

令和 5 年 6 月 14 日現在

機関番号：21401

研究種目：基盤研究(C) (一般)

研究期間：2019～2022

課題番号：19K12066

研究課題名(和文) 耳から耳へ音空間を伝える收音・再生システムの開発

研究課題名(英文) Development of sound signal acquisition/reproduction system that conveys sound field from ear to ear

研究代表者

渡辺 貫治 (Watanabe, Kanji)

秋田県立大学・システム科学技術学部・准教授

研究者番号：20452998

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：音空間の收音・再生を小規模なシステムで実現するための検討を行った。本研究では、ダミーヘッド録音信号を聴取者に合った信号に変換する個人化を行うシステムを構築した。システムでは、両耳間レベル差(ILD)に基づき音信号の到来方向を推定する。さらに、到来方向に対応した聴取者のILD・ITDとなるように録音信号を調整・置換することで個人化を行う。客観評価及び主観評価によって、1音源であれば方向の推定が可能であること、聴取者による個人化の効果が得られることを示した。また、並行してマイクロホンアレイによる小規模な收音システムについても検討し、約2 mのアレイで実現可能であることを示した。

研究成果の学術的意義や社会的意義

本研究のシステムは、ある聴取者が空間で得られる体験を、別の聴取者にあたかも遠隔地の空間にいるかのような高い臨場感を与えることができるものである。すなわち、バーチャルリアリティの発展や遠隔コミュニケーションシステムの高度化への貢献が期待できる。同様なシステムの例はあるものの、聴覚による音空間の手がかりを活用したものの例は少ないため、本研究によって得られた知見は従来のシステムにも反映され得るものであり、学術的にも意義があると考えている。

研究成果の概要(英文)：An investigation was conducted to realize sound acquisition and reproduction in a sound field with a small-scale system. In this study, we constructed a individualization processing that converts a dummy head recording signal into a signal suitable for a listener. The system estimates the direction of arrival of the sound signal based on the interaural level difference (ILD). In addition, individualization is performed by adjusting and replacing the recorded signals so that the listener's ILD and interaural time difference (ITD) correspond to the direction of arrival. By objective evaluation and subjective evaluation, it was shown that it is possible to estimate the direction for one sound source, and that the effect of individualization can be obtained although it depends on the listener. In parallel, we also investigated a small-scale sound acquisition system using a microphone array, and showed that it is possible to implement it with an array of about 2 m.

研究分野：音響学

キーワード：バーチャルリアリティ バイノーラル 收音 頭部伝達関数 両耳間レベル差 両耳間時間差

1. 研究開始当初の背景

(1) 聴取者に実際には音源が存在しない場所から音が聞こえるように感じさせるなど、仮想的な音空間を提示する装置である聴覚ディスプレイは、音響バーチャルリアリティの実現に必要不可欠である。例えば、会議やコンサートを遠隔地においても高い臨場感で体感できる、テレイグジスタンスへの応用が考えられる。

(2) コンサートなど実在する空間の音情報を提示するには、元の空間で音情報を收音するシステムも必要である。従来検討されている收音システムでは、マイクロホンアレイを用い音源位置による信号の相違などを検出して音情報を空間的に分離する、“空間サンプリング”を行う方法が多い。研究代表者もビームフォーミングによる方向に関する空間サンプリングを行う收音・再生システムの検討(基盤研究(C)(15K00280))を行ってきた。このようなシステムではマイクロホン間の信号の相違を用いるため、波長の長い信号に対してはマイクロホンの間隔を大きくする必要があり、原理的にマイクロホンアレイが巨大になるという問題がある。

(3) ヒトが両耳に入射した音から空間を知覚していることを考えると、頭部の大きさに比べて巨大なアレイは必ずしも必要ではないと思われる。また、従来の收音・再生システムでは、提示する両耳信号の算出に音源から両耳までの音響伝達関数(頭部伝達関数)を用いることが多い。頭部伝達関数(head-related transfer function, HRTF)は音源位置と個人に依存して異なるが、個人ごとに全周方向に対して測定するのは非現実的であるため、聴取者個人の頭部伝達関数を用いず両耳信号を個人に合わせる、個人化の必要がある。代表者は、両耳間レベル差(interaural level difference, ILD)と両耳間時間差(interaural time difference, ITD)を個人に合わせることで、他人の頭部伝達関数を個人化可能であることを示した(文献)。そのことを踏まえると、聴取者以外(例えばダミーヘッド)の両耳信号のレベル差・時間差を聴取者に合わせることで、あたかも聴取者の両耳で聴取しているかのような音信号を合成できると考えられる。

2. 研究の目的

(1) 本研究では、マイクロホンアレイのような大規模な收音システムに代わる、小規模なシステムの構築を目標とした。ヒトが頭部の大きさ程度の收音システムで音源方向を識別していることから、聴覚の機能を活用したシステムが実現できると考え、検討を行うこととした。特に、ダミーヘッド録音信号の両耳間レベル差・時間差を聴取者のものとなるよう調整・置換することで、個人化が可能であることを示すことを目的とした。

(2) 收音システムとしては、音源方向が未知であることが想定されるため、收音信号(例えばダミーヘッド録音信号)から音源方向を推定できることが重要である。そこで、両耳信号の差に基づく推定法についても検討を行うことを目的とした。

(3) (1)(2)の検討を踏まえて、收音から再生までを行う一連の処理を実装し、システムの評価を行い、提案法の有効性を示すことを目的とした。

3. 研究の方法

(1) 図1に、本研究で実装した收音・個人化・再生システムの概要を示す。本研究では、ダミーヘッドで録音された信号に対し、到来方向別に分離した上で聴取者のILD、及びITDとなるように合わせることで個人化を行う。ダミーヘッドのHRTFはあらかじめ測定などの手段で取得しておくことを想定する。聴取者のILD、ITDもあらかじめHRTFを測定して算出しておくか、代表者の知見(文献)に基づき、身体特徴量から推定しておくこととする。また、到来方向の推定もILDを手掛かりに行うこととする。

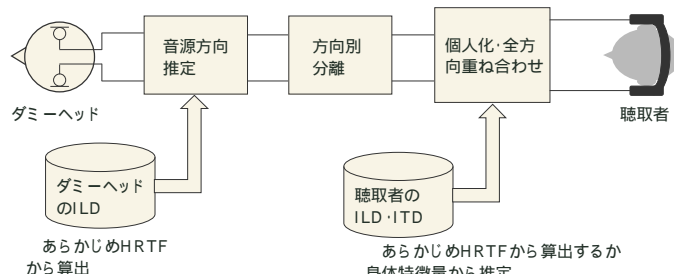


図1. 提案する收音・個人化・再生システムの概要

(2) 提案システムの有効性を検証するため、ILDに基づく到来方向推定、及び方向別の分離が可能かを客観評価によって検討した。また、ILD・ITDによる個人化の効果を主観評価によって検討した。

(3) 検討を進めていく中で、現状ではILDによる到来方向別の分離が多音源に対して困難であ

ったため、小規模のマイクロホンアレイによる収音の可能性についても並行して実施した。

4. 研究成果

(1) 図2は、音源方向とある周波数のILDの対応を示している。ダミーヘッドによるバイノーラル録音信号から得られた両耳信号のレベル差とILDが交差する点が音源方向として推定される。この例だけでは候補は2方向あることとなるが、他の周波数に対しても同様に候補を推定することで、多数決で音源方向を決定することができる。図3は、一音源の信号に対して音源方向を推定した結果であり、正解の音源方向(横軸)に対して推定方向(縦軸)がほぼ一致していることが見て取れる。すなわち、システムの方角推定については、一定の成果が得られたと言える。しかしながら、複数音源については、推定精度が悪かったため、今後の課題である。

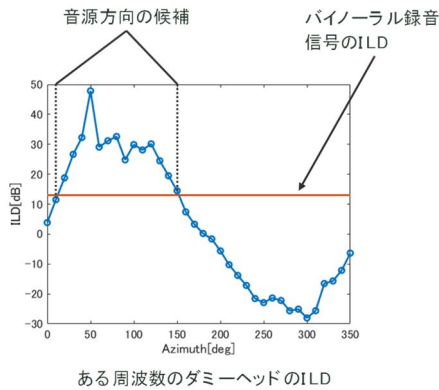


図2.音源方向とILD対応の例

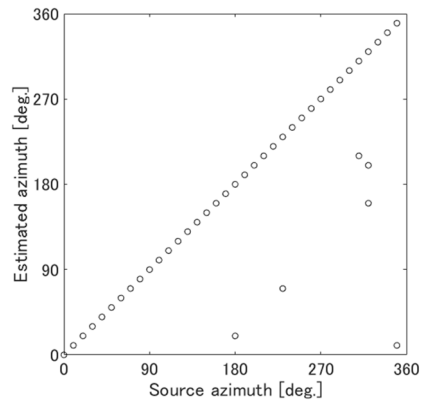


図3.音源方向推定結果の例

(2)1音源については方向推定が可能であることが示されたため、個人化の効果を検証した。バイノーラル録音信号のILD・ITDを、推定された音源方向に対応する聴取者のILD・ITDとなるように調整・置換することで、個人化した音信号を作成した。作成した音信号、聴取者本人のHRTF、ダミーヘッドのHRTFのいずれかを積み込んだ音信号を刺激音として、相違が知覚できるかを比較する主観評価実験を行った。実験では聴取者本人のHRTFを積み込んだ音(“own”)をリファレンスとし、提案法によって個人化した音(“individualized”),ダミーヘッドのHRTFを積み込んだ音(“dummy”),未処理の音信号(“anchor”)を、定位感、音色、総合的な印象について5段階(1:全く似ていない~5:ほぼ同じ)で評価させた。図4に全被験者(5名)の平均値を示す。提案法はダミーヘッドの結果とほぼ同じであり、この結果からは有効性を確認することはできなかった。一方、図5はある被験者の結果で、ダミーヘッドの評点よりも個人化した場合の評点が高いつけられており、被験者によっては提案法の効果が見られた。また、図4を見ると、音色(“quality”)の評価が定位感(“localization”)より低くなっており、総合的な印象が音色の再現精度によって決められていると思われる。本システムの個人化に用いているILD・ITDは主に定位の手がかりとして重要であるが、音色に対する考慮が不十分であると言える。したがって、今後は音色も含めた個人化について検討することが課題である。

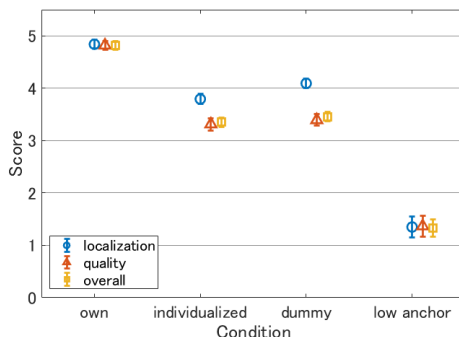


図4.主観評価の結果(全被験者の平均)

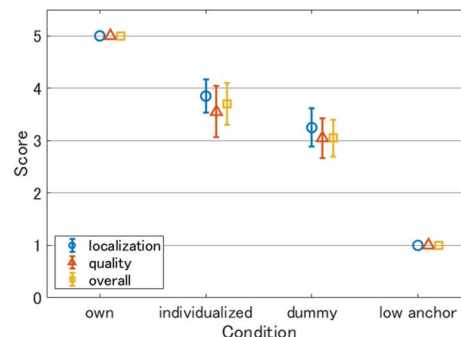


図5.主観評価の結果(効果が見られた被験者)

(3) 複数音源に対する到来方向別の分離が困難であることがわかったため、代表者が以前実施したビームフォーミングによる収音・再生システム(基盤研究(C)(15K00280))による方法に

についても検討を行った。この方法ではマイクロホンアレイを用いるため、背景で述べたシステムの大規模化が課題であった。一方、聴覚の性質を踏まえると、アレイが巨大化する低域については方向知覚の手がかりが少ないと仮定すれば、ビームフォーミングが不要であることが考えられる。そこで、ビームフォーミング処理を適用しない、あるいは小規模なアレイでビームフォーミングを行うことが聴感上影響を及ぼすかを主観評価によって検証することとした。本研究のシステムでは、可聴域内の 150 ~ 19200 Hz の帯域を 7 分割して処理を行っている。その内、低域の 2 帯域 (150 ~ 600 Hz) に対してビームフォーミング処理を適用しないこととし、3 帯域目 (600 ~ 1200 Hz) についてはより高域のアレイでビームフォーミングを行うこととした。実験では、小規模なアレイで得られる両耳信号が、全ての帯域に対して処理を行った場合とで聴感上相違があるかを比較実験によって検証した。また、高域のアレイを低域の処理に適用するとビームが太くなるため、ゲイン調整も適用することとした。刺激音は音源信号をホワイトノイズとし、全帯域にビームフォーミング処理を適用した場合を reference とし、1200 Hz までビームフォーミングを適用しない場合 (“condition0”), 600 Hz まで適用せず、かつ 600 ~ 1200 Hz に高域のアレイを使用した場合 (“condition1”), さらにビームが太くなることを考慮したゲイン調整を行った場合 (“condition2”) の 3 条件で行った。図 6 は、5 名の被験者の結果である。音源方向に対して、相違を知覚した回数をプロットしている。赤い横線は、 χ^2 検定に基づき有意に相違が区別できるかどうかの基準であり、その基準を超えている、すなわち相違を知覚できた場合を塗りつぶして示している。結果から、全くビームフォーミングを行わない場合は相違が知覚されていることがわかる。低域のビームフォーミング処理を簡易化した場合も、側方 ($\pm 90^\circ$) に近い方向では相違が知覚されているが、正面 (0°) から斜め方向については相違が知覚されておらず、小規模なアレイでも收音が可能であることが示唆された。本システムの仕様では、1200 Hz 以上のアレイで約 2 m の大きさであるので、比較的現実的なサイズで実現可能であると言える。

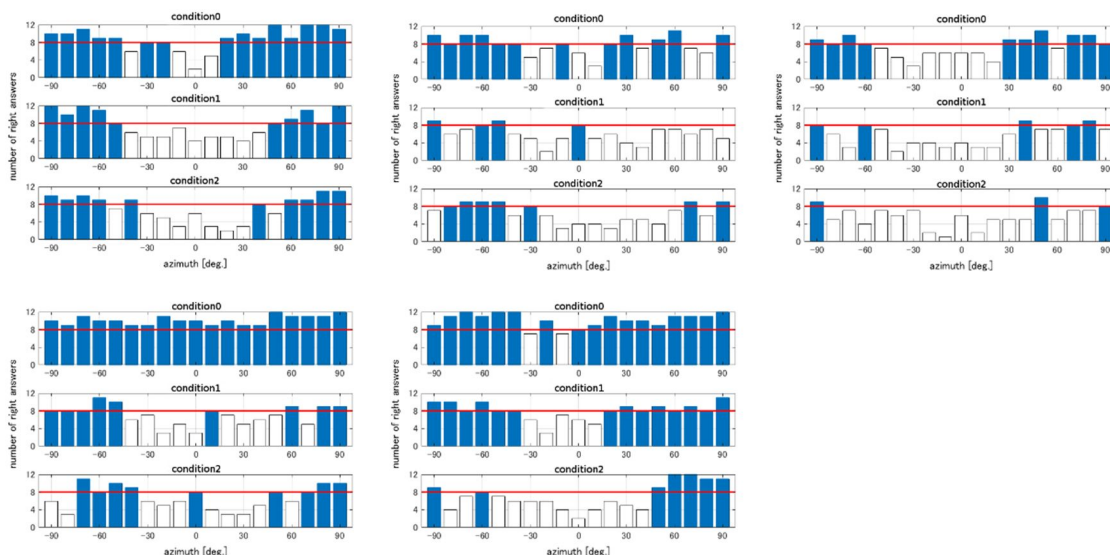


図 6 . 主観評価の結果 (被験者 1 ~ 5)

(4) 以上の検討から、本研究の目的が完全に達成できたとは言えないものの、以下の成果が得られたと言える。

- ・ ダミーヘッド録音信号に対する、ILD に基づく音信号の到来方向別の推定が可能である。ただし、方向別の信号分離について課題が残されている。
- ・ ダミーヘッド録音信号に対する、ILD・ITD による個人化については、聴取者によるものの一定の効果があると言える。さらに効果を向上させるためには、音色を含めた個人化の方法を検討する必要がある。
- ・ 聴覚の性質を考慮することで、マイクロホンアレイによる收音システムの小型化が実現可能であることを示した。本研究の個人化と組み合わせることが今後の課題である。

(5) 本研究の主目的とは別に、ビームフォーミングによって音信号を到来方向別に分離する精度を向上させるため、サイドローブ抑圧の検討も行い、得られた成果を公表した。また、将来的な收音信号の伝送・符号化を狙いとした、異なる方向の音源同士が及ぼす空間的なマスキング効果の検討も行い、得られた成果を公表した。

< 引用文献 >

Kanji Watanabe, Kenji Ozawa, Yukio Iwaya, Yo[^]iti Suzuki, and Kenji Aso, “Estimation of interaural level difference based on anthropometry and its effect on sound localization,” J. Acoust. Soc. Am., Vol. 122(5), pp. 2832-2841(2007).

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計5件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 (2)渡邊貫治, 日高幸祐, 西口正之, 高根昭一, 安倍幸治
2. 発表標題 バイノーラル收音再生システムにおける両耳間レベル差及び時間差の個人化に関する一検討
3. 学会等名 日本音響学会
4. 発表年 2021年

1. 発表者名 渡邊貫治, 西口正之, 高根昭一, 安倍幸治
2. 発表標題 音空間收音における低域のビームフォーミング処理の有無の聴感への影響
3. 学会等名 電子情報通信学会
4. 発表年 2020年

1. 発表者名 山中佑馬, 西口正之, 渡邊貫治, 高根昭一, 安倍幸治
2. 発表標題 量子化雑音の知覚に対する空間的マスキング効果の影響
3. 学会等名 電子情報通信学会
4. 発表年 2020年

1. 発表者名 渡邊貫治, 西口正之, 高根昭一, 安倍幸治
2. 発表標題 音空間收音において低域に適用するビームフォーミング処理の聴感への影響
3. 学会等名 日本音響学会
4. 発表年 2020年

1. 発表者名 山中佑馬, 西口正之, 渡邊貫治, 高根昭一, 安倍幸治
2. 発表標題 空間的マスクング効果に基づく3D音響符号化に関する検討
3. 学会等名 日本音響学会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	西口 正之 (Masayuki Nishiguchi) (90756636)	秋田県立大学・システム科学技術学部・教授 (21401)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------