

軸対称型收音装置を用いた高精度・高感性バーチャル音空間 創成システムの構築 (082102005)

Development of accurate sound space recording and reproduction system
using spherical microphone array with numerous microphones

研究代表者

坂本修一 東北大学電気通信研究所

Shuichi Sakamoto Research Institute of Electrical Communication, Tohoku University

研究期間 平成 20 年度～平成 23 年度

概要

本研究では、バーチャル收音・再生手法の実現を目指し、解決すべき課題として、1) 軸対称型複数マイクロホン收音装置の形状最適化、2) 收音装置収録音からの頭部伝達関数 (head related transfer function : HRTF) 合成手法の決定、3) 收音装置により得られた音情報のネットワーク配信手法の検討の 3 つのサブテーマを柱として技術開発を進める。

項目 1) と 2) は密接に関連する技術課題であり、計算機シミュレーションに基づき形状の検討や HRTF 合成手法の検討を行うほか、実際に何種類か軸対称型複数マイクロホン收音装置を試作し、最適な形状、HRTF 合成手法を決定する。その後、各マイクロホンで得られる信号の差異を分析し、項目 3) に示したネットワーク配信に適した信号への変換手法を開発する。

これら開発された技術を統合し、最終年度には実際にシステム構築を行い、性能を評価する。

Abstract

In this project, we have developed a system to acquire 3D sound space information that can transmit accurate sound space information to a distant place using a microphone array on a human-head-sized solid sphere with numerous microphones on its surface. The system comprises an object with numerous microphones called Symmetrical object with ENchased Zillion microphones (SENZI). During the project, we investigated following three sub-topics: 1) the requirement and specification of microphone array, 2) how to synthesize sound information for each individual listener, and 3) how to transmit recorded signals through network. As the results, we succeed in realization of real-time system and demonstrated that developed system can synthesize 3D sound space with high precision.

1. まえがき

これまでの通信システムは、音情報を高精度で感度高く收音・再生するという観点では、極めて不十分である。とくに音空間の高精度收音技術は、先進的な技術が多く開発されている音空間の高精度再生技術に比べて研究が少なく、收音される音空間の精度も不足している。

音空間情報は臨場感などの高次感性情報通信において重要な情報であり、この高精度收音技術は次世代通信システム構築のための基盤技術として必須となるものである。

2. 研究内容及び成果

本研究ではこのような現状を打破するため、「再生すべき音空間における音源位置の情報なし」に、「遠隔地にいる聴取者に対し頭部運動も含め忠実に臨場感高く音空間を再現」するシステムを構築した。具体的には、多数のマイクロホン有する軸対称型收音装置を利用したバーチャル收音・再生手法を提案し、その実現に向けて、1) 軸対称型複数マイクロホン收音装置の形状最適化、2) 收音装置収録音からの頭部伝達関数 (head related transfer function : HRTF) 合成手法の決定、3) 收音装置により得られた音情報のネットワーク配信手法の検討、の 3 つのサブテーマを設定して研究を推進し、その知見に基づいて実際にシステムの構築を行った。

以下では、具体的に得られた成果を項目ごとに説明する。

2. 1. 軸対称型複数マイクロホン收音装置の形状、および、HRTF 合成手法の検討

本提案システムでは、收音装置の形状により各マイクロホンから得られる音信号が異なり、その差は合成される聴取者の HRTF の精度にも影響するため、どのような形状の收音装置を構築するかは大きな課題となる。また、同一

形状の收音装置であっても、マイクロホンの設置位置により得られる音信号が異なるため、マイクロホンの位置の検討も重要である。

そこで、様々な形状の收音装置を計算機上で構築し、それぞれの装置に設置したマイクロホンから得られる音信号の特性を検討した。結果の一部を図 1 に示す。最終的に、最も基本的な形状として球形をベースとし、以下に示す 3 項目にしたがって收音装置の形状を決定するとよいことが明らかとなった。

- 1) 收音装置に配置するマイクロホンの数を十分に狭くする。
- 2) 使用する HRTF の間隔 (制御音源方向の間隔) を、5 度間隔以内にする。
- 3) 收音装置の大きさを、合成対象となる聴取者の頭部の大きさと同じにする。

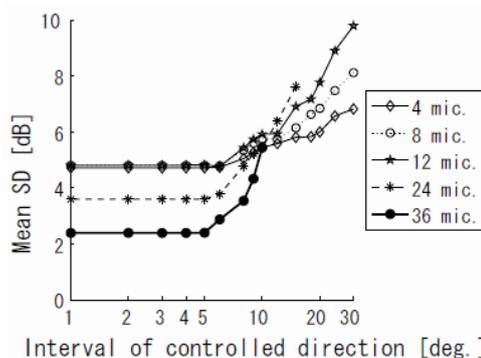


図 1 再現音空間精度にマイクロホン数と制御音源方向数が及ぼす影響

2. 2. 收音装置により得られた音情報のネットワーク配信手法の検討

マイクロホンの個数が確定したことからそれに基づきデータ量圧縮手法の適用可能性を検討した。隣接するマイクロホン間の収録信号の相関を調べたところ、音源方向からマイクロホンまでの距離差が小さいような状況では相関係数が 0.99 程度であるのに対し、音源の位置が変化すると相関値も大きく変化することが明らかとなった。したがって、組み合わせるマイクロホンのペアを考慮することで高い圧縮効率での情報圧縮の可能性が示された。一方で、MP3 等の単純な音信号圧縮処理を個々のマイクロホンから得られた信号 (750 kbps) に対して行った場合、かなりの情報圧縮が可能なもの、位相情報がくずれるため、再現音空間の劣化の程度を考慮する必要が示された。

2. 3. バーチャル音空間創成システムの構築

理論的な検討結果に基づき実際にシステム試作を行った。できるだけ面密度が一定になるようなマイクロホン配置として、直径 17 cm の球体表面上に正二十面体を規範として 252 個のマイクロホンを設置した。構築した收音装置による音空間再現精度を計算機シミュレーションにより検証 (図 2) し、特に 10 kHz までであれば、高い精度で音空間を再現できることが明らかとなった。

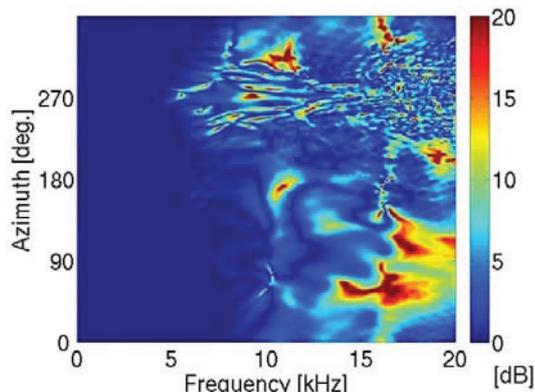


図 2 構築した收音装置を用いて再現される音空間の精度

この試作された收音装置を用いて実際にシステム全体を構築した。試作された收音装置からは、量子化ビット数 16 bit、サンプリング周波数 48 kHz の音信号を出力する。この信号をリアルタイムで周波数分析し、各聴取者に適した重み係数 (複素数) をかけ、最終的には両耳用 2 ch の量子化ビット数 16 bit、サンプリング周波数 48 kHz として出力することになる。必要となる演算量が大きくなることから、システムの演算性能と処理すべき演算量の検討に基づき、日本 NI 社の FPGA ボード PXIe-7965R を 3 枚使用して、実装することとした。構築したシステムの全体像と收音装置の拡大図を図 3 に示す。

構築システムでは、各 FPGA ボードで以下のようにタスク分配をして処理を行った。FPGA ボード 1 では、球状マイクロホンの 252 ch の音信号の收音処理のみを担当する。FPGA ボード 2 では、FPGA ボード 1 で得られた音データを受け取り、各チャンネルで音データに対して 512 ポイントの窓かけし、FFT を行って周波数成分に分解する。この際には、音データを 50 % ずつオーバーラップさせることで、連続した音データを不整合なく処理している。FPGA ボード 3 では、事前に算出した各聴取者に適した重み係数を FPGA ボード 2 で得られた周波数成分にかけて足しあわせ、2 ch 用の信号にして逆 FFT し、聴取者に提示する。なお、FPGA ボード 3 では、あらかじめ頭部

位置に応じた複数の重み係数を格納しておき、聴取者の頭部運動に応じた重み係数を切り替えることにより、頭部運動に対応した最適な音データの出力が可能となる。

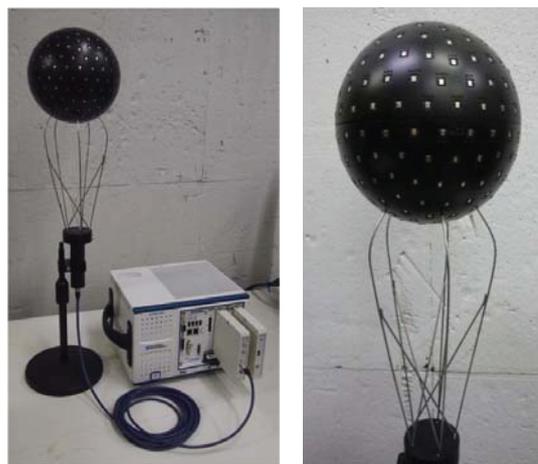


図 3 構築したバーチャル音空間創成システムの全体像 (左) と收音装置として構築した 252 ch マイクロホンアレイ (右)

3. むすび

本研究で実現したシステムは、遠隔地にいる不特定多数の聴取者に対して、空間だけでなく時間をも超えて高精度な音空間情報を再現できるという、既存の收音技術を置き換えるポテンシャルを秘めているものである。

今後はシステムの精緻化、高速化を進めることと合わせ、ネットワーク上で信号を遅延なく通信するといった実装の部分を行うことで、近未来には社会実験などにより大規模なシステム評価が行えるものと確信している。

【誌上発表リスト】

- [1] S. Sakamoto, R. Kadoi, S. Hongo and Y. Suzuki, "SENZI and ASURA: New high-precision sound-space sensing systems based on symmetrically arranged numerous microphones," Proc. Second International Symposium on Universal Communication (ISUC2008), 429-434 (2008 年 12 月 15~16 日)
- [2] S. Sakamoto, S. Hongo and Y. Suzuki, "A new comprehensive 3D sound-space sensing and reproduction method by a symmetrically and densely arranged microphone array" (invited lecture), Proc. the 10th Western Pacific Acoustics Conference (WESPAC X), 0245 (2009 年 9 月 21 日)
- [3] S. Sakamoto, J. Kodama, S. Hongo, T. Okamoto, Y. Iwaya and Y. Suzuki, "Effects of microphone arrangements on the accuracy of a spherical microphone array (SENZI) in acquiring high-definition 3D sound space information", Book Chapter of International Workshop on the Principles and Applications of Spatial Hearing (IWPASH), 314-323 (2011 年 3 月)

【受賞リスト】

- [1] 松永純平, 日本 NI アプリケーションコンテスト 2011 学生部門優秀賞, "252 ch 球状マイクロホンアレイを用いた高精度收音システム", 2011 年 11 月 29 日

【本研究開発課題を掲載したホームページ】

<http://www.ais.riec.tohoku.ac.jp/lab/budget/index-j.htm>
1