

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 15 日現在

機関番号：13901

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25280060

研究課題名(和文)空間を超越するインタラクティブ聴覚拡張システムの研究

研究課題名(英文)Interactive auditory extension system including without spatioal limitation

研究代表者

武田 一哉 (Takeda, kazuya)

名古屋大学・情報科学研究科・教授

研究者番号：20273295

交付決定額(研究期間全体)：(直接経費) 13,100,000円

研究成果の概要(和文)：空間を超越するインタラクティブ聴覚拡張システムを提案する。本研究では、音響空間情報の取得、取得信号の分析・加工、立体的音場再生の高臨場感音響通信技術の三要素を統一的に研究し、効果的に音響空間印象を変化させる方法の理論的枠組みを確立する。提案システムでは、利用者がインタラクティブかつ効率的に音響空間印象を操作することが可能となるように、少数の変数で音響空間印象の操作を行うための手法を提案する。本研究における成果の応用の一例として、研究対象の要素技術を統合し、ウェブ上でデモシステムを公開し、音響空間印象のインタラクティブな加工受聴が可能であることを示した。

研究成果の概要(英文)：Interactive auditory extension system which can adaptively control the perceived impression of audio signals is proposed. In this study, we uniformly study about the three elements in the high realistic audio communication techniques; acquisition of spatial audio information, analysis and signal processing of obtained signals, three-dimensional sound reproduction. Unlike conventional impression control systems, which controls two or more dimensional parameters, the proposed system allows users to search for their optimal localization by manipulating a single control knob, interactively and effectively. We also create a demonstration of the proposed system on the web to give an example of the application of the proposed system.

研究分野：情報学

キーワード：音場再生 リア収録 空間受音 アレイ信号処理 潜在次元 音源信号分離 立体音響 インタラクティブ 狭工

1. 研究開始当初の背景

特定の音場における音響情報を、空間的な印象を含めて取得・伝送・再生する「高臨場感音響通信」の実現が求められ、要素技術の研究が進められているが、「受聴者がインタラクティブに高臨場感音響通信系全体を調整」し、自分が望む空間音響印象(音源毎の強度、音源位置分布、残響感等)を創出する技術の研究は立ち遅れている。

音響空間情報の取得、信号の分析・加工、立体的音場再生の三技術は独立して研究されることが多く、その視点も、音響デバイス、信号処理、聴覚心理といった異なる領域に跨っていた。このため、「受聴者がインタラクティブに高臨場感音響通信系全体を調整し、自分が望む空間音響印象(音源毎の強度、音源位置、残響感)を創出する」技術のように、高臨場感音響通信系全体を一つの応用システムとして理解・定式化するための研究が立ち遅れている。

2. 研究の目的

実環境下において取得された空間的な音響情報を、インタラクティブに加工受聴可能な聴覚拡張システムを実現することを目指す。音響信号空間の「潜在次元(=特性行列の主要固有値数)」に着目して、高臨場感音響通信技術の三要素(音響空間情報の取得、取得信号の分析・加工、立体的音場再生)を統一的に研究し、少数の変数で効果的に音響空間印象を変化させる方法の理論的枠組みを確立する。本提案の目的は、空間音響情報の「取得、分析・合成、再生」分野における第一線の研究者が協力して、「取得、分析・加工、再生」に跨る空間音響情報処理を統一的に議論する枠組みを構築することであり、単に新しい応用システムを作ることに留まらない。

潜在次元に基づき、空間音響情報の表現能力を議論する方法は、高臨場感音響通信を構成する様々なシステムや要素技術の評価や定式化に応用され、統合的な研究を大きく前進さ

せる。「インタラクティブ聴覚拡張」システム自体も広い応用を持ち、実用研究段階に進むことが可能になる。

本研究により、「インタラクティブ聴覚拡張システム」の実現可能性が示されるとともに、以下の2つの知見が得られる。

- (1) 所与の条件の下で、高臨場感音響通信系が生成しうる空間音響情報の範囲を数理的に規定する方法。
- (2) (1)により与えられる範囲において、主観的な印象と対応づけつつ、効率的に(少ない変数を用いて、幅広い範囲に)空間音響情報を変化させる方法。

3. 研究の方法

統括班、受音班、分析班、再生班の4つのグループを構成し、(1)音響空間情報の取得、(2)音源推定、(3)音場操作、(4)立体的空間印象の付加、(5)システム化の5つの課題に全体として取り組む。

(1) 音響空間情報の取得

空間内の任意の2点間の音響伝達特性を全て計測することは不可能であり、限られた音源/受音位置間の伝達特性群(部分行列)から、空間伝達行列を推定する必要がある。拡散受音を応用して効率的な(潜在次元数が高い)部分行列を取得する理論を確立する。

(2) 音源推定

複数点で受音された信号群から、音源信号群とその位置を推定する必要がある。原音源を忠実に「分離」推定するのではなく、受聴者がインタラクティブに操作可能な自由度に対応する数の仮想音源群への「分解」法を確立する。

(3) 音場操作

空間伝達行列に交代行列や拡大行列を順次作用させる「空間伝達行列の派生操作」を、少数の変数でパラメトリックに表現する必要がある。

(4) 立体的空間印象の付加

希望する音響空間に、より近い空間印象を生成する頭部伝達特性を少ない変数の調整で選択可能にする必要がある。多様な個人や空間の頭部伝達特性を派生させる方法を確立する。

(5) システム化

上述の結果を統合し、音場操作を準リアルタイムに実行するシステムを実装する必要がある。

4. 研究成果

(1) 空間伝達行列を単一の変数でパラメトリックに操作する手法の実現

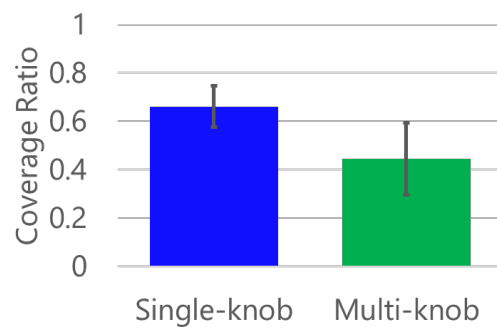
本研究の目的システムを実現するためには、仮想的な空間印象を創成する空間伝達行列 H を少数の変数でパラメトリックに表現し、システムの利用者が手軽に H を操作・変更する仕組みを実現する必要がある。 H を操作・変更することで利用者とシステム中の各音源の間の仮想的な伝達特性が変化し、利用者に対する聴覚的な印象を変化させることが可能となる。

本研究では、空間伝達行列 H を構成する膨大な空間中から、主要な部分空間を取得する手法を提案し、その有用性を確認した。提案手法では、 H の空間を構成するベクトル f を聴覚印象ベクトルと呼称し、聴覚印象ベクトル同士の類似度を各要素に持つ行列を以下のように設計する。

$$K = \begin{bmatrix} k(1,1) & \cdots & k(1,J) \\ \vdots & \ddots & \vdots \\ k(J,1) & \cdots & k(J,J) \end{bmatrix}$$

ここで $k(i,j)$ は聴覚印象ベクトル f_i, f_j 同士の類似度を表し、 J は H の主要な部分空間を構成するベクトル数を表す。より効率的に H の主要な部分空間を表現するためには、空間を構成するベクトル同士の類似度が低く、より多様なベクトルにより空間が構成されていることが望ましいと考えられるため、提案手法では、行列 K が可能な

限り対角となり、異なる聴覚印象ベクトル間の類似度が低減するように H の部分空間を選択する。その後、選択された部分空間を構成する聴覚印象ベクトルをパラメトリックに切り替えることで、聴覚印象ベクトルを指定するための単一の変数を利用して、空間伝達行列をより効率的に操作・変更することができるようになる。



上図に、実際に被験者がシステムを利用した際の主観評価結果の結果を示す。実験では、提案した単一変数による空間伝達行列操作手法(single-knob)と、多変数を利用して空間伝達行列の各要素を自由に操作することのできる手法(Multi-knob)を比較した。被験者にはシステムを利用して自由に空間伝達行列を操作させ、空間伝達行列が為す空間のうち実際に被験者が体験することのできた聴覚印象の割合(カバー率)を評価した。図から、提案した Single-knob 手法のほうが、実際の利用において、空間伝達行列のより広い範囲をカバーすることが可能なが分かり、提案手法の有効性が確認された。

(2) 事前知識を利用した楽曲音源分離技術

目的とするシステム的主要な応用先は、すでに幅広く普及している楽曲 CD などのステレオ音源であると推定される。そこで、楽曲信号を構成する音源に関する事前知識を利用することで、ある程度高精度に楽曲の音源分離を行う手法を提案した。

提案手法では、事前知識として音源ごとの左右の音圧差の情報を持つ行列 A と時間周波数領域における各音源のパワーの情報を利用する。これらの情報は楽曲信号のミキシングの仮定で取得することができ、また、情報を適切に圧縮

することで楽曲信号に比べて十分小さな容量で実現することができることを確認した。

提案手法では、楽曲のミキシングの過程を以下のような式で表現できると仮定する。

$$X=AS$$

但し、 X はステレオ楽曲信号、 S が楽曲を構成する音源信号により構成される行列である。利用する信号はステレオ 2 チャンネルの信号であるため、ある音源に注目するとそれ以外の 1 音源は混合過程の逆行列を作用させることで完全に抑圧することができる。そこで、各時間周波数において目的音以外の最大パワーを持つ信号を選択し、その音源を抑圧する。その後、事前情報として与えた各音源の時間周波数領域でのパワーを利用してウィナーフィルタを構成することで、音源分離精度を向上させる。

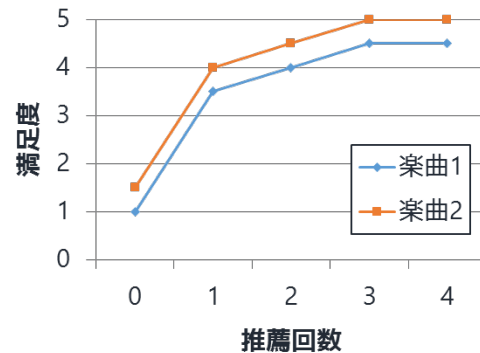
(3) 利用者の好みを反映した空間伝達行列推薦手法

提案システムで利用することのできる空間伝達行列は非常に膨大であるため、実際の利用を考えた際には、たとえ少数のパラメトリックな選択を可能にしたとしても、利用者が自身の望む空間伝達関数を容易には探索することができないことが予想される。そこで、利用者の過去の操作履歴を利用することで、各利用者に最適な空間伝達行列を推薦する仕組みを提案した。

提案法では、利用者がシステムを利用した際に、各自の好みなどの情報をフィードバックすることが可能となるようにシステムを組み換え、各利用者のフィードバック情報を収集する。その後得られたフィードバック情報に基づき、混合ガウスモデルを利用して空間伝達行列が好まれる確率をモデル化し、確率密度が高くなる空間伝達行列を利用者に推薦する。システムによる推薦とユーザからのフィードバックを繰り返すことで、徐々にシステムが利用者に最適化されていくこととなる。

実際に被験者により提案システムを利用した際の空間伝達行列推薦回数とその時の満足度を次の図に示す。図から、更新が行われるごと

にユーザの満足度が徐々に向上し収束していくことが確認される。この結果から提案する推薦手法が有用なものであることが確認される。また、

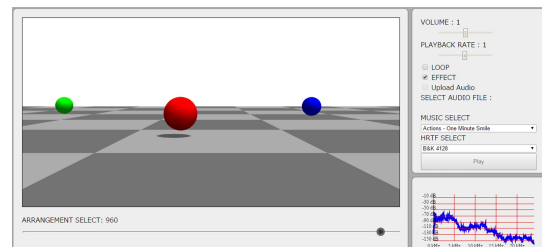


推薦を行わない際の満足度が非常に低いことも確認され、推薦手法が提案手法に対して非常に重要であることも確認された。

(4) 実際のシステムの実装

(1), (2)の研究成果を統合し、実際にシステムを作成し、ウェブ上で利用することができるように公開を行った。

(<http://sp.m.is.nagoya-u.ac.jp/SLP/>)



5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

(雑誌論文)(計 7 件)

- (1) 大谷健登, 小松達也, 近藤多伸, 西野隆典, 武田一哉, “相補ウィナーフィルタを利用した残響抑圧手法に対するフィルタ係数推定手法,” 電子情報通信学会論文誌 A, 査読有, J98-A(2), 2015.2.
- (2) 羽田陽一, 古家賢一, 小山翔一, 丹羽健太, “球面調和関数展開に基づく 2 種類の超接話マイクロホンアレイ,” 電子情報通信学会論文誌 A, 査読有, J97-A(4), 2014.4.

- (3) Yusuke Mizuno, Kazunobu Kondo, Takanori Nishino, Norihide Kitaoka, and Kazuuya Takeda, "Effective frame selection for blind source separation based on frequency domain independent component analysis," IEICE Transactions on Fundamentals, 査読有, E97-A, 2014, P784-P791.
- (4) 丹羽健太, 江崎知, 日岡裕輔, 西野隆典, 武田一哉, "空間相関行列の固有値分布に着目した音源別距離推定," 電子情報通信学会論文誌, 査読有, J97-A, 2014, P68-P76
- (5) Kenta Niwa, Yusuke Hioka, Ken'ichi Furuya, and Yoichi Haneda, "Diffused sensing for sharp directive beamforming," IEEE Transactions on Audio, Speech and Language Processing, 査読有, 21 巻, 2013, P2346-P2355, DOI: 10.1109/TASL.2013.2274695.
- (6) Kenta Niwa, Yusuke Hioka, Sumitaka Sakauchi, Ken'ichi Furuya, and Yoichi Haneda, "An estimation method of sound source orientation using eigenspace variation of spatial correlation matrix," IEICE Transactions on Fundamentals, 査読有, E96-A, 2013, P1831-P1839.
- (7) Yusuke Hioka, Ken'ichi Furuya, Kazunori Kobayashi, Kenta Niwa, and Yoichi Haneda, "Underdetermined sound source separation using power spectrum density estimated by combination of directivity gain," IEEE Transactions on Audio, Speech and Language Processing, 査読有, 21 巻, 2013, P1240-P1250, DOI: 10.1109/TASL.2013.2248715.
- [学会発表] (計 36 件)
- (1) Maureen and Yoichi Haneda, "Linear array aperture extrapolation based on a spherical wave assumption," 12th Western Pacific Acoustics Conference 2015 (WESPAC2015), 2015.12.5, Singapore.
- (2) Tomomi Suzuki, Yoshio Ishiguro, Takanori Nishino, and Kazuuya Takeda, "Integration of acoustic information in Google Street View using a spherical microphone array," AUN/SEED-Net Regional Conference for Computer and Information Engineering 2015 (RCCIE 2015), 2015.10.2, Hanoi.
- (3) 武田一哉, 西野隆典, 丹羽健太, 羽田陽一, 猿渡洋, 西村竜一, "音響空間の聴覚的操作 ~ 超臨場化と個性化 ~," 電子情報通信学会応用音響研究会, 2015.7.3, 電気通信大学.
- (4) 西村竜一, 榎本成悟, 加藤宏明, "複数マイクを用いた周波数領域における超解像度の試み," 日本音響学会春季研究発表会, 2015.3.18, 中央大学
- (5) Kento Ohtani, Tatsuya Komatsu, Takanori Nishino, and Kazuuya Takeda, "Adaptive dereverberation method based on complementary Wiener filter and modulation transfer function," REVERB Workshop, 2014.5.10, Italy.
- (6) Yoichi Haneda, Ken'nich Furuya, Shoichi Koyama, Kenta Niwa, "Close-talking spherical microphone array using sound pressure interpolation based on spherical harmonic expansion," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014.5.4, Italy.
- (7) Tomo Miyauchi, Daichi Kitamura, Hiroshi Saruwatari, Satoshi Nakamura, "Depth estimation of sound images using directional clustering and activation-shared nonnegative matrix factorization," 2014 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2014), 2014.3.2, U.S.A.

- (8) Ryouichi Nishimura, Seigo Enomoto, Parham Mokhtari, Hironori Takemoto, "Application of Three-Dimensional Audio to Copyrighted Multimedia Contents," International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2013.10.16-2013.10.18, China.
- (9) Kenta Niwa, Yusuke Hioka, Kazunori Kobayashi, Ken'ichi Furuya, and Yoichi Haneda, "Evaluation of microphone array based on diffused sensing with various filter design methods," 21th European Signal Processing Conference (EUSIPCO 2013), 2013.9.9-2013.9.13, Morocco.
- (10) Tatsuya Komatsu, Takanori Nishino, Gareth Peters, Tomoko Matsui, and Kazuya Takeda, "Modeling head-related transfer functions via spatial-temporal Gaussian process," 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), 2013.5.28, Canada.

6. 研究組織

(1) 研究代表者

武田 一哉 (TAKEDA, Kazuya)
名古屋大学・大学院情報科学研究科・教授
研究者番号: 20273295

(2) 研究分担者

羽田 陽一 (HANEDA, Yoichi)
電気通信大学・大学院情報理工学研究科・教授
研究者番号: 80647496

西村 竜一 (NISHIMURA, Ryouichi)
国立研究開発法人情報通信研究機構・
ソーシャルイノベーションユニット耐震災
ICT 研究センター・主任研究員
研究者番号: 30323116

西野 隆典 (NISHINO, Takanori)
三重大学・大学院工学研究科・准教授
研究者番号: 40329769

猿渡 洋 (SARUWATARI, Hiroshi)
東京大学・大学院情報理工学系研究科・教授
研究者番号: 30324974