

空間スペクトルを用いた時間断続信号の検出における  
主成分分析と周波数分析の比較評価\*

☆加藤通朗, 杉本侑哉, 牧野昭二, 山田武志, 北脇信彦 (筑波大)

## 1 はじめに

会議や打合せなどの、会議録を収録したアーカイブを効率的に再生するためには、誰が、いつ、どのように発話しているかという発話状況を自動的に検出し、タグ付けしておくことが有効である。我々はこのような効率再生を実現する一環として、音響信号のみを収録したアーカイブ中の発話状況を判定する試みを行っている。

発話状況を判定するにあたって、まず「連続信号」、「時間断続信号」、「無信号」という音響イベントを検出することを考える。連続信号は、ある方向から時間的に連続して音響信号が発生するような状況である。同様に時間断続信号は、相槌のように発話の時間長が短く、かつそれらが時間的に断続しているような音響信号が、ある方向から到来している状況である。我々はこれまで、空間スペクトルに対して主成分分析を適用することにより、時間断続信号を検出する手法<sup>[1]</sup>、周波数分析を用いて同様に時間断続信号を検出する手法<sup>[2]</sup>をそれぞれ提案している。

本稿では、二つの手法を、同様の音声に適用した場合の結果を比較し、両者の特徴を考察する。

## 2 提案手法

本章では、空間スペクトルを求める方法として MUSIC を説明する。その後、我々が提案している時間断続信号を検出する手法について、主成分分析法と周波数分析法について述べる。

### 2.1 MUSIC による空間スペクトルの算出

会議用テーブルの中央にマイクロホンアレーを設置することを想定し、各マイクロホンで収録した音響信号に対して MUSIC (Multiple Signal Classification) 法<sup>[3]</sup>を適用し、空間スペクトルを算出する。本稿では、文献[2]と同様に、MUSIC 法を周波数領域に拡張

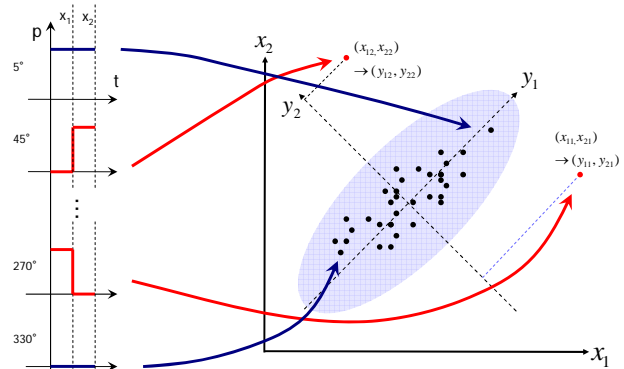


Fig.1 主成分スコアによる  
時間断続信号の検出

した手法<sup>[4]</sup>を用いる。

### 2.2 主成分分析法

MUSIC で求めた空間スペクトルに対して主成分分析法を適用し、時間断続信号を検出する<sup>[1]</sup>。連続する 2 フレームにおいて求めた空間パワースペクトルを考える。

このとき、方向ごとの 2 フレームのパワーの変化を観察することにより、信号が連続的(有音または無音)であるのか、断続的であるのかを判別できる。すなわち、1 番目のフレームのパワーを  $x_1$  軸、2 番目のフレームのパワーを  $x_2$  軸としたとき、Fig.1 のような分布になると考えられる。連続信号、及び無信号については  $45^\circ$  方向を中心に分布し、時間断続信号については、それから外れた位置に分布する。

提案手法では、このような分布に対し、主成分分析法による外れ値検出アルゴリズム<sup>[5]</sup>を適用する。具体的には、主成分分析後の上位の主成分スコアを確認することによって、連続信号、時間断続信号および無信号の状態を識別する。実際には 2 フレームではなく、連続する任意のフレーム長での分析を行う。

主成分分析法は、ある方向では時間断続信号、他のある方向では連続信号、その他のほとんどの方向で無信号であることを利用し、上位の主成分スコアの分布の特徴により検出

\*Comparison between Principal Component Analysis and Frequency Analysis Applied for Scattered Signal Detection for Spatial Spectrum, by Michiaki KATOH, Yuya SUGIMOTO, Shoji MAKINO, Takeshi YAMADA, and Nobuhiko KITAWAKI (University of Tsukuba).

する。このため、分布の判別に、上位3つ程度の主成分スコアのパターンを判別すればよいというメリットがある。一方、フレームごとに独立した主成分分析法を行うため、各方向の信号のパワーの時間変化によっては、想定するスコアのパターンが異なる可能性がある。

### 2.3 周波数分析

2.1節で述べた MUSICにより求まる空間スペクトルから、方向ごとに連続する  $L$  フレームに対して離散フーリエ変換を適用する<sup>[2]</sup>。これは、各方向の音響信号のパワーの時間変化を周波数分析していることに相当する。周波数分析では、このようにして求めた振幅スペクトルの形状から、連続信号、時間断続信号、および無信号のいずれであるかを判定する。周波数分析法では、各方向での判別は方向ごとに独立に行われるため、主成分分析法での結果が異なる場合でも判別可能であると考えられる。しかしながら、正解パターンの周波数スペクトルとのパターンマッチングが必要となり、これは主成分分析法での判別に比べ、複雑になる。

### 3 実験による両手法の比較

本稿での実験に用いる、収録した音声の状況および周波数分析のパラメータについては、文献[2]と同様、テーブル上に置いた直径20cmのマイクロホンアレーを用いて収録したインタビュー音声を利用している。主成分分析法は文献[2]と同様に、0.5[sec]のフレームを30フレームずつ取り出して実施している。周波数分析法についても、30フレームずつを1024点のFFTにより周波数スペクトルを算出している。

Fig.2は、上段より連続信号と時間断続信号を含むフレーム群(30フレーム)の(1)空間スペクトル、および180°方向の話者における(2)主成分分析法、(3)周波数分析法の結果である。主成分分析法では、例えば①のように音声が続いて発せられる状況では第1主成分スコアが反応し、②のように断続発話の部分では第2~3主成分スコアが反応しているため、これらの判別が可能であると考えられる。一方、周波数分析の結果では、連続信号部分では低周波スペクトルが強く、高周波成分が弱いいため、同様にこれらのパターンマッ

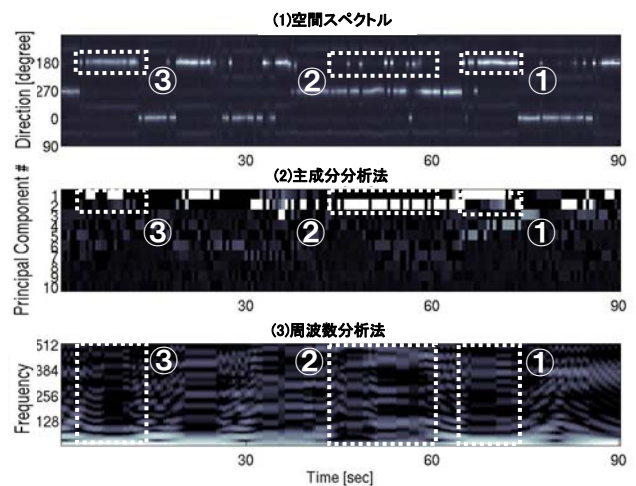


Fig.2 ある音声区間での主成分分析と周波数分析の結果

グによる判別が可能であると考えられる。

また③の箇所に注目すると、①と同様に連続信号が強い箇所であるが、主成分分析法では第1主成分スコアが期待するように反応していない箇所が見られる。これは、同じ連続信号区間であっても、主成分分析法を実施した時、第1主成分の固有ベクトルの向きが逆になる場合があるため、結果として主成分スコアも負の値になっているためである。

### 4 おわりに

本稿では、空間スペクトルを用いた時間断続信号の検出手法として、主成分分析法と周波数分析法の2手法を取り上げ、両者の比較を行った。主成分分析法では、上位数個の主成分スコアで断続信号の検出ができるというメリットがあるが、符号の反転の問題があることが分かった。周波数分析法では、このような問題はないが、断続信号検出に向けては、主成分分析と比較してパターンマッチングが複雑になると考えられる。

#### 参考文献

- [1] 加藤他, 信学技法, EA2010-47, pp.25-30, 2010.
- [2] 杉本他, 音講論 3-P-7(c), 2011年3月.
- [3] Schmidt *et al*, IEEE Trans. on antennas and propagation, Vol.AP-34, No.3, pp.276-280, 1986.
- [4] Asano *et al*, EURASIP journal on applied signal processing, pp.1727-1738, 2004.
- [5] Jolliffe *et al.*, *Principal Component Analysis*, 2<sup>nd</sup> Edition, Springer, 2001.