accurate only in x-direction. To be suitable for daily use the system must not use markers. Therefore, a webcam based head tracker was developed in cooperation with the faculty of computer science at TU Dresden.

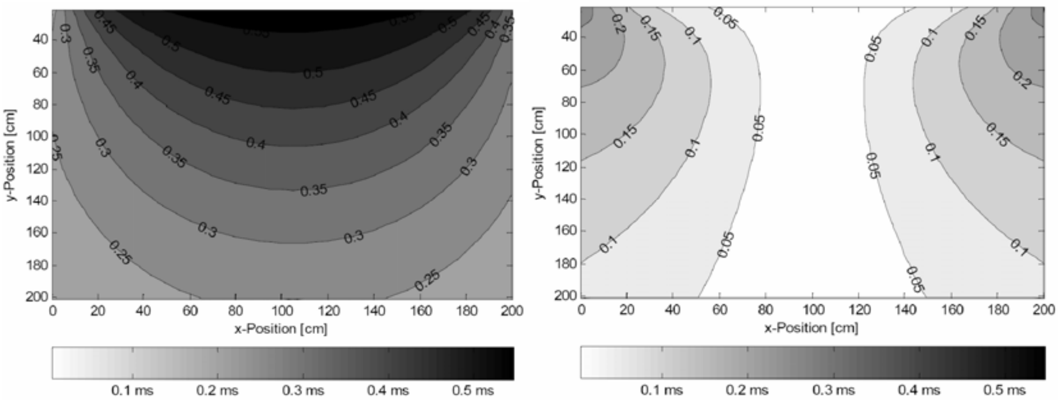


Fig. 9: Absolute error of the delay between the loudspeaker signals if the listener position is detected wrong; left: x_{ERR} = 10 cm to the left, right: y_{ERR} = 10 cm to the fore.

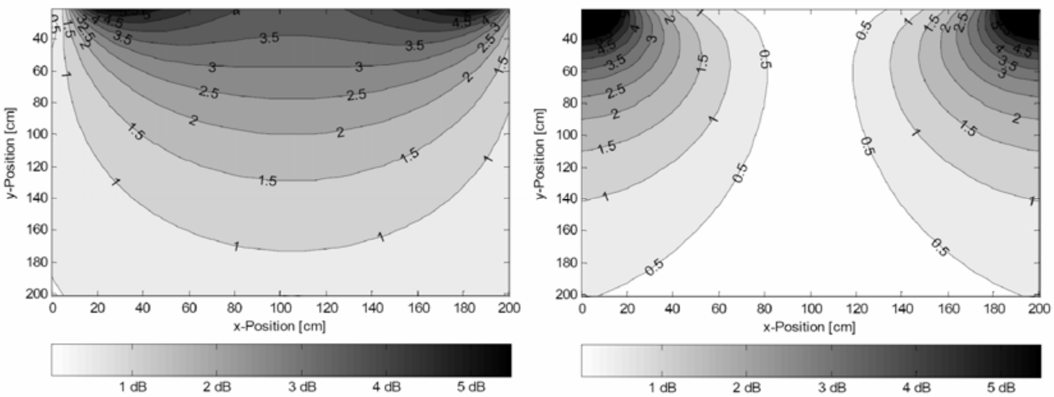


Fig. 10: Absolute error of the amplitude difference between loudspeaker signals if the listener position is detected wrong; left: x_{ERR} = 10 cm to the left, right: y_{ERR} = 10 cm to the fore.

CONLUSION

The localization in a stereophonic system with adaptively adjusted “sweet spot” was investigated on the basis of a binaural model. The results indicate that an adaptive adjustment of the signals relative to the center of the listeners head improves the localization over the whole listening area. This is independent of the recording technique. Still some localization error remains. This remaining error can be estimated and compensated for using the binaural model.

The ability to compensate also for head rotation, elevation effects or coloration in reverberant environments will be topics of further research.

REFERENCES

[1] Braasch, J.: Modelling of binaural hearing, Communication Acoustics (J. Blauert ed.), Springer Verlag, 2005, S. 75-103.

[2] Bauer, B., Broadening the Area of Stereophonic Perception, Journal of the Audio Engineering Society, Vol. 8, No. 2, 1960, S. 91-94.

[3] Kyriakakis, C. et al., Signal Processing, Acoustics, and Psychoacoustics for High Quality Desktop Audio, Journal of Visual Communication and Image Representation, Vol. 9 No. 1, 1998, S. 51-61.

[4] Merchel, S., Groth, S., Analysis and Implementation of a Stereophonic Play Back System for Adjusting the “Sweet Spot” to the Listener’s Position, Proceedings of 126th AES Convention, Munich, Germany, 2009.

[5] Lipshitz, S. P.: Stereo Microphone Techniques: Are the Purists Wrong?, Journal of the Audio Engineering Society, Vol. 34, No. 9, 1986, S. 716-744