ARCHIVES OF ACOUSTICS
**33**, 1, 7–10 (2008)

# CONVERSION OF STEREO RECORDING TO 5.1 FORMAT USING HEAD-RELATED TRANSFER FUNCTIONS

Paweł HEN, Maurycy KIN, Przemysław PLASKOTA

Wroclaw University of Technology
Institute of Telecommunications, Teleinformatics and Acoustics
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: przemyslaw.plaskota@pwr.wroc.pl

(*received June 15, 2007; accepted November 30, 2007*)

The paper presents the conversion of stereo recordings into multi-channel format 5.1 by means of the HRTF filtering. An algorithm of additional channels preparation (a central one and two surround ones) using various filters created on the base of HRTF was tested. The filter bank was prepared on the basis of impulse responses given by users. The algorithm uses the impulse response convolution of the filter with a processed signal. The convolution is realized in frequency domain, i.e. Fourier transforms of the processed signals are multiplied. Recordings created in such way were tested for the features of sound color distortion and spaciousness of the audio-scene in frontal plane.

**Keywords:** HRTF, sound recording.

## 1. Introduction

The number of users of surround sound systems increases rapidly. Unfortunately, there are not so many recordings in 5.1 format and quite a number of classical stereo CDs still exist. Any amplifier of a home theater system has a possibility of stereo reproduction for all speakers. It uses simple mixing-up algorithms based on creating back signals from frontal ones, and sometimes exploring time-delays for spaciousness improvement. The effectiveness of such systems is not satisfying, so there is a need to set up research on methods of the process of mixing-up. One of the methods presented in this paper is based on the HRTF filtering [2, 4].

The influence of a head, torso and pinna on sound spectrum is an additional cue which together with basic cues such as Interaural Time Difference and Interaural Level Difference enables to localize a sound source. This paper describes a computer program which uses various sets of HRTFs. When filtering, the system uses the symmetry of processing: a complete set of HRTFs contains filter banks measured for the left and right ear for one azimuth angle (left-back – 125°) and filters obtained for the angle at

the opposite side of the head (right-back – 235°). These exemplary azimuth angles are a compromise between the set-up of surround speakers suggested by Dolby Laboratories and the listener's localization of sound sources.

The program calculates the filter's coefficients necessary to process the recording on the base of impulse responses. The coefficients measured for the angle of $\theta = 0°$ are used to create the central channel signal. Filters obtained for $\theta = 125°$ are used for surround signals – left-back, and similarly for $\theta = 235°$ – right-back. For signal processing, the method of the FFT convolution is used [5]. This allows effective filtering, even in real time. For processing particular recordings, the set of FFTs is used which calculates a FFT with the Mixed-Radix algorithm [6]. Moreover, the procedure calculates only the real part of a FFT which reduces the number of calculations by half.

## 2. Processing the recording

Converting stereo recordings into 5.1 format starts with filtering the original signal. In the next step, filtered audio tracks are processed in order to remove some artifacts of filtering. The sixth channel signal (for subwoofer) is created artificially. The subwoofer signal is a sum of the original left and right signals reduced by 3 dB, and the cut-off frequency of a low-pass filter is about 120 Hz. This process is illustrated in Fig. 1.

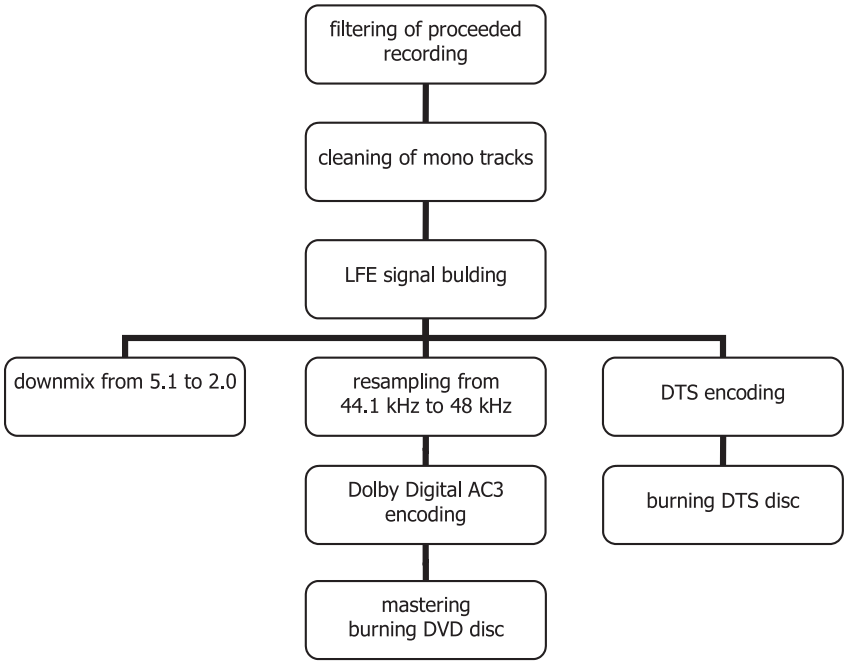


Fig. 1. Block diagram of mix-up process steps.

The central channel is made of filtered signals reduced down by 6 dB. The rear channels are created by summing up filtered front channels multiplied by particular coefficients, i.e. the left back contains two channels: filtered left front channels are multiplied by 0.65 and filtered right front channels are multiplied by 0.45. These coefficients determine which signal's amplitude is the most important and dominates in a track (see Fig. 2). This process is carried out upon separated left and right signals of a stereo recording. The last step is coding the original recording into 5.1 format [1, 3].

This process also allows a mix-down operation for the previously enhanced recording, i.e. from stereo to 5.1. This is done to compare the AES/ITU recommended mixing-down algorithm and other algorithms commonly used in commercial devices with the one involving HRTF proposed by the authors. The main aim of a mixing-down operation is to check the compatibility between the classical stereo and the surround sound format, as analogous to mono-stereo compatibility.

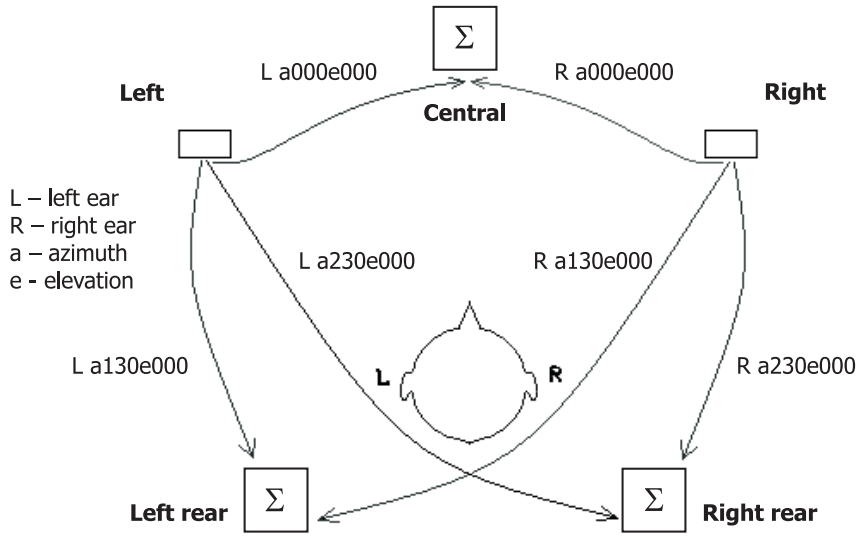


Fig. 2. Signal processing diagram for filtering procedure.

## 3. Listening impressions

Three of the authors tested this algorithm separately and had common impressions. The first important feature of the system was a great influence of musical style on perceived spaciousness: classical music allowed to achieve more effective results than pop-music, what was mainly caused by a greater amount of headroom. Moreover, changes to the color of natural instruments sound were noticeable for pop-music mainly in the mid-frequency range. This was not observed when a mix-down process was provided by a commonly used device compliant with standard ITU algorithms. It was noticed,

that when musical material was listened via headphones, the addition of a central signal caused a better boost in central direction than in all other sides. Moreover, listeners had an impression to be more surrounded by the sound coming from the back. Concluding, the headphone presentation of a mix-down made from a previously created 5.1 format, provided a closed audiosphere around the listener. Spaciousness was satisfying because there were no noticeable changes in the original set of the scene.

When playing back the 5.1 format, the method of producing surround by hardware means was compared with the described HRTF based method, and the main conclusion was that listeners were surrounded by the soundsphere coming from every direction (an impression similar to a strongly diffused field), and the sound was rich, enhanced by the proposed algorithm. The scene remained unchanged and the possibility of a volume control for every channel allowed to satisfy special requirements. Noticeable sound coloration was not observed.

## 4. Conclusion

Concluding, the comparison between classical hardware way of mixing-up with the one obtained via the use of the proposed HRTF based algorithm showed, that the system described in this paper gave better results of spatial impression. It can be said that the HRTF filtering applied in our algorithm is useful for a mixing process of standard stereo recordings up to 5.1, and down back to stereo for headphone exposure. The algorithm is a very useful tool for creating sound space during the process of mixing-up a stereo recording into 5.1 format for the DVD Audio or DTS standards.

## References

[1] *ATSC Standard: Digital Audio Compression (AC-3), Revision A* – Advanced Television Systems Committee 1750 K Street, N.W. Suite 1200, Washington, D.C. 2006.

[2] CHENG C. I, WAKEFIELD G. H., *Introduction to head-related transfer functions (HRTFs): representations of HRTFs in time, frequency, and space*, J. Audio Eng. Soc., **51**, 9 (2003).

[3] ITU-R Recommendation BS.775-1 *Multichannel stereophonic sound system with and without accompanying picture*, 1994.

[4] KLEINER M., DALENBACK, B. I. SVENSSON P., *Auralisation – an overview*, Journal of the Audio Engineering Society, **41**, 11, 861–875, (1993).

[5] OPPENHEIM A. V., SCHAFER R. W., *Digital signal processing*, Englewood Cliffs, NJ, Prentice-Hall, Inc. 1975.

[6] SINGLETON R. C., *An algorithm for computing the mixed radix fast Fourier transform*, IEEE Trans. Audio Electroacoust., **AU-17**, p. 93, June 1969.