BINAURAL TECHNIQUE

Dorte Hammershøi, Henrik Møller

Department of Acoustics, Aalborg University
Fredrik Bajers Vej 7 B4, DK-9220 Aalborg Ø, Denmark
e-mail: acoustics@acoustics.dk

Abstract

The term binaural technique is used as a headline for methods for sound recording, synthesis and reproduction, where the sound pressure signals recorded (or synthesized) are the eardrum signals of the listener, and where successful reproduction is achieved, if these are truly reproduced in the ears of the listener. Numerous applications to these methods exist of which some are introduced. Further, the lecture included main results of the investigations that are the current basis for the utilization of the binaural technique at Aalborg University and possibly elsewhere, including investigations on sound transmission in the ear canal, measurements of HRTFs (head-related transfer functions), calibration of headphones, and localization experiments (in real life and with binaural recordings from real heads and artificial heads). Please see [1] for references.

0. Introduction

The fundamental idea of binaural technique is based on the fact that we create our auditory impression on the basis of only two inputs, namely the sound pressures at our two eardrums. If these are recorded – literally – in the ears of the listener, and reproduced exactly when played back, then the same listening experience – including all spatial aspects and spectral properties (timbre) – is obtained.

Binaural signals may also be synthesized by computer when the sound transmission from the source to the listener's ears has been thoroughly mapped – in principle – for all possible placements of the sound source. It can then be controlled by the computer where the listener will hear the sound, and sound sources can be placed at any place spatially. This technique is obvious for virtual reality application, but many other areas of applications exist.

The definition of binaural technique may be extended more generally to binaural technology, as done by Blauert [2]: *"Binaural technology is a body of methods that involves the acoustic input signals to both ears of the listener for achieving practical purposes, for example, by recording, analyzing, synthesizing, processing, presenting, and evaluating such signals".*

Basic principles

The methods required for obtaining true binaural signals and reproducing them correctly are presented in summary in the following. For a thorough review of spatial hearing, see Blauert [3] and for a detailed introduction to the fundamentals of binaural technique, see Møller [4].

Recording

With binaural recording, the sound is recorded in the ears of the listener, though usually replaced by a mannikin with the relevant acoustical properties similar to those of a human. The technique is therefore commonly referred to as the *artificial head recording technique* or the *dummy head recording technique*. Playback is most often done by headphones, because they offer a complete channel separation and acoustical isolation from the surroundings, thus enabling the optimal control of the reproduction situation.

When the binaural recording is made, the spatial properties of the exposing sound field are transformed into two electrical signals representing the input to the left respectively right ear. For some purposes the cumbersome recordings at the eardrums may be replaced by recordings elsewhere in the external ear, where the full spatial information is maintained. The correct reproduction will then be obtained, if the reproduction chain is calibrated accordingly. This is done by reproducing the sound pressures at the exact same physical point as the recording was made. Thus, if the recording isn't made deeply in the ear canal, neither the headphone equalization requires measurements close to the eardrum.

Binaural synthesis

The artificial head serves to perform the alteration of the sound field that the listener would have caused in the recording situation. When placed in a free sound field, the listener will obstruct an incoming sound wave, and ears, head and body will cause a linear filtering of the sound signal. This filtering is described by the *Head-related Transfer Functions*, abbreviated HRTFs.

The general definition of the HRTF is given as the ratio of two sound pressures, both measured in the free field:

$$HRTF(\phi,\theta) = \frac{sound\ pressure\ at\ listener's\ ear}{sound\ pressure\ at\ head\ center\ position}(\phi,\theta)$$

where (ϕ,θ) indicate the angle of incidence of the sound wave. The complete HRTF is then the two ratios for left respectively right ear (Figure 1). The similarities and differences in the sound signals reaching our two ears are the cues that enable us to localize the sound source by our hearing.

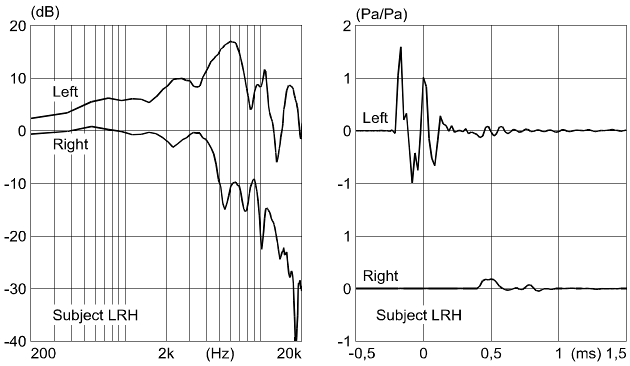

Figure 1*.Head-related transfer function (HRTF) in frequency and time domains (sound from the left for subject LRH, measured at the blocked entrance, after Møller et al. [5]). It is obvious in both domains that the sound reaching the ear closest to the sound source (left ear) is stronger than the sound reaching the ear opposite to the sound source (right ear). In the frequency domain representation it can also be seen that the sound transmission to the left ear exceeds 0 dB which is mainly due to pressure buildup and diffraction around ear, head and body. The frequency response for the right ear sound transmission resembles the characteristics of a lowpass filter, which is due to the shadowing effect of the head. In the time domain it can be seen that the sound at the closest ear (left) arrives some hundred microseconds earlier than the sound at the opposite ear (right).*

Note that in the general definition absolutely all properties of the sound transmission (frequency dependent amplitude, absolute phase, arrival time, etc.) are included. Any descriptor of localization cues, such as the interaural differences in time (ITD), phase (IPD), level (ILD) or in intensity (IID), or monaural cues, group delay, etc. *is maintained* in the HRTF. The HRTF thus represent completely and uniquely the sound transmission for the particular angle.

In binaural synthesis the filtering which in the real life listening situation is carried out by ear, head and body is done electronically. Filters representing the HRTFs are typically implemented by means of digital signal processing and the binaural signals are thereby generated artificially by a computer.

Applications

In the following the principles of some "classical" areas of application of binaural technique (mostly synthesis) are introduced.

Room simulation

Room simulation systems are usually developed for assistance in architectural design and acoustical adjustment of concert halls, auditoria, theaters etc. Rearrangements of walls, orchestras or audience can be listened to before making alterations in the real world. Listening conditions at critical places in the concert hall and the speech intelligibility of the auditoria can be checked by actual listening.

The simulation system can't only be used to suggest physical changes in the room. Using summation, an unlimited number of sound sources can – in principle – be simulated in the computer model, and the system can be used for adjusting any multi channel public address or electro-acoustical enhancement system prior to the costly setup.

The room simulation system can also be used in various contexts, where standardized listening conditions are needed. This could for instance be for the evaluation of different loudspeakers, where the loudspeaker producer can avoid building a rather costly listening environment, or for mobile monitoration environments for studio recording and mixing.

Binaural mixing consoles

The usual way to make recordings in the studios is to record each channel at a time, and then afterwards mix it into the intended stereo perspective using time and intensity differences. With a binaural mixing console usual recordings are transferred into binaural signals in the mixing phase, and the binaural mixing console can be regarded as an *"electronic artificial head"*.

If an artificial head is used for the recording, then the physical setup in the recording studio defines the position of the sound source for the playback situation. These recordings and ordinary recordings processed by means of the binaural mixing console can be mixed together to binaural signals.

Surround sound by headphones

The use of multichannel sound reproduction systems is increasing, e. g. the HDTV and DVD formats, where five or more channels are proposed. The compatibility with headphone playback isn't trivial. Using binaural synthesis the five channels can be transformed into a two channel signal for headphone playback.

The idea isn't at all new, and has also been proposed for headphone playback of traditional stereo recordings, where the same compatibility problem exists.

Communication systems

The use of binaural technique in the context of communication systems is most often proposed for situations, where the receiving part of the information needs to pay attention to more than one communication channel at a time. The idea is to spatially arrange the communication channels around the listener, and in this way enable him or her to distinguish between the different channels as in daily life. The advantage of the spatially arranged channels is that our hearing can focus on one of several sound sources in noisy environments (the *cocktail-party effect*), if the sounds are heard from different directions.

Implementations of this kind are seen within military research institutions to be used between pilots and air control towers. More general kinds of teleconferencing systems also exist.

3D auditory displays

The idea of the 3D auditory displays is to create spatially arranged auditory icons – sometimes denoted *earcons* – for situations, where it is either natural to have audible information passed, or where for instance the visual information channel is fully occupied. The sound of a fuel tank drying out is for instance a natural and important message for a pilot, and the sound of a printer could be a way to transfer the graphical interfaces of modern computers into something blinds can use.

Earcons can also be used for transfer of information which has no origin in the auditive world. Examples are sonar information, where the distance information is passed and used for instance as navigation tool for blind persons, or the association of sounds to stock exchange information (e. g. high and low pitched sounds for rates going up and down).

Obviously, earcons serve a purpose even when not arranged spatially, but the 3D auditory displays may offer an environment, where the user may effectively monitor more activities at the same time.

Virtual Reality

A modern word, for something that has been known and used for ages, is Virtual Reality (VR). The distinction between VR systems and other sound simulation systems can sometimes be difficult to see, but the ultimate aim is to provide a person stimuli for his (or her) senses so convincing that he feels like being somewhere else. The world he is experiencing might only exist as a computer model, and the computer controls various kinds of manmade machinery, which stimulates different senses (almost always vision).

References

[1] Hammershøi, D. and Møller, H. *"Methods for Binaural Recording and Reproduction"*, ACUSTICA–acta acustica (in press).

[2] Blauert, J. *"An Introduction to Binaural Technology"*, In Gilkey and Anderson: *"Binaural and Spatial Hearing in Real and Auditory Environments"*, Lawrence Erlbaum Associates, Publishers, Mahwah, New Yersey, 1997.

[3] Blauert, J. *"Spatial hearing. The psychophysics of human sound localization"*, The MIT Press, UK, 1997.

[4] Møller, H. *"Fundamentals of binaural technology"*, Applied Acoustics Vol. 36, 1992, pp. 171-218.

[5] Møller, H., Sørensen, M. F., Hammershøi, D. and Jensen, C. B. *"Head-related transfer functions measured on human subjects"*, J. Audio Eng. Soc., Vol. 43 (1995), pp. 300-321.