

La technologie binaurale : bases scientifiques et domaines d'application génériques

J. BLAUERT

Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract:

Binaural Technology: Scientific bases and generic application areas

- The head-and-ears array: Physics of binaural technology.
Application areas: Binaural recording and authentic reproduction, binaural measurement and evaluation, binaural simulations and displays.
- The auditory system: Psychophysics of binaural hearing.
Application areas: spatial hearing, binaural psychophysical descriptors, binaural signal enhancement.
- The central nervous system: Psychology of binaural hearing.
Application areas: knowledge-based systems, multimodal systems.

1. Introduction

L'homme, comme la plupart des vertébrés, a deux oreilles positionnées à (environ) égale hauteur des deux côtés de la tête. Physiquement, les deux oreilles et la tête forment une structure d'antenne, montée sur une base mobile. L'antenne reçoit des ondes mécano-élastiques ("acoustiques") du milieu dans lequel il est immergé - normalement l'air. Les deux ondes reçues et enregistrées par les deux oreilles sont l'entrée physiologique adéquate d'un système sensoriel spécifique, le système auditif.

Le système auditif transforme chacune des deux ondes en trains d'impulsions nerveuses - après avoir effectué une décomposition spectrale courante en multiple canaux fréquentiels - parmi d'autres prétraitements.

Les trains d'impulsions nerveuses multi-canaux de chaque oreille sont ensuite combinés d'une manière sophistiquée pour générer finalement un "motif d'activité binaurale" quelque part dans le système auditif. Ce motif d'activité binaurale - très probablement en combinaison avec des motifs d'activité monaurale rendus individuellement par les canaux auditifs de chaque oreille - forme l'entrée auditive en direction du système nerveux central - un puissant ordinateur biologique multi-tâches en parallèle avec une mémoire énorme, diverses interfaces et ports d'entrée-sortie.

En sortie, le système nerveux central délivre un monde perceptif individuel et, éventuellement, des commandes nerveuses pour déclencher et contrôler des expressions motrices spécifiques.

Il va sans dire qu'un certain nombre de contraintes doit être pris en compte pour que ce récit soit valable. Par exemple, les ondes acoustiques doivent être dans le domaine d'audibilité quant au domaine fréquentiel et à l'intensité, le système auditif doit être opératif, et le système nerveux central doit être dans un mode conscient, prêt à accepter et interpréter l'information auditive.

En outre, il paraît censé de supposer que de multiples rétroactions sont impliquées dans le processus d'enregistrement, de traitement et d'interprétation des signaux acoustiques: entre les modules du système auditif ainsi qu'entre le système auditif et le système nerveux central. Des rétroactions évidentes du système nerveux central vers le système de positionnement moteur de la structure oreilles-tête peuvent aussi être observées chaque fois que le système nerveux central produit des mouvements de la tête de recherche de position.

Bien que l'homme puisse entendre avec une oreille seulement - soi-disant écoute monaurale - l'écoute avec deux oreilles en ordre de fonctionnement est nettement supérieure. Ce fait peut être mieux compris en considérant le rôle biologique de l'écoute, à savoir:

C'est la fonction biologique de l'écoute d'obtenir des informations sur l'environnement, particulièrement sur les positions spatiales et les trajectoires de sources sonores et leur état d'activité. De plus, pour plusieurs espèces (et quelquefois aussi entre espèces), la communication interindividuelle est effectuée éminemment de manière acoustique.

En conséquence, l'écoute est indispensable dans cette chaîne acoustique de communication, où le système nerveux central déchiffre la signification comme encodée dans des signaux acoustiques par d'autres systèmes nerveux centraux.

L'écoute binaurale fonctionne significativement mieux que la monaurale pour ce qui concerne les tâches suivantes:

- localisation de sources simple ou multiple et, en conséquence, formation d'une perspective auditive et/ou d'une impression de salle auditive;
- séparation de signaux venant de multiples sources sonores incohérentes répandues spatialement ou, sous certaines restrictions, de sources cohérentes;
- renforcement de signaux d'une source choisie parmi un mélange de signaux de sources incohérentes, ainsi que renforcement de signaux directs (non réfléchis) de sources dans un environnement réverbérant.

Evidemment les caractéristiques performantes de l'écoute binaurale forment un défi pour les ingénieurs en terme d'application technologique. Dans ce contexte, une soi-disant "Technologie Binaurale" s'est développée depuis plus de trois décennies. Une définition opérationnelle de cette Technologie Binaurale peut être lue comme suit:

La Technologie Binaurale est un corps de méthodes pour l'accomplissement de buts pratiques - ayant à faire avec les signaux d'entrée acoustiques aux deux oreilles d'un auditeur, par exemple, pour l'enregistrement, l'analyse, la synthèse, le traitement, la présentation et l'évaluation de tels signaux.

La Technologie Binaurale a récemment gagné en force vive économique, à la fois pour elle-même ainsi qu'en tant que technologie permettant des applications plus complexes. Une industrie spécialisée pour la Technologie Binaurale se développe rapidement.

Ce chapitre a pour objectif de donner un aperçu sur ce processus sortant et de rapporter brièvement les bases sur lesquelles cette technologie repose - i.e. sur ses fondements expérimentaux et théoriques.

Comme mentionné plus haut, trois "modules" sont engagés fondamentalement dans la réception, perception et interprétation de signaux acoustiques par les humains, la structure oreilles-tête, le système auditif et le système nerveux central. La Technologie Binaurale fait usage des connaissances des principes de fonctionnements de tout ceux-ci.

Dans les trois sections suivantes, les fonctions spécifiques de chacun des trois modules sont passées en revue dans la lumière de leur application pour la Technologie Binaurale.

2. La structure oreilles-tête: Physique de l'écoute binaurale

L'ensemble oreilles-tête a une structure d'antenne avec des caractéristiques de transmission très complexes et spécifiques. Puisqu'il s'agit d'une structure physique

et que la propagation du son est un processus linéaire, l'ensemble peut être considéré comme un système linéaire. En prenant l'onde incidente sonore comme entrée et les signaux de pression sonore devant les deux tympans comme sortie, il est correct de décrire le système comme un ensemble de deux filtres auto-ajustés connectés à la même entrée, dont les fonctions de transfert dépendent de l'orientation géométrique de l'onde sonore par rapport à la structure oreilles-tête.

Physiquement, ce comportement s'explique par des résonances dans le système oreille externe - conduit auditif externe - tympan. Les résonances sont excitées différemment lorsque l'onde sonore arrive de différentes directions et/ou avec différentes courbures du front d'onde. Les fonctions de transfert résultantes sont généralement différentes pour les deux filtres, causant ainsi des différences "interaurales" des signaux de pression sonore aux deux tympans.

Puisque les distorsions linéaires qui sont superposées sur l'onde sonore par les deux "filtres auditifs" sont très spécifiques par rapport aux paramètres géométriques de l'onde sonore, il est vraisemblable d'affirmer que le système oreilles-tête code l'information sur les positions des sources sonores dans l'espace, relativement à ce système, en attributs temporel et spectral des signaux aux deux tympans et en leurs différences interaurales.

Toutes les manipulations que le système oreilles-tête applique aux signaux sonores sont purement physiques et linéaires. Il est alors évident qu'elles puissent être simulées. En effet, il existe une importante branche de la Technologie Binaurale qui tente justement de le faire.

Il paraît censé à ce point de débiter la discussion sur la technologie par la plus proche et cependant très importante application de la Technologie Binaurale, à savoir, la reproduction auditive authentique.

La reproduction auditive authentique est achevée lorsqu'un auditeur entend exactement la même chose dans une situation de reproduction que ce qu'il/elle entendrait dans un champ sonore original - ce dernier existant à une autre période et/ou un autre lieu.

Comme hypothèse de travail, la Technologie Binaurale a l'habitude de débiter avec la supposition qu'un auditeur entend la même chose dans une situation de reproduction lorsque les signaux aux deux tympans pendant la reproduction sont exactement les mêmes que ceux qui existeraient dans le champ sonore original.

En terme de Technologie Binaurale ce but est achevé au moyen de soi-disant Têtes Artificielles qui sont des répliques de têtes naturelles en termes acoustiques, i.e. elles réalisent deux filtres auto-ajustés comme des têtes naturelles.

Les Têtes Artificielles, en combinaison avec des équipements de reproduction ajustés de manière adéquate, sont l'instrumentation de base pour un certain nombre d'applications économiquement séduisantes. L'équipement de reproduction, qui est additionnellement nécessaire, se compose usuellement d'écouteurs. Cependant, sous certaines conditions spécifiques, restrictives, des haut-parleurs peuvent aussi être utilisés. Dans ce contexte, une première catégorie d'applications peut être résumée sous les rubriques suivantes.

Enregistrement binaural et reproduction authentique

Ces applications exploitent la possibilité de la Technologie Binaurale d'archiver un champ sonore d'une manière perceptivement authentique - et de le rendre accessible à l'écoute à volonté - par exemple pour le divertissement, l'éducation, l'instruction, la recherche scientifique, la documentation, la surveillance et le télémonitoring.

Il est important de noter que les enregistrements binauraux peuvent être comparés en séquence consécutive (par exemple comparaison A/B), ce qui est souvent impossible au cours de situations sonores originales.

Puisque les signaux de pression sonore aux deux tympans sont l'entrée physiologique adéquate du système auditif, ils sont souvent considérés comme la base de mesure et d'évaluation auditivement adéquate dans un sens physique et/ou auditif.

Par conséquent, nous avons une catégorie d'applications de plus, à savoir:

Mesure et évaluation binaurale

Dans le cas physique, des procédures basées sur la physique sont utilisées, dans le cas auditif, des auditeurs humains servent comme instruments de mesure et d'évaluation.

Des applications courantes couvrent des domaines comme le contrôle du bruit, la conception d'environnement acoustique, l'évaluation de la qualité du son (par ex. pour la technologie de la parole, l'architecture acoustique et la conception de produit sonore), les mesures spécifiques sur des systèmes téléphoniques, sur des écouteurs, sur des protections auditives personnelles et sur les prothèses auditives.

Puisque les têtes artificielles, fondamentalement, sont juste un moyen spécifique de réalisation d'un ensemble de filtres linéaires, on peut penser à d'autres manières de réaliser de tels filtres, par ex. électroniquement. Pour plusieurs applications cela

ajoute des degrés de liberté supplémentaires car les filtres électroniques peuvent être contrôlés à volonté dans un large domaine.

Cette idée conduit à une catégorie d'applications de plus comme suit:

Simulation et présentations binaurales

Il existe plusieurs applications courantes dans cette catégorie, avec le potentiel d'un nombre toujours croissant. La liste suivante en fournit des exemples: mixage binaural, simulation de salle binaurale, effets sonores avancés (par ex. pour les jeux électroniques), délivrance d'indications d'orientation spatiale auditive (par ex. dans un cockpit, pour l'aveugle), présentation auditive de données complexes, représentation auditive dans les systèmes de téléconférence, téléprésence et téléopérations.

Toutes les applications dont nous venons de discuter jusqu'à présent sont basées sur la délivrance des deux signaux de pression sonore aux tympans d'être humains, ou sur l'utilisation de tels signaux pour des mesures ou des applications. Elles sont basées sur notre connaissance du fonctionnement du système oreilles-tête, i.e. sur la physique de l'écoute binaurale.

Nous allons maintenant poursuivre avec la section suivante, dans laquelle il s'agit du traitement du signal derrière le tympan et de ses applications techniques potentielles.

3. Le système auditif: Psychophysique de l'écoute binaurale

Comme déjà mentionné plus haut, le système auditif convertit les ondes sonores incidentes en trains d'impulsions nerveuses qui sont ensuite traitées d'une manière sophistiquée. Parmi d'autres choses que nous savons des expérimentations physiologiques voici les suivantes:

Les signaux sont décomposés en bandes spectrales qui sont maintenues dans tout le système. Autocorrélation des signaux de chacune des oreilles aussi bien qu'intercorrélation des signaux des deux oreilles sont accomplies. Des inhibitions spécifiques et des effets d'excitation sont intensément présents.

Des modèles du fonctionnement du système auditif prennent en compte notre connaissance de sa physiologie, mais sont usuellement principalement orientés sur la modélisation des résultats psychoacoustiques. La plupart des modèles ont une architecture conduite en signal (bottom-up). En sortie sont fournis des motifs courants d'activité nerveuse qui présentent des caractéristiques

correspondant à des évidences psychoacoustiques et/ou permettant l'explication des caractéristiques du fonctionnement binaural.

Puisque la psychoacoustique - du moins dans le sens classique - tente de créer des expériences d'écoute d'une façon "quasi-objective", les observations psychoacoustiques peuvent être - en règle générale - principalement associées avec des traitements du système auditif et non du système nerveux central.

Nous allons maintenant considéré la question de savoir si nos connaissances physiologiques et psychoacoustiques telles que manifestées dans des modèles du système auditif binaural peuvent être appliquées en terme de Technologie Binaurale. Les applications de cette sorte auraient à faire usage des motifs d'activité binaurale.

Il existe en effet une variété de possibilités d'application suivant cette ligne. On peut penser au système auditif - dans le sens utilisé dans cet article - comme une interface spécifique au système nerveux central qui extrait et réhausse certains attributs des ondes acoustiques pour une évaluation additionnelle.

Des algorithmes de traitement du signal tels qu'observés dans le système auditif peuvent être aussi bien utilisés dans des systèmes techniques pour simuler les caractéristiques de fonctionnement de l'écoute binaurale. Une première catégorie d'applications de cette sorte concerne

L'écoute spatiale

Des algorithmes analogues à ceux du système auditif peuvent décoder l'information des signaux d'entrée d'oreille, ce qui permet d'évaluer la position spatiale de sources sonores. Ils peuvent en outre être utilisés pour la prédiction de la manière dont les humains forment les positions et étendues spatiales de leurs événements auditifs, dont ils établissent une perspective auditive, et dont ils suppriment échos et réverbération, par exemple, le soi-disant effet de préséance (precedence effect).

Des domaines d'applications typiques sont: des chercheurs de position de source, des outils pour l'évaluation de l'acoustique architecturale et des systèmes sonores (par exemple, appareils de mesures de l'effet d'espace, détecteurs d'échos), des outils pour l'évaluation des environnements virtuels auditifs et pour la recherche psychoacoustique.

Il existe des caractéristiques perceptuelles supplémentaires des événements auditifs - aux côtés de la position et l'étendue spatiale - qui sont basées sur une information binaurale plutôt que monaurale. Suivant l'usage dans le domaine de la création de produit sonore, ils peuvent être nommés

Critères psychoacoustiques binauraux

Parmi ceux-ci sont le loudness binaural, la hauteur binaurale, le timbre binaural et la "bien-sonnée" (Wohlklang) binaurale.

Des algorithmes tirés des modèles auditifs binauraux peuvent être utilisés pour générer des estimations de ces critères.

Il existe une demande croissante de tels outils, comme par exemple dans le domaine de l'évaluation de la qualité sonore.

Cependant, le champ d'application le plus tentant pour les modèles auditifs binauraux concerne la capacité de l'écoute binaurale de traiter selectivement des signaux provenant de différentes sources, et de réhausser l'un parmi les autres. Un titre pour cette catégorie pourrait être:

Réenforcement du signal binaural

Un terme bien connu dans ce contexte est le soi-disant "effet de cocktail-party", dénotant qu'avec l'aide de l'écoute binaurale, les humains peuvent se concentrer sur un orateur en présence d'autres orateurs interférants.

Il a en outre été établi qu'avec l'écoute binaurale un signal et bruit donné peuvent être mieux séparés qu'avec l'écoute monaurale.

Des modèles auditifs binauraux peuvent aider à simuler ces capacités en fournissant des interfaces permettant une meilleure séparation d'un mélange de sources sonores.

Les possibilités d'applications sont nombreuses, telles que des outils pour l'édition d'enregistrement binauraux, des interfaces pour des prothèses auditives avec traitement de signal, des appareils de reconnaissances de la parole et les téléphones avec haut-parleur et microphone intégrés. En général, ils peuvent être utilisés pour construire de meilleurs "microphones" pour des conditions acoustiques défavorables.

Les indications fournies par des modèles du système auditif - contenues dans les motifs d'activité binaurale - doivent être évaluées conséquemment. La section suivante traite de ce problème.

4. Le système nerveux central: Psychologie de l'écoute binaurale

Jusqu'à la sortie du système auditif, des modèles supposent généralement un traitement conduit par signal (bottom-up) se terminant par des motifs d'activité binaurale. Le système nerveux central utilise ces motifs comme entrée.

L'évaluation est usuellement considérée comme un traitement conduit par hypothèse (top-down). Le système nerveux central, en suivant cette ligne de pensée, pose des hypothèses en terme de motifs attendus et essaie de confirmer ces hypothèses.

Au cours de la pose de ces hypothèses, le système nerveux central réfléchit cognitivement, i.e. avec sa connaissance et conscience de la situation courante et du monde en général. De plus, il prend en compte les entrées d'autres sens, comme par exemple l'information visuelle ou tactile.

Après avoir posé une certaine hypothèse, le système nerveux central peut renvoyer aux modules plus périphériques pour inciter et contrôler des tests d'hypothèses optimaux. Il peut, par exemple, induire des mouvements de la structure oreilles-tête ou influencer le traitement de décomposition spectrale dans le système auditif.

Deux exemples peuvent aider à illustrer la structure des problèmes qui surviennent ici du point de vue technologique.

- Dans une "situation de cocktail-party", un auditeur humain peut suivre un orateur et ensuite, immédiatement, dériver son attention sur un autre. Une prothèse auditive avec traitement de signal devrait être capable d'effectuer la même chose, contrôlée délibérément par son utilisateur.

Un instrument de mesure pour évaluer la qualité acoustique d'une salle prendra certainement en compte des indices psychoacoustiques tels que l'effet d'espace auditif, la réverbérance, la transparence auditive etc.. Cependant, l'impression générale d'espace et de qualité qu'un auditeur développe dans une salle, peut être cogérée par des indications visuelles, par la spécificité de la représentation, par l'attitude des auditeurs et par des facteurs tels que la mode ou le goût, parmi d'autres choses.

Il ne fait aucun doute que l'engagement du système nerveux central dans le traitement d'évaluation ajoute une considérable quantité de "subjectivité" à l'écoute binaurale, ce qui pose de sérieux problèmes à la Technologie Binaurale.

Les ingénieurs en particulier, comme la plupart des scientifiques, sont entraînés à traiter avec des objets comme étant indépendants de l'observateur (hypothèse de "l'objectivité") et préfèrent négliger les phénomènes qui ne peuvent être mesurés ou évalués d'une manière strictement "objective".

Il ont de plus tendance à croire que tout problème peut être compris en le séparant en plusieurs parts et en analysant chacune de ces parts séparément.

Au niveau du système nerveux central, cependant, on traite avec des "percepts", i.e. des objets résultants de la perception qui n'existent pas en tant qu'entités séparées, mais en tant que part d'un rapport sujet-objet (percevant-percept). Percevant et percept interagissent et peuvent tous les deux varier considérablement pendant le processus de perception.

Par exemple, un percept auditif peut changer lorsque l'auditeur concentre son attention sur un composant spécifique tel que le son d'un instrument particulier parmi l'orchestre. De même, l'attitude d'un percevant vis à vis d'un percept peut varier au cours de séries expérimentales.

Il devrait de plus être noté que les auditeurs écoutent normalement dans un mode "gestalt", i.e. ils perçoivent globalement plutôt que segmentairement. Une analyse du type de l'ingénierie courante peut écarter complètement des caractéristiques relevantes.

En conséquence, nous voyons dans le futur de la Technologie Binaurale que des modèles psychologiques seront exploités et implémentés, quoique non - pendant quelque temps - sous la forme de calculs massifs biologiques en parallèle comme pour le système nerveux central.

Déjà, des discussions sont en cours et de premiers exemples de combinaisons entre systèmes-experts et autres systèmes basés sur connaissance (knowledge-based systems) avec des têtes artificielles, des présentations auditives et des modèles du système auditif.

Si nous pensons à des applications telles que des interfaces homme/machine complexes, des systèmes multi-média, des environnements virtuels interactifs et des systèmes de téléopérations, il devient clair que la Technologie Binaurale conventionnelle doit être combinée avec, ou intégrée dans, des systèmes capables d'effectuer des décisions et de contrôler des actions d'une manière intelligente.

Avec cette vision en tête il est censé d'énoncer que la Technologie Binaurale en est encore à son premier stade. Plusieurs challenges technologiques relevantes et opportunités d'affaires sont au devant.