

Title	多点受信信号を用いた残響除去法に関する基礎的研究
Author(s)	水町, 光徳
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1039">http://hdl.handle.net/10119/1039</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

# 多点受信信号を用いた残響除去法に関する基礎的研究

水町 光徳

北陸先端科学技術大学院大学 情報科学研究科

1997年2月14日

キーワード： 残響除去、マイクロホンアレー、スペクトルサブトラクション.

## 1 はじめに

現在、音声認識システムの実用化には大きな関心が寄せられている。しかし、雑音や残響が存在する実使用環境においては、現在の音声認識装置では、その性能が大幅に低下することが報告されている [1]。

よって、本研究では、実環境における音声認識装置の実用化を支援できるような、音声認識のプリプロセッサとしての残響除去法を提案する。

## 2 残響除去アルゴリズム

本研究では、3本の無指向性マイクロホンを等間隔直線配置したマイクロホンアレーを使用する。そして、まず初めに、3本のマイクロホンのうち両端の2本のマイクロホンで受信した信号を利用し、中央のマイクロホン位置における残響成分を解析的に推定する。

次に、3本のマイクロホンのうち残りの1本(中央)のマイクロホンで受信した信号、すなわち、直接音と残響成分との合成信号から、先に求めた残響成分のみを減算することにより残響除去を行なう。

### 2.1 残響除去の推定について

本研究では、残響成分の抽出に、今田らによる2波形分離アルゴリズム [2] を応用する。その2波形分離アルゴリズムの特徴は、正中面方向とそれ以外の1方向から到来する2つ

の信号しか存在しない場合に、2本のマイクロホンで構成されるマイクロホンアレーで受音した信号を用いて、正中面方向から到来する信号を消去し、もう一方の信号のみを抽出することである。

そこで、2波形分離アルゴリズムにより構成されるマイクロホンアレーの指向特性を調査した結果、1 kHz程度までの低中周波数成分に関しては、正中面方向付近から到来する信号以外をほぼ等感度で抽出できることが分かった。

よって、本研究では、正中面方向から直接音が到来すると仮定した場合、それ以外の全方向から到来するであろう残響成分の抽出を行なうための残響成分抽出アルゴリズムを作成した。それは、今田らの2波形分離アルゴリズムを改良したもので、その原理については以下に解説する。

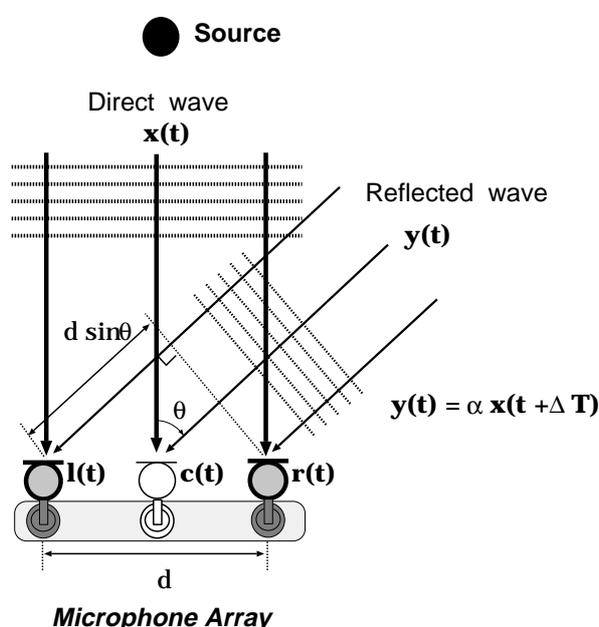


図 1: 信号抽出に関する解説図

まず、図 1 に示すように、正中面方向から直接音  $x(t)$ 、正中面方向に対して角度  $\theta$  方向に残響成分  $y(t)$  が到来するとした場合、マイクロホンアレーの左端、中央、右端のマイクロホンで受音される信号をそれぞれ  $l(t)$ 、 $c(t)$ 、 $r(t)$  とすると

$$\begin{cases} l(t) = x(t) + y(t - \delta) \\ c(t) = x(t) + y(t) \\ r(t) = x(t) + y(t + \delta) \end{cases} \quad (1)$$

となる。ここで、上式の  $\delta$  は、抽出したい信号  $y(t)$  が左右両端のマイクロホンに到来する際の時間差の  $1/2$  である。

ここで、抽出したい信号  $y(t)$  と左右両端のマイクロホンで受音できる信号  $l(t)$ 、 $r(t)$  との関係を、正中面方向から到来する信号  $x(t)$  を用いずに表現することを考えて

$$g(t) = \{l(t + \tau) - l(t - \tau)\} - \{r(t + \tau) - r(t - \tau)\} \quad (2)$$

と  $g(t)$  を定める。そして、 $y(t)$ 、 $g(t)$  をフーリエ変換した信号をそれぞれ  $Y(\omega)$ 、 $G(\omega)$  とすると

$$G(\omega) = 4Y(\omega) \sin \omega \delta \sin \omega \tau \quad (3)$$

となり、実際の信号の到来方向（時間差  $2\delta$  方向）を知ることができれば、マイクロホンアレーの焦点（抽出したい信号が到来する方向、時間差  $2\tau$  方向）をそれと一致させることにより、抽出したい目的信号  $y(t)$  の中央のマイクロホン位置における推定値  $\hat{y}(t)$  は

$$\hat{y}(t) = \begin{cases} \text{IFFT} \left[ \frac{G(\omega)}{4 \sin^2 \omega \tau} \right], & 4 \sin^2 \omega \tau > 0.02 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

により計算できる。

なお、マイクロホンアレーの焦点の決定方法に関しては、2つのマイクロホンで受音した、目的信号の各マイクロホンへの到来時間差に基づいた方向推定法を応用した。

## 2.2 残響成分の除去について

残響成分の抽出の次段階として、直接音と残響成分の合成信号である受音信号から、先に推定した残響成分の減算を考えなければならない。

そこで、本残響除去法が音声認識システムのプリプロセッサを前提としていることから、周波数領域において振幅スペクトルの減算のみを行なう。なお、その演算には Boll によって提案されたスペクトルサブトラクション [3] を利用する。

なお、本研究では、より高精度の残響成分除去を目指し、フレーム処理を導入し、サブトラクション係数に関しては、そのフレームにおける残響成分と前フレームで求めた直接音との相関値を利用して決定するようにした。これにより、音声などの時変性の信号に対する残響除去の精度が向上することが期待できる。

## 3 残響除去実験

残響除去実験を行なうための残響のある音声データとしては、ATR 音声データベースのクリーンな音声データ（単母音 /a/、単語 /buri/）を残響のある実音場にスピーカを用いて放射し、マイクロホンアレーで収音したものを使用した。

そして、残響除去を行なった結果、100 Hz から 6 kHz までの音声帯域におけるスペクトル歪 (SD) がどれくらい改善されたかによって本残響除去法の評価を行なう。そこで、音声区間における SD の平均改善量を表 1 に示している。

表 1: 音声帯域セグメンタル SD の音声区間平均改善量

データ	音声帯域セグメンタル SD の平均改善量
/ a /	3.18 [dB]
/ bunri /	1.29 [dB]

一般に、残響除去は非常に困難な問題であり、王らは残響除去を行なった結果が SD で約 1 dB しか改善されなくても残響除去の効果はあると述べているため [4]、本残響除去法は、音声認識のための残響除去プリプロセッサとしての素質は十分に備えていると考えられる。

## 4 おわりに

本研究では、3チャンネルの等間隔直線配置マイクロホンアレーを用い、その両端のマイクロホン対で受音した信号に含まれる、到来方向に関する情報を利用した残響除去法を提案した。

そして、残響除去実験を行なった結果、本方法は母音区間では有効であると言えるが、単語などの子音区間では、残響除去を行なうとスペクトルの歪みが増すという問題が存在することも分かった。この問題点については、今後、更なる検討が必要である。

## 参考文献

- [1] 飛田、管村: “音声認識における周囲環境の影響,” 日本音響学会誌, vol. 51, no. 4, pp. 331-335 (1995)
- [2] 今田 稔一郎: “複数受音波形による音源分離に関する研究,” JAIST 修士論文 (1996)
- [3] S. F. Boll: “Suppression of Acoustic Noise in Speech Using Spectral Subtraction,” IEEE Trans. ASSP-27 (2) (1979)
- [4] 王、板倉: “多数マイクロホン帯域包絡推定による残響音声の回復 — マイクロホンの配置及び雑音の影響 —,” 音講論集, 1-3-6 (1989.9)