

日本音響学会 2017 年 秋季研究発表会に参加して

安 枝 和 哉
Kazuya YASUEDA

情報メディア学専攻博士課程 2 年

1. はじめに

私は、2017 年 9 月 25 日から 27 日に愛媛大学で開催された日本音響学会 2017 年秋季研究発表会に参加し、「複数スピーカアレーを用いた分解信号の再合成によるエリア再生」という題目で発表を行った。

2. 研究背景と目的

近年、音声を特定の領域のみに再生する技術の研究が進められている。本研究では、入力信号を複数の信号に分解し、空間上で再合成することにより、エリア再生を実現する手法を提案する。図 1 に局所的再合成のイメージ図を示す。

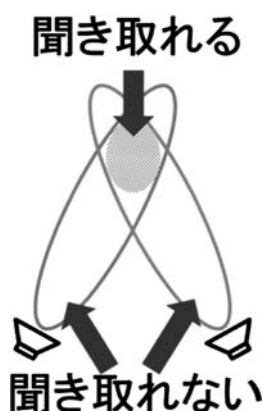


図 1 局所的再合成のイメージ図

3. 入力信号の分解

3.1 ランダムベクトル分解法

入力信号を複数のサブ信号に分解する手法として、ランダムベクトル分解法を用いる。入力信号 $s(n)$ は一次独立な雑音 $v(n)$ と係数 a との線型結合

で表現することができる。

$$s(n) = \sum_{k=1}^K a_k v_k(n) \quad \cdots (1)$$

ここで、 K はフレーム長、 n はサンプルである。 v を行列とし、その逆行列を s にかけることで a を求めることができる。

3.2 最小二乗法の適用と差分信号

ランダムベクトル分解法を用いて作成したサブ信号は使用するランダムベクトルに依存するため、サブ信号のパワーが大きくなることがある。この対策として、式 (2) の最小二乗法を用いて、少ない本数のランダムベクトルを用いる手法が提案されている。

$$\mathbf{a} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{s} \quad \cdots (2)$$

$$\begin{cases} s_1(n) = \sum_{k=1}^K a_k v_k(n) \\ s_2(n) = \sum_{k=\frac{K}{2}+1}^K a_k v_k(n) \end{cases} \quad \cdots (3)$$

しかし、最小二乗法を適用することで、誤差が生じ、サブ信号を再合成しても元の入力信号に戻らないことが考えられる。また、実環境での実装を考えると、少しの設置誤差が音質に大きな影響を及ぼすことが考えられる。そこで、入力信号から 2 つのサブ信号 s_1, s_2 を減算した差分信号 s_3 を、3 つめのサブ信号として再生する。

$$s_3(n) = s(n) - (s_1(n) - s_2(n)) \quad \cdots (4)$$

4. サブ信号の再生

作成したサブ信号を空間上の特定のエリアのみで再合成することで、特定のエリアへの再生を実現する。本研究ではサブ信号を空間上で再合成するための再生手法として、遅延和法 (DS) と多点制御法 (MPCM) を用いた。

4.1 複数スピーカアレーの配置

図 2 にスピーカアレーの配置を示す。左右 2 本のスピーカアレーでそれぞれランダムベクトル分解により分解した信号 $s_1(n), s_2(n)$ を再生し、中央のスピーカアレーで差分信号 $s_3(n)$ を再生する。1 つのスピーカアレーは 16 ch のスピーカで構

成される。両側のスピーカアレーの角度は30 [deg] とし、中央のスピーカアレーの中心を $x=0$, $y=0$ とすると、 $x=0$, $y=1.3995$ の点で再合成されることが期待できる。

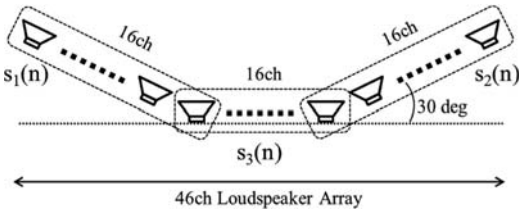


図2 スピーカ配置と入力信号

4.2 多点制御法によるエリア再生

多点制御法は、空間上に設置した制御点での音圧特性を伝達関数の逆フィルタを構成することにより各制御点での音圧を制御する。本研究では、スピーカ数と制御点数が一致しないため、伝達関数の逆行列は打ち切り特異値分解による一般化逆行列として求めた。

図3に中央のアレーの制御点配置を示す。3つのアレーの中心線が交わる位置を応答制御点とし、抑圧制御点の間隔は0.02 [m] とした。制御点数は162点である。左右のアレーも同様の制御点配置として再生を行った。スピーカ間隔は0.05 [m] である。

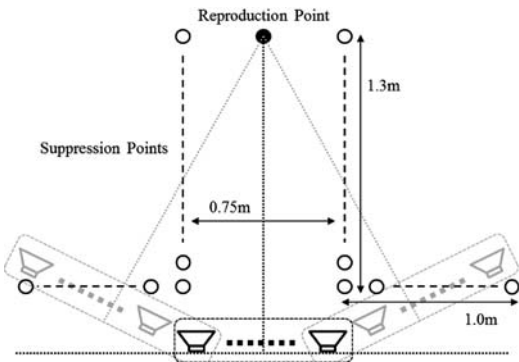


図3 多点制御法の制御点配置

5. 計算機シミュレーション

計算機シミュレーションを用いて、各手法で再生

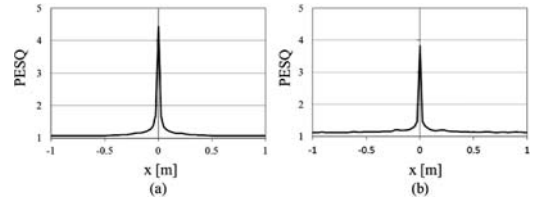


図4 PESQ 値 ((a) DS (b) MPCM)

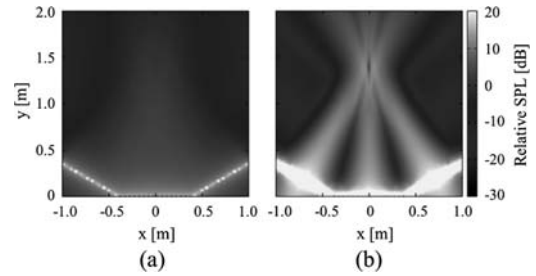


図5 相対音圧レベル ((a) DS (b) MPCM)

したときの品質を PESQ (Perceptual Evaluation of Speech Quality) によって評価した。また、相対音圧レベルにより各再生手法の抑圧性能を確認した。

図4に $y=1.4$ [m] 地点の PESQ 値、図5に相対音圧レベルを示す。図4の横軸は x 座標であり、 $x=0$ がスピーカアレーの中心である。

図4より、どちらの手法においても、 $x=0$ において PESQ 値が高くなっており、サブ信号が再合成され、音質が良くなっていることが確認できる。また、 $x=0$ 以外の点では PESQ 値が低くなっており、交点のみで再合成されていることが確認できる。

音圧レベルについては、図5より DS は空間内においてほぼ一様な音圧で再生されている。一方、MPCM では、抑圧制御点を設置した横方向において音圧が低下していることが確認できる。

6. おわりに

発表に参加し、多くの方々から意見を頂き、大変参考になりました。研究や発表に対して多大なご指導を頂いた片岡章俊教授に深く感謝いたします。