Proceedings of the Institute of Acoustics

FOUNDATIONS AND APPLICATION OPPORTUNITIES OF BINAURAL TECHNOLOGY

Jens Blauert

Ruhr-Universitaet Bochum, D-44780 Bochum, Germany

Humans, like most vertebrates, have two ears that are positioned at about equal height at the two sides of the head. Physically, the two ears and the head form an antenna system, mounted on a mobile base. This antenna system receives elastomechanical (acoustic) waves of the medium in which it is immersed, usually air. The two waves received and transmitted by the two ears are the physiologically adequate input to a specific sensory system, the auditory system.

The peripheral parts of the auditory system transform each of the two waves into neural spike trains, after having performed a running spectral decomposition into multiple frequency channels, among other preprocessing. The multi-channel neural spike trains from each of the two ears are then combined in a sophisticated way to generate a running "binaural-activity pattern" somewhere in the auditory system. This binaural-activity pattern, most probably in combination with monaural-activity patterns rendered individually by each ear's auditory channels, forms the auditory input to the cortex, which represents a powerful biologic multi-purpose parallel computer with a huge memory and various interfaces and in- and output ports. As an output, the cortex delivers an individual perceptual world and, eventually, neural commands to trigger and control specific motoric expressions.

It goes without saying that a number of constraints must to hold for this story to be true. For example, the acoustic waves must be in the range of audibility with respect to frequency range and intensity, the auditory system must be operative, and the cortex must be in a concious mode, ready to accept and interpret auditory information. Further, it makes sense to assume that multiple sources of feedback are involved in the processes of reception, processing and interpretation of acoustic signals. Feedback clearly occurs between the modules of the subcortical auditory system, and between this system and the cortex. Obvious feedback from higher centers of the central nervous system to the motoric positioning system of the ears-and-head array can also be observed whenever position-finding movements of the head are induced.

Although humans can hear with one ear only - so called monaural hearing - hearing with two functioning ears is clearly superior. This fact can best be appreciated by considering the biological role of hearing. Specifically, it is the biological role of hearing to gather information about the environment, particularly about the spatial positions and trajectories of sound sources and about their state of activity. Further, it should be recalled in this context that interindividual communication is predominantly performed acoustically, with brains deciphering meanings as encoded into acoustic signals by other brains.

In regard of this generic role of hearing, the advantage of binaural as compared to monaural hearing stands out clearly in terms of performance, particulary in the following areas (Blauert 1983.)

(i) localization of single or multiple sound sources and, consequently, formation of an auditory perspective and/or an auditory room impression;

(ii) separation of signals coming from multiple incoherent sound sources spread out spatially or, with some restrictions, coherent ones;
(iii) enhancement of the signals from a chosen source with respect to further signals from incoherent sources, as well as enhancement of the direct (unreflected) signals from sources in a reverberant environment.

It is evident that the performance features of binaural hearing form a challenge for engineers in terms of technological application. In this context a so-called "Binaural Technology" has evolved during the past three decades, which can operationally be defined as follows.

Binaural Technology is a body of methods that involve the acoustic input signals to both ears of the listener for achieving practical purposes, e.g., by recording, analyzing, synthesizing, processing, presenting and evaluating such signals.

Binaural Technology has recently gained in economic momentum, both on its own and as an enabling technology for more complex applications. A specialized industry for Binaural Technology is rapidly developing. It is the purpose of this chapter to take a brief look at this exciting process and to reflect on the bases on which this technology rests, i.e., on its experimental and theoretical foundations. As has been discussed above, there are basically three "modules" engaged in the reception, perception and interpretation of acoustical signals: the ears-and-head array, the subcortical auditory system, and the cortex. Binaural Technology makes use of knowledge of the functional principles of each. In the following three sections, particular functions of these three modules are reviewed in the light of their specific application in Binaural Technology.

1. THE EARS-AND-HEAD ARRAY: PHYSICS OF BINAURAL HEARING

The ears-and-head array is an antenna system with complex and specific transmission characteristics. Since it is a physical structure and sound propagation is a linear process, the array can be considered to be a linear system. By taking an incoming sound wave as the input and the sound pressure signals at the two eardrums as the output, it is correct to describe the system as a set of two self-adjusting filters connected to the same input. Self-adjusting, in the sense used here, means that the filters automatically provide transfer functions that are specific with regard to the geometrical orientation of the wavefront relative to the ears-and-head array.

Physically, this behavior is explained by resonances in the open cavity formed from pinna, ear canal and eardrum, and by diffraction and reflection by head and torso. These various phenomena are excited differently when a sound wave impinges from different directions and/or with different curvatures of the wavefront. The resulting transfer functions are generally different for the two filters, thus causing "interaural" differences of the sound-pressure signals at the two eardrums. Since the linear distortions superimposed upon the sound wave by the two "ear filters" are very specific with respect to the geometric parameters of the sound wave, it is not far from the mark to say that the ears-and-head system encodes information about the position of sound sources in space, relative to this antenna system, into temporal and spectral attributes of the signals at the eardrums and into their interaural differences. All manipulations applied to the sound signals by the ears-and-head array are purely physical and linear. It is obvious, therefore, that they can be simulated. As a matter of fact, there is one important branch of Binaural Technology that attempts to do just this.

It makes sense at this point to begin the technological discussion with the earliest, and still a very important application, of Binaural Technology, namely, authentic auditory reproduction. Authentic auditory reproduction has been achieved when listeners hear exactly the same in a reproduction situation what they would hear in an original sound field, the latter existing at a different time and/or location. As a working hypothesis, Binaural Technology begins with the assumption that listeners hear the same in a reproduction situation as in an original sound field when the signals at the two ear-drums are exactly the same during reproduction as in the original field. Technologically, this goal is achieved by means of so-called artificial heads which are replicas of natural heads in terms of acoustics, i.e. they realize two self-adjusting ear filters like natural heads.

Artificial heads, in combination with adequate playback equipment, are a basic instrumentation for a number of economically appealing applications. The playback equipment, needed for this application, is usually based on headphones. Yet, under specific, restricted conditions, loudspeakers can also be used. A first category of application in this context is subsumized under the following section.

Binaural Recording and Authentic Reproduction

These applications exploit the capability of Binaural Technology to archive the sound field in a perceptually authentic way, and to make it available for listening at will, e.g., in entertainment, education, instruction, scientific research, documentation, surveillance, and telemonitoring. It should be noted here that binaural recordings can be compared in direct sequence (e.g., by A/B comparison), which is often impossible for the original sound situations. Since the sound-pressure signals at the two ear-drums are the physiologically adequate input to the auditory system, they are considered the basis for auditory-adequate measurement and evaluation, both in a physical and/or auditory way (Blauert & Genuit, 1993.) Consequently, we have the further application category discussed below.

Binaural Measurement and Evaluation

In physical binaural measurement, physically based procedures are used, whereas in the auditory case, human listener serve as measuring and evaluating instruments. Current applications of binaural measurement and evaluation can be found in areas such as noise control, acoustic-environment design, sound-quality assessment (for example, in speech-technology, architectural acoustics and product-sound design,) and in specific measurements on telephone systems, headphones, personal hearing protectors, and hearing aids (Blauert, Els, and Schroeter, 1980, Schroeter, 1986, Schroeter & Poesselt, 1986.) For some applications, scaled-up or scaled-down artificial heads are in use, for instance, for the evaluation of architectural scale models (Els & Blauert, 1985, 1986, Xiang & Blauert, 1991, 1993.)

Since artificial heads, basically, are just a specific way of implementing a set of linear filters, one may think of other ways of realizing such filters, e.g., electronically. For many applications this adds additional degrees of freedom, as electronic filters can be controlled at will over a wide range. This idea leads to yet another category of application, as follows.

Binaural Simulation and Displays

There are many current applications in binaural simulation and displays, with the potential of an ever-increasing number. The following list provides examples: binaural mixing (Poesselt, Schroeter, Opitz, and Divenyi, 1986,) binaural room simulation (Lehnert & Blauert, 1989, 1992,) advanced sound effects (for example, for computer games), provision of auditory spatial-orientation cues (e.g., in the cockpit or for the blind), auditory display of complex data, and auditory representation in teleconference, telepresence and teleoperator systems.

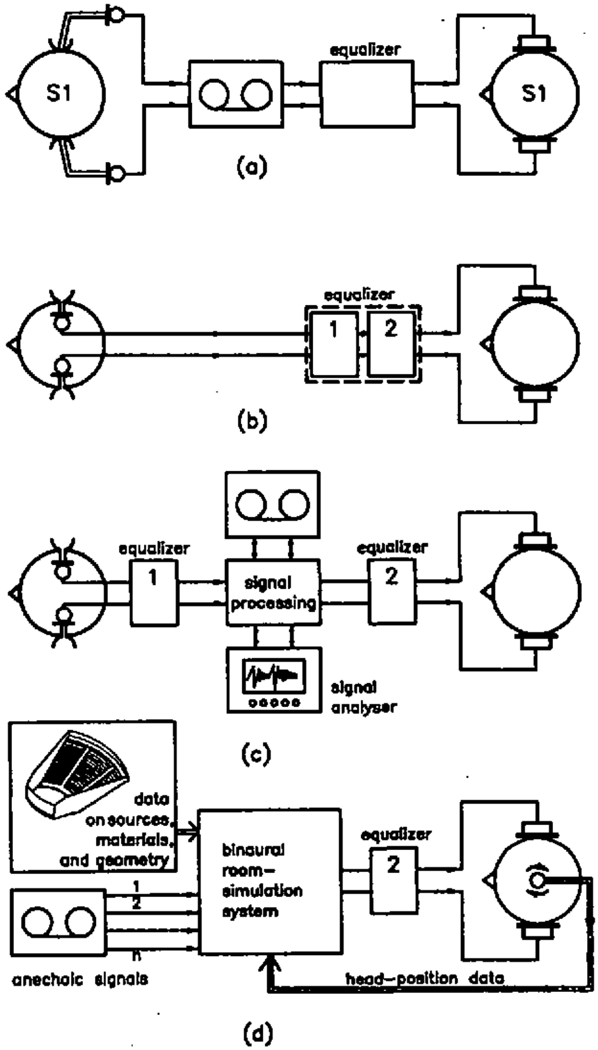

Fig.1: Binaural-Technology equipment of different complexity: (a) probe-microphone system on a real head, (b) artificial-head system, (c) artificial-head system with signal-processing and signal-analysis capabilities, (d) binaural room-simulation system with head-position tracker for virtual-reality applications.

Fig. 1, by showing Binaural-Technology equipment in an order of increasing complexity, is meant to illustrate some of the ideas discussed above. The most basic equipment is obviously the one shown in panel (a). The signals at the two ears of a subject are picked up by (probe) microphones in a subject's ear canal, then recorded, and later played back to the same subject after appropriate equalization. Equalization is necessary to correct linear distortions, induced by the microphones, the recorder, and the headphones, so that the signals in the subject's ear canals during the playback correspond exactly to those in the pick-up situation. Equipment of this kind is adequate for personalized binaural recordings. Since a subject's own ears are used for the recording, maximum authenticity can be achieved.

Artificial heads (panel b) have practical advantages over real heads for most applications; for one thing, they allow for auditory real-time monitoring of a different location. One has to realize, however, that artificial heads are usually cast or designed from a typical or representative subject. Their directional characteristics will thus, in general, deviate from those of an individual listener. This fact can lead to a significant decrease in perceptual authenticity. For example, errors such as sound coloration or front-back confusion may appear. Individual adjustment is only partly possible, namely, by equalizing the headphones specifically for each subject. To this end, the equalizer may be split into two components, a head equalizer (1) and a headphone equalizer (2). The interface between the two allows some freedom of choice. Typically, it is defined in such a way that the artificial head features a flat frequency response either for frontal sound incidence (free-field correction) or in a diffuse sound field (diffuse-field correction). The headphones must be equalized accordingly. It is clear that individual adjustment of the complete system, beyond a specific direction of sound incidence, is impossible in principle, unless the directional characteristics of the artificial head and the listener's head happen to be identical.

Panel (c) depicts the set-up for applications were the signals to the two ears of the listener are to be measured, evaluated and/or manipulated. Signal-processing devices are provided to work on the recorded signals. Although real-time processing is not necessary for many applications, real-time play back is mandatory. The modified and/or unmodified signals can be monitored either by a signal analyser or by binaural listening.

The most complex equipment in this context is represented by panel (d). Here the input signals no longer stem from a listener's ears or from an artificial head, but have been recorded or even generated without the participation of ears or ear replicas. For instance, anechoic recordings via conventional studio microphones may be used. The linear distortions which human ears superimpose on the impinging sound waves, depending on their direction of incidence and wave-front curvature, are generated electronically via a so-called ear-filter bank (electronic head). To be able to assign the adequate head-tranfer function to each incoming signal component, the system needs data of the geometry of the sound field. In a typical application, e.g. architectural-acoustics planning, the system contains a sound-field simulation based on data of the room geometry, the absorption features of the materials implied, and the positions of the sound sources and their directional characteristics. The output of the sound-field modeling is fed into the electronic head, thus producing so-called binaural impulse responses. Subsequent convolution of these impulse responses with anechoic signals generates binaural signals as a subject would observe in a corresponding real room. The complete method is often referred to as binaural room simulation.

To give subjects the impression of being immersed in a sound field, it is important that perceptual room constancy is provided. In other words, when the subjects move their heads around, the perceived auditory world should nevertheless maintain its spatial position. To this end, the simulation system needs to know the head position in order to be able to control the binaural impulse responses adequately. Head position sensors have therefore to be provided. The impression of being immersed is of particular relevance for applications in the context of virtual reality.

All of the applications discussed in this section are based on the provision of two sound-pressure signals to the ear-drums of human beings, or on the use of such signals for measurement and application. They are built on our knowledge of what the ears-and-head array does, i.e., on our understanding of the physics of the binaural transmission chain in front of the eardrum. We shall now proceed to the next section, which deals with the signal processing behind the eardrum and its possible technical applications.

2. THE SUBCORTICAL AUDITORY SYSTEM: PSYCHOPHYSICS OF BINAURAL HEARING

As mentioned above, the subcortical auditory system converts incoming sound waves into neural spike trains which are then processed in a very sophisticated way. Among the things that we know from physiological experiments are the following. The signals are decomposed into spectral bands that are maintained throughout the system. Autocorrelation of the signals from each of the ears, as well as cross-correlation of the signals from both ears, are performed. Specific inhibition and excitation effects are extensively present.

Models of the function of the subcortical auditory system take our knowledge of its physiology into account, but are usually oriented primarily towards the modeling of psychoacoustic findings. Most models have a signal-driven, bottom-up architecture. As an output, a (running) binaural-activity pattern is rendered that displays features corresponding to psychoacoustic evidence and/or allows for the explanation of binaural performance features. Since psychoacoustics, at least in the classical sense, attempts to design listening experiments in a "quasi-objective" way, psychoacoustic observations are, as a rule, predominantly associated with processes in the subcortical auditory system.

There seems to be the following consensus among model builders. A model of the subcortical auditory system must at least incorporate three functional blocks to simulate binaural performance in the areas as listed above (Fig.2).
(i) a simulation of the functions of the external ear, including head (skull), torso, pinnae, ear canal, and eardrum; plus, eventually, the middle ear;
(ii) a simulation of the inner ears, i.e. the cochleae, including receptors and first neurons; plus a set of binaural processors to identify interaurally correlated contents of the signals from the two cochleae and to measure interaural arrival-time and level differences; along with, eventually, additional monaural processors.
(iii) a set of algorithms for final evaluation of the information rendered by the preceding blocks with respect to the specific auditory task to be simulated.

The first block corresponds to the head-and-ears array as discussed in the preceding section, with the exception of the middle ear. As a matter of fact, detailed modeling of the middle ear is deemed unnecessary in current Binaural Technology. The middle ear is approximated by a linear time-invariant bandpass, thus neglecting features such as the middle-ear reflex. Nevertheless, more elaborate models of the middle ear were readily available from literature, if needed, (Hudde, 1983, 1983a, 1994, Blauert, Hudde, and Letens, 1987.)

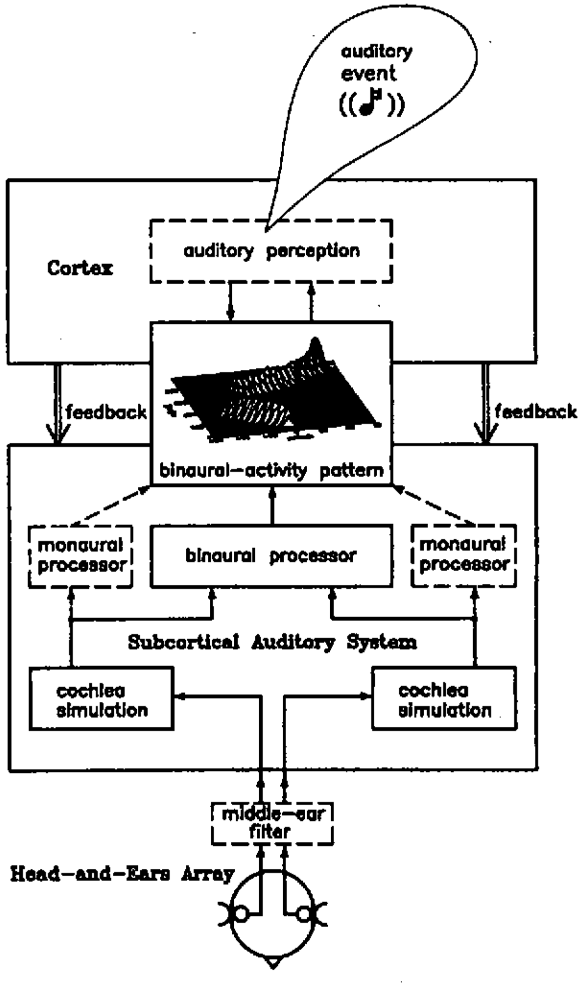

Fig.2: Architecture for an application oriented model of binaural hearing: Binaural signals as delivered by the ear-and-head array (or its electronic simulation) are fed into a model of the subcortical auditory system, implying simulation of the function of the cochleae and of binaural interaction as essential modules. The interface between the subcortical auditory model and the evaluation stages on top of it is provided by a running binaural-activity pattern.

The second block includes two essential modules, cochlea simulation and simulation of subcortical binaural interaction. They will now be discussed in this order. The cochlea model simulates two primary functions, namely, a running spectral analysis of the incoming signals, and a transformation of the (continuous) mechanical vibrations of the basilar membrane into a (discrete) nerve-firing pattern: physiological analog-to-digital conversion. In doing so, it has to be considered that both spectral selectivity and A/D conversion depend on the signal amplitude, i.e., behave nonlinearly. The simplest approximation for the spectral selectivity to be simulated is by means of a bank of adjacent band-pass filters, each, for example, of critical bandwidth. This realization is often used when computing speed is more relevant than precision. More detailed modeling is achieved by including the spectrally-selective excitation at each point of the basilar membrane. The amplitude dependence of excitation and selectivity can optionally be included into the model by simulating active processes, which are supposed to be part of the functioning of the inner ear.

A more precise simulation of the physiological A/D conversion requires a stochastic receptor-neuron model to convert movement of the basilar membrane into neural-spike series. Such models have indeed been implemented for simulations of some delicate binaural effects. However, for practical applications, it is often not feasible to process individual neural impulses. Instead, one can generate deterministic signals that represent the time function of the firing probability of a bundle of nerve fibers. For further simplification, a linear dependence of the firing probability on the receptor potential is often assumed. The receptor potential is sufficiently well described for many applications by the time function of the movement of the basilar membrane, half-wave rectified and fed through of first order low-pass with a 800 Hz cut-off frequency. This accounts for the fact that, among other things, in the frequency region above about 1.5 kHz, binaural interaction works on the envelopes rather than on the fine structure of the incoming signals.

With regard to the binaural processors, the following description results from work performed in the author's lab at Bochum (e.g., Lindemann, 1986, 1986a, Gaik, 1993.) First, a modified, interaural running-cross-correlation function is computed, based on signals originating at corresponding points of the basilar membranes of the two cochlea simulators , i.e., points which represent the same critical frequency. The relevance of cross-correlation to binaural processing has been assumed more than once and is, moreover, physiologically evident. A Bochum modification of cross-correlation consists in the employment of a binaural contralateral inhibition algorithm. Monaural pathways are further included in the binaural processors to allow for the explanation of monaural-hearing effects.

Some details of the binaural processors are given in the following. The first stage of the processor is based on the well known coincidence-detector hypothesis. A way to illustrate this is by assuming two complementary tapped delay lines - one coming from each ear - whose taps are connected to coincidence cells which fire on receiving simultaneous exitation from both side's delay lines. It can be shown that this stage renders a family of running interaural cross-correlation functions as output. Thus we arrive at a three-dimensional pattern (interaural arrival-time difference, critical-band frequency, cross-correlation amplitude) which varies with time and can be regarded as a running binaural-activity pattern. The generation of the running cross-correlation pattern is followed by application of a mechanism of contralateral inhibition based on the following idea. Once a wavefront has entered the binaural system through the two ears, it will consequently give rise to an activity peak in the binaural pattern. Consequently, inhibition will be applied to all other possible positions of activity in each band where excitation has taken place. In each band where signals are received, the first incoming wavefront will thus gain precedence over possible activity being created by later sounds which are spectrally similar to the first incoming wavefront, such as reflections. The actual amount of inhibition is determined by specific weights which vary as a function of position and time, such as to fit psychoacoustical data. Inhibition may, for example, continue for a couple of milliseconds and then gradually die away until it is triggered again. Using this concept as well as specific algorithm of contralateral inhibition, in combination with the inclusion of monaural pathways into the processor, the processing of interaural level differences by the binaural system is properly modeled at the same time. For certain combinations of interaural arrival-time and interaural level differences, e.g. "unnatural" ones, the model will produce multiple peaks in the inhibited binaural activity pattern, thus predicting multiple auditory events - very much in accordance with the psychoacoustical data (Gaik & Wolf, 1988.)

To deal with the problem of natural interaural level differences being much higher at high frequencies than at low ones, the binaural processors must be adapted to the external-ear transfer functions used in the model. To this end, additional inhibitory weighting is implemented on the delay lines of the coincidence networks in such a way that the binaural processors are always excited within their "natural" range of operation. This additional weighting is distributed along the delay lines. The complete set of binaural processors can, thus, be conceptualized as an artificial neural network, more specifically, as a particular kind of time-delay neural network. The adaptation of this network to the particular set of external-ear transfer functions used is accomplished by means of a supervised learning procedure.

The output of the binaural processor, a running binaural-activity pattern, is assumed to be interfacing to higher nervous centers for evaluation. The evaluation procedures must be defined with respect to the actual, specific task required. Within the scope of our current modeling, the evaluation process is thought of in terms of pattern recognition. This concept can be applied when the desired output of the model system is a set of sound-field parameters, such as the number and the positions of the sound source, the amount of auditory spaciousness, reverberance, coloration etc. Also, if the desired output of the model system is processed signals, such as a monophonic signal which has been improved with respect to its S/N ratio, the final evaluative stage may produce a set of parameters for controlling further signal processing.

Pattern-recognition procedures have so far been projected for various tasks in the field of sound localization and spatial hearing, such as lateralization, multiple image phenomena, summing localization, auditory spaciousness, binaural signal enhancement, and parts of the precedence effect (see Blauert and Col, 1992, for cognitive components of the presedence effect.) Further, effects such as binaural pitch, dereverberation and/or decoloration are within the scope of the model.

We shall now consider the question of whether the physiological and psychoacoustic knowledge of the subcortical auditory system, as manifested in models of the kind decribed above, can be applied for Binaural Technology. Since we think of the subcortical auditory system as a specific front-end to the cortex that extracts and enhances certain attributes from the acoustic waves for further evaluation, signal-processing algorithms as observed in the subcortical auditory system may certainly be applied in technical systems to simulate performance features of binaural hearing. Progress in signal-processor technology makes it feasable to implement some of them on microprocessor hardware for real-time operation. Consequently, a number of interesting technical applications have come into reach of today's technology. A first category is concerned with spatial hearing, as decribed below.

Spatial Hearing

Auditory-like algorithms may decode information from the input signals to the ear that allows assessment of the spatial position of sound sources. They may further be used for predictions of how humans form the positions and spatial extents of their auditory events, how they establish an auditory perspective, and how they suppress echoes and reverberance. Typical applications are: source-position finders, tools for the evaluation of architectural acoustics and sound systems (such as spaciousness meters, echo detectors, and precedence indicators,) tools for the evaluation of auditory virtual environments and for psychoacoustic research. There are further perceptual features of auditory events, besides position and spatial extent, which are based on binaural rather than monaural information. Following a usage in the field of product-sound design, they may be called binaural psychoacoustic descriptors, as discussed in the next section.

Binaural Psychoacoustic Descriptors

Binaural psychoacoustic descriptors include binaural loudness, binaural pitch, binaural timbre and binaural sensory-consonance. Algorithms taken from binaural auditory models may be used to generate estimates of these descriptors. There is an increasing demand for such tools, e.g., in the area of sound-quality evaluation. The most tempting field of application for binaural auditory models, however, concerns the ability of binaural hearing to process signals from different sources selectively, and to enhance one of them with regard to the others. A key word for this area of application could be binaural signal enhancement.

Binaural Signal Enhancement

A well-known term in the context of binaural signal enhancement is the so-called "cocktail-party effect," denoting that, with the aid of binaural hearing, humans can concentrate on one talker in the presence of competing ones. It has further been established that with binaural hearing a desired signal and noise can be separated more effectively than with monaural hearing. Binaural auditory models may help to simulate these capabilities by providing front-ends that allow for better separation of a mix of sound sources. In a specific Bochum version of a so-called "cocktail-party" processor, i.e., a processor to enhance speech in a "cocktail-party" situation, the binaural processor of the auditory model is used to control a Wiener filter (Bodden & Blauert, 1992, Bodden, 1993).) This is accomplished by first identifying the position of a desired talker in space, and then estimating its S/N ratio with respect to competing talkers and other noise signals. The system performs its computation within critical-bands. In the case of two competing talkers, the desired signal can be recovered to reasonable intelligibility, even when its level is 15 dB lower than that of the competing one. Application possibilities for this kind of systems are numerous, such as tools for editing binaural recordings, front ends for signal-processing hearing aids, speech-recognizers and hands-free telephones. In general, binaural signal enhancement may be used to build better "microphones" for acoustically adverse conditions. As stated above, the cues provided by models of the subcortical auditory system, and contained in binaural-activity patterns, must consequently be evaluated in adequate ways. The next section deals with this problem.

3. THE CORTEX: PSYCHOLOGY OF BINAURAL HEARING

Most models of the subcortical auditory system assume a bottom-up, signal-driven process up to their output, the running binaural-activity pattern. The cortex, consequently, takes this pattern as an input. The evaluation of the binaural-activity pattern can be conceived as a top-down, hypothesis-driven process. According to this line of thinking, cortical centers set up hypotheses, e.g., in terms of expected patterns, and then try to confirm these hypotheses with appropriate means, e.g., with task-specific pattern-recognition procedures. When setting up hypotheses, the cortex reflects on cognition, namely, on knowledge and awareness of the current situation and the world in general. Further, it takes into account input from other senses, such as visual or tactile information. After forming hypotheses, higher nervous stages may feed back to more peripheral modules to prompt and control optimum hypothesis testing. They may, for example, induce movement of the ears-and-head array or influence the spectral decomposition process in the subcortical auditory system.

The following two examples help to illustrate the structure of problems that arise at this point from a technological point of view. First, in a "cocktail-party" situation a human listener can follow one talker and then, immediately, switch his attention to another. A signal-processing hearing aid should be able to do the same thing, deliberately controlled by its user. Second, a measuring instrument to evaluate the acoustic quality of concert halls will certainly take into account psychoacoustic descriptors like auditory spaciousness, reverberance, auditory transparency, etc. However, the general impression of space and quality that a listener develops in a room, may be codetermined by visual cues, by the specific kind of performance, by the listener's attitude, and by factors like fashion or taste, among other things.

There is no doubt that the involvement of the cortex in the evaluation process adds a considerable amount of "subjectivity" to binaural hearing, which poses serious problems to Binaural Technology. Engineers, as most scientists, are trained to deal with the object as being independent of the observer (assumption of "objectivity") and prefer to neglect phenomena that cannot be measured or assessed in a strictly "objective" way. They further tend to believe that any problem can be understood by splitting it up into parts, and analyzing these parts separately. At the cortical level, however, we deal with percepts, i.e. objects that do not exist as separate enteties, but as part of a subject-object (perceiver-percept) relationship. It should also be noted that listeners normally listen in a "gestalt" mode, i.e., they perceive globally rather than segmentally. An analysis of the common engineering type may thus completely miss relevant features.

Perceiver and percept interact and may both vary considerably during the process of perception. For example, the auditory events may change when listeners focus on specific components such as the sound of a particular instrument in an orchestra. Further, the attitude of perceivers towards their percepts many vary in the course of an experimental series, thus leading to response modification.

A simple psychological model of the auditory perception and judgment process, shown in Fig.3, will now be used to elaborate on the variance of listeners' auditory events in a given acoustic setting and the variance of their respective responses. The schematic symbolizes a subject in a listening experiment. Sound waves impinge upon the two ears, are preprocessed and guided to higher centers of the central nervous system, where they give rise to the formation of an auditory event in the subject's perceptual space. The auditory event is a percept of the listener being tested, i.e., only he/she has direct access to it. The rest of the world is only informed about the occurence of the said percept, if the subject responds in such a way as to allow conclusion to be made from the response to the percept (indirect access). In formal experiments the subject will usually be instructed to respond in a specified way, for example by formal judgement on specific attributes of the auditory event. If the response is a quantitative desriptor of perceptual attributes, we may speak of measurement. Consequently, in listening experiments, subjects can serve as an instrument for the measurements of their own perception, i.e., as both the object of measurement and the "meter". The schematic in Fig.3 features a second input into both the auditory-perception and the judgement blocks where "response-moderating factors" are fed in to introduce variance to the perception and judgement processes.

Following this line of thinking an important task of auditory psychology can be to identify such response-moderating factors and to clarify their role in binaural listening. Many of these factors represent conventional knowledge or experience from related fields of perceptual acoustics, e.g., noise- and speech-quality evaluation. It is well known that the judgements of listeners on auditory events may depend on the cognitive "image" which the listeners have with respect to the sound sources involved (source-related factors). It may happen, for instance, that the auditory events evoked by sources that are considered agressive (e.g., trucks), are judged louder than those from other sources (e.g., passenger cars) - given the same acoustical signals. The "image" of the source in the listeners' minds may be based, among other things, on cues from other senses (e.g., visual) and/or on prior knowledge. Situative factors are a further determinant in this context, i.e., subjects judge an auditory event bearing the complete (multi-modal) situation in mind in which they occur. Another set of factors is given by the individual characteristics of each listener (personal factors), for example his/her subjective attitude towards a specific sound phenomenon, an attitude that may even change in the course of an experiment. Response-moderating factors that draw upon cognition tend to be especially effective when the sounds listened to transmit specific information, i.e., act as carriers of meaning. This is obvious in the case of speech sounds, but also in other cases. The sound of a running automobile engine, for instance, may signal to the driver that the engine is operating normally.

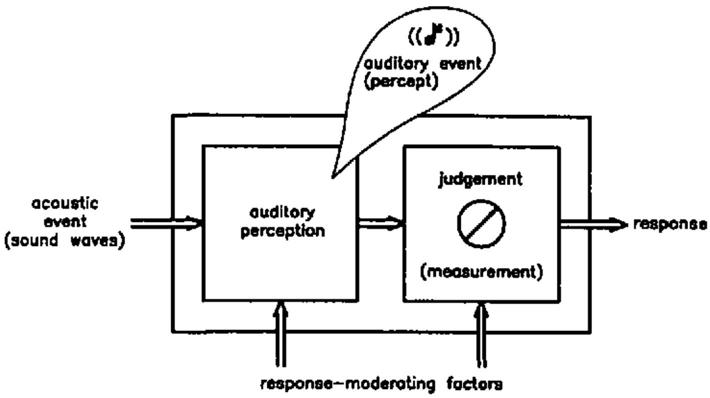

Fig.3: Schematic of a subject in a listening experiment. Perception as well as judgement are variant, as modeled by the assumption of response-moderating factors.

The fact that response moderating factors do not only act on judgements but also on the process of perception itself, may seem to be less obvious at a first glance, but is, nevertheless, also conventional wisdom. We all know that people in a complex sound situation have a tendency to miss what they do not pay attention to and/or do not expect to hear. There is psychoacoustical evidence that, e.g., the spectral selectivity of the cochlea is influenced by attention. At this point, the ability to switch at will between a global and an analytic mode of listening, should also be noted. It is commonly accepted amongst psychologists that percepts are the result of both the actual sensory input at a given time and of expectation.

If we want to built sophisticated Binaural-Technology equipment for complex tasks, there is no doubt that psychological effects have to be taken into account. Let us consider, as an example, a binaural-surveillance system for acoustic monitoring of a factory floor. Such a system must know the relevance and meaning of many classes of signals and must pay selective attention to very specific ones, when an abnormal situation has been detected. A system for the evaluation of acoustic qualities of spaces for musical performances must detect and consider a range of different shades of binaural signals, depending on the kind and purpose of the performances. It might even have to take into account the taste of the local audience or that of the most influential local music reviewer. An intelligent binaural hearing aid should know to a certain extent, which components of the incoming acoustic signals are relevant to its user, e.g., track a talker who has just uttered the users name.

As a consequence, we shall see in the future of Binaural Technology that psychological models will be exploited and implemented technologically, though, may be not, for a while, in the form of massively parallel biologic computing as in the cortex. There are already discussions about and early examples of combinations of expert systems and other knowledge-based systems with artificial heads, auditory displays and auditory-system models. When we think of applications like complex human/machine interfaces, multi-media systems, interactive virtual environments, and teleoperation systems, it becomes obvious that conventional Binaural Technology must be combined with, or integrated into, systems that are able to make decisions and control actions in an intelligent way. With this view in mind it is clear that Binaural Technology is still in an early stage of development. There are many relevant technological challenges and business opportunities ahead.

4. ACKNOWLEDGEMENT

Many of the ideas discussed in this chapter have evolved from work at the author's lab. The author is especially indebted to his doctoral students over the years who have helped to provide a constant atmosphere of stimulating discussion. A list of completed doctoral dissertations which are correlated to this chapter, along with references to tagged publications in English, is given in the following.

Bodden, Markus: Binaurale Signalverarbeitung: Modellierung der Richtungserkennung und des Cocktail-Party-Effektes (Binaural signal processing (Modeling of direction finding and of the cocktail-party effect), Bochum 1992, (Bodden & Blauert, 1992, Bodden, 1993, Blauert, Bodden, and Lehnert, 1993)

Col, Jean-Pierre: Localisation auditiv d'un signal et aspects temporels de l'audition spatiale (Auditory localization of a signal and temporal aspects of spatial hearing), Marseille 1990, (Blauert & Col, 1992)

Els, Hartmut, Ein Meßsystem fuer die akustische Modelltechnik (A measuring system for the technique of acoustic modeling), Bochum 1986, (Els & Blauert, 1985, 1986)

Gaik, Werner: Untersuchungen zur binauralen Verarbeitung kopfbezogener Signale (Investigations into binaural signal processing of head-related signals), Bochum 1990, (Gaik & Wolf, 1988, Gaik, 1993)

Hudde, Herbert, Messung der Trommelfellimpedanz des menschlichen Ohres bis 19kHz (Measurement of eardum impedance up to 19 kHz), Bochum 1980, (Hudde 1983, 1983a)

Lehnert, Hilmar: Binaurale Raumsimulation: Ein Computermodell zur Erzeugung virtueller auditiver Umgebungen (Binaural room simulation: A computer model for generation of virtual auditory environments), Bochum 1992, (Lehnert & Blauert, 1989, 1992, Blauert, Bodden, and Lehnert, 1993)

Letens, Uwe: Ueber die Interpretation von Impedanzmessungen im Gehoergang anhand von Mittelohr-Modellen (Interpretation of impedance measurements in the ear canal in terms of middle-ear models), Bochum 1988, (Blauert, Hudde, and Letens, 1987)

Lindemann, Werner, Die Erweiterung eines Kreuzkorrelationsmodells der binauralen Signalverarbeitung durch kontralaterale Inhibitionsmechanismen (Extension of a cross-correlation model of binaural signal processing by means of contralateral inhibition mechanisms), Bochum 1985, (Lindemann, 1986, 1986a)

Pompetzki, Wulf: Psychoakustische Verifikation von Computermodellen zur binauralen Raumsimulation (Psychoacoustical verification of computer-models for binaural room simulation), Bochum 1993

Poesselt, Christoph: Einfluss von Knochenschall auf die Schalldaemmung von Gehoerschuetzern (Influence of bone conduction on the attenuation of personal hearing protectors), Bochum 1986, (Schroeter & Poesselt, 1986)

Schlichthaerle, Dietrich, Modelle des Hoerens - mit Anwendungen auf die Hoerbarkeit von Laufzeitverzerrungen (Models of hearing - applied to the audibility of arrival-time distortions), Bochum 1980

Schroeter, Juergen, Messung der Schalldaemmung von Gehoerschuetzern mit einem physikalischen Verfahren- Kunstkopfmethode (Measurement of the attenuation of personal hearing protectors by means of a physical technique - dummy-head method), Bochum 1983, (Blauert, Els, and Schroeter, 1980, Schroeter, 1986)

Slatky, Harald: Algorithmen zur richtungsselektiven Verarbeitung von Schallsignalen eines binauralen Cocktail-Party-Prozessors (Algorithms for direction-selective processing of sound signals by means of a binaural cocktail-party processor), Bochum 1993

Wolf, Siegbert: Lokalisation von Schallquellen in geschlossenen Raeumen (Localisation of sound sources in enclosed spaces), Bochum 1991

Xiang, Ning: Mobile Universal Measuring System for the Binaural Room-Acoustic-Model Technique, Bochum 1991, (Xiang & Blauert, 1991, 1993)

REFERENCES

Blauert, J. & Col, J.-P. (1992), "A Study of Temporal Effects in Spatial Hearing," in: Cazals, Y. et. al. (eds), Auditory Psychology & Perception, Pergamon Press, 531-538, GB-Oxford, 1992

Blauert, J. & Genuit, K. (1993), "Sound-Environment Evaluation by Binaural Technology: Some Basic Consideration," J. Acoust. Soc. Japan 14, 139-145

Blauert, J. (1983), "Psychoacoustic Binaural Phenomena," in: Klinke, R., Hartmann, R. (eds.), Hearing - Psychological Bases and Psychophysics, Springer-Verlag, D-Berlin, USA-New York NY

Blauert, J., Bodden, M. & Lehnert, H. (1993), "Binaural Signal Processing and Room Acoustics," IEICE Transact. Fundamentals E75 (Japan), 1454-1458

Blauert, J., Els, H., Schroeter, J. (1980), "A Review of the Progress in External Ear Physics Regarding the Objective Performance Evaluation of Personal Ear Protectors," Proc. Inter-Noise'80, Vol.II, 653-658, Noise-Control Found., USA-New York NY

Blauert, J., Hudde, H. & Letens, U. (1987), "Eardrum-Impedance and Middle-Ear Modeling," in: J.L. Bento Coheljo (ed.), Proc. Symp. Fed. Europ. Acoust. Soc., FASE, 125-128, P-Lisboa

FOUNDATIONS & APPLICATION OPPORTUNITIES OF BINAURAL TECHNOLOGY

Bodden, M. & Blauert, J. (1992), "Separation of Concurrent Speech Signals: A Cocktail-Party Processor for Speech Enhancement," Proc. ESCA Workshop on: Speech Processing in Adverse Conditions, 147-150, F-Cannes/Mandelieu, ISSN 1018-4554

Bodden, M. (1993), "Modeling Human Sound Source Localization and the Cocktail-Party-Effect," Acta Acustica 1, 43-55.

Els, H., Blauert, J. (1985), "Measuring Techniques for Acoustic Models - Upgraded," Proc. Internoise'85, Schriftenr. Bundesanst. Arbeitsschutz, Vol. Ib 39/II, 1359-1362, D-Dortmund

Els, H., Blauert, J. (1986), "A Measuring System for Acoustic Scale Models, 12th Int. Congr. Acoust.," Proc. Vancouver Symp. Acoustics & Theatre Planning for the Performing Arts, 65-70, CND-Toronto

Gaik, W. & Wolf. S. (1988), "Multiple Images: Psychological Data and Model Predictions," in: Basic Issues of Hearing, H. Duifhius, J.W. Horst & H.P. Wit (eds.), Academic Press, GB-London, 386-393

Gaik, W. (1993), "Combined Evaluation of Interaural Time and Intensity Differences: Psychoacoustic Results and Computer Modeling," J. Acoust. Soc. Am., 94, 98-110.

Hudde, H. (1983), "Measurement of Eardrum Impedance of Human Ears," J. Acoust. Soc. Am. 73, 242-247

Hudde, H. (1983a), "Estimation of the Area Function of Human Ear Canals by Sound-Pressure Measurements," J. Acoust. Soc. Am. 73, 24-31

Hudde, H. (1994), "Measurement-Related Modeling of Normal and Reconstructed Middle Ears," Acta Acustica, submitted

Lehnert, H. & Blauert, J. (1992)," Principles of Binaural Room Simulation," J. Appl. Acoust. 36," 259-291

Lehnert, H., Blauert, J. (1989) , "A Concept for Binaural Room Simulation," Proc. IEEE-ASSP Workshop on Application of Signal Processing to Audio & Acoustics, USA-New Paltz NY

Lindemann, W. (1986), "Extension of a Binaural Cross-Correlation Model by Means of Contralateral Inhibition. I. Simulation of Lateralization of Stationary Signals," J. Acoust. Soc. Am. 80, 1608-1622.

Lindemann, W. (1986a), "Extension of a Binaural Cross-Correlation Model by Means of Contralateral Inhibition. II. The Law of the First Wave Front," J. Acoust. Soc. Am. 80, 1623-1630.

Poesselt, Ch., Schroeter, J., Opitz, M., Divenyi, P.L. (1986), "Generation of Binaural Signals for Research and Home Entertainment," Proc. 12th Int. Congr. Acoust. Vol. I, B1-6, CND-Toronto

Schroeter, J. & Poesselt, Ch., "The Use of Acoustical Test Fixures for the Measurement of Hearing-Protector Attenuation, Part II: Modeling the External Ear, Simulating Bone Conduction, and Comparing Test Fixure and Real-Ear Data," J. Acoust. Soc. Am. 80, 505-527

Schroeter, J. (1986), "The Use of Acoustical Test Fixures for the Measurement of Hearing-Protector Attenuation, Part I: Review of Previous Work and the Design of an Improved Test Fixure," J. Acoust. Soc. Am. 79, 1065-1081

Xiang, N., & Blauert, J. (1991), "A Miniature Dummy Head for Binaural Evaluation of Tenth-Scale Acoustic Models," J. Appl. Acoust. 33, 123-140

Xiang, N. & Blauert, J. (1993), "Binaural Scale Modelling for Auralization and Prediction of Acoustics in Auditoria," J. Appl. Acoust. 38, 267-290