

平成 30 年 5 月 17 日現在

機関番号：33919

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00235

研究課題名(和文)高臨場感音場創生のための多方向同時バイノーラル録音システム

研究課題名(英文)Omni-binaural recording system for high realistic sound field creation

研究代表者

西野 隆典(Nishino, Takanori)

名城大学・都市情報学部・教授

研究者番号：40329769

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：360°画像に代表される多視点映像へ立体音響信号を付与するための全方向バイノーラル録音システムについて検討を行った。本課題では、耳甲介腔にみたてた半球状のくぼみを複数有する球状マイクロホンバッフルを提案し、このバッフルの音響解析と、各くぼみに配置したマイクロホンから得られる音響信号の統合方法について検討を進めた。音響解析においては、くぼみの数による音響特性の変化が小さいことを数値解析により明らかとした。音響信号の統合方法では、時間領域信号を加重平均する手法を提案し、所望する音響信号が得られることを確認した。

研究成果の概要(英文)：An omni-binaural recording system for 360-degree video was proposed. The proposed system consists of a spherical microphone baffle and a method of integrating recorded signals. This spherical microphone baffle has six hemispherical hollows. Each hollow is modeled on a cavity of concha. Acoustical characteristics of the proposed baffle were obtained by a numerical analysis, and spectral differences between two and six hollows were insignificant. Moreover, a generating of binaural signals for arbitrary listening direction was proposed. The proposed method was based on the weighted sum method and able to generate a desired binaural signals.

研究分野：音響工学，音信号処理

キーワード：立体音響 バイノーラル録音 多視点映像 くぼみ付き球状マイクロホンバッフル

1. 研究開始当初の背景

高臨場感音場再生技術は、任意の音響空間の情報を観測し、別空間で再現すること目的とした技術であり、遠隔地間コミュニケーションや仮想現実などでの利用が想定されている。この技術において、音源信号と音源位置を効率的に収録することが、品質の良い臨場感通信の実現や、技術の普及に必要不可欠である。任意の空間での音響収録技術である立体音響収録技術では、バイノーラル録音システムや、方向受音可能なマイクロホンアレイシステム[1][2]などが提案され、特にバイノーラル録音技術は、安価な録音システムの登場も相まって、様々なバイノーラル録音コンテンツが利用可能である。しかし、これらの録音技術は聴取者の向きや位置、および音源位置が変化しない静的な条件で用いられることが多い。一方で、動画像分野においては、360°映像や全天球画像と呼ばれる、視点を自由に変えることが可能な映像コンテンツが登場し、町並みや建物内の様子を、あたかもその場にいるように表示させることが可能となってきている[3]。高臨場感再生の実現のためには、このような全天球画像に対し、音情報を付与するための技術が必要であると言える。

2. 研究の目的

本研究課題の目的は、全天球画像に対応した立体音を収録、再生するシステムの検討である。従来技術の一つであるバイノーラル録音・再生は、人間の頭部形状を模擬した疑似頭を収録したい環境に設置し、両耳に配置したマイクロホンで収録を行うものであるが、再生時に聴取者が聴取方向を変化させることは考慮されていない。そのため、任意の聴取方向を想定し、従来型のバイノーラル録音システムを、全方向対応へ拡張する必要がある。バイノーラル録音は、人頭と耳介による音響特性の変化を反映したものであることから、全方向バイノーラル録音システム(omni-binaural recording system)においては、システムに複数の耳を搭載することで解決できると考えた。

そこで本課題では、これまでに検討を進めてきた小型球状マイクロホンバッフル[4]に、複数の耳を設置することで、全方位バイノーラル録音システムの実現を図った。

3. 研究の方法

本課題は、小型球状マイクロホンバッフルを基にした全方位バイノーラル録音システムの開発と、開発したシステムで収録した音を任意の聴取方向に対応した立体音信号へと統合する方法の2課題により実施した。

- (1) これまでに提案した小型球状マイクロホンバッフルは、直径10 cmで、左右に直径2 cmの半球状のくぼみを有する形状をしている。これまでの検討結果から、

このバッフルは小型ではあるが、音像位置や空間的拡がりを受取者に提示することが可能であることが示されている。本研究課題では、このバッフルに4つのくぼみを追加することで、全方位バイノーラル録音への対応を図った。

- (2) 再生時には、複数のくぼみから得られた信号を、左右2チャンネルの信号へと統合する必要がある。この統合においては、低計算コスト、高速処理を課題として、時間領域において音信号を加重平均する手法について検討を進めた。全天球画像との利用を想定し、音源が視線方向に位置する場合と、視線方向外に位置する場合との2条件にて評価を行った。

4. 研究成果

本課題により得られた成果は、6つのくぼみを有する球状マイクロホンバッフルが全方位バイノーラル録音システムとして利用可能なこと、ならびに時間領域加重平均による音信号の統合方法が有効に機能することの2点である。

- (1) これまでに提案したくぼみ付きマイクロホンバッフルに、4つのくぼみを新たに追加したバッフルを検討した(図1)。球の直径は10 cm、各くぼみは半球状で直径が2 cmと従来のバッフルと同じ仕様とした。各くぼみの内に無指向性のマイクロホンを設置することで、簡易バイノーラル録音が可能となる。初めに両バッフルの音響伝達特性を数値解析により求め、その差を比較した。図2に両バッフル間のスペクトル特性の差を示す。結果から、両バッフル間ではスペクトルの差は、 ± 2 dBに収まっており、聴感上の差異は少ないと考えた[5][6]。

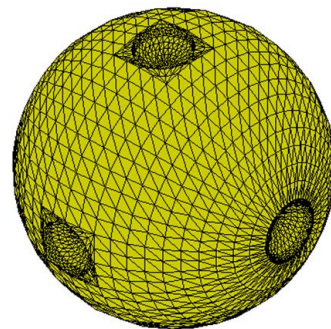


図1 6つのくぼみを有する球状マイクロホンバッフル[5][6]

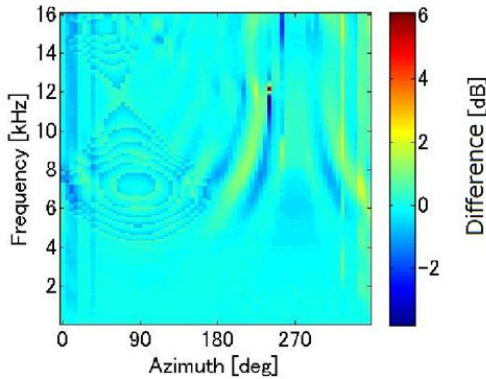


図2 くぼみの個数(2 と 6)の違いによるスペクトルの差[5][6](横軸：音源方位、縦軸：周波数、色はスペクトルの差(dB)に対応)

- (2) 6つのくぼみに配置したマイクロホンで得られた信号を、ヘッドホン等で聴取するために2チャンネル信号へと統合する必要がある。本課題では、頭部伝達関数の補間問題を参考に、課題解決を図った[6]。頭部伝達関数の補間問題においては、左右の伝達特性の初期遅延時間を揃えた上で、時間領域で線形補間を行うことが、精度や処理量の面で最も有効であることが示されている。本課題では、球状バッフルの直径が人頭の約半分であり、左右での遅延時間が人頭の場合よりも短いことを利用し、初期遅延時間の差を考慮せずに次式のように時間領域で加重平均をとることとした。

$$s_R(t) = (1-w_1) x_i(t) + w_1 x_j(t)$$

$$s_L(t) = (1-w_2) x_m(t) + w_2 x_n(t)$$

なお、この式では、加重平均の重みは左右の信号で異なること、また統合に用いるマイクロホンも左右で異なることを仮定している。

この式における重みを、線形関数、べき関数、3次スプライン関数、三角関数により生成した。評価は、音源が視線方向に存在する場合(音源方向は 0°)と、視線方向に存在しない場合の2条件にて行った。音源が視線方向にある場合の結果を図3に示す。この結果から、3次スプライン関数、三角関数を用いた場合に、平均スペクトル歪がともに3.6 dBと最も精度の良い統合となった。続いて、音源が視線方向外にある場合での結果では、3次スプライン関数を用いた場合の平均スペクトル歪が4.1 dB、三角関数を用いた場合の平均スペクトル歪が4.0 dBとなり、両関数を用いた場合に最も精度の良い統合となった。

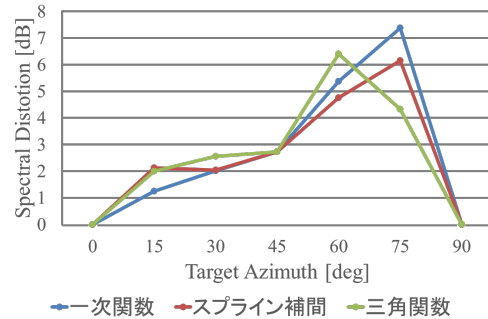


図3 音源が視線方向にある場合の重み係数決定関数の違いによる統合精度[6](横軸：音源方位、縦軸：スペクトル歪(dB))

実験により得られた歪値は、音情報のみかつ音像位置が静止しているような状況では、定位精度の劣化につながると考えられる。しかし、映像と合わせることや、音像位置を聴取者が変化させ得ることを考慮すれば、聴感上許容できる歪であると考えられる。映像と合わせた評価を行うため、7台の小型カメラを搭載したテスト機を作成し、評価を行った結果、聴感上問題の無いことが確認された[7]。ただし、このテスト機は球状ではなく、提案形状のシステムを直接評価するものではないため、今後さらなる検証が必要である。

<引用文献>

- [1] Ken Farrar, Soundfield microphone, Wireless World, 48-50, 1979
- [2] Juha Merimaa, Applications of a 3-D microphone array, Proc. AES Convention, 11 pages, 2002
- [3] Google, Google Earth VR <https://vr.google.com/earth/> (accessed May 1, 2018)
- [4] 中桐大志、山村俊貴、西野隆典、成瀬央、武田一哉、くぼみ付き球状マイクロホンバッフルを用いたバイノーラル録音の検討、日本音響学会 2015 年春季研究発表会、2015 年 3 月
- [5] Taishi Nakagiri, Toshiki Yamamura, Takanori Nishino, Hiroshi Naruse, Kazuya Takeda, Development and evaluation of spherical microphone baffle with two hollows for binaural recording, Proc. WESPAC2015, 2015 年 12 月
- [6] 今村健人、複数くぼみつきマイクロホンバッフルを用いた 360°映像のための立体音生成に関する研究、三重大学大学院工学研究科修士学位論文、2018 年 3 月(指導教員：西野隆典)
- [7] Takanori Nishino, Kento Imamura, Taishi Nakagiri, Kenta Niwa,

Development of a cylindrical omni-binaural recording system for 360-degree videos, Acoustical Science and Technology, Vol.39, Issue 4, 2018 (印刷中)

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 3 件)

- (1) Takanori Nishino, Kento Imamura, Taishi Nakagiri, Kenta Niwa, Development of a cylindrical omni-binaural recording system for 360-degree videos, Acoustical Science and Technology, 査読有、Vol.39, Issue 4, 2018 (印刷中)
- (2) 大谷健登、丹羽健太、西野隆典、武田一哉、畳み込み雑音除去自己符号化器と対数周波数領域スペクトル特徴を用いた楽曲音源強調、電子情報通信学会論文誌、査読有、Vol.J101-D, No.3, 2018, pp.615-627
- (3) 朝倉大智、西野隆典、成瀬央、運筆音を用いた記述者識別、日本音響学会誌、査読有、Vol.74, No.1, 2018, pp.14-16

〔学会発表〕(計 31 件)

- (1) 朝倉大智、西野隆典、成瀬央、署名運筆音を用いた記述者識別の検討、電子情報通信学会応用音響研究会、2017年10月21日、牛岳温泉リゾート(富山)
- (2) 今村健人、西野隆典、成瀬央、全方位映像に対応した立体音生成のための音響信号統合方法の検討、電子情報通信学会応用音響研究会、2017年10月21日、牛岳温泉リゾート(富山)
- (3) 大谷健登、丹羽健太、西野隆典、武田一哉、楽器音生成過程を考慮した畳み込みニューラルネットワークに基づく楽曲音源強調、日本音響学会2017年秋季研究発表会、2017年9月26日、愛媛大学(松山)
- (4) 朝倉大智、西野隆典、成瀬央、運筆音を用いた記述者識別における記述者モデル数の影響の調査、日本音響学会2017年秋季研究発表会、2017年9月26日、愛媛大学(松山)
- (5) 朝倉大智、西野隆典、成瀬央、運筆音による記述者識別における記述者モデル数の影響の調査、平成29年度電気・電子・情報関係学会東海支部連合大会、2017年9月7日、名古屋大学(名古屋)
- (6) 今村健人、西野隆典、成瀬央、複数くぼみを持つマイクロホンバツフルにおける立体音響信号の統合の検討、平成29年度電気・電子・情報関係学会東海支部連合大会、2017年9月7日、名古屋大学(名古屋)

- (7) 今村健人、西野隆典、成瀬央、全方位立体集音のためのマイクロホンバツフルの設計と評価、日本音響学会2017年春季研究発表会、2017年3月15日、明治大学(川崎)
- (8) 朝倉大智、西野隆典、成瀬央、運筆音による記述者識別のための識別モデルの検討、日本音響学会2017年春季研究発表会、2017年3月15日、明治大学(川崎)
- (9) Takahiro Kuriwaki, Takanori Nishino, Hiroshi Naruse, Drum sound onset detection based on class separation using deep neural network, The Fifth Joint Meeting ASA and ASJ, 査読有、2016年12月2日、Honolulu(Hawaii, USA)
- (10) Takanori Nishino, Kenta Niwa, Omni-binaural recording system for 360° video, The Fifth Joint Meeting ASA and ASJ, 査読有、2016年12月2日、Honolulu(Hawaii, USA)
- (11) Tomomi Suzuki, Takanori Nishino, Yoshio Ishiguro, Kazuya Takeda, Sound enhancement system using selective binary filtering, The Fifth Joint Meeting ASA and ASJ, 査読有、2016年12月2日、Honolulu(Hawaii, USA)
- (12) Daichi Asakura, Takanori Nishino, Hiroshi Naruse, Writer recognition with a sound in hand-writing, The Fifth Joint Meeting ASA and ASJ, 査読有、2016年12月2日、Honolulu(Hawaii, USA)
- (13) Kento Imamura, Takanori Nishino, Taishi Nakagiri, Hiroshi Naruse, Evaluation of spherical microphone baffle with six hollows for omni-binaural recording, The Fifth Joint Meeting ASA and ASJ, 査読有、2016年11月29日、Honolulu(Hawaii, USA)
- (14) 栗脇隆宏、西野隆典、成瀬央、DNNによる二値分類に基づく混合音楽信号中のスネアドラムの発音時刻検出、日本音響学会2016年秋季研究発表会、2016年9月15日、富山大学(富山)
- (15) 朝倉大智、西野隆典、成瀬央、運筆音を用いた記述者認識の検討、日本音響学会2016年秋季研究発表会、2016年9月15日、富山大学(富山)
- (16) 西野隆典、頭部伝達関数って何だろう?、日本音響学会2016年秋季研究発表会、招待講演、2016年9月14日、富山大学(富山)
- (17) 今村健人、西野隆典、成瀬央、空撮映像への音響情報付与のための指向性マイクロホンシステムの検討、平成28年度電気・電子・情報関係学会東海支部連合大会、2016年9月12日、豊田工業高等専門学校(豊田)

- (18) 朝倉大智、西野隆典、成瀬央、運筆音による記述者認識の検討、平成 28 年度電気・電子・情報関係学会東海支部連合大会、2016 年 9 月 12 日、豊田工業高等専門学校（豊田）
- (19) 栗脇隆宏、西野隆典、成瀬央、DNN を用いた混合音楽信号中のドラムスの発音時刻検出の検討、平成 28 年度電気・電子・情報関係学会東海支部連合大会、2016 年 9 月 12 日、豊田工業高等専門学校（豊田）
- (20) 鈴木友美、西野隆典、石黒祥生、戸田智基、武田一哉、選択型傾聴システムのための音信号提示手法の検討、電子情報通信学会応用音響研究会、2016 年 3 月 28 日、別府国際コンベンションセンター（別府）
- (21) 関翔悟、西野隆典、戸田智基、武田一哉、非負値テンソル因子分解を用いた劣決定ステレオチャンネル音源分離、日本音響学会 2016 年春季研究発表会、2016 年 3 月 9 日、桐蔭横浜大学（横浜）
- (22) Taishi Nakagiri, Toshiki Yamamura, Takanori Nishino, Hiroshi Naruse, Kazuya Takeda, Development and evaluation of spherical microphone baffle with two hollows for binaural recording, 12th Western Pacific Acoustics Conference 2015 (WESPAC2015), 2015 年 12 月 9 日、Singapore (Singapore)
- (23) Ryohei Kudo, Takanori Nishino, Hiroshi Naruse, Identification for octave chords based on probability models in frequency domain, 12th Western Pacific Acoustics Conference 2015 (WESPAC2015), 2015 年 12 月 8 日、Singapore (Singapore)
- (24) Takahiro Kuriwaki, Takanori Nishino, Hiroshi Naruse, Automatic transcription of drums based on arising of power in particular frequency band, 12th Western Pacific Acoustics Conference 2015 (WESPAC2015), 2015 年 12 月 8 日、Singapore (Singapore)
- (25) Tomomi Suzuki, Yoshio Ishiguro, Takanori Nishino, Kazuya 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 年 10 月 2 日、Hanoi (Vietnam)
- (26) 中桐大志、西野隆典、成瀬央、バイノーラル録音用くぼみ付き球状マイクロホンバッフルの性能評価、平成 27 年度電気・電子・情報関係学会東海支部連合大会、2015 年 9 月 28 日、名古屋工業大学（名古屋）
- (27) 栗脇隆宏、西野隆典、成瀬央、固有振動数検出に基づく音楽信号中のドラムスの発音時刻推定、平成 27 年度電気・電子・情報関係学会東海支部連合大会、2015 年 9 月 28 日、名古屋工業大学（名古屋）
- (28) 栗脇隆宏、西野隆典、成瀬央、固有振動数検出に基づく混合音楽信号中のドラムスの発音時刻推定、日本音響学会 2015 年秋季研究発表会、2015 年 9 月 17 日、会津大学（会津若松）
- (29) 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、招待講演、2015 年 7 月 3 日、電気通信大学（調布）
- (30) 栗脇隆宏、西野隆典、成瀬央、特定周波数帯域におけるパワーの立ち上がりに着目したドラムスの自動採譜の検討、音学シンポジウム 2015、2015 年 5 月 24 日、電気通信大学（調布）
- (31) 大谷健登、鈴木友美、丹羽健太、西野隆典、石黒祥生、武田一哉、音像空間配置のインタラクティブな制御手法、音学シンポジウム 2015、2015 年 5 月 23 日、電気通信大学（調布）

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

360° video with binaural sound (システムデモ動画)

<https://youtu.be/VF2pHLZClw8>

6. 研究組織

(1) 研究代表者

西野 隆典 (NISHINO, Takanori)
名城大学・都市情報学部・教授
研究者番号：40329769

(2) 研究分担者

(3) 連携研究者

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

(4) 研究協力者