

雑音除去音声の品質評価*

- 線形遅延和アレーにおけるマイクロホン配置の検討 -

◎吉國信太郎 水町光徳 二矢田勝行 (九工大)

1. はじめに

雑音環境下で目的の方向の音のみを取り出す手法として、二次音圧傾度型や干渉管型の指向性マイクロホンを用いる方法があるが、高度な設計ノウハウや加工技術が必要でコストがかかることや、一度設計すると特性が固定されてしまうという問題がある。更に、目的音方向を変化させるためには、受音器の向きを移動させる必要がある。そこで、本研究では、線形マイクロホンアレーを使用して指向性を制御することにする。また、人間の視覚(目)は正面を向いており、意識も自然と正面方向に集中する。そこで、正面方向の音を聴くことを目的としたノイズキャンセリングヘッドホンを作成することを目標とし、本稿では、遅延和アレーの目的音方向を正面方向に固定する。

本研究は、線形遅延和アレーにおけるマイクロホンを不等間隔で配置することにより、等間隔に配置した場合と比較し、客観的及び主観的な歪み低減量にどの程度差異があるのかを評価することを目的とする。評価方法は、客観評価尺度として信号対雑音比(SNR)、主観評価尺度として5段階の平均オピニオン評点(MOS)を採用する。

2. 線形遅延和アレー

2.1 遅延和(Delay-and-Sum;DS)アレーの原理

Fig.1 に、DS アレー処理の原理を示す[1]。Fig.1 に示した各遅延器は、次式で表わされる遅延量 D_i を各受信信号に与える。

$$D_i = D_0 - (i - 1)\tau_L \quad i=1,2,\dots,M \quad (1)$$

$$\tau_L = (d \sin \theta_L) / c \quad (2)$$

但し、 θ_L は指向特性の主ローブ方向を表し、以降目的方向と呼ぶ。 d はマイクロホン間隔、

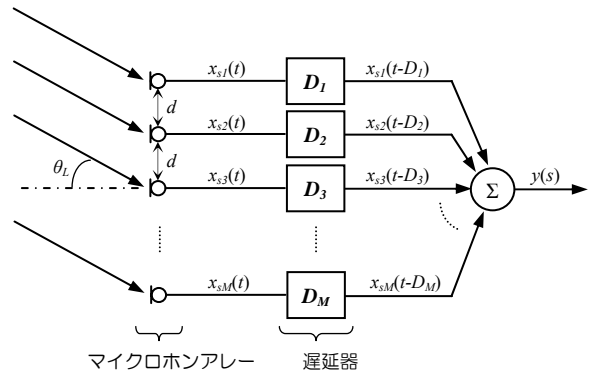


Fig.1 線形遅延和アレー処理

c は音速、 D_0 は固定遅延量である。

このとき、目的信号が θ_L 方向から到来していると仮定すると、各マイクロホンで受信された各受信信号 $x_{si}(t)$ は、

$$x_{si}(t) = x_{s1}(t - (i - 1)\tau_L) \quad i=1,2,\dots,M \quad (3)$$

と表わされ、遅延 D_i を付加した信号 $x_{si}(t - D_i)$ は、式(1)(3)より、

$$\begin{aligned} x_{si}(t - D_i) &= x_{s1}(t - (i - 1)\tau_L - D_i) \\ &= x_{s1}(t - D_0) \end{aligned} \quad (4)$$

となる。この式から分かるように、各遅延器からの出力は全て同一の信号になっており、この遅延付加操作により θ_L 方向から到来する信号の位相差が補正され、同相化されることになる。そして、これらの信号を加算することで、目的方向 θ_L から到来する音は強調される。一方、 θ_L とは異なった方向 θ_N から到来する音は、 τ_L とは異なる時間差 τ_N を持って受信される。従って、式(1)で表される D_i による遅延付加操作では、信号は同相化されず、これを加算しても強調効果は小さい。結果として、遅延和アレーは目的方向 θ_L に対して感度が高い指向特性を形成する。

なお、本研究では、目的信号を正面方向に固定しているため、各受信信号に遅延を付加する必要はない。

2.2 等間隔 DS アレーの指向特性

Fig.2 に、目的方向 $\theta_L = 0^\circ$ 、マイクロホン間隔 $d = 5.55\text{cm}$ 、マイクロホン数 $M = 4$ の場合の等間隔 DS アレーの指向特性を示す。

Fig.2 から分かるように、上記の設定では、 $\pm 30^\circ$ より外側で、周波数が 5kHz 以上の領域では空間エイリアシングが発生してしまい、目的方向以外の信号まで強調されてしまう。マイクロホン間隔を狭くすると、空間エイリアシングが発生する周波数は上昇するが、指向性が緩く(主ローブ幅が大きく)になってしまう。このように、エイリアシングが起こらない周波数と主ローブ幅の間にはトレードオフの関係が成り立ち、単一の等間隔 DS アレーではそれらを両立させることは難しい。

3. 3ⁿ型線形遅延和アレー

本研究では、2.2 で述べた問題点を解決すべく、マイクロホンを 3^n ($n=0,1,2$) 間隔で配置し、周波数帯域を複数に分割して、使用するマイクロホンを切り替える方式の 3^n 間隔 DS アレーを提案する。

Fig.3 に、本研究で提案する 8ch 3^n 間隔 DS アレー(8ch 3^n アレー)、比較対象とする 4ch 等間隔 DS アレー(4ch DS アレー)、8ch 等間隔 DS アレー(8ch DS アレー)、のマイクロホン配置図を示す。次に、8ch 3^n アレーにおける周波数帯域の分割方法を Tab.1 に、指向特性を Fig.4 に示す。

Fig.4 から分かるように、周波数帯域を複数に分割することにより、 $\pm 30^\circ$ より外側で、周波数が 5kHz 以上の領域で空間エイリアシングを発生させることなく、目的方向の信号のみを強調することが可能となる。本研究では、正面方向の音声信号を目的信号に設定しており、正面方向以外の雑音を除去することを目的としたため、このような配置を選択した。また、2ⁿ(オクターブ)間隔[2]では周波数帯域を分割する際、音声信号の周波数帯での分割数が多くなってしまい、音声信号をターゲットにした場合では、 3^n 間隔の方が性能が良いという仮定の下、アレー間隔に 3^n を採用した。

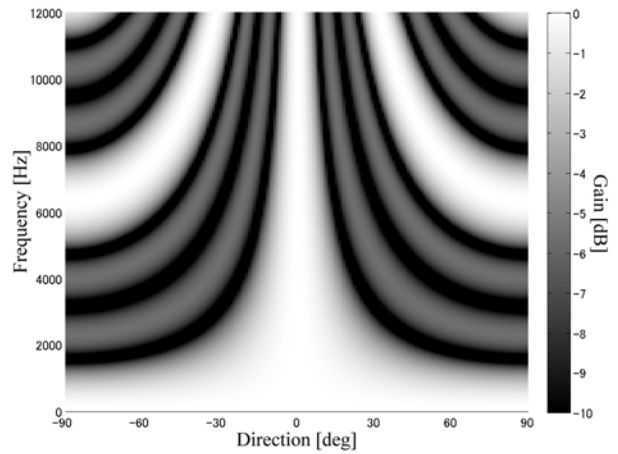


Fig.2 遅延和アレーの指向特性 (4ch 等間隔)

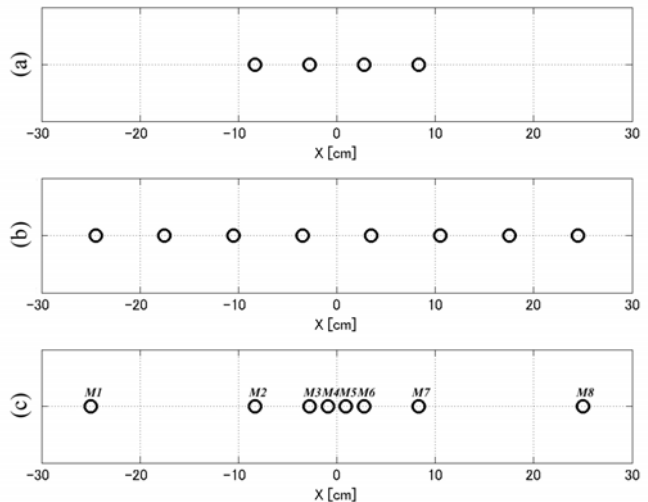


Fig.3 線形遅延和アレーのマイクロホン配置

- (a) マイクロホン間隔 5.55cm 等間隔配置、マイクロホン数 4
- (b) マイクロホン間隔 7cm 等間隔配置、マイクロホン数 8
- (c) マイクロホン間隔 3^n 型不等間隔配置、マイクロホン数 8

Tab.1 周波数帯域の分割

周波数帯域	マイクロホン間隔	使用するマイクロホン
$\sim 1700\text{Hz}$	16.65cm	$M1, M2, M7, M8$
$1700\text{Hz} \sim 5000\text{Hz}$	5.55cm	$M2, M3, M6, M7$
$5000\text{Hz} \sim$	1.85cm	$M3, M4, M5, M6$

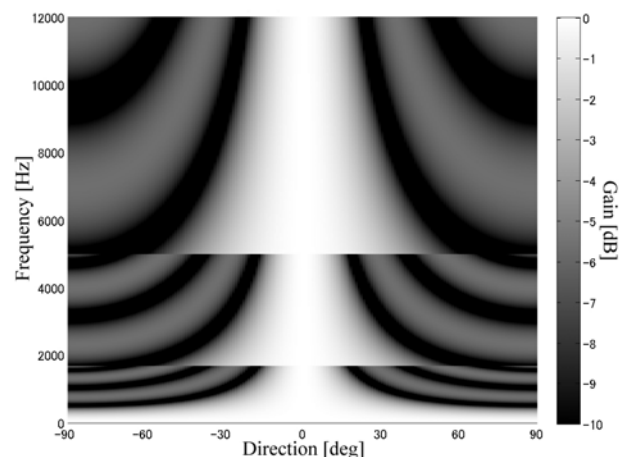


Fig.4 遅延和アレーの指向特性 (8ch 3^n 型)

4. 雑音除去音声の歪み評価

4.1 実験条件

本研究では、Fig.3 に示す 3 種類の線形アレーの雑音除去能力を比較するため、Tab.2 の実験条件の下、処理前と各アレーで処理後の信号を客観的及び主観的に評価する。評価方法は、客観評価尺度として信号対雑音比(SNR)、主観評価尺度として 5 段階の平均オピニオン評点(MOS)を採用した[3]。

SNR による評価では、各アレーでの処理後信号の SNR から処理前信号の SNR を引いた SNR 改善量を求めた。

主観評価では、雑音信号が 60°方向から到来する場合に、被験者へ各刺激音をランダムに提示して、「1：非常に悪い～5：非常に良い」の 5 段階 MOS を求めた。刺激音は、(未処理信号+各アレーでの処理後信号)×雑音の種類×観測信号の SNR で計 80 種類である。被験者は各刺激音を 6 回ずつ評価したため、各被験者の実験総数は 480 である。また、被験者は本研究室の男性 15 名である。

4.2 実験結果

・客観評価 (SNR)

雑音信号が工場 A の場合で観測信号の SNR が-10[dB]の場合の各アレーによる SNR 改善量を Fig.5 に示す。

Fig.5 より、8ch 3ⁿアレーは雑音信号の入力角度が 0°～20°の場合、雑音除去能力は他のアレーに劣るが、入力角度が増すにつれて他のアレーよりも高くなることが分かった。次に、主観評価による歪み評価結果と比較するために、雑音信号の入力角度が 60°の場合の各アレーでの SNR 改善量を雑音の種類ごとに Fig.6 に示す。Fig.6 より、SNR 評価では本研究で提案した 8ch 3ⁿアレーは、等間隔 DS アレー以上の雑音除去能力があることが分かった。

・主観評価 (MOS)

平均オピニオン評点で SNR ごとに求めた MOS の平均値を Fig.7 に示す。Fig.7 より、MOS 評点では、工場 C と駅以外の

Tab.2 実験条件

目的信号	数字読み上げ音声 (英語・女性)
雑音信号 (特徴・エネルギーが集中している帯域)	工場×3種類 A: 定常・13kHz 以下 B: 定常・3kHz 以下 C: 非定常・1kHz 以下
	展示会場×1種類 (定常・人の声・500Hz 以下)
	駅のホーム×1種類 (非定常・人の声・1kHz 以下)
目的信号の方向	正面方向 (0°)
雑音信号の方向	0°～90° (10°きざみ)
観測信号の SNR	-5,-10,-15,-20 [dB]

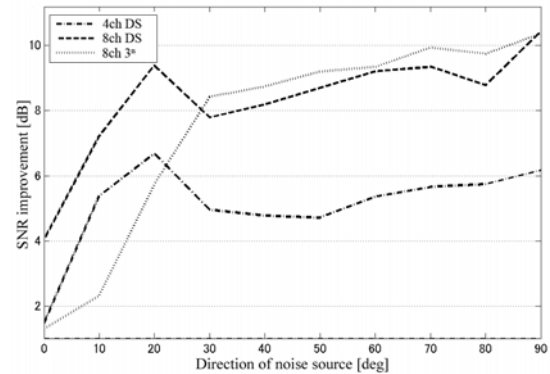


Fig.5 各アレーでの SNR 改善量 (工場 A)

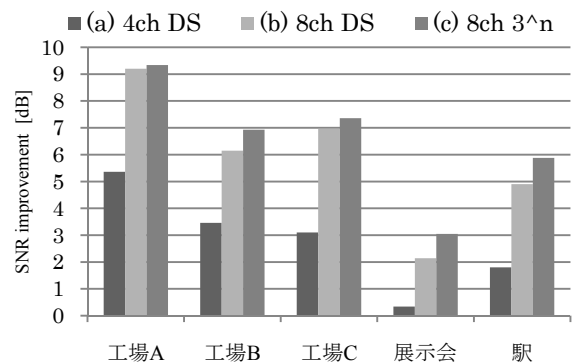


Fig.6 各アレーにおける SNR 改善量 (60°の場合)

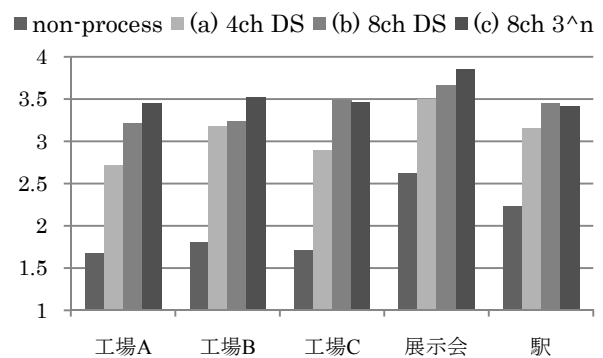


Fig.7 MOS 評点の結果

雑音に対して、8ch 3ⁿアレー処理後の信号の評点が高いという結果が得られた。雑音工場 C と駅の場合、8ch 3ⁿアレー処理後の信号の評点は 4ch DS アレーより高いが、8ch DS アレーとは同等であると分かった。この原因については、4.3 の考察で詳しく述べることにする。

4.3 考察

・SNR改善量のみを考慮した場合

等間隔アレーには、「①マイクロホン間隔を広くとると、主ローブ幅は狭くなるが、信号の入力角度が大きくなると空間エイリアシングが発生してしまう。」一方で、「②マイクロホン間隔を狭くとると、空間エイリアシングが発生する周波数は高くなるが、主ローブの幅は広がってしまう。」という問題点がある。本研究で提案した 8ch 3rd アレーはそれらの問題点を解決するべく、各周波数帯域で使用するマイクロホンを切り替える方式を採用した。そのために雑音の入力角度が浅い場合、十分に雑音を抑圧することができない。したがって、Fig.5 において雑音信号の入力角度が 0°~20°の場合に SNR があまり改善されなかった。しかし本研究では、正面方向以外の雑音を除去することを目的としている。将来ヘッドホンで利用する際、精度が良すぎると少し視線をずらしただけで目的信号まで聞こえなくなってしまうのはいけないので、目で捉えることが出来る範囲の雑音はあまり除去できなくても良い。

また、雑音信号の入力角度が 60°の場合、雑音の種類に関係なく 8ch 3rd アレーの雑音除去能力が一番高いという結果が得られた。これは、周波数が 12kHz 以下の領域で雑音信号の入力角度によってエイリアシングが発生しないように配置を工夫したためであり、8ch n³ アレーの有効性が証明できた。

・SNR改善量とMOS評点を考慮した場合

次に、SNR の改善量と MOS 評点の関係を比較する。雑音が工場 A の場合、SNR 改善量は 8ch DS アレーと 8ch 3rd アレーに有意な差は見られないが、MOS 評点では 8ch 3rd アレーのほうが評点は高かった。一方、工場 C の場合、SNR 改善量は 8ch 3rd アレーの方が高いにも関わらず、MOS 評点では 8ch DS アレーと差が見られなかった。工場 A は白色性ノイズであるが工場 C は 1kHz 以下にエネルギーが集中しているノイズでさらに非定常性のノイズであった。目的信号である女性の声は 3kHz 以下

にエネルギーが集中しており、工場 C の方が低い周波数に大きなパワーを持っている。更に非定常性ノイズであることも考慮すると、工場 C の場合に MOS 評点において、8ch 3rd アレーと 8ch DS アレーに差が見られなかったのは、同時マスキングと継時マスキングの影響によるものであると思われる。また、駅のノイズは、非定常であり、1kHz 以下にエネルギーが集中していて、更に人の声が多く含まれていた。このためにマスキングの影響を受けて、MOS 評点において、8ch 3rd アレーと 8ch DS アレーに差が見られなかったと思われる。このように目的信号と雑音信号の周波数領域のエネルギー分布が似ている場合、SNR は改善できていても人が聞いてみると差が感じられない場合もあるということがわかった。

以上より、雑音の種類によって SNR は改善できても人の耳で判断するとあまり改善が見られない場合があるので、そこをいかに克服するかが今後の課題である。

5. まとめ

本研究で提案した 8ch 3rd アレーは、SNR 改善量だけで判断するとその有効性は証明できたが、MOS 評点まで含めるとまだまだ改善の余地がある。目的信号を音声とし、人が聴いて判断する場合、人間の聴覚的な特徴まで考慮してビームフォーマを設計する必要がある。

今後の課題として、マイクロホン配置の更なる検討、ビームフォーマのアルゴリズムの改善、2nd, 4th 間隔 DS アレーとの性能の評価などを行っていきたい。

参考文献

- [1] 大賀寿朗 他, 音響システムとデジタル信号処理, 電子情報通信学会, 1995.
- [2] J. L. Flanagan et. al., "Autodirective Microphone Systems," *Acoustica*, vol. 73, pp. 58-71, 1991.
- [3] 伊藤憲三 他, "音声のデジタル符号化方式の客観的品質評価尺度の検討," *信学論(A)*, vol. J66-A, no. 3, pp. 274-281, 1983.