

多重線形部分空間解析法とスペクトル減算法を用いた広帯域雑音除去

Reduction of Broadband Noise in Speech Signals by a Combined Multilinear Subspace Analysis and Spectral Subtraction Method

○吉井 達也¹ 保谷 哲也²
Tatsuya Yoshii¹ Tetsuya Hoya²

Abstract: In this paper, we report the results of simulation study for the denoising of noisy speech signals using a combined N-mode singular value decomposition(N-mode SVD) and the spectral subtraction method based upon minimum statistics(MS).

1. 雑音付加音声におけるテンソルの導入

多重線形部分空間解析法では、テンソルを導入する^[1]。テンソルとは、簡易的にはプログラミング上での多次元配列を指し、通常の行列は2次のテンソルと見なすことができる。本研究では、3次のテンソル $\mathcal{D} \in \mathbb{R}^{L \times M \times N}$ を用いる。Dの要素 lmn と音声信号 $sp \in \mathbb{R}^{P \times 2}$ の要素 pq の関係式は以下のように与えられる。

$$l + L(m - 1) + LM(n - 1) = p + P(q - 1) \quad (1)$$

次に、テンソルDから展開された3つの行列 $\mathcal{D}_1(L \times MN)$, $\mathcal{D}_2(M \times NL)$, $\mathcal{D}_3(N \times ML)$ は次のように与えられる。

$$\mathcal{D}_1(l, m + M(n - 1)) = \mathcal{D}(l, m, n) \quad (2)$$

$$\mathcal{D}_2(m, n + N(l - 1)) = \mathcal{D}(l, m, n) \quad (3)$$

$$\mathcal{D}_3(n, l + L(m - 1)) = \mathcal{D}(l, m, n) \quad (4)$$

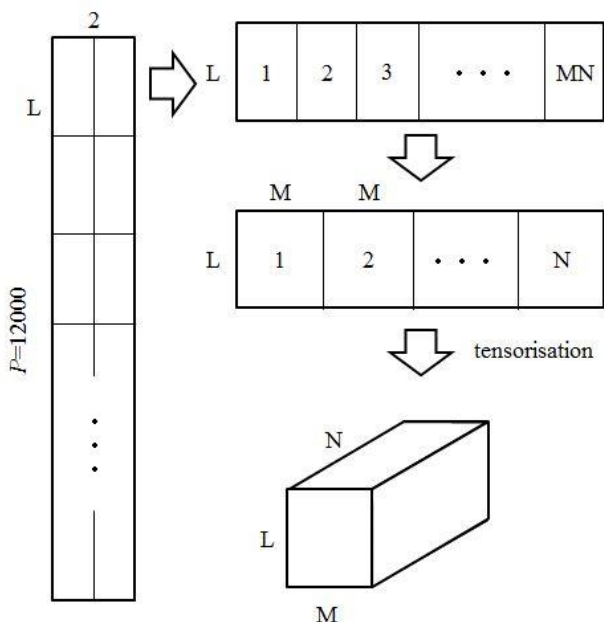


図1：行列のテンソル化

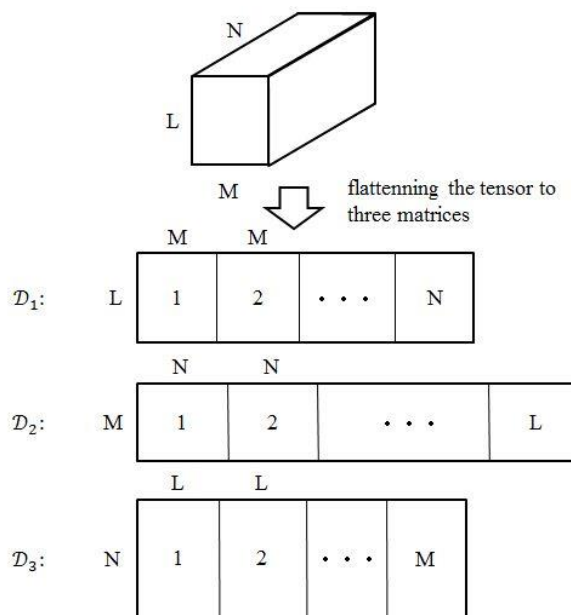


図2：データテンソルDの展開図

2.N-mode SVD および MS法の組み合わせによる雑音除去

本研究における雑音除去は Vasilescu らがSVDを多次元データに用いる手法として発展させた、N-mode 直交繰り返しアルゴリズム(OIA)^[1]とスペクトル減算法の一種の minimum statistics(MS)法^[2]を組み合わせで行う。具体的な手順は以下の通りである。

1)音声データに対し、N-mode OIA を施す。その際の各パラメータの設定値を実験では、

$$L = 30, M = 30, R_1 = R_2 = R_3 = 20$$

とした。

2)N-mode OIA を施した音声データの振幅(左チャンネル)の2乗の和の平均を ave とする。

3)さらに、音声データ 12000 サンプルを 1000 サンプル毎に分け、それぞれの振幅(左チャンネル)の2乗の和の平均を ave_1, \dots, ave_{12} とする。

4) ave_1 から順に ave と大小を比較し、 ave より大きい最初の ave_i を求める。

1：日本理工・院・数学 2：日大理工・教員・数学

5)次は ave_{12} から同様に ave と大小を比較し、 ave より大きい最初の ave_j を求める。

6) $ave_1 \sim ave_{i-1}$ と $ave_j \sim ave_{12}$ の範囲の音声データ、雑音データ、雑音付加データにN-mode OIA を施し、 $ave_i \sim ave_{j-1}$ の範囲にはMS を施す。

具体例を図3に示す。図3に示されるように、区間 ave_2 と ave_3 との境界および ave_{11} と ave_{12} の境界において全体の平均値 ave との大小関係が逆転していることから、 $ave_3 \sim ave_{11}$ の範囲(サンプル数では 2001~11000)では MS 法を、他の範囲ではN-mode OIA を用いる。

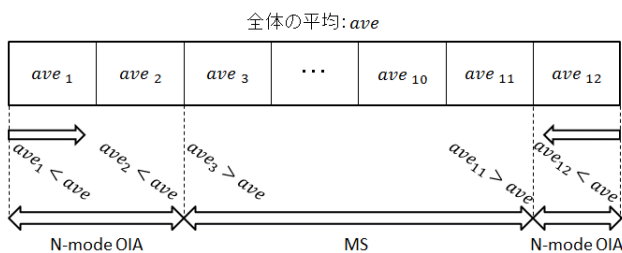


図3：N-mode OIA およびMS 法による組み合わせの一例

3.雑音除去実験

実験では、被験者3人(男性2人、女性1人)で、それぞれ「おはようございます」、「こんにちは」、「こんばんは」という文を3回ずつ録音した。その際、サンプリング周波数 44.1 kHz, 量子化ビット数 16 ビット, ステレオで wav ファイルとして記録し、サンプリング周波数を 8kHz に落とした。次に、計算の便宜上、音声データのサンプル数を 12000 に統一した。

雑音は白色雑音を用いた。また、白色雑音は MATLAB の randn 関数によって生成したものを用いた。

N-mode OIA において、パラメータが5つあるが、以下のように設定した。

$$L = 30, M = 30, R_1 = R_2 = R_3 = 2, 4, \dots, 30$$

4.実験結果

2で述べたN-mode OIA およびMS 法を組み合わせさせた雑音除去法による実験結果の一部を図4~図8に示す。図で示されるように、N-mode OIA のみでは、音声部分の雑音が残っているのが明らかであり、また、MS のみでは、音声の特に最初の部分が欠落しているのが分かる。しかしながら、これら2つを組み合わせさせた場合は、これらの欠点が緩和されているのが分かる。

学術講演会当日では、上記のような視覚的な結果のみではなく、客観的に結果を考察するために、segmental signal to

noise ratio(SNR)^[3]等を用いた評価基準による結果について報告する予定である。

5.結論

N-mode OIA のみ、もしくはMS のみの場合よりも、2つを組み合わせることで、4で示したような2つの欠点を緩和することが確認できた。今後の課題としては、与えられた入力信号に対してN-mode OIA とMS のどちらを用いるかをより効果的に決定するような評価関数について再考することが挙げられる。

参考文献

- [1]“Multilinear Analysis of Image Ensembles : Tensorfaces,” M. A. O. Vasilescu, and D.Terzopoulos,, Proc. ECCV-02, Copenhagen, Denmark, May 2002, pp.447-460
- [2] “Spectral subtraction based on minimum statistics,” R. Martin, Proc.EUSIPCO-94, pp. 1182-1185, Edinburgh, Sep. 1994.
- [3] “Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model,” T. Lotter and P. Vary, EURASIP J. Applied Signal Process., vol. 2005, no. 7, pp. 1110-1126, Jan. 2005.

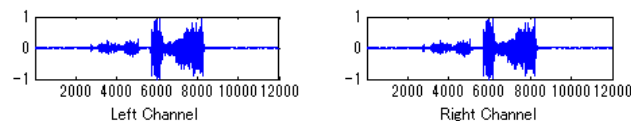


図4：オリジナル音声波形

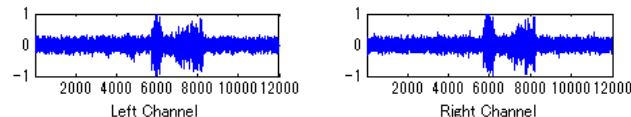


図5：雑音付加音声波形

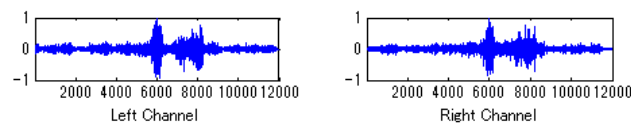


図6：N-mode OIAによる雑音除去音声波形

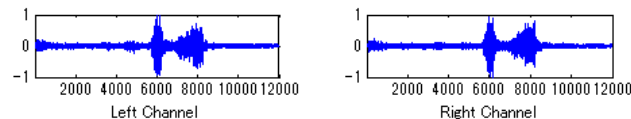


図7：MSによる雑音除去音声波形

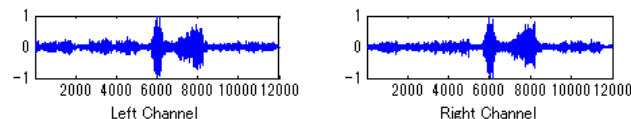


図8：N-mode OIAとMSの組み合わせによる雑音除去音声波形