

K-040

高詳細音響属性の抽出による音空間編集システム構築に向けて

Toward an editable audio-space system using high-resolution sound properties

岡本 拓磨^{*†}, 岩谷 幸雄^{*‡}, 鈴木 陽一^{*‡}Takuma Okamoto^{*†}, Yukio Iwaya^{*‡} and Yōiti Suzuki^{*‡}

1 はじめに

近い将来のセンサネットワークでは、多数のマイクロホンをいたるところに配置することが可能となり音空間を丸ごと集音できることが可能になると考えられ、この利点を生かした音場提示手法が期待される。

音波は原音たる時間波形が音圧として音源から放射され、壁等による反射や回折を繰り返しながら音空間を形成する。これらの物理現象を厳密に再生する方法の一つとして、Kirchhoff-Helmholtz(KH)型音場再現[1]がある。これは、音場を囲む境界面における音圧と音圧勾配を記録、再生することで境界面内の音場が補償されるという原理に基づく。しかし、KH法では、音源や反射特性などは有機的に結合したままであり、音空間をコンテンツとして考えた場合、その特性を変更することは難しい。

これに対して我々は、超臨場感音空間の創成のため、現実の音空間はもちろんのこと、空間属性を自由に変更できるエディタブルオーディオシステムを提案する。目指すエディタブルオーディオシステムとは、収録段階で音空間を属性に分解し、その属性を自由に編集可能にするシステムである。具体的な属性としては、音源位置、原音波形、放射指向特性、反射音、残響などの音場と音源の構成要素を考えている。

このシステムが実現すれば、音空間の任意の音属性を忠実に提示することや、他の音空間の属性と融合した新しい音空間の特性をもった提示も可能となり、従来の音場再現では表現し得なかった臨場感や迫真性を与える音空間提示メディアが実現できる。

本稿では、音場・音源の属性を抽出する技術として音源位置情報、原音波形、放射指向特性抽出の概要を述べる。

2 包囲型マイクロホンアレイの構築

音空間の音情報を記録するために、床を除く壁面と天井の5面に格子状(50 cm 間隔)に計157 chのマイクロホン(B&K Type 4951)で構成した包囲型マイクロホンアレイを構築した(Fig 1)。各壁面とマイク

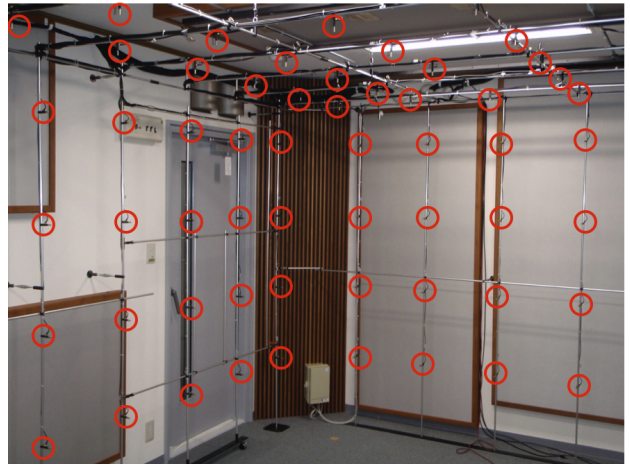


Fig 1: 包囲型マイクロホンアレイの概観 : マイクロホン

ロホン間は30 cmの間隔を持つ。クロックジェネレータ、A/D&D/A、4台のマシンをシンクロさせ、48 kHz、16 bit リニア PCMの157 ch同時収録を実現している。このアレイにより、室内の音情報を余すところなく収録可能となる。

3 音空間属性値の抽出

3.1 音源属性抽出の原理

包囲型マイクロホンアレイの各マイクロホン i で時刻 n に記録された音信号を $x_i(n)$ とする。この音信号 $x_i(n)$ と各マイクロホンの位置情報を用いて音場と音源の属性を抽出する。エディタブルオーディオシステムとして自由に属性値として編集を可能とするためには、記録された音信号 $x_i(n)$ から音源位置情報を3次元位置座標 (x, y, z) 、原音情報を信号 $s(n)$ 、音源から各マイクロホンまでの放射指向特性を伝達関数(=インパルス応答) $h_{Di}(n)$ 、反射音成分を室内インパルス応答 $h_{Ri}(n)$ として抽出することが必要であり、本研究の目的である。本研究では放射指向特性 $h_{Di}(n)$ を考慮するため、音源と各マイクロホン間の伝達関数 $h_i(n)$ を $h_i(n) = h_{Di}(n) + h_{Ri}(n)$ としてモデル化している。

*東北大学電気通信研究所

†東北大学大学院工学研究科

‡東北大学大学院情報科学研究科

3.2 音源位置情報の推定

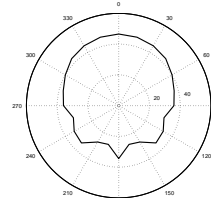
室内空間における音源位置推定では、壁面からの反射、回折音が推定精度を劣化させるという問題がある。到来方向推定においてはその対策法として空間平均化法 [2] が提案されている。我々は Multiple Signal Classification (MUSIC)[3] と空間平均化法を組み合わせた反射音に頑健な Re-Arrangement and Pre-smoothing for MUSIC (RAP-MUSIC)[4] を提案し、高精度な 3 次元音源位置座標抽出に成功した。

3.3 原音波形の推定

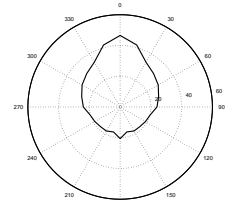
室内で収録した音信号から残響を除去し、高精度に原音を抽出する従来方法として線形予測ブラインド残響除去である Linear-predictive Multi-input Equalization (LIME) アルゴリズム [5] があるが、高いサンプリング周波数では正常に動作しないという問題があった。これに対して我々は LIME を改良し、広帯域信号でも原音情報の抽出を可能とする White-LIME[6] を考案した。

3.4 放射指向特性の推定

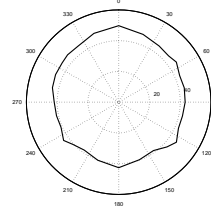
抽出した位置情報 ($\hat{x}, \hat{y}, \hat{z}$) と原音情報 $\hat{s}(n)$ を用いて放射指向特性の抽出を行う。抽出の原理は、まず受信信号 x_n と $\hat{s}(n)$ の逆畳み込みから音源とマイクロホン間のインパルス応答 $\hat{h}_i(n)$ の抽出を行う。その上で、抽出したインパルス応答のうち、初期反射までの時間のインパルス応答を切り出し、放射指向特性 $\hat{h}_{D_i}(n)$ として抽出する。包囲型マイクロホン 157 ch のうち、床からの高さ 1.0 m にあるマイクロホン 28 チャンネルのインパルス応答を原音 (音楽信号: サンプリング周波数 44.1 kHz, 2.7 s) に畳み込んだ観測信号から White-LIME によって原音を抽出した。抽出した原音の SDR は 57.3 dB と非常に高い値であった。抽出した原音を用いて各チャンネルのインパルス応答を抽出したところ、SDR の平均値は 62.6 dB と極めて高い値であり、インパルス応答をほぼ正確に抽出することができた。抽出したインパルス応答から放射指向特性の抽出を行う。実験環境では壁面とマイクロホン間は 30 cm であるため、サンプリング周波数 44.1 kHz では抽出したインパルス応答の約 78 タップを切り出して、距離補正を施す。無響室での測定結果、室内での測定結果、抽出した放射指向特性の周波数応答を 1/3 オクターブバンド解析した結果を Fig 2 に示す。結果より、切り出したインパルス応答により放射指向特性を模擬できていることが確認できる。今後はロバストな原音抽出の検討を行い、実環境での放射指向特性抽出を実現する予定である。



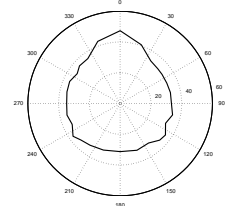
(a1) 無響室での測定結果



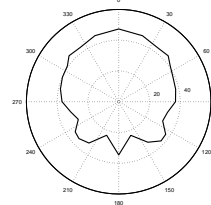
(a2) 無響室での測定結果



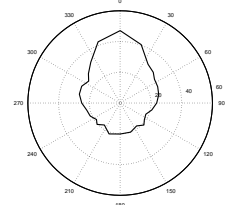
(b1) 室内での測定結果



(b2) 室内での測定結果



(c1) 抽出結果



(c2) 抽出結果

Fig 2: 方向ごとの放射指向特性の違いによるスピーカの音圧分布 (a1 ~ c1 : 2k Hz, a2 ~ c2 : 8 kHz)

4 おわりに

音源属性を自由に変更、編集、提示可能であるエディタブルオーディオシステムの提案と、包囲型マイクロホンを用いた音源属性の抽出に関する概要、放射指向特性の抽出に関する検討について述べた。

謝辞

本研究は東北大学大学院工学研究科電気情報系 GCOE プログラム (CERIES) の助成による。エディタブルオーディオについて貴重なコメントをいただいた元東北大学 (現 JAIST) の宮内良太博士に感謝する。

参考文献

- [1] *ACUSTICA*, 85(1), 78-87, (1999).
- [2] *IEEE Trans. ASSP*, 33(4), 806-811, (1985).
- [3] *J. Acoust. Soc. Am.*, 78(5), 1508-1518, (1985).
- [4] *Acoust. Sci. & Tech.*, 28(3), 181-189, (2007).
- [5] *IEEE Trans. ASLP*, 15(2), 430-440, (2007).
- [6] 音講論 (春), 675-676, (2009).