

独立成分分析に基づく実環境下での
信号分離と脳波解析に関する研究

石橋 孝昭

目次

第1章 序論	1
第2章 ブラインド信号分離と独立成分分析	7
2.1 緒言	7
2.2 ブラインド信号分離	8
2.2.1 フィルタ	9
2.2.2 マイクロホンアレー	10
2.2.3 Spectral Subtraction	13
2.2.4 SAFIA	14
2.3 独立成分分析	17
2.3.1 Kullback-Leibler 情報量	20
2.3.2 最急降下法	22
2.3.3 Natural Gradient アルゴリズム	26
2.3.4 Negentropy	26
2.3.5 中心化と白色化	28
2.3.6 FastICA アルゴリズム	29
2.4 独立成分分析による瞬時混合信号の分離	31
2.4.1 Natural Gradient による分離シミュレーション	32
2.4.2 FastICA による分離シミュレーション	34
2.5 結言	34
第3章 瞬時混合過程における独立成分分析の不定性問題の解決	36
3.1 緒言	36
3.2 独立成分分析の限界	37
3.2.1 スケールの不定性	37

3.2.2	成分置換	38
3.3	不定性問題の解決	39
3.3.1	分割信号の定義	39
3.3.2	分割信号の原信号と伝達関数による表現	40
3.3.3	分割信号の物理的解釈	41
3.3.4	スケールの不定性の解消	42
3.4	スケールの回復と原信号数の推定シミュレーション	43
3.4.1	原信号数が混合信号数と等しい場合	44
3.4.2	原信号数が混合信号数より少ない場合	44
3.4.3	原信号数が混合信号数より多い場合	48
3.5	結言	48
第4章	実環境下での音響信号の分離と目的音声の選択	52
4.1	緒言	52
4.2	周波数領域独立成分分析	54
4.2.1	周波数領域独立成分分析	54
4.2.2	周波数領域 FastICA アルゴリズム	56
4.3	周波数領域独立成分分析の不定性問題の解決と目的信号の選択	58
4.3.1	スケールと位相の不定性の解決	59
4.3.2	伝達関数の推定と成分置換の修正原理	61
4.3.3	実環境下の音響特性	63
4.3.4	周波数による修正能力の違い	64
4.3.5	統合的成分置換修正法	67
4.3.6	多音源多マイクロホン環境での成分置換修正法	68
4.3.7	目的音声の選択	69
4.4	音源分離実験	70
4.4.1	バーチャルルームでの音源分離評価実験	71
4.4.2	実環境下での雑音除去実験	73
4.4.3	実環境下での音源分離実験	74
4.4.4	目的音声選択実験	77

4.4.5	多音源多マイクロホン環境での音源分離実験	79
4.5	結言	79
第 5 章	単一試行視覚認識時脳波の特徴抽出と変動識別	82
5.1	緒言	82
5.2	脳波研究の歴史と脳波の概要	84
5.2.1	脳波研究の歴史	84
5.2.2	国際脳波学会連合標準電極配置法	85
5.2.3	脳波の計測	86
5.2.4	視覚性オドボール課題実験	87
5.2.5	視覚誘発電位	88
5.2.6	事象関連電位	89
5.3	独立成分分析を用いた脳波の特徴抽出	91
5.3.1	独立成分分析の脳波解析への適用	91
5.3.2	事象関連電位と閃光刺激に関連した成分の抽出	94
5.3.3	原信号数の推定と思考に関連した成分の抽出	97
5.4	単一試行視覚認識時脳波の変動識別実験	104
5.4.1	事象関連電位と閃光刺激に関連した成分に基づく脳波変動識別	104
5.4.2	思考に関連した成分に基づく脳波変動識別	105
5.5	結言	107
第 6 章	結論	109
	謝辞	113
	参考文献	114

第1章 序論

我々の身の回りの情報や信号のほとんどは、我々が必要としている本質的な情報と、そうでない雑音とが混合している。人間は、このような混合信号から、本質的な情報のみを取り出すことができる。例えば、複数の話者が同じ部屋で同時に発話している状況で観測される音声信号は各人の話が混じり合っているが、人間はこの混合信号から個々の話者の話を分離できる。このことは、カクテルパーティー効果 (cocktail-party effect)、カフェテリア効果 (cafeteria effect)、聖徳太子効果などと呼ばれる音源分離である。このように、雑音が混入された混合信号から、本質的な信号を抽出することが、信号処理や情報処理の大きな目的である。

目的信号の抽出や雑音の除去に関しては、フィルタ (Filter) 処理やマイクロホンアレー (Microphone-array) 処理等が研究されている。これらの従来の方法では、目的信号の特徴や到来方向などの先験的な知識が必要である。そこで近年、本質的な信号である原信号と、原信号に雑音が混合されるまでの過程を未知として、原信号を抽出する研究が進められている。このように、観測した混合信号のみを用いて元の原信号を復元する手法をブラインド信号分離 (BSS: Blind Source Separation) という。

ブラインド信号分離問題を解く方法として、独立成分分析 (ICA: Independent Component Analysis) [1]～[8] が脚光を浴びている。独立成分分析は、観測された混合信号が互いに統計的に独立な原信号の重ね合わせであるとの仮定の下で、混合信号から独立な成分を分解する統計的手法である。一方、主成分分析 (PCA: Principal Component Analysis) では信号のパワーに重点をおいており、因子分析 (FA: Factor Analysis) では信号にガウス性を仮定することが多い。これらに対して、独立成分分析では原信号に独立性のみを仮定しているだけなので、原信号のパワーの大きさに依存せず、原信号のガウス性を仮定する必要が無いため、原信号の制約が少なく非定常信号も分離できる特徴を持っている。したがって現在では、音声信号の分離、脳のデータ解析、通信での雑音除去、画像の分解、機械振動の解析、さらには時系列予測や社会学、心理

学のデータにも応用され始めている。

本論文では、非定常信号を分離できる独立成分分析に基づいて、実環境で実際に観測された混合信号から本質的な信号を抽出することを目的とする。具体的には、複数の話者が発声した音声信号を分離することと、人間の頭皮で観測される脳波信号から我々が必要とする信号を抽出することを目的として研究し、その結果を報告する。

まず最初に、独立成分分析によって得られる分離信号には、スケールの不定性 (Scaling) と成分置換 (Permutation) の問題が残ることを述べ、これらの問題を解消する方法を説明する。スケールの不定性問題は分割信号 (Decomposed signal) [9], [10] を導入して解消できて、分割信号は原信号と伝達関数の積で表現できることを導出する [11] ~ [13]。また、分割信号のスケールは伝達関数によって決まるため、スケールの不定性は自ずと解消できることを述べる。さらに、原信号数が混合信号数より少ない場合、分割信号を適用することで簡単に原信号数の推定できることを述べる [14]。

次に、実環境下での音声信号に対する音源分離問題については、観測される混合信号が原信号の畳み込みとして得られるため、独立成分分析は一般的に周波数領域で適用されることが多い [9], [10], [15] ~ [22]。これは、時間領域での畳み込みが周波数領域では積で表現できるため、混合過程に時間遅れや反射の無い瞬時混合過程のアルゴリズムを拡張して利用できるからである。周波数領域の独立成分分析 (FDICA: Frequency Domain Independent Component Analysis) は、原信号のスペクトルを各々の周波数帯域で推定する手法であり、各周波数帯域での原信号のスペクトルは良好に推定できる。

しかし、それらの推定スペクトルを時間領域に変換するとき、スケールの不定性と成分置換の問題が立ちはだかることになる。スケールの不定性は、推定信号のスペクトルの振幅や位相のスケールが原信号と異なることであり、成分置換は推定信号のスペクトルの推定順序が原信号の順序と異なることである。したがって、現実の応用に際しては、これらの問題を解決しなければならない。

スケールと位相の不定性の問題は、分割スペクトル (Decomposed spectrum) を導入することで解消できる。このとき、分割スペクトルのスケールは伝達関数によって決まる [9], [10]。一方、成分置換の問題は、分離信号の類似度に基づく方法 [9], [10] や、分離行列の推定結果を隣接する周波数の分離行列の初期値に利用する方法 [23]、混合行列の周波数軸上での連続性を用いた方法 [24] が提案されている。また、Hyvärinen に

よる FastICA アルゴリズム [15], [25], [26] は、雑音に比べて非ガウス性の高い話者音声
を最初に出力する性質があるため、成分置換問題はある程度緩和される。しかし、こ
れらを実環境下で適用しても実用的な結果を得るのは難しい。他に、マイクロホンア
レーとしての指向特性を考慮した方法 [27], [28] が提案されているが、計算量の問題な
ど改善すべき点が残っている。話者音声の尖度が雑音に比べて高いことを利用した方
法 [29] は、原理的に、他人の発話が干渉音となる場合は適用できない。伝達関数のゲ
インや位相のそれぞれに基づいた方法 [11], [12], [30], [31] は、後述するように、ゲイン
のみや位相のみを用いても成分置換は正しく修正できない。

そこで本論文では、実環境下で有効に機能する雑音除去システムの構築を目的とし
て、周波数領域の独立成分分析における成分置換問題を解決するとともに、音源とマ
イクロホンの相対的な位置関係が分かっている状況では目的音声の選択を可能にする
アルゴリズムを述べる。まず、実環境におけるゲインと位相の音響伝達特性を詳細に
調べて、その特性が周波数帯域やマイクロホン間隔、音源からマイクロホンまでの距
離に依存するため、ゲインや位相を単独に利用しただけでは成分置換は修正できない
ことを述べる [32], [33]。次に、音響伝達特性のゲインと位相の双方を組み合わせた成
分置換の修正法を述べる。具体的には、ブラインド信号分離では音源からマイクロホ
ンまでの距離は未知であるため、提案法では実環境の音響特性をもとに、周波数帯域
でゲインと位相を使い分けて成分置換を修正する [34]。最後に、音源とマイクロホン
の相対的な位置情報が得られない場合に目的音声を簡単に選択する方法について述べ
る [12]。

さらに、頭皮で観測される脳波 (EEG: Electroencephalogram) 信号の解析につい
ては、近年では生体計測技術の進歩により脳波計だけでなく、脳磁計 (MEG: Magne
toencephalogram)、磁気共鳴画像 (MRI: Magnetic Resonance Imaging)、機能的磁
気共鳴画像 (fMRI: Functional Magnetic Resonance Imaging)、陽電子放出断層撮影
法 (PET: Positron Emission Tomography) などの装置を用いて、脳や体を痛めずに
非侵襲で脳神経活動を計測することが可能になっている。そして、このような計測装
置を利用した臨床診断などが行われている。

しかし、脳波には、眼電図 (EOG: Electrooculogram) や瞬目 (blink)、筋電図 (EMG:
Electromyogram)、心電図 (ECG: electrocardiogram) 等のアーチファクト (artifact)

が含まれていて、我々の必要とする情報が隠されている。したがって、感覚刺激に関連した注意や認識などの活動によって変動する事象関連電位（ERP: Event Related Potential）を見るためには、同一条件下で複数回試行されて計測した脳波を加算平均（averaging）する処理が必要となる。加算平均するためには、音や光などの外部刺激によって、特定の脳内活動を誘発する必要がある。しかし、このような脳活動は厳密に同一の時間構造で誘発されているとは限らないため、加算平均は本来着目しようとしている信号を歪ませる可能性もある。また、外部刺激による誘発を行うことが難しい場合、単一試行波形には雑音によって乱されていて、そのままの波形を解析することは困難である。

これまでに、独立成分分析を適用して、脳内の誘発電位（evoked potential）、自発電位（spontaneous potential）、心拍運動（heart rate）や眼球運動（eye movement）、電源の干渉（interferences）などの成分を計測信号から分解する研究が進められている [35]～[42]。このとき、頭皮で観測される脳波として得られる脳の活動は、脳の潜在的な成分の混合信号と考えられる。この状況は、完全にブラインド信号分離の概念と一致する。また、原信号が発生して観測されるまでの時間は非常に短いと考えられるため、反射や残響を考慮しない時間領域の独立成分分析（TDICA: Time Domain Independent Component Analysis）を適用することになる。このように、独立成分分析によって脳内の本質的な信号を得ることで、脳活動の興味深い情報を明らかにすることができる。

本論文では、実際に観測された脳波に対して独立成分分析を適用して、単一試行データから特徴を抽出することと、その変動を識別することを目的とする。具体的には、視覚認識時の応答脳波に対して特徴を抽出するために、視覚性オドボール課題実験によって得られる観測信号に独立成分分析を適用する。このようにして、事象関連電位や閃光刺激に対する応答、思考に対する応答と考えられる成分を抽出して、その特徴から、脳波の変動を判断する方法を提案する。このとき、脳内の原信号数が完全に未知であるため、脳波の主要な信号数を推定して、その数に基づいて独立成分分析を用いることにより、さらに明確な脳波の特徴抽出を行う。

本論文の構成は以下の通りである。

第2章では、ブラインド信号分離の基礎概念を明確にするため、最初に、目的信号の強調と雑音除去の代表的な手法について述べる。次に、独立成分分析について概説

して、その後、独立成分分析の代表的な手法である Natural Gradient アルゴリズムと FastICA アルゴリズムについて説明する。最後に、それぞれのアルゴリズムは、観測されるまでに反射や時間遅れのない環境で得られる混合信号から、スケールの不定性と成分置換の問題を除いて、原信号を復元できることをシミュレーションにより確認する。

第3章では、まず、独立成分分析によって得られる分離信号には、スケールの不定性と成分置換の問題が残ることを述べる。次に、スケールの不定性問題を解消するために、分割信号を導入する。このとき、分割信号は原信号と伝達関数の積によって決まることを証明し、スケールは伝達関数で決まることを述べる。最後に、観測信号に時間遅れや反射が無く、暗騒音の無い環境下で、分割信号が原信号と伝達関数の積で推定されることをシミュレーションによって確認する。さらに、原信号数が混合信号数より少ない場合、分割信号を用いることで原信号数が容易に推定できることを述べる。

第4章では、実環境下で発話された音声信号に対して、独立性分析に基づく原信号の推定方法と目的音声の選択方法について述べる。まず、実環境下での音声信号は畳み込まれて観測されるため、独立成分分析を周波数領域へ拡張する方法について概説する。この処理によって、時間領域の畳み込み混合を周波数領域では積として扱うことができることになる。そして、周波数領域の独立成分分析では、スケールと位相の不定性と成分置換の問題が生じるため、単純に原信号は推定できないことを述べる。その後で、このスケールと位相の不定性と成分置換の解決法を提案する。さらに、推定された原信号から我々の必要とする目的信号を選択する方法を提案する。ここで提案する手法については、実環境下で実際に話者が発声した音声の混合信号を用いた音源分離実験によって、その有効性を検証する。

第5章では、人間の頭皮で実際に観測された単一試行視覚認識時脳波信号から、独立成分分析を用いて人間の考えを識別する方法について述べる。脳波信号は一般になじみがないため、まず、脳波研究の歴史や脳波計測方法、視覚誘発電位、事象関連電位について簡単に説明する。次に、脳波信号に独立成分分析を適用した特徴抽出法を提案する。このとき、事象関連電位や閃光刺激に対する応答と考えられる成分を抽出する。その後、脳波の主要な信号数を推定することで、より明確な特徴を抽出できる方法を提案する。以上の提案法は、単一試行脳波に対する実験により有効性が確認さ

れ，これまで同一条件下で複数試行した後にしか得られなかった脳波の特徴を，単一試行で抽出することが可能となったことを述べる．さらに，従来の脳波変動識別法と比較して高い精度で識別できることを述べる．

第6章は，本論文の総括である．

第2章 ブラインド信号分離と独立成分分析

2.1 緒言

我々の身の回りの信号には、本来の情報と雑音が混合している．人間はこのような多様な情報を見分け、聞き分けることができる．このように、雑音が混入された信号から、本質的な信号を抽出することが、信号処理や情報処理の大きな目的である．目的信号の強調や雑音の除去処理については、現在までに、多くの方法が研究されている．

さらに近年では、本質的な信号である原信号と、原信号に雑音が混合されるまでの過程が未知である状況で、原信号を抽出する研究が進められている．このように、観測した混合信号のみを用いて元の原信号を分離して復元する手法を、ブラインド信号分離（BSS: Blind Source Separation）という．このブラインド信号分離問題を解く方法の一つに、独立成分分析がある（ICA: Independent Component Analysis）[1]～[8]．独立成分分析は、原信号が互いに統計的に独立であると仮定するだけで、混合信号から原信号を推定する統計的手法である．また、定常信号だけでなく非定常信号も分離できるという特徴を持っているため、多くの分野での応用が期待されている．

独立成分分析は、音声信号の分離だけでなく、現在では多くの分野で研究が進められている．例えば、頭皮で観測される脳波（EEG: Electroencephalogram）として得られる脳の電位は、おそらく、脳の潜在的な成分の混合信号として得られる．この状況は、完全にブラインド信号分離の概念と一致していて、独立成分分析によって脳内の潜在的な成分を得ることができる．さらに、独立成分分析は特徴抽出においても応用されている．すなわち、通信、画像処理、音声信号のデジタル信号処理における圧縮やノイズ除去に対して研究されている．その上、金融の変動における隠れた要素の発見に対しても研究されている．

そこで本章では、ブラインド信号分離の基礎概念を明確にするため、最初に、目的

信号の強調と雑音除去の代表的な手法について述べる。その後、独立成分分析の概要と、その代表的なアルゴリズムである Natural Gradient と、FastICA について説明する。Natural Gradient アルゴリズムは Kullback-Leibler 情報量の最小化に基づく方法で、全ての原信号を同時に推定し、白色化を必要としない。一方、FastICA アルゴリズムは非ガウス性の最大化に基づく方法で、原信号を一つずつ推定し、収束が速いという特徴を持つ。

本章の構成は以下の通りである。まず、2.2 では、ブラインド信号分離を実現するための信号処理について概説する。次に、2.3 で独立成分分析の概要を述べて、Natural Gradient アルゴリズムと FastICA アルゴリズムについて解説する。これらのアルゴリズムによって、反射や時間遅れがないシミュレーション環境における混合信号から、原信号を完全に復元できることを 2.4 で確認する。2.5 は本章のまとめである。

2.2 ブラインド信号分離

原信号や伝達関数が未知であるときに、観測した混合信号のみを用いて元の原信号を分離して復元する手法を、ブラインド信号分離（BSS: Blind Source Separation）という。ブラインド信号分離は、多くの人が同じ場所で同時に発話している場合であっても、人間は必要とする音声を聞き分ける脳内の信号処理である。このことは、カクテルパーティー効果（cocktail-party effect）、カフェテリア効果（cafeteria effect）、聖徳太子効果などと呼ばれていて、雑音除去や目的信号強調の分野で多くの研究が行われている。

具体的には、 M 個のセンサが異なる位置に配置されている下で、 N 個の原信号が同時に発生している状況を考える。このとき、原信号を $s_n(t)$ ($n = 1, \dots, N$)、センサで観測される混合信号を $x_m(t)$ ($m = 1, \dots, M$) と表記すると、混合信号 $x_m(t)$ は

$$x_m(t) = \sum_{n=1}^N a_{mn}s_n(t) \quad (2.1)$$

のように $s_n(t)$ を重み付けした和で表現できる。ここに、 a_{mn} は原信号からセンサまでの伝達関数に相当するパラメータである。原信号から混合信号が生成される過程を混合過程（mixing process）という。

観測された混合信号 $x_m(t)$ のみを用いて、例えば

$$u_n(t) = \sum_{m=1}^M w_{nm} x_m(t) = \hat{s}_n(t) \quad (2.2)$$

のように、原信号の推定値 $\hat{s}_n(t)$ を求めることがブラインド信号分離の目的である。ここに、 $u_n(t)$ は分離信号、 w_{nm} は分離荷重である。混合信号 $x_m(t)$ から分離信号 $u_n(t)$ を出力するまでの過程を分離過程 (separated process) という。もし伝達関数 a_{mn} が既知であれば、簡単に分離荷重 w_{nm} が求まるので、混合信号 $x_m(t)$ から元の原信号 $s_n(t)$ を推定することが可能である。しかしながら、実際には a_{mn} が未知であるため、 w_{nm} の推定は非常に困難な問題となる。

以降では、ブラインド信号分離を解くための代表的な手法について概説する。

2.2.1 フィルタ

目的信号や雑音が特定の周波数成分を持つ場合、その周波数帯域を選択的に通過させる処理を行えば、簡単に目的信号の抽出ができる [43]。

低域通過フィルタ (low-pass filter) は、低い周波数は通過させて高い周波数の正弦波成分を減衰させる。理想的な低域通過フィルタは、遮断周波数 (cut-off frequency) より低い周波数の信号を全て通過させ、それよりも高い周波数の信号は全て阻止する。すなわち、遮断周波数より低い周波数帯域では入力信号と同じ信号を出力して、遮断周波数よりも高い周波数では出力が得られない。また、高域通過フィルタ (high-pass filter) は、低域通過フィルタとは反対に、遮断周波数より低域の信号は阻止して、それよりも高域の信号は全て通過させる。

実際のフィルタでは、通過域 (pass-band) と阻止域 (stop-band) の間にある遷移域は、理想的な場合と異なり急激に変わることはない。したがって、遮断周波数は振幅応答が最大値よりも 3[dB] 下がったところの周波数と定められている。3[dB] 下がる点 (half-power point) は、信号のパワーが半分になるということを意味している。

低域通過フィルタと高域通過フィルタを組み合わせることによって、帯域阻止フィルタ (band-stop filter) と帯域通過フィルタ (band-pass filter) が実現できる。帯域阻止フィルタは、Fig. 2.1(a) のように、低域通過フィルタと高域通過フィルタを並列接続して作られ、特定の周波数帯域の信号成分だけを除去する。一方、帯域通過フィル

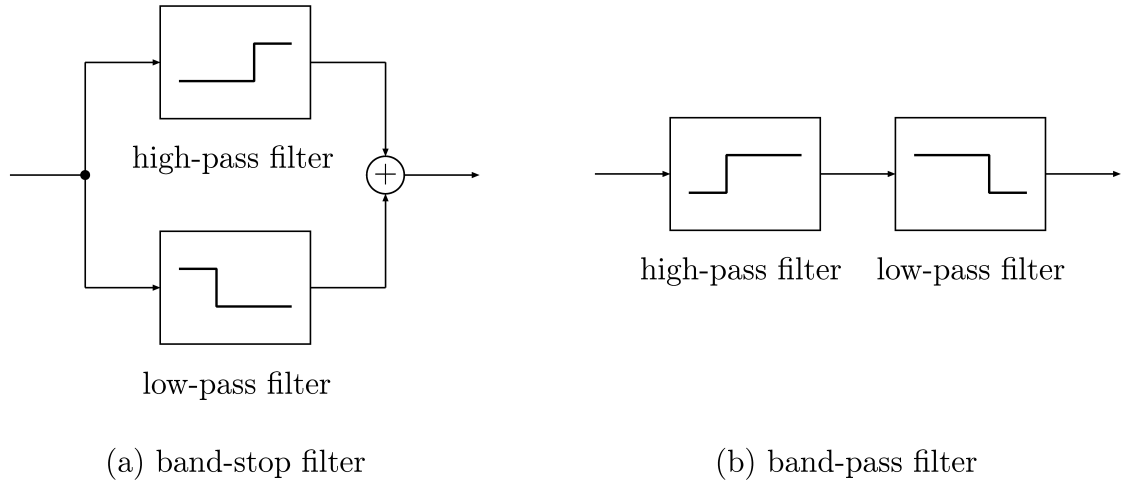


Fig. 2.1: Band-stop filter and band-pass filter.

タは Fig. 2.1(b) のように、低域通過フィルタと高域通過フィルタの縦続接続することで作られ、特定の周波数の信号成分だけを通過して抽出する。

以上のフィルタ処理によって、目的信号の周波数成分に通過域を持ち、雑音の周波数成分に阻止域を持つ特性のフィルタを作成することで、目的信号のみを抽出することができる。他にも、観測信号と目的信号の平均二乗誤差が最小になるようにフィルタ係数を求めるウィナーフィルタ (Wiener filter) [44] や、音声信号に自己回帰モデルに基づくカルマンフィルタ (Kalman filter) [45] を用いて目的信号を強調する方法もある。

2.2.2 マイクロホンアレー

遅延和型マイクロホンアレー (DS: Delay-and-Sum microphone-array) 処理は、複数のマイクロホンを用いて、各マイクロホンの観測信号に遅延器を付加した後に総和をとることで、目的信号方向 (look direction) に鋭い指向性を作成する処理である [46]～[49]。遅延和型マイクロホンアレーの原理を Fig. 2.2 に示す。等間隔で直線上に配列されたマイクロホンアレーで観測された信号 $x_m(t)$ は、Fig. 2.2 のように、各遅延器により、以下で表される遅延量 D_m だけ遅れる。

$$D_m = D_0 - (m - 1)\tau_L \quad (2.3)$$

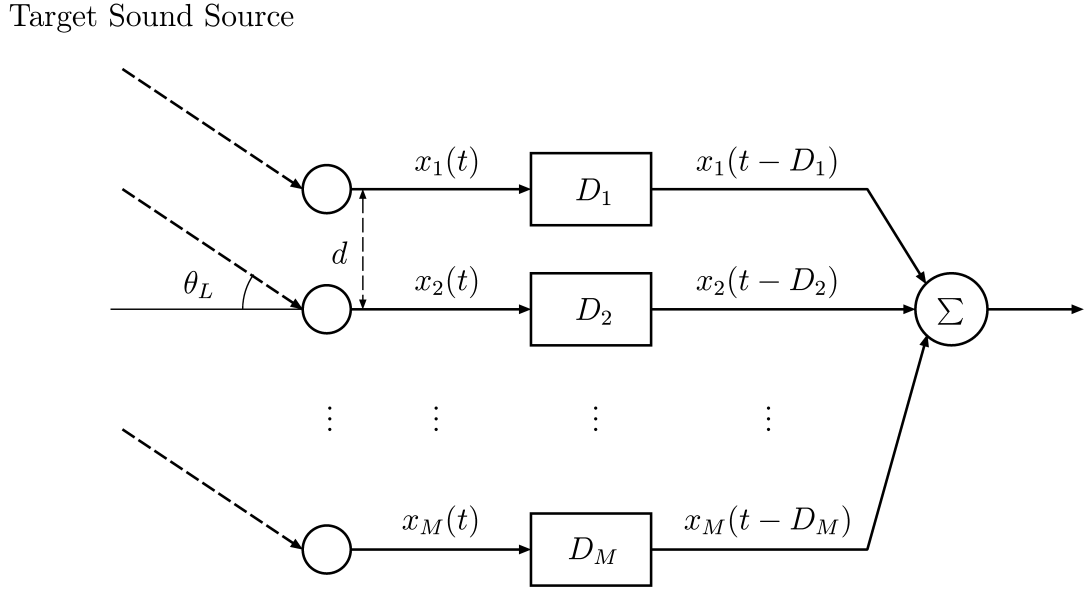


Fig. 2.2: Delay-and-Sum microphone-array.

ここに、 D_0 は D_m の値が小さ過ぎて遅延特性をデジタルフィルタで実現するときの精度が低下することを防ぐための固定遅延量、 τ_L は目的信号が各マイクロホンに到達するまでの時間差で

$$\tau_L = \frac{(m-1)d \sin \theta_L}{c} \quad (2.4)$$

のように計算できる．ここに、 θ_L は目的信号の到来方向、 d はマイクロホン間隔、 c は音速である．

このとき、目的信号は θ_L から到来すると仮定すると、各マイクロホンで観測される信号 $x_m(t)$ は

$$x_m(t) = x_1(t - (m-1)\tau_L) \quad (2.5)$$

と表される．各観測信号 $x_m(t)$ に遅延量 D_m を付加した信号 $x_m(t - D_m)$ は

$$x_m(t - D_m) = x_1(t - (m-1)\tau_L - D_m) = x_1(t - D_0) \quad (2.6)$$

のように求められる．Eq. (2.6) から分かるように、各遅延器からの出力信号 $x_m(t - D_m)$ は、チャンネル番号 m に因らず、 $x_1(t)$ に遅延量 D_0 を付加した同一の信号となっている．言い換えると、この遅延操作によって、 θ_L 方向から到来する信号の時間差が補正

Noise Source

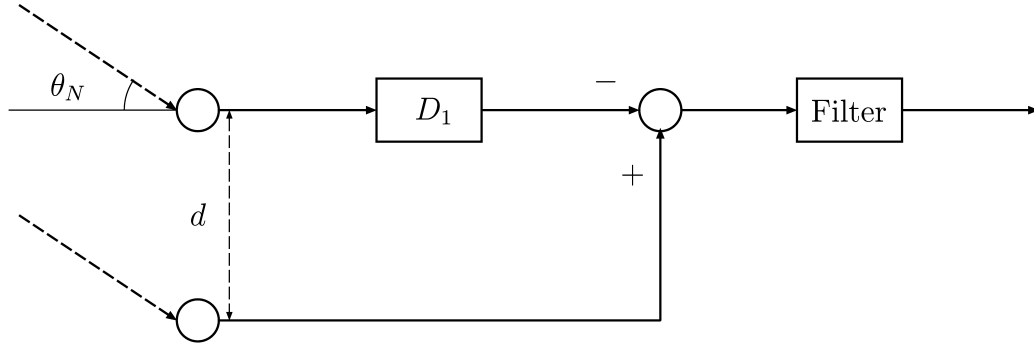


Fig. 2.3: Adaptive microphone-array.

され、同相化されることが分かる。したがって、同相化された信号を加算することで、目的信号の方向 θ_L から到来する信号は強調されることになる。

一方、 θ_L とは異なる方向 θ_N から到来する信号は、 τ_L とは異なる時間差 τ_N を持って観測される。したがって、Eq. (2.6) の遅延操作では同相化されず、遅延器の出力は時間的にずれた波形になり、これを加算しても信号を強調することはできない。以上のことから、遅延和型マイクロホンアレーは、目的信号の方向に対して感度が高い指向性を形成する。

適応型マイクロホンアレー (Adaptive microphone-array) 処理は、マイクロホンアレーの観測信号に遅延器を付加した後に減算を行うことで、周囲から到来する雑音方向を検出して、その方向にマイクロホン数 -1 の死角を形成する処理である [50], [51]。このとき、死角は周囲の雑音環境に応じて自動的に形成することができる。適応型マイクロホンアレーの基本原理を Fig. 2.3 に示す。雑音は未知の方向 θ_N から到来し、二つのマイクロホンで観測される。このとき、二つの観測信号には、 $\tau_N = d \sin \theta_N / c$ の時刻のずれが生じる。そこで、早い時刻で観測された雑音に $D_1 = \tau_N$ の遅延を付加すれば、二つのマイクロホンで観測された雑音は同相化される。その後、これらを減算すれば、雑音は消去されることになる。適応型マイクロホンアレー処理において、遅延時間 τ_N や雑音の到来方向 θ_N を知っている必要はなく、減算出力のパワーが最小になるように遅延量 D_1 を求めればよい。

一方、目的信号は雑音方向とは異なる方向から到来するため、上述の遅延操作では

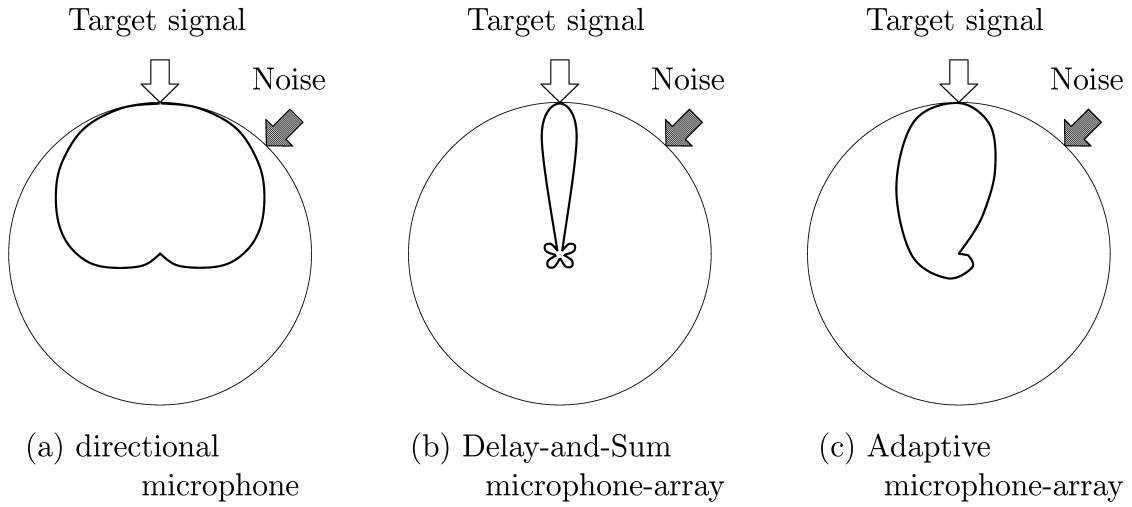


Fig. 2.4: The directivity response patterns.

同相化されず，減算によって消去されることはない．ただし，この処理の結果，目的信号が変形してしまうため，減算した後に補正フィルタによって補正する必要がある．

以上のマイクロホンアレー処理によって形成される指向性パターンを Fig. 2.4 に示す [52]．(a) は通常の単一指向性マイクロホンの指向性パターンである．目的信号方向の反対方向には感度が低いため，雑音が目的信号方向から $\pm 90^\circ$ 以上の方向から到来する場合，良好な雑音低減効果を有する．しかし，目的信号方向から $\pm 90^\circ$ 以内の方向に対しては感度が高く，図のように雑音が存在する場合，雑音低減効果は期待できない．(b) は遅延和型マイクロホンアレー処理によって形成された超指向性パターンである．目的信号の方向に鋭い指向性ビームを形成できるため，良好に雑音を低減することができる．(c) は適応型マイクロホンアレー処理によって形成された指向性パターンである．適応型マイクロホンアレーでは，目的信号方向に対して感度を高める超指向性とは対照的に，雑音方向に死角を持つ指向性を形成する．

2.2.3 Spectral Subtraction

スペクトルサブトラクション (SS: Spectral Subtraction) は雑音が重畳した混合信号のスペクトルから，雑音信号だけが存在する区間を利用して雑音スペクトルを推定し，それを引くことで目的信号を強調する方法である [53]．すなわち，騒音環境で観

測される混合信号 $x(t)$ は

$$x(t) = s(t) + \xi(t) \quad (2.7)$$

のように、目的信号 $s(t)$ と雑音 $\xi(t)$ が重畳していると考える．このとき、目的信号と雑音は無相関と仮定される．スペクトル領域 (spectrum domain) もしくはパワースペクトル領域 (power spectrum domain) で、観測信号のスペクトル $|x(\omega)|$ から、あらかじめ推定していた雑音のスペクトル $|\hat{\xi}(\omega)|$ を

$$|s(\omega)|^2 = |x(\omega)|^2 - \alpha |\hat{\xi}(\omega)|^2 \quad (2.8)$$

のように減算して、目的信号のスペクトル $|s(\omega)|$ を推定すれば、目的信号が復元されることになる．ここに、 α は減算する雑音成分を調節する係数である [54]．

スペクトルサブトラクションは定常雑音に対して有効であり、演算量が少ないため容易に実装できる．しかしながら、雑音の変動に脆弱であり、雑音成分の引き過ぎによって人工的雑音であるミュージカルノイズ (Musical noise) が発生しやすいという問題がある．

また、2チャンネルスペクトルサブトラクションは、目的信号方向に死角を作成し、雑音スペクトルを推定して、推定スペクトルを減算する方法である．目的信号の方向が既知である場合や、目的信号の方向が正確に推定できる場合には、高い精度で目的信号が推定できる [55], [56]．

2.2.4 SAFIA

SAFIA (sound source Segregation based on estimating incident Angle of each Frequency component of Input signals Acquired by multiple microphones) は、同時に発生された二つの信号であっても、同時刻、同一周波数上には二つの信号が重ならない性質を利用して、混合信号から目的信号を分離する手法である [57]～[60]．このとき、原信号は、音声などのように、調波構造を持つ信号、すなわち、基本周波数とその整数倍の周波数においてピークを持つスペクトル構造であると仮定する．このような信号に対して、周波数分解能を適切に定めれば、各周波数成分にはいずれかの音源だけが含まれていると近似できる [57]．

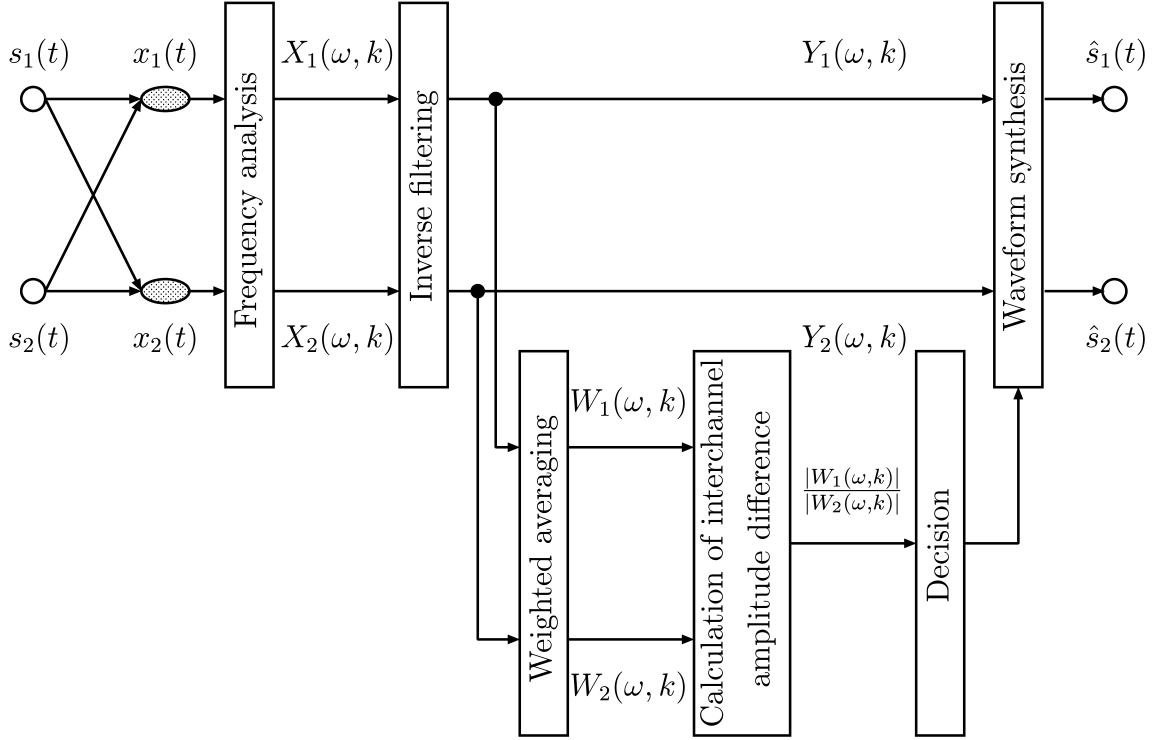


Fig. 2.5: SAFIA (sound source Segregation based on estimating incident Angle of each Frequency component of Input signals Acquired by multiple microphones).

SAFIA の処理手順を Fig. 2.5 に示す. 以降では SAFIA の処理手順に従って, その概要を述べる. SAFIA では, Fig. 2.6 のように, 目的音源はマイクロホン 2 に比べてマイクロホン 1 に近く, 雑音源はマイクロホン 1 に比べてマイクロホン 2 に近いと仮定する. さらに, 目的音と雑音は, 調波構造を持つ信号であると仮定する. 2 チャンネル入力された混合信号 $x_1(t)$, $x_2(t)$ を離散フーリエ変換した周波数成分を, それぞれ, $X_1(\omega)$, $X_2(\omega)$ と表記する. $X_1(\omega)$, $X_2(\omega)$ のチャンネル間の到達レベル差 $\Delta A(\omega)$ を

$$\Delta A(\omega) = 20 \log_{10} \frac{|X_1(\omega)|}{|X_2(\omega)|} \quad (2.9)$$

のように定義する. 各周波数成分における 2 チャンネル間の到達レベル差は, 上述の仮定より, 単独の原信号による到達レベル差であると考えられる. したがって, この到達レベル差から, 各周波数成分の到来方向が推定できる. さらに, それを用いることで, 各周波数成分が目的信号であるかノイズであるかを判定することができる. すなわち, ここでは目的信号がマイクロホン 1 に近いと仮定するため, $\Delta A(\omega)$ が正である帯域は目

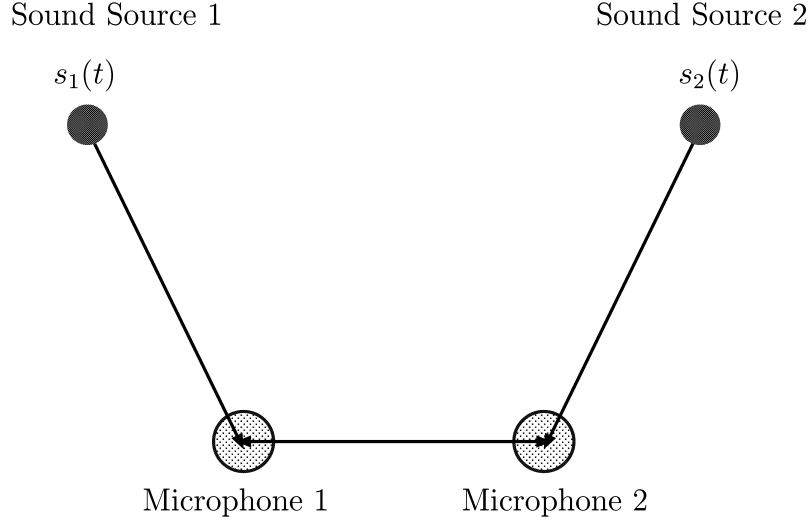


Fig. 2.6: Placement of sound sources and microphones.

的信号の帯域と判定でき，目的信号のスペクトル成分 $\hat{S}_1(\omega)$ と雑音のスペクトル成分 $\hat{S}_2(\omega)$ は

$$\hat{S}_1(\omega) = \begin{cases} X_1(\omega) & \text{if } \Delta A(\omega) \geq \beta \\ \alpha X_1(\omega) & \text{if } \Delta A(\omega) < \beta \end{cases} \quad (2.10)$$

$$\hat{S}_2(\omega) = \begin{cases} X_2(\omega) & \text{if } \Delta A(\omega) \geq \beta \\ \alpha X_2(\omega) & \text{if } \Delta A(\omega) < \beta \end{cases} \quad (2.11)$$

のように推定される [59]. ここに， α は雑音成分と判定された帯域に乗算する係数， β は判定レベルの閾値である．推定された目的信号のスペクトル $\hat{S}(\omega)$ を逆フーリエ変換して時間信号に戻すことで，目的信号は回復される．同様に，位相差 $\Delta\psi(\omega)$ を用いた手法は以下のように定式化される [59].

$$\hat{S}_1(\omega) = \begin{cases} X_1(\omega) & \text{if } \Delta\psi(\omega) \geq \beta \\ \alpha X_1(\omega) & \text{if } \Delta\psi(\omega) < \beta \end{cases} \quad (2.12)$$

$$\hat{S}_2(\omega) = \begin{cases} X_2(\omega) & \text{if } \Delta\psi(\omega) \geq \beta \\ \alpha X_2(\omega) & \text{if } \Delta\psi(\omega) < \beta \end{cases} \quad (2.13)$$

目的信号とノイズが Fig. 2.6 のようにマイクロホンに対して対称に配置している場合，目的信号のチャンネル間レベル差分布 $\Delta A_S(\omega)$ とノイズのチャンネル間レベル差分布 $\Delta A_N(\omega)$ の平均は 0 となるため，Eq. (2.10),(2.11) で $\beta = 0$ と設定すればよい．しかしながら，目的信号とノイズがマイクロホンに対して非対称に配置される場合， β は， $\Delta A_S(\omega) > \beta > \Delta A_N(\omega)$ を満たすように設定する [60]. さらに， α については， α が小さいほど雑音成分が抑圧されるが，目的音声の周波数成分まで出力されないため，

条件によってはミュージカルノイズが生じることがある。復元音声の音質を向上するためには、目的音声区間を検出して、目的音声区間では α を大きくして音声品質を向上し、目的音声区間でないときには α を小さくすることでSN比が改善される[60]。

2.3 独立成分分析

独立成分分析 (ICA: Independent Component Analysis) は、原信号が統計的に独立という仮定の下で、観測された混合信号から原信号を推定するブラインド信号分離の手法の一つで、エンジン音のような定常信号だけでなく、話者音声や音楽などの非定常信号も分離できるという特徴を持っている[1]～[8]。また、原信号の推定だけでなく、観測信号の隠れた構造を見つけるための特徴抽出としても利用されている。

最も基本的な独立成分分析のモデルは Fig. 2.7 のように示される。 M 個のセンサで観測された既知の混合信号 (mixture signal) は $\mathbf{x}(t) = [x_1(t), \dots, x_m(t), \dots, x_M(t)]^T$ とベクトル表記して

$$\mathbf{x}(t) = A\mathbf{s}(t) \quad (2.14)$$

のように表現される。ここに、 $\mathbf{s}(t) = [s_1(t), \dots, s_n(t), \dots, s_N(t)]^T$ は N 個の未知の原信号 (source signal)、 A は原信号からセンサまでの減衰係数 a_{mn} を要素とする未知の混合行列 (mixing matrix)、 T は転置を表す。ここでは、原信号が各センサに到達するまでの時間遅れや反射などの影響を無視して、 a_{mn} を定数として扱う。すなわち、単純に原信号 $s_n(t)$ を a_{mn} 倍して同時刻 t で瞬時に足し合わせている。したがって、このように生成される混合信号は瞬時混合 (instantaneous mixture) 信号または空間的混合 (spatial mixture) 信号と呼ばれ、Eq. (2.14) は瞬時混合過程と呼ばれる。

混合信号 $\mathbf{x}(t)$ から、分離信号 (separated signal) $\mathbf{u}(t) = [u_1(t), \dots, u_n(t), \dots, u_N(t)]^T$ を

$$\mathbf{u}(t) = W\mathbf{x}(t) \quad (2.15)$$

のように変換する。ここに、 W は要素に w_{nm} を持つ分離行列 (separated matrix) である。

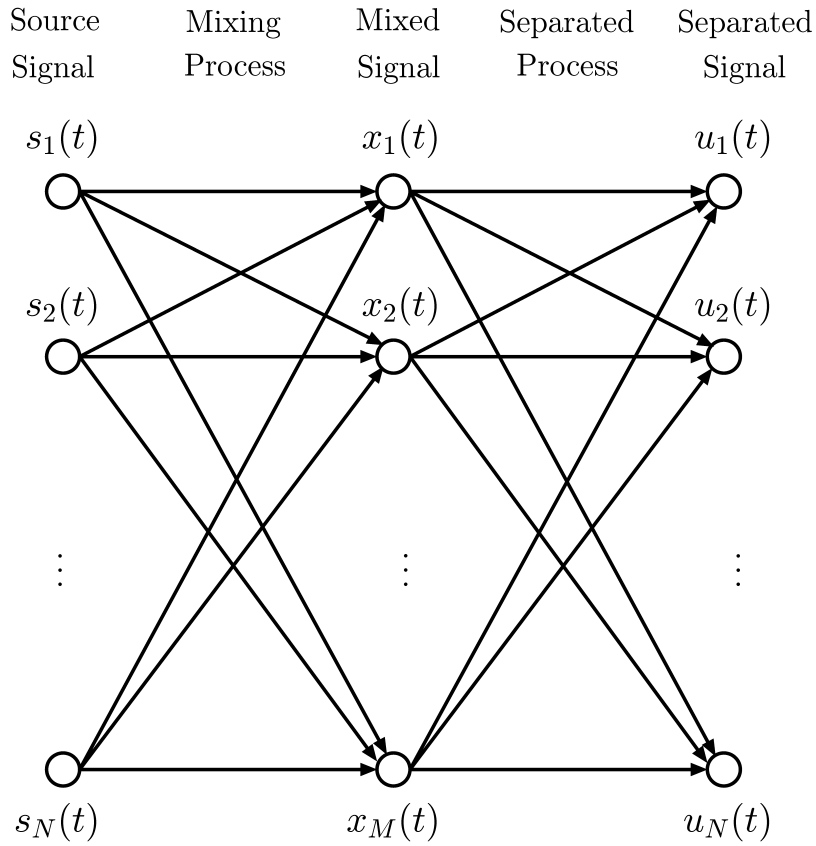


Fig. 2.7: ICA (Independent Component Analysis).

独立成分分析によって，分離行列 W に混合行列 A の逆行列 A^{-1} を推定すれば

$$\mathbf{u}(t) = W\mathbf{x}(t) = W\mathbf{A}\mathbf{s}(t) = A^{-1}\mathbf{A}\mathbf{s}(t) = \mathbf{s}(t) \quad (2.16)$$

となって，原信号が完全に復元できたことになる．しかしながら，混合行列 A は未知であるため，混合信号 $\mathbf{x}(t)$ のみを手掛かりとして，分離行列 W を推定しなければならない．

独立成分分析では，各信号を確率変数として扱う．したがって以降では，時刻のパラメータ t を省略して議論する．また本章では，原信号数と混合信号数が等しい最も簡単な $N = M$ のモデルを用いて説明する．

原信号 s_n は互いに統計的に独立と仮定する．すなわち， s_n の同時確率密度関数 (Joint Probability Density Function) $p(s_1, \dots, s_N)$ は

$$p(s_1, \dots, s_N) = \prod_{n=1}^N p(s_n) \quad (2.17)$$

のように、 s_n の周辺確率密度関数 (Marginal Probability Density Function) $p(s_n)$ の積で表現できると仮定する. この仮定は、多くの場合は非現実的な仮定でなく、実際には厳密に独立である必要はないとされている. 一方, Eq. (2.14) によって変換された混合信号 x_m の同時確率密度関数 $p(x_1, \dots, x_M)$ は

$$p(x_1, \dots, x_M) \neq \prod_{m=1}^M p(x_m) \quad (2.18)$$

のように独立性が失われる. したがって, Eq. (2.15) によって推定される分離信号 u_n の同時確率密度関数 $p(u_1, \dots, u_N)$ が

$$p(u_1, \dots, u_N) = \prod_{n=1}^N p(u_n) \quad (2.19)$$

のように、周辺確率密度関数 $p(u_n)$ の積で表現できて統計的に独立になれば、 u_n は原信号 s_n とみなすことができる.

このように、原信号の独立性のみを手掛かりとして、原信号を推定する統計的手法を独立成分分析という. 独立成分分析を実現するため、これまでに多くのアルゴリズムが提案されている. 具体的には、原信号の非ガウス性を利用し全ての原信号を同時に復元する方法として、分離信号の相互情報量を最小化する方法、最尤法 (Maximum likelihood method) を用いた方法、高次キュムラントを用いる方法などがある [61]~[63]. また、原信号を一つずつ復元する方法として、尖度 (Kurtosis) を用いる方法、Negentropy に基づく方法などがある [25], [26]. また、原信号の非定常性に基づく方法や信号の時間相関に基づく方法がある [9], [10], [64]. その他にも、原信号の確率密度関数を推定して独立成分分析を行う方法や幾何学的な観点から解く方法なども研究されている [65]~[67].

しかしながら、分離信号には、以下のようにスケールの不定性と成分置換の問題が残る. 独立成分分析は混合信号 \mathbf{x} から統計的に独立となるような分離信号 \mathbf{u} を推定するだけなので

$$W\mathbf{A} = P\mathbf{D} \quad (2.20)$$

のように、スケールの不定性と成分置換の問題が残る. ここに、 P は各行各列で値が 1 の要素を一つ持ち、それを除いた全ての要素での値が 0 である置換行列、 $\mathbf{D} = \text{diag}[d_1, \dots, d_N]$ は対角行列で、その要素 d_n は振幅の係数を表す. したがって、分離信

号 u_n には d_n 倍された原信号が推定され、分離信号の順番と原信号の順番は必ずしも一致しないことになる。これらの問題の解消法については次章で詳しく述べるが、式 (2.14) で得られる瞬時混合信号の場合、スケールの不定性は原信号の分散を 1 と仮定すればよいし、成分置換の問題は原信号が出力される順番が異なるだけで、原信号は推定できている。したがって、瞬時混合過程に対する独立成分分析では、この不定性はさほど大きな問題にならない。

以降では、時間領域の独立成分分析について、Kullback-Leibler 情報量の最小化に基づく Natural Gradient アルゴリズムと、非ガウス性の最大化に基づく FastICA のアルゴリズムを述べる。

2.3.1 Kullback-Leibler 情報量

Natural Gradient [63] は、独立性の評価に Kullback-Leibler 情報量 (Kullback-Leibler divergence) を用いて、分離信号の Kullback-Leibler 情報量が最小になるように分離行列 W を推定する逐次更新アルゴリズムである。したがって、分離信号は全て同時に推定されることになる。以降では、Kullback-Leibler 情報量と Natural Gradient アルゴリズムについて概説する。

分離信号 $\mathbf{u} = [u_1, \dots, u_N]^T$ の Kullback-Leibler 情報量 $I_{KL}(\mathbf{u})$ は

$$I_{KL}(\mathbf{u}) = \int p(\mathbf{u}) \log \frac{p(\mathbf{u})}{\prod_{n=1}^N p(u_n)} d\mathbf{u} \quad (2.21)$$

$$= \sum_{n=1}^N \mathcal{H}(u_n) - \mathcal{H}(\mathbf{u}) \quad (2.22)$$

のように表記される。ここに、 $\mathcal{H}(u_n)$ は Entropy で

$$\mathcal{H}(u_n) = - \int p(u_n) \log p(u_n) du \quad (2.23)$$

のように定義され、情報のあいまいさを表す。

Kullback-Leibler 情報量は、2 変数の場合のみ、相互情報量 (Mutual Information) $I_{MI}(\mathbf{u})$ と一致して

$$I_{KL}(\mathbf{u}) = \mathcal{H}(u_1) + \mathcal{H}(u_2) - \mathcal{H}(u_1, u_2) = I_{MI}(\mathbf{u}) \quad (2.24)$$

となる。このことは、Fig. 2.8 のように示される [68]。

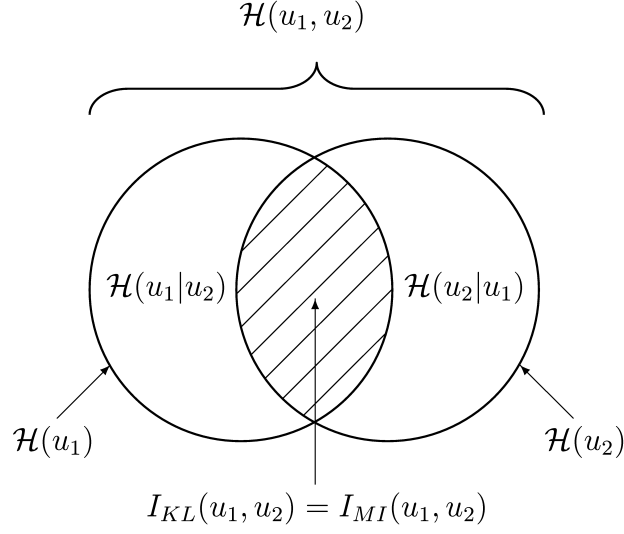


Fig. 2.8: Kullback-Leibler divergence and mutual information in the case of two variables.

しかし、例えば3変数の場合、Kullback-Leibler 情報量は

$$I_{KL}(\mathbf{u}) = \mathcal{H}(u_1) + \mathcal{H}(u_2) + \mathcal{H}(u_3) - \mathcal{H}(u_1, u_2, u_3) \quad (2.25)$$

と表現される。一方、相互情報量 $I_{MI}(\mathbf{u})$ は

$$\begin{aligned}
 I_{MI}(\mathbf{u}) &= I(u_1, u_2) - I(u_1, u_2|u_3) \\
 &= \{\mathcal{H}(u_1) + \mathcal{H}(u_2) - \mathcal{H}(u_1, u_2)\} \\
 &\quad - \{\mathcal{H}(u_1|u_3) + \mathcal{H}(u_2|u_3) - \mathcal{H}(u_1, u_2|u_3)\} \\
 &= \{\mathcal{H}(u_1) + \mathcal{H}(u_2) - \mathcal{H}(u_1, u_2)\} \\
 &\quad - \{\{\mathcal{H}(u_1, u_3) - \mathcal{H}(u_3)\} - \{\mathcal{H}(u_2, u_3) - \mathcal{H}(u_3)\}\} \\
 &\quad + \{\mathcal{H}(u_1, u_2, u_3) - \mathcal{H}(u_3)\}\} \\
 &= \mathcal{H}(u_1) + \mathcal{H}(u_2) + \mathcal{H}(u_3) - \mathcal{H}(u_1, u_2) \\
 &\quad - \mathcal{H}(u_2, u_3) - \mathcal{H}(u_1, u_3) + \mathcal{H}(u_1, u_2, u_3) \quad (2.26)
 \end{aligned}$$

となる。したがって、3変数の場合、Kullback-Leibler 情報量と相互情報量は一致しない。すなわち、Fig. 2.9(a) に示すように、Kullback-Leibler 情報量は2つ以上の変数が共有する情報量である。一方、相互情報量は Fig. 2.9(b) のように全ての変数が共有している情報量である。したがって、Fig. 2.10 の場合、各変数は独立ではないが、相互

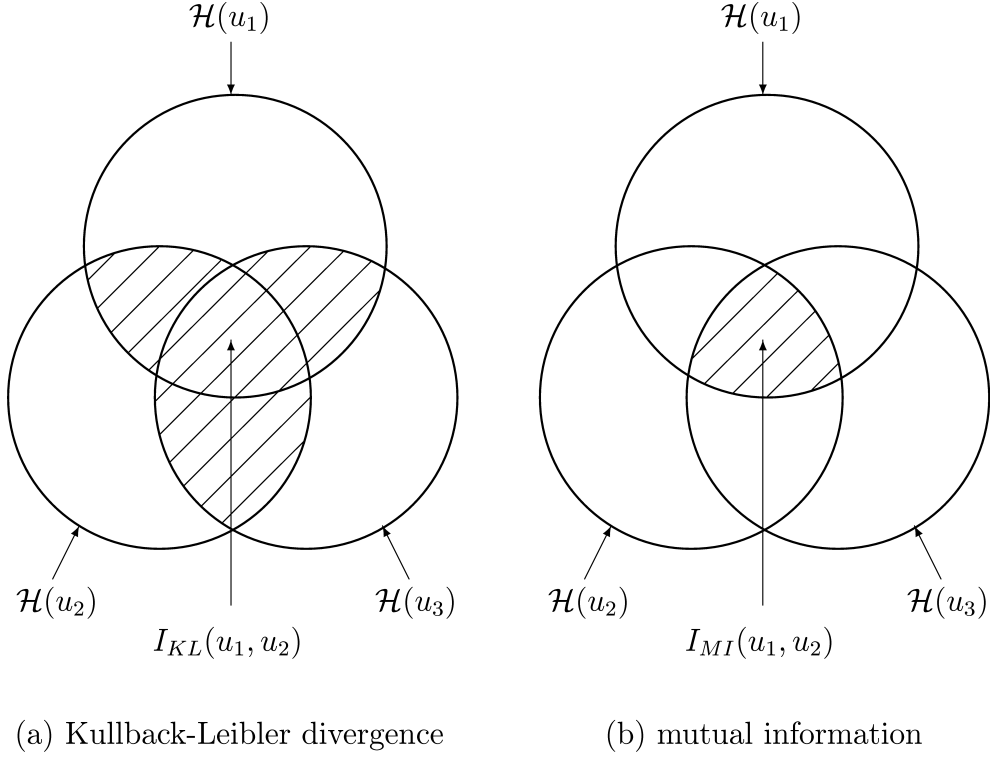


Fig. 2.9: Kullback-Leibler divergence and mutual information in the case of three variables.

情報量は0となる．これに対して，Kullback-Leibler 情報量は全ての確率変数が独立になったときにのみ0になることが分かる．以上のことから，Kullback-Leibler 情報量は多変数の独立の尺度として有効である [7]．

2.3.2 最急降下法

Kullback-Leibler 情報量 $I_{KL}(\mathbf{u})$ が0となるときの，分離信号 \mathbf{u} は互いに統計的に独立になる．したがって，分離信号 \mathbf{u} を独立にするには

$$W = \arg \min_W I_{KL}(\mathbf{u}) \quad (2.27)$$

となる分離行列 W を推定することになる．実際には，Kullback-Leibler 情報量の最小値を，最急降下法によって

$$W + \Delta W = W - \eta \frac{\partial I_{KL}(\mathbf{u})}{\partial W} \quad (2.28)$$

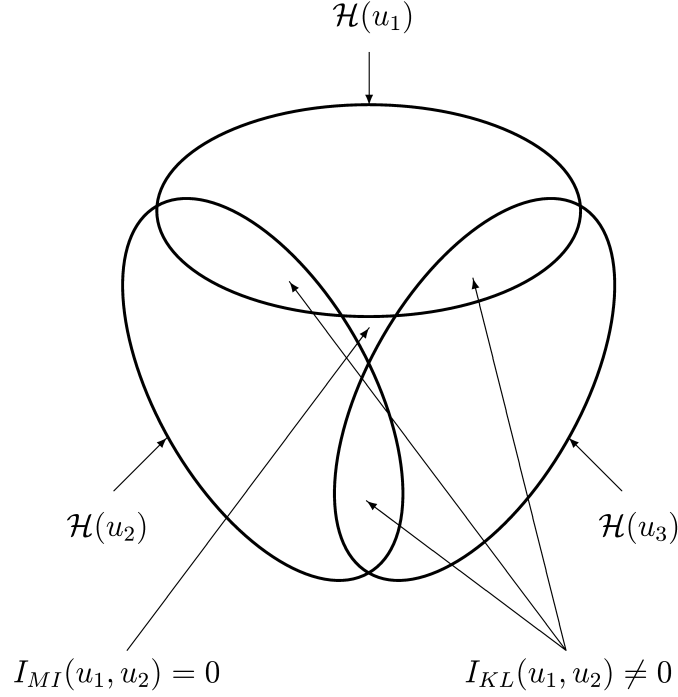


Fig. 2.10: Not independent 3 variables.

のように求める。ここに、 η は学習係数 (learning coefficient) である。このとき、確率変数 $p(\mathbf{u})$ は

$$p(\mathbf{u}) = \frac{1}{|\frac{\partial \mathbf{u}}{\partial \mathbf{x}}|} p(\mathbf{x}) = \frac{1}{|W|} p(\mathbf{x}) \quad (2.29)$$

のように変数変換されるので、 $\mathcal{H}(\mathbf{u})$ は

$$\begin{aligned} \mathcal{H}(\mathbf{u}) &= - \int p(\mathbf{u}) \log p(\mathbf{u}) d\mathbf{u} \\ &= - \int \frac{1}{|W|} p(\mathbf{x}) \log \left\{ \frac{1}{|W|} p(\mathbf{x}) \right\} d\mathbf{u} \\ &= - \int p(\mathbf{x}) \log \left\{ \frac{1}{|W|} p(\mathbf{x}) \right\} d\mathbf{x} \\ &= \int p(\mathbf{x}) \log |W| d\mathbf{x} - \int p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} \\ &= \log |W| \int p(\mathbf{x}) d\mathbf{x} + \left\{ - \int p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} \right\} \\ &= \log |W| + \mathcal{H}(\mathbf{x}) \end{aligned} \quad (2.30)$$

のように変換できる。したがって、 $I_{KL}(\mathbf{u})$ は

$$I_{KL}(\mathbf{u}) = \sum_{n=1}^N \mathcal{H}(u_n) - \log |W| - \mathcal{H}(\mathbf{x}) \quad (2.31)$$

のように求められるので, Eq. (2.28) の $\frac{\partial I_{KL}(\mathbf{u})}{\partial W}$ は

$$\frac{\partial I_{KL}(\mathbf{u})}{\partial W} = \frac{\partial}{\partial W} \left\{ \sum_{n=1}^N \mathcal{H}(u_n) - \log |W| - \mathcal{H}(\mathbf{x}) \right\} \quad (2.32)$$

$$= \frac{\partial}{\partial W} \sum_{n=1}^N \mathcal{H}(u_n) - \frac{\partial}{\partial W} \log |W| \quad (2.33)$$

と書き改めることができる. また, エントロピー $\mathcal{H}(u_n)$ は

$$\mathcal{H}(u_n) = - \int p(u_n) \log p(u_n) du_n = -E[\log p(u_n)] \quad (2.34)$$

のように表記できるので, Eq. (2.33) の第1項の各要素について

$$\begin{aligned} \frac{\partial}{\partial w_{nm}} \mathcal{H}(u_n) &= -E \left[\frac{\partial}{\partial w_{nm}} \log p(u_n) \right] \\ &= -E \left[\frac{d \log p(u_n)}{du_n} \frac{du_n}{dw_{nm}} \right] \\ &= -E \left[\frac{d \log p(u_n)}{du_n} x_m \right] \\ &= E[\varphi_n(u_n) x_m] \end{aligned} \quad (2.35)$$

のように求められる. ここに, $\varphi_n(u_n)$ は

$$\varphi_n(u_n) = - \frac{d \log p(u_n)}{du_n} \quad (2.36)$$

である. したがって, Eq. (2.33) の第1項は

$$\frac{\partial}{\partial W} \sum_{n=1}^N \mathcal{H}(u_n) = E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{x}^T] \quad (2.37)$$

となる. ここに, $\frac{\partial}{\partial W}$ と $\boldsymbol{\varphi}(\mathbf{u})$ は, それぞれ以下のように定義される.

$$\frac{\partial}{\partial W} = \begin{bmatrix} \frac{\partial}{\partial w_{11}} & \cdots & \frac{\partial}{\partial w_{1N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial w_{N1}} & \cdots & \frac{\partial}{\partial w_{NN}} \end{bmatrix} \quad (2.38)$$

$$\boldsymbol{\varphi}(\mathbf{u}) = [\varphi_1(u_1), \dots, \varphi_N(u_N)]^T \quad (2.39)$$

一方, Eq. (2.33) の第2項については, W の逆行列 W^{-1} の nm 成分 $[W^{-1}]_{nm}$ と, W の余因子行列 \tilde{W}_{mn} の関係が

$$[W^{-1}]_{nm} = \frac{1}{|W|} \tilde{W}_{mn} \quad (2.40)$$

となることと、行列式の余因子展開が

$$|W| = \sum_{n=1}^n w_{nm} \tilde{W}_{nm} \quad (2.41)$$

のように求められることから

$$\frac{\partial}{\partial w_{nm}} \log |W| = \frac{1}{|W|} \frac{\partial |W|}{\partial w_{nm}} = \frac{1}{|W|} \tilde{W}_{nm} = [W^{-1}]_{mn} = [W^{-T}]_{nm} \quad (2.42)$$

のように書き改めることができる．したがって、 $\frac{\partial}{\partial W} \log |W|$ は

$$\frac{\partial}{\partial W} \log |W| = W^{-T} \quad (2.43)$$

となる．以上のことから、Eq. (2.28) の第2項について

$$\begin{aligned} \frac{\partial I_{KL}(\mathbf{u})}{\partial W} &= \frac{\partial}{\partial W} \sum_{n=1}^N \mathcal{H}(u_n) - \frac{\partial}{\partial W} \log |W| \\ &= E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{x}^T] - W^{-T} \\ &= E[\boldsymbol{\varphi}(\mathbf{u})(W^{-1} \mathbf{u})^T] - IW^{-T} \\ &= E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T W^{-T}] - IW^{-T} \\ &= E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T - I] W^{-T} \end{aligned} \quad (2.44)$$

のように導出され、Eq. (2.28) から

$$W + \Delta W = W - \eta E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T - I] W^{-T} \quad (2.45)$$

の逐次更新式が得られる．このとき、 $\varphi_n(u_n)$ は、確率密度関数 $p(u_n)$ が未知であるため、求めることができない．しかし、 u_n の分布がラプラス分布のようなスーパーガウシアン (Super-Gaussian)，すなわち、裾が長くピークが尖った分布の場合

$$\varphi_n(u_n) = \tanh(u_n) \quad (2.46)$$

を選択し、一方、一様分布のようなサブガウシアン (Sub-Gaussian)，すなわち、裾が短くピークが滑らかである分布の場合

$$\varphi_n(u_n) = u_n^3 \quad (2.47)$$

を選択することで、安定して分離できる．

2.3.3 Natural Gradient アルゴリズム

最急降下法では、 W の空間を正規直交空間、すなわち、ユークリッド空間 (Euclidean space) での勾配を求めている。しかし、 W の空間は正規直交空間であるとは限らないため、ユークリッド空間からリーマン空間 (Riemann space) へ拡張した Natural Gradient アルゴリズムが提案されている。これは、自然勾配法とも呼ばれている。

リーマン空間での勾配 $\tilde{\Delta}W$ とユークリッド空間での勾配 ΔW の関係は

$$\tilde{\Delta}W = \Delta W W^T W \quad (2.48)$$

のように得られることから、リーマン空間で Kullback-Leibler 情報量の最急降下方向は

$$\frac{\partial I_{KL}(\mathbf{u})}{\partial W} W^T W = E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T - I] W^{-T} W^T W \quad (2.49)$$

$$= E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T - I] W \quad (2.50)$$

のように探索されることになる。したがって、Natural Gradient アルゴリズムは

$$W + \tilde{\Delta}W = W - \eta E[\boldsymbol{\varphi}(\mathbf{u}) \mathbf{u}^T - I] W \quad (2.51)$$

のように定式化される。

2.3.4 Negentropy

FastICA アルゴリズムは、独立性の評価に非ガウス性に着目し、非ガウス性を測る尺度に Negentropy を用いて、分離信号の Negentropy が最大となるような分離荷重ベクトル \mathbf{w}_n を一つずつ推定する手法である。したがって、分離信号は非ガウス性の高い原信号から順に、一つずつ推定されることになる。以降では、Negentropy と FastICA アルゴリズムについて概説する。

Negentropy は情報量の Entropy に基づいて定義される。 u_n の Entropy $\mathcal{H}(u_n)$ を再度表記する。

$$\mathcal{H}(u_n) = - \int p(u_n) \log p(u_n) du_n \quad (2.52)$$

Entropy は情報理論の基本概念で、確率変数の持つ情報の曖昧さと解釈できる。変数がより確率的、つまり、推測が困難な場合、Entropy は大きくなる。逆に、推測が容

易な場合, Entropy は小さくなる. 情報理論の基本的結果から, 同じ分散を持つ確率変数の中で最も Entropy が大きいのは, 確率変数がガウス分布になるときである. つまり, ガウス分布は他のどんな分布よりも確率的である. 逆に, 確率変数がある値に集中した分布を持つ場合には, Entropy は小さくなる. このことは, Entropy が非ガウス性の尺度として利用できることを意味している.

式 (2.52) の Entropy に基づいて, 常に非負の値をとり, 確率変数がガウス分布であるときのみ 0 の値をとるようにした統計量を Negentropy という. Negentropy $J(u_n)$ は

$$J(u_n) = \mathcal{H}(u_{gauss}) - \mathcal{H}(u_n) \quad (2.53)$$

と定義される. ここに, u_{gauss} は u_n と等しい平均と分散を持つガウス分布の確率変数である. しかし, Negentropy $J(u_n)$ を非ガウス性の尺度として利用するには, 確率変数 u_n の確率密度関数を推定する必要がある, その計算量は膨大となる. そこで, Negentropy を直接求める代わりに, 近似する方法を考える.

古典的な Negentropy の近似は, 確率変数 u_n の平均は 0 で分散は 1 と仮定して

$$J(u_n) \propto \frac{1}{12}E[u_n^3]^2 + \frac{1}{48}\kappa(u_n)^2 \quad (2.54)$$

のように高次統計量で与えられる [69]. ここに, $\kappa(u_n)$ は u_n の Kurtosis で

$$\kappa(u_n) = E[u_n^4] - 3\{E[u_n^2]\}^2 \quad (2.55)$$

のように 2 次のモーメントと 4 次のモーメントで定義される.

しかし, このように高次統計量を使った近似は, 異常値に対して敏感であり, 頑健性に問題がある. この問題を避けるため, Hyvärinen は新しい近似

$$J(u_n) \propto \sum_{i=1}^p \alpha_i \{E[F_i(u_n)] - E[F_i(\nu)]\}^2 \quad (2.56)$$

を提案している [70]. ここに, α_i は正の定数, ν は平均が 0 で分散が 1 のガウス変数, $F_i(\cdot)$ は 2 次的でない関数 (non-quadratic function) である. 式 (2.56) は Negentropy を正確に近似していない場合でも, 常に非負であり, 確率変数 u_n がガウス分布のときのみ 0 になるという点で, Negentropy の性質を反映している.

ここで、非線形関数として

$$F(u_n) = \log \cosh(u_n) \quad (2.57)$$

$$F(u_n) = -\exp(-u_n^2/2) \quad (2.58)$$

を選ぶことで、式 (2.56) は

$$J(u_n) \propto \{E[F(u_n)] - E[F(\nu)]\}^2 \quad (2.59)$$

となり、一つの関数で近似できる。また、式 (2.59) は、より頑健な近似となっており、非ガウス性の尺度として有効となる。したがって、式 (2.59) を最大化することで、非ガウス性の最も高い信号が推定されることになる。

2.3.5 中心化と白色化

FastICA を適用する場合には、そのアルゴリズムで混合信号の平均が 0 で分散が 1 と仮定しているため、中心化 (Centering) と白色化 (Whitening) の前処理が必須となる。

中心化は、混合信号 \mathbf{x} の平均を 0 にすることであり、中心化後の混合信号 \mathbf{x}' は

$$\mathbf{x}' = \mathbf{x} - E[\mathbf{x}] \quad (2.60)$$

のように計算する。

この中心化した \mathbf{x}' をもとに白色化を行う。白色化は、混合信号 \mathbf{x}' を無相関で分散が 1 となるベクトル $\tilde{\mathbf{x}}$ に変換することである。言い換えれば、 $\tilde{\mathbf{x}}$ の共分散行列 $E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T]$ が

$$E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = I \quad (2.61)$$

となるように変換することである。ここに、 I は単位行列である。白色化を行う方法の一つに固有値分解法が知られている [71]。固有値分解法は、 \mathbf{x}' の共分散行列 $E[\mathbf{x}'\mathbf{x}'^T]$ を

$$E[\mathbf{x}'\mathbf{x}'^T] = \Phi\Lambda\Phi^T \quad (2.62)$$

と分解する．ここに、 $\Lambda = \text{diag}[\lambda_1, \dots, \lambda_N]$ 、 $\Phi = [\phi_1, \dots, \phi_N]$ で、 λ_n と ϕ_n はそれぞれ共分散行列 $E[\mathbf{x}'\mathbf{x}'^T]$ の固有値と固有ベクトルである．固有値分解法で求めた Φ と Λ を用いて、白色化した混合信号 $\tilde{\mathbf{x}}$ は

$$\tilde{\mathbf{x}} = \Lambda^{-\frac{1}{2}} \Phi^T \mathbf{x}' = Q \mathbf{x}'$$

と求められる．ここに、 Q は白色化行列である．このようにして、混合信号は無相関化され、分散が1になる．

2.3.6 FastICA アルゴリズム

FastICA アルゴリズムは、非ガウス性の最大化に基づいて原信号を推定する方法である．すなわち、非ガウス性の尺度に Negentropy を用いて、 u_n の分散が1という制約条件の下で、その最大値を求める条件付き最適化問題となる．ここで、Negentropy の近似値 $J(\mathbf{w}_n)$ は

$$J(\mathbf{w}_n) = \left\{ E[F(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[F(\nu)] \right\}^2 \quad (2.63)$$

のように求められる．また、制約条件は

$$E[u_n^2] = E[\mathbf{w}_n^T \tilde{\mathbf{x}} \tilde{\mathbf{x}}^T \mathbf{w}_n] = \mathbf{w}_n^T E[\tilde{\mathbf{x}} \tilde{\mathbf{x}}^T] \mathbf{w}_n = \|\mathbf{w}_n\|^2 = 1 \quad (2.64)$$

のように表現できる．したがって、ラグランジュの未定乗数法より

$$L(\mathbf{w}_n) = \left\{ E[F(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[G(\nu)] \right\}^2 - \beta \{\|\mathbf{w}_n\|^2 - 1\} \quad (2.65)$$

と表記できる．したがって、分散が1の制約条件下で Negentropy の最大値を求めることは、Eq. (2.65) の微分値

$$\frac{\partial L(\mathbf{w}_n)}{\partial \mathbf{w}_n} = 2E[\tilde{\mathbf{x}} f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - 2\beta \mathbf{w}_n = 0 \quad (2.66)$$

の解を求めることに帰着される．ここに、 β は未定乗数、 $f(\cdot)$ は非線形関数 $F(\cdot)$ の導関数で、Eq. (2.57) と Eq. (2.58) のとき、それぞれ

$$f(u_n) = \tanh(u_n) \quad (2.67)$$

$$f(u_n) = u_n \exp(-u_n^2/2) \quad (2.68)$$

のように計算できる.

Eq. (2.66) から得られる

$$E[\tilde{\mathbf{x}}g(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n = 0 \quad (2.69)$$

の解を求めるため, Newton 法を用いると, 更新後の分離荷重 \mathbf{w}_n^+ は

$$\begin{aligned} \mathbf{w}_n^+ &= \mathbf{w}_n - \left\{ \frac{\partial}{\partial \mathbf{w}_n} (E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n) \right\}^{-1} (E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n) \\ &= \mathbf{w}_n - \left\{ E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta I \right\}^{-1} (E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n) \\ &\approx \mathbf{w}_n - \left\{ E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta I \right\}^{-1} (E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n) \\ &= \mathbf{w}_n - \left\{ E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] I - \beta I \right\}^{-1} (E[\tilde{\mathbf{x}}g(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n) \\ &= \mathbf{w}_n - \frac{E[f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n}{E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta} \end{aligned} \quad (2.70)$$

のように逐次的に求められる. ここに, $f'(\cdot)$ は $f(\cdot)$ の導関数で, Eq. (2.67) と Eq. (2.68) の場合, それぞれ

$$f'(u_n) = 1 - \tanh^2(u_n) \quad (2.71)$$

$$f'(u_n) = (1 - u_n^2) \exp(-u_n^2/2) \quad (2.72)$$

となる. Eq. (2.70) は

$$\begin{aligned} \left\{ \beta - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \right\} (\mathbf{w}_n^+ - \mathbf{w}_n) &= E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n \\ \left\{ \beta - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \right\} \mathbf{w}_n^+ &= E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n \\ \mathbf{w}_n^+ &= \frac{1}{\beta - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})]} \left\{ E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n \right\} \end{aligned} \quad (2.73)$$

のように解くことができる. そして, 制約条件 $\|\mathbf{w}\|^2 = 1$ を満たすように

$$\mathbf{w}_n = \frac{\mathbf{w}_n^+}{\|\mathbf{w}_n^+\|} \quad (2.74)$$

$$= \frac{\frac{1}{\beta - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})]} \left\{ E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n \right\}}{\frac{1}{\beta - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})]} \|E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n\|} \quad (2.75)$$

と規格化される. したがって, FastICA アルゴリズムの荷重更新式は

$$\mathbf{w}_n^+ = E[\tilde{\mathbf{x}}f(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[f'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n \quad (2.76)$$

と導出される.

以上のことから、原信号の一つを推定する場合のFastICA アルゴリズムは、Eq. (2.76) で逐次的に荷重更新し、Eq. (2.74) で規格化すればよいことになる。このときの収束条件は、更新前後の分離荷重 \mathbf{w}_n の向きが一致すること、つまり

$$|\mathbf{w}_{n,\text{old}}^T \mathbf{w}_{n,\text{new}}| \simeq 1 \quad (2.77)$$

を満たすことである。ここに、添え字の old と new はそれぞれ更新前後を表す。ここで得られた分離荷重 \mathbf{w}_n を用いて、 $u_n = \mathbf{w}_n^T \tilde{\mathbf{x}}$ と計算すれば、分離信号 u_n は原信号の一つが推定されることになる。

二つ以上の原信号を推定する場合、上述の FastICA アルゴリズムを適用して、原信号を一つ一つ推定すればよい。このとき、二つ以上の分離荷重が同じ値をとる可能性があるため、それぞれの分離荷重を直交化して無相関化する。具体的には、二つ以上の分離荷重が得られたとき、グラムシュミットの直交化法により

$$\mathbf{w}_n = \mathbf{w}_n - \sum_{i=1}^{n-1} \mathbf{w}_i^T \mathbf{w}_n \mathbf{w}_i \quad (2.78)$$

のように、それぞれの分離行列を直交化させて、再度、Eq. (2.74) に代入し規格化する。このようにして、全ての分離荷重 \mathbf{w}_n ($n = 1, \dots, N$) を推定すれば、分離行列 W は $W = [\mathbf{w}_1, \dots, \mathbf{w}_N]^T$ のように求められて、全ての分離信号 \mathbf{u} は Eq. (2.15) のように推定する。

FastICA アルゴリズムは分離信号を一つずつ推定するデフレーション (deflation) 法であり、非ガウス性の高い原信号から推定されやすい特徴を持っている。

2.4 独立成分分析による瞬時混合信号の分離

独立成分分析を用いたブラインド信号分離について、本章では、Natural Gradient アルゴリズムと FastICA アルゴリズムについて述べた。ここでは、2 音源 2 マイクロホンの環境の下で、瞬時混合信号から原信号を推定するシミュレーションによって、それぞれのアルゴリズムの有効性を検証する。

原信号は音声信号を用いて、新聞記事読み上げ音声コーパス [72] から 10 パターン (男性 5 パターン、女性 5 パターン) の話者の発話音声と、Ambient Noise Database [73] から 2 パターンの駅構内の騒音を、それぞれ原信号 $s_1(t)$ と $s_2(t)$ とした。音声信号は、

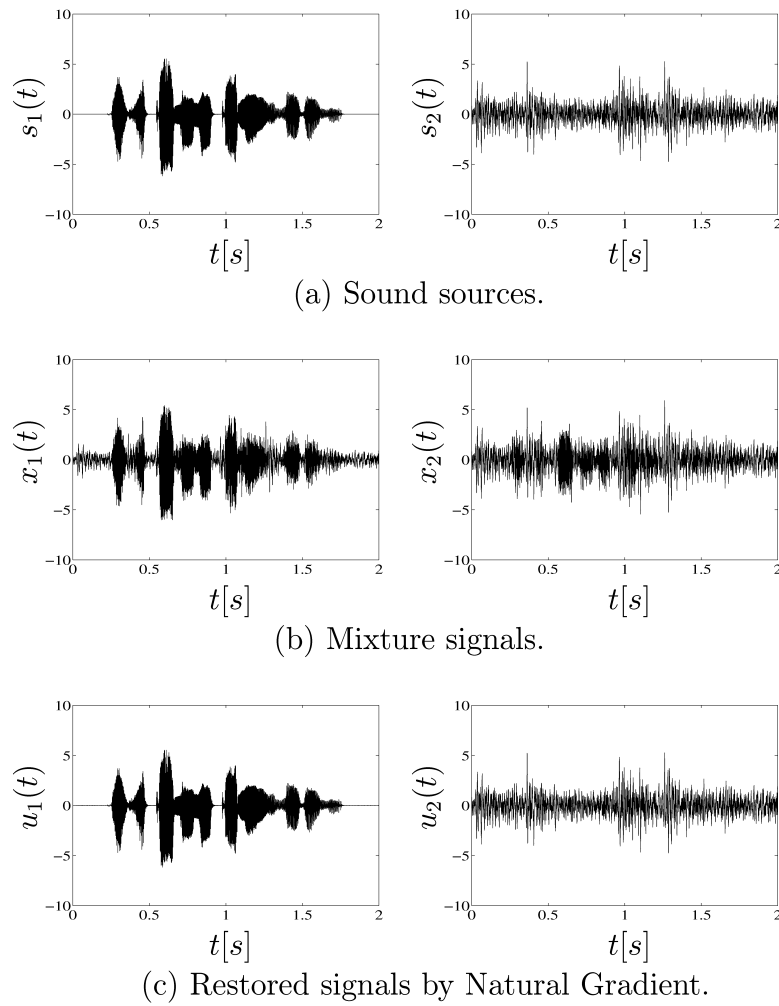


Fig. 2.11: simulation result using Natural Gradient algorithm.

サンプリング周波数を 8000[Hz], 分解能を 16[bit] で, 平均は 0, 分散は 1 とした. また, 混合行列 A は対角成分が 1 で非対角成分が 0.4~0.6 の乱数として, 合計 20 組の混合信号を生成した.

2.4.1 Natural Gradient による分離シミュレーション

Natural Gradient アルゴリズムでは, 分離行列 W の各要素 w_{mn} の初期値は $-1 \sim 1$ の乱数, 非線形関数を $\varphi_n(u_n) = \tanh(u_n(t))$, 最大繰り返し回数を 100, 学習係数は $\eta = 0.1$ とした.

この条件の下で, 推定される分離信号の一例を Fig. 2.11 に示す. Fig. 2.11 の (a) は

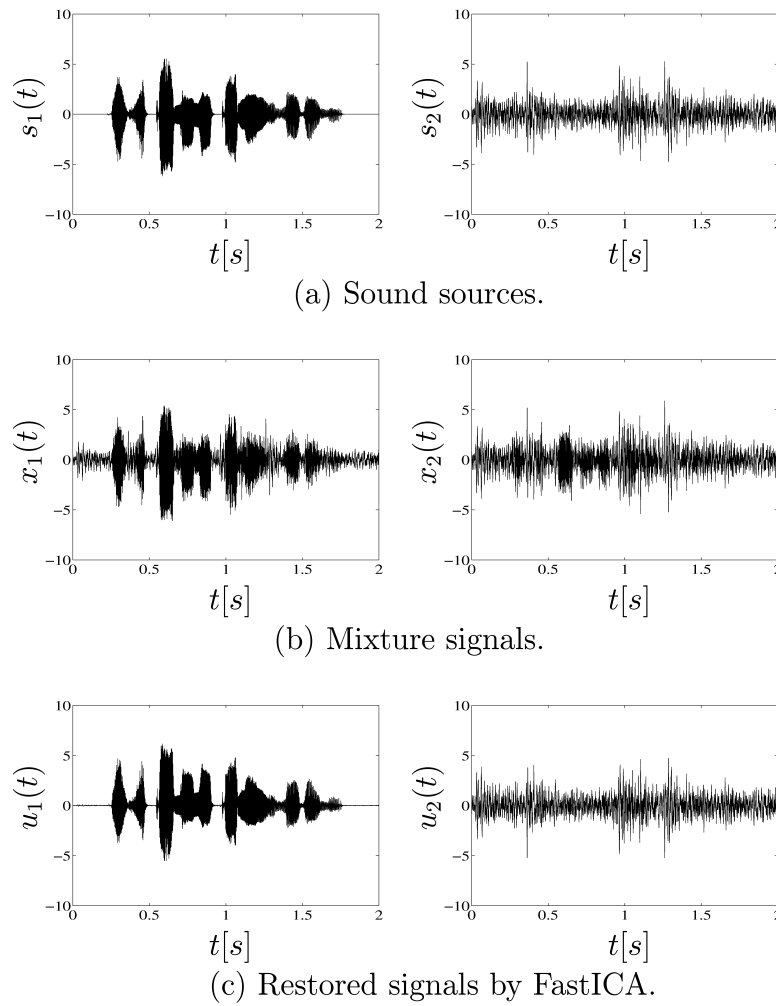


Fig. 2.12: simulation result using FastICA algorithm.

原信号, (b) は瞬時混合信号, (c) は Natural Gradient アルゴリズムを用いて推定した分離信号である. この結果から, 分離信号は完全に原信号を復元できていることが分かる.

Fig. 2.11 では, 原信号の符号と分離信号の符号は一致し, 原信号の順序と分離信号の順序は一致していることが分かる. しかし, アルゴリズムにおける分離行列の初期値に依存してこれらの符号や順序が変わり, スケールの不定性と成分置換の問題が残ることも確認した. 以上のことから, 時間領域での Natural Gradient アルゴリズムは, スケールの不定性と成分置換の問題は残るものの, それらは大きな問題とならないため, 原信号を復元できていることを確認した.

2.4.2 FastICA による分離シミュレーション

FastICA アルゴリズムでは、初期荷重はノルムが1の乱数、非線形関数を $f(u_n(t)) = \tanh(u_n(t))$ 、最大繰返し回数を100、誤差を0.000001とした。

このときの推定信号を Fig. 2.12 に示す。(a) は原信号、(b) は瞬時混合信号、(c) は推定された分離信号である。推定された分離信号は、原信号と符号が反転しているが、原信号を復元できていることが分かる。また、FastICA アルゴリズムでは、非ガウス性の高い発話音声が最初に推定されやすいことも確認した。

2.5 結言

観測した混合信号のみを用いて、未知の原信号と伝達関数を推定するブラインド信号分離問題において、本章では、その問題を解くための代表的な手法について概説した。その中で、独立成分分析は原信号が互いに統計的に独立であると仮定の下で、混合信号から原信号を推定する統計的手法であり、定常信号だけでなく非定常信号も分離できる特徴を持っていることを述べた。

次に、独立成分分析を実行するため、Natural Gradient アルゴリズムと FastICA アルゴリズムについて述べた。Natural Gradient アルゴリズムは、独立性の尺度に Kullback-Leibler 情報量を用いて、それが最小になるように分離する逐次更新アルゴリズムで、分離信号を全て同時に出力する。また、FastICA アルゴリズムは、非ガウス性の最大化に基づく手法で、その尺度に Negentropy を用いて、Negentropy が最大となるように分離するアルゴリズムである。FastICA アルゴリズムは分離信号を一つずつ推定するデフレーション法であり、収束が速く、非ガウス性の高い信号から推定されやすい特徴を持っている。

最後に、Natural Gradient アルゴリズムと FastICA アルゴリズムの分離性能を調べ、スケールの不定性と成分置換の問題を除いて原信号を復元できることを瞬時混合信号によるシミュレーションによって確認した。さらに、成分置換の問題はアルゴリズムの初期値に依存するが、FastICA アルゴリズムでは非ガウス性の高い信号から出力される頻度が高いことを確認した。

本章では瞬時混合における独立成分分析について述べてきたが、実際の雑音除去や

特徴抽出に用いる場合、独立成分分析をそのまま適用できない。すなわち、実環境下で音声信号を分離して雑音除去に用いる場合、観測信号は音源と伝達関数が畳み込まれて観測されたため、瞬時混合モデルでは正確に分離できない。また、脳波解析における特徴抽出に用いる場合、脳内の原信号は完全に未知であるため、分離信号にどのような物理的意味があるのかは分からない。したがって以降では、独立成分分析の不定性問題の解消法を述べた後、実環境下で観測された音声信号の雑音除去と、実際に観測された脳波の特徴抽出について、独立成分分析を適用するときの問題点を明らかにして、その問題の解決方法を提案する。

第3章 瞬時混合過程における独立成分分析の不定性問題の解決

3.1 緒言

観測した混合信号のみを用いて原信号と伝達関数を推定するブラインド信号分離において、原信号の統計的独立性に基づく独立成分分析は精度良く原信号を推定できる。さらに、独立成分分析では原信号が互いに統計的に独立という制約のみを必要としているため、人の発話した音声信号のような非定常信号も容易に分離できる。他にも、脳のデータ解析や通信での雑音除去、画像の分解、機械振動の解析だけでなく、時系列予測や社会学、心理学のデータにも応用され、観測信号の隠れた構造を見つけるための特徴抽出として利用されている。

前に述べたように、独立成分分析によって推定される分離行列が混合行列の逆行列となれば、原信号の完全な復元が可能である。しかしながら、独立成分分析は、単に混合信号から互いに統計的に独立となる分離信号を抽出するだけなので、出力される分離信号には曖昧さが残ることになる。すなわち、スケールの不定性 (Scaling) と成分置換 (Permutation) の問題である。

スケールの不定性問題は分割信号 (Decomposed signal) を導入して解消できる [9], [10]。この分割信号は原信号と伝達関数の積で表現できる。また、分割信号のスケールは伝達関数によって決まるため、スケールの不定性は自ずと解消できる [11]~[13]。さらに、原信号数が混合信号数より少ない場合、分割信号を適用することで簡単に原信号数の推定が可能である [14]。

本章の構成は以下の通りである。まず、**3.2** では、独立成分分析のスケールの不定性と成分置換の問題について詳しく説明する。次に、分割信号を導入したスケールの不定性の解消法を **3.3** で述べる。このとき、分割信号は原信号と伝達関数の積で一意に表現でき、分割信号のスケールは伝達関数によって定められることを導出する。**3.4**

では、原信号数と混合信号数が異なる条件下で、原信号数と観測信号数が等しいと仮定して独立成分分析を用いた原信号の推定シミュレーションを行う。シミュレーション結果から、独立成分分析による分離信号にスケールを回復した後の推定信号は、原信号数が混合信号数より同じか少ないときには原信号に一致し、原信号数が混合信号数より多いときには非ガウス性の高い原信号の特徴を抽出することを述べる。さらに、原信号数が混合信号数より少ない場合、分割信号を適用することで原信号の数が簡単に推定できることを述べる。3.5 は本章のまとめである。

3.2 独立成分分析の限界

独立成分分析によって推定される分離信号には、スケールの不定性と成分置換の問題が残る。すなわち、独立成分分析では、分離行列 W は混合行列 A の逆行列 A^{-1} とはならず

$$WA = PD \quad (3.1)$$

のようにスケールの不定性と成分置換の問題が残して推定される。ここに、 P は各行各列で値が1の要素を一つ持ち、それを除いた全ての要素での値が0である置換行列、 $D = \text{diag}[d_1, \dots, d_N]$ は対角行列で、その要素 d_n は振幅の係数を表す。したがって、分離信号 u_n には d_n 倍された原信号が推定され、分離信号の順番と原信号の順番は必ずしも一致しないことになる。

以降では、これらの不定性問題について詳しく説明する。

3.2.1 スケールの不定性

スケールの不定性は、推定された分離信号が $\mathbf{u} = D\mathbf{s}$ のように、原信号 s_n が任意の定数 d_n 倍されて、 $u_n = d_n s_n$ として推定される問題である。この原因は以下のように説明できる。

混合信号は Eq. (2.14) のように $\mathbf{x} = A\mathbf{s}$ と表されるが、混合行列 A と独立成分 \mathbf{s} の両方が未知であるため、原信号を $D\mathbf{s}$ 、混合行列を AD^{-1} とした場合でも

$$\mathbf{x} = (AD^{-1})D\mathbf{s} = A\mathbf{s} \quad (3.2)$$

のように全く同じ混合信号が得られることになる。この状況の下で、混合行列 AD^{-1} の逆行列が分離行列 W として得られた場合、分離信号 \mathbf{u} は

$$\begin{aligned}\mathbf{u} &= W\mathbf{x} \\ &= (AD^{-1})^{-1}(AD^{-1})D\mathbf{s} \\ &= D\mathbf{s}\end{aligned}\tag{3.3}$$

となり、 D 倍された信号が推定されることになる。

このとき、分離信号 u_n の同時確率密度関数 $p(u_1, \dots, u_N)$ は

$$\begin{aligned}p(u_1, \dots, u_N) &= \prod_{n=1}^N p(u_n) \\ &= \prod_{n=1}^N p(d_n s_n)\end{aligned}\tag{3.4}$$

のように、周辺確率密度関数 $p(u_n)$ の積で表現できる。このことが、スケールが不定となる原因である。

3.2.2 成分置換

独立成分分析法では、単に混合信号を統計的に独立となる信号に分離するだけなので、分離信号 \mathbf{u} は $u_1 = s_1, u_2 = s_2, \dots$ と推定されるとは保証されず、 $u_1 = s_2, u_2 = s_1, \dots$ のように入れ替わって推定されることも起きる。この問題を成分置換という。この原因は以下のように説明できる。

各行各列で値が1の要素を一つ持ち、それを除いた全ての要素での値が0である置換行列 P は、 $PP = I$ のように単位行列となることより、混合信号 $\mathbf{x} = A\mathbf{s}$ は、原信号を $P\mathbf{s}$ 、混合行列を AP とした場合でも

$$\begin{aligned}\mathbf{x} &= A\mathbf{s} \\ &= APP\mathbf{s}\end{aligned}\tag{3.5}$$

のように全く同じになる。この状況の下で、混合行列 AP の逆行列が分離行列 W として得られた場合、分離信号 \mathbf{u} は

$$\mathbf{u} = (AP)^{-1}\mathbf{x}$$

$$\begin{aligned}
 &= (AP)^{-1}APP\mathbf{s} \\
 &= P\mathbf{s}
 \end{aligned} \tag{3.6}$$

となり，分離信号 \mathbf{u} は信号源 \mathbf{s} が入れ替わって推定される．

このとき，分離信号 u_n の同時確率密度関数 $p(u_1, \dots, u_N)$ は

$$\begin{aligned}
 p(u_1, \dots, u_N) &= \prod_{n=1}^N p(u_n) \\
 &= \prod_{n=1}^N p(s_n)
 \end{aligned} \tag{3.7}$$

のように，周辺確率密度関数 $p(u_n)$ の積で表現できる．このことが，成分置換が起きる原因である．

スケールの不定性と成分置換の問題は同時に生じる．以降では，これらの不定性問題の解消法について述べる．

3.3 不定性問題の解決

本節では，分割信号を導入してスケールの不定性を解消する方法を述べる．このとき，分割信号は原信号と伝達関数の積で表現できることを導出して，その物理的解釈を与える．さらに，この物理的な解釈に基づいて，分割信号にはスケールの不定性がないことを明らかにする．

3.3.1 分割信号の定義

分割信号 $\mathbf{v}_n = [v_{n1}, \dots, v_{nM}]^T$ は，分離行列 W の逆行列 W^{-1} を用いて

$$\mathbf{v}_n(\omega, k) = W^{-1}[0, \dots, 0, u_n, 0, \dots, 0]^T \tag{3.8}$$

と定義される [9], [10]．

これらの分割信号の和は

$$\sum_{n=1}^N \mathbf{v}_n = \mathbf{x} \tag{3.9}$$

のように混合信号と等しくなる．これは，分割信号のスケールが混合信号のスケールと等しいところまで戻されることを意味している．

3.3.2 分割信号の原信号と伝達関数による表現

本節では簡単のため、原信号数と混合信号数がともに二つである場合を考える。このときの分割信号は

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = W^{-1} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} \quad (3.10)$$

$$\mathbf{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = W^{-1} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} \quad (3.11)$$

のように求められる。

まず、成分置換が無くスケールの不定性のみが生じる場合、Eq. (3.1) の $WA = PD$ は、その右辺 P が単位行列となって、分離行列 W と混合行列 A の関係は

$$WA = D \quad (3.12)$$

と表現されることになる。したがって、分離信号 \mathbf{u} は、 $\mathbf{u} = W\mathbf{x} = WAs = D\mathbf{s}$ となり

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} d_1 s_1 \\ d_2 s_2 \end{bmatrix} \quad (3.13)$$

と表される。つまり、分離信号 \mathbf{u} の各成分 u_n ($n = 1, 2$) は、それぞれ原信号 s_n ($n = 1, 2$) を d_n 倍したものになる。

このときの分割信号 \mathbf{v}_1 は

$$\begin{aligned} \mathbf{v}_1 &= \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = W^{-1} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} = AD^{-1} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} = \frac{1}{d_1} \begin{bmatrix} a_{11}u_1 \\ a_{21}u_1 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}s_1 \\ a_{21}s_1 \end{bmatrix} \end{aligned} \quad (3.14)$$

となり、原信号と伝達関数の積として生成されることが導出できる。同様に、分割信号 \mathbf{v}_2 は

$$\begin{aligned} \mathbf{v}_2 &= \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = W^{-1} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = AD^{-1} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = \frac{1}{d_2} \begin{bmatrix} a_{12}u_2 \\ a_{22}u_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{12}s_2 \\ a_{22}s_2 \end{bmatrix} \end{aligned} \quad (3.15)$$

となり，原信号と伝達関数の積として生成されることが導出できる [11]～[13].

次に，成分置換とスケールの不定性の両方が有る場合，すなわち

$$WA = PD \quad (3.16)$$

となる場合を考える．この場合，分離信号 \mathbf{u} は $\mathbf{u} = W\mathbf{x} = W\mathbf{A}\mathbf{s} = P\mathbf{D}\mathbf{s}$ より，最終的に

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} d_2 s_2 \\ d_1 s_1 \end{bmatrix} \quad (3.17)$$

と表現される．つまり，分離信号 u_1 と u_2 は，対応する原信号が s_2 と s_1 に入れ替わって，それぞれ d_2 倍と d_1 倍されたものとなる．

このときの分割信号 \mathbf{v}_1 は

$$\begin{aligned} \mathbf{v}_1 &= \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = W^{-1} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} = AD^{-1}P \begin{bmatrix} u_1 \\ 0 \end{bmatrix} = \frac{1}{d_2} \begin{bmatrix} a_{12}u_1 \\ a_{22}u_1 \end{bmatrix} \\ &= \begin{bmatrix} a_{12}s_2 \\ a_{22}s_2 \end{bmatrix} \end{aligned} \quad (3.18)$$

となり，原信号と伝達関数の積として生成されることが導出できる．同様に，分割信号 \mathbf{v}_2 は

$$\begin{aligned} \mathbf{v}_2 &= \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = W^{-1} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = AD^{-1}P \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = \frac{1}{d_1} \begin{bmatrix} a_{11}u_2 \\ a_{21}u_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}s_1 \\ a_{21}s_1 \end{bmatrix} \end{aligned} \quad (3.19)$$

となり，原信号と伝達関数の積として生成されることが導出できる [11]～[13].

以上の導出によって，分割信号 v_{nm} は原信号 s_n と伝達関数 a_{mn} の積で表現できて，その組み合わせは排他的一意に決定される．

3.3.3 分割信号の物理的解釈

分離信号 u_n に成分置換が無い場合，Eq. (3.14) で対になって出力される分割信号は原信号 s_1 に対応し， v_{11} は伝達関数 a_{11} を通過してセンサ 1 で観測された原信号 1 で

あり、 v_{12} は伝達関数 a_{21} を通過してセンサ 2 で観測された原信号 1 であることを意味している。同様に、Eq. (3.15) で対として出力される分割信号は原信号 $s_2(\omega, k)$ に対応し、 $v_{21}(\omega, k)$ は伝達関数 a_{12} を通過してセンサ 1 で観測された原信号 2 であり、 v_{22} は伝達関数 a_{22} を通過してセンサ 2 で観測された原信号 2 であることを意味している。

一方、分離信号 u_n に成分置換が有る場合、Eq. (3.18) で対になって出力される分割信号は原信号 s_2 に対応し、 v_{11} は伝達関数 a_{12} を通過してセンサ 1 で観測された原信号 2 であり、 v_{12} は伝達関数 a_{22} を通過してセンサ 2 で観測された原信号 2 であることを意味している。同様に、Eq. (3.19) で対になって出力される分割信号は原信号 s_1 に対応し、 v_{21} は伝達関数 a_{11} を通過してセンサ 1 で観測された原信号 1 であり、 v_{22} は伝達関数 a_{21} を通ってセンサ 2 で観測された原信号 1 であることを意味している。

以上のことから、分割信号 v_{nm} は、成分置換が無いときと有るときで、対応する原信号と伝達関数は異なるが、それぞれどれか一つの原信号がどれか一つの伝達関数を通過して観測された信号である。

3.3.4 スケールの不定性の解消

スケールの不定性は、分割信号を利用することで自ずと解決できる。

まず、成分置換が無い場合、分離信号 u_n は、Eq. (3.13) のように原信号 s_n を定数倍 d_n した信号 $d_n s_n$ として得られるため、スケールは不定となる。一方、分割信号 \mathbf{v}_n は、Eq. (3.14), Eq. (3.15) に示すように、原信号と伝達関数の積として定式化される。したがって、分割信号のスケールは伝達関数の値に一致するため、スケールの不定性が解消される。

また、成分置換が有る場合、分離信号 \mathbf{u}_n は、Eq. (3.17) の $u_1 = d_2 s_2$ と $u_2 = d_1 s_1$ のように原信号を定数倍して成分が入れ替わって得られるため、スケールは不定となる。一方、分割信号は、Eq. (3.18), Eq. (3.19) に示すように、成分置換の問題は残るが、原信号と伝達関数の積として定式化される。このことは、スケールの不定性が解消されたことを意味している。

以上のように、成分置換の有無に関わらず、分割信号は原信号と伝達関数の積で一意に表現でき、スケールの不定性問題は自ずと解消できることが分かる。以上の処理の流れを Fig. 3.1 に示す。

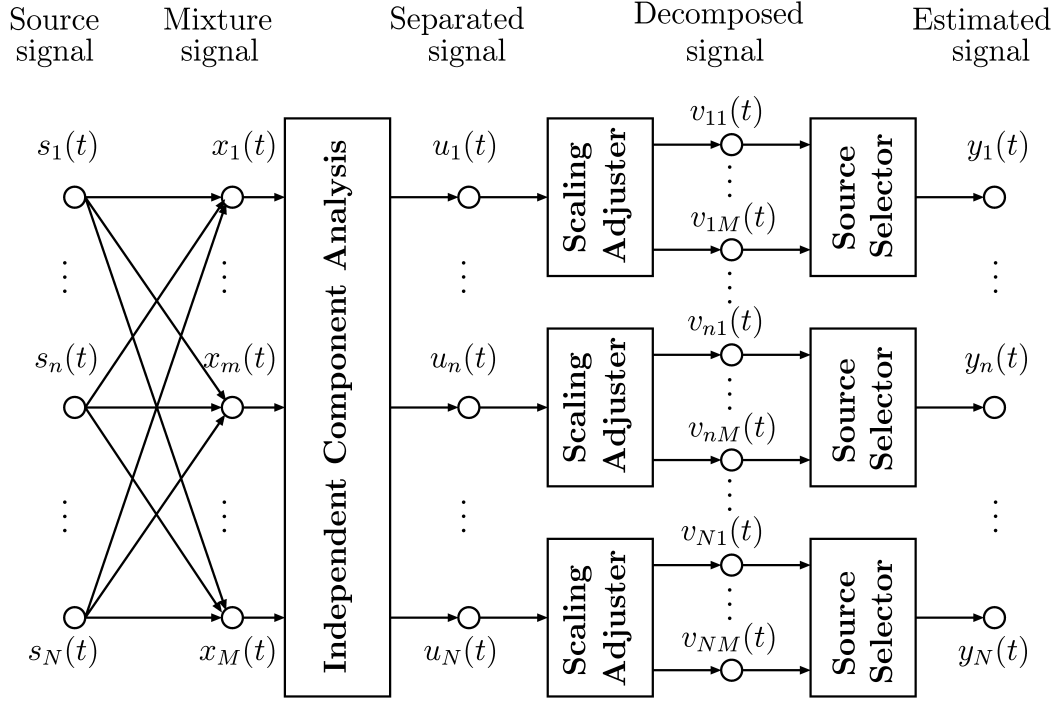


Fig. 3.1: Signal flow from source signals to estimate signals.

3.4 スケールの回復と原信号数の推定シミュレーション

独立成分分析を用いて信号分離を行う場合、原信号数は完全に未知であるため、一般に原信号数は観測信号数と等しいと仮定して解くことが多い。このとき、分割信号 $v_n(t)$ は原信号 $s_n(t)$ と伝達関数 a_{mn} の積として与えられ、 n 番目の原信号 $s_n(t)$ が m 番目のセンサで観測されたときの値である。本節では、分割信号が分離信号のスケールを回復し、さらに、分割信号を適用することで簡単に原信号数の推定ができることをシミュレーションにより確認する。

ここでは、原信号に音声データを用いる。また、混合信号数は5とする。したがって、原信号数 N と観測信号数 M が等しく $N = M = 5$ と仮定して、独立成分分析を用いて原信号の推定を行う。原信号は、新聞記事読み上げ音声コーパス [72] の話者の発話音声と、Ambient Noise Database [73] の工場内の突発音、駅構内の騒音、自動車のエンジン音を用いた。音声信号は、サンプリング周波数を 8000[Hz]、分解能を 16[bit] とした。混合行列は 0.5~1 の乱数で与えた。また、最終的な推定信号 $y_n(t)$ は

$$y_n(t) = \frac{1}{M} \sum_{m=1}^M v_{nm}(t) \quad (3.20)$$

のように、平均した値を出力した。この条件の下で、 $N = 5$ のように $N = M$ の場合、 $N = 4, 2$ のように $N < M$ の場合、 $N = 6, 10$ のように $N > M$ の場合の分離信号について検証する。

3.4.1 原信号数が混合信号数と等しい場合

原信号数と混合信号数が等しく、 $N = M = 5$ のときの結果を Fig. 3.2 に示す。このとき、原信号は $s_1(t)$ に女性の発話、 $s_2(t)$ に男性の発話、 $s_3(t)$ に工場内の突発音、 $s_4(t)$ に駅構内の騒音、 $s_5(t)$ に自動車エンジンの定常雑音とした。Fig. 3.2(a) は原信号 $s_n(t)$ 、(b) は瞬時混合信号 $x_m(t)$ 、(c) は分離信号 $u_n(t)$ 、(d) はスケールを回復した後の推定信号 $y_n(t)$ である。原信号数と観測信号数が等しい場合、分離信号のスケールを回復した後の推定信号は、完全に原信号を復元できていることが分かる。

3.4.2 原信号数が混合信号数より少ない場合

次に、原信号数が観測信号数より少ない場合、すなわち $N = 4, M = 5$ のときの出力信号について調べた。原信号は上述の $s_n(t)$ ($n = 1, \dots, 4$) のときの結果を Fig. 3.3 に示す。Fig. 3.3(d) から分かるように、スケールを回復した推定信号は、四つ全ての原信号と、ほぼ 0 に近いスケールの信号が一つ出力された。すなわち、存在していない原信号に対する推定信号は、振幅が 0 の信号として出力されることが分かる。

さらに、 $N = 2, M = 5$ の場合、すなわち、原信号が上述の $s_n(t)$ ($n = 1, 2$) のときの結果を Fig. 3.4 に示す。Fig. 3.4(d) は、Fig. 3.3 と同様に、存在していない原信号に対する推定信号は、振幅が 0 の信号として出力されることが確認できる。したがって、二つの原信号が復元され、残りは振幅が 0 の信号が出力されている。

以上の結果から、原信号数がセンサ数より同じか少ない場合、独立成分分析を適用した後にスケールを回復させることで、推定信号は全ての原信号を出力することが明確になった。

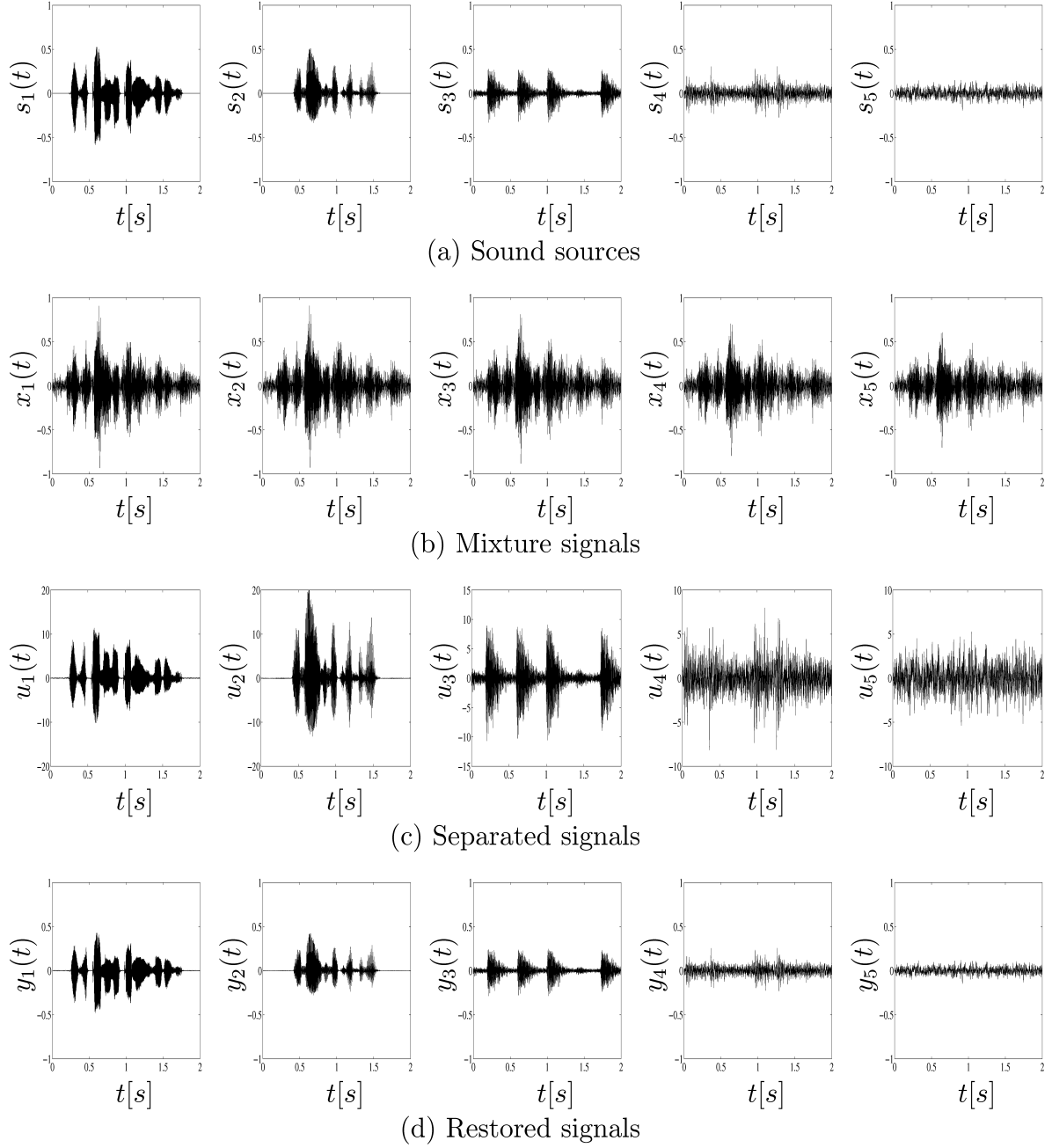


Fig. 3.2: Estimated signals in the case of 5 source signals and 5 mixture signals.

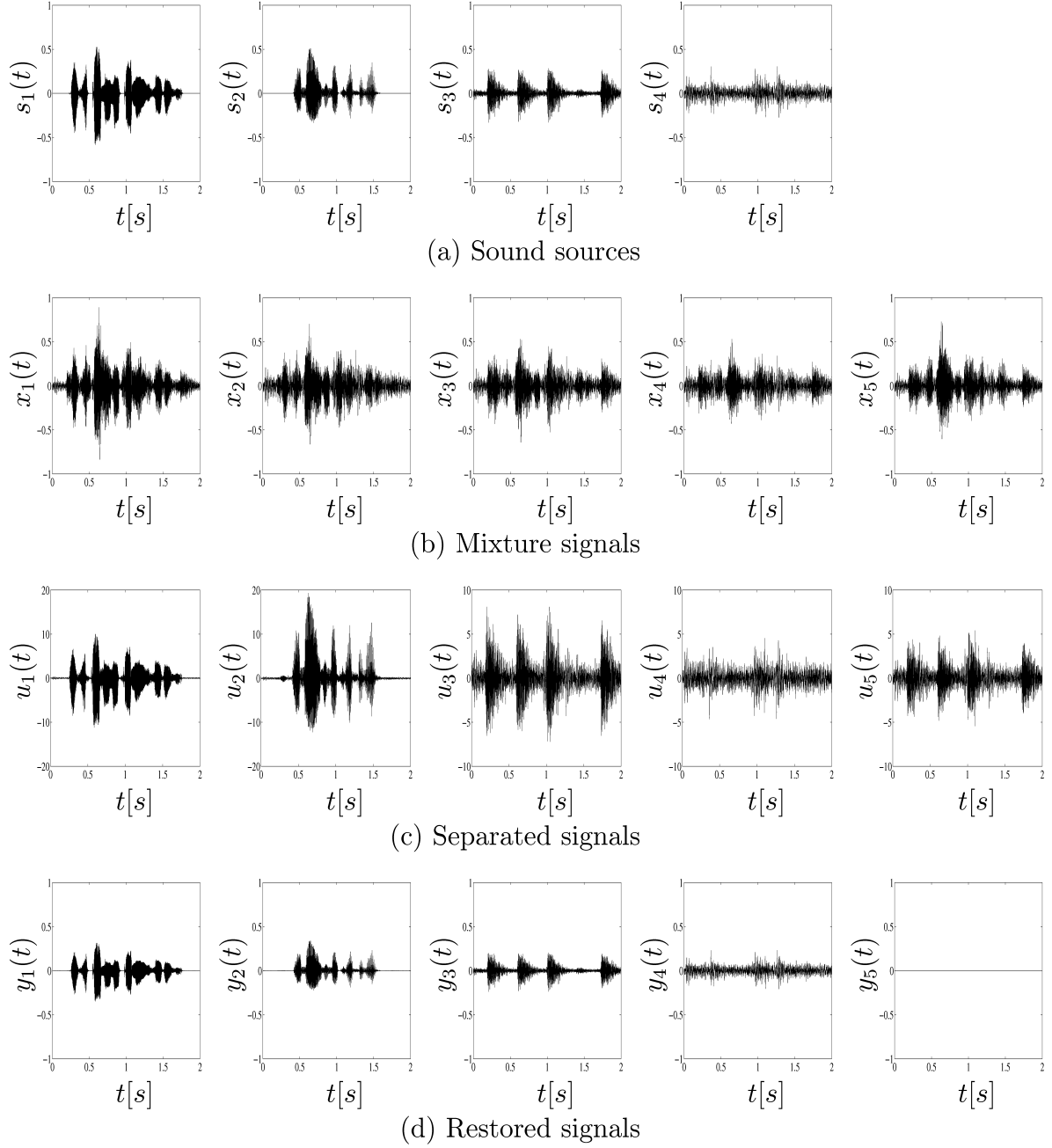
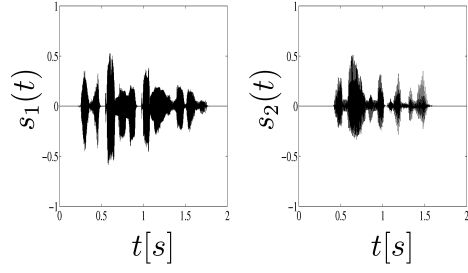
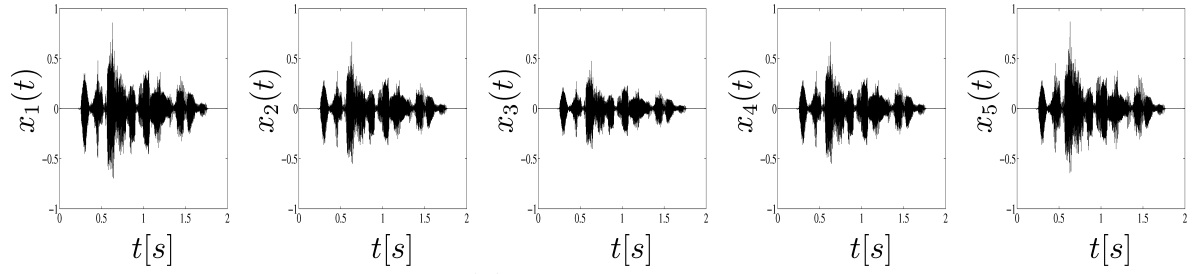


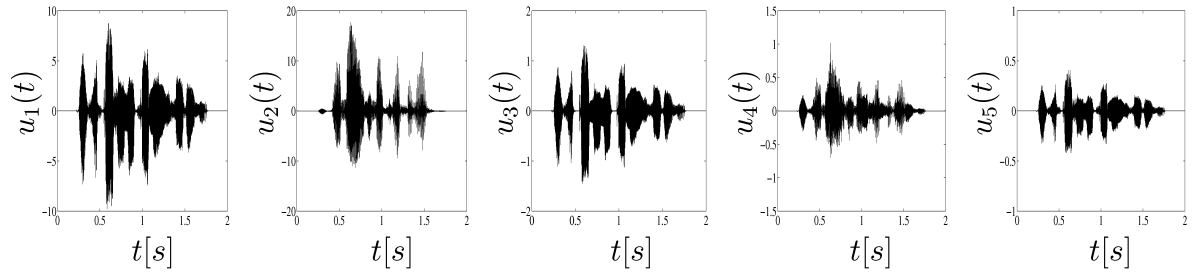
Fig. 3.3: Estimated signals in the case of 4 source signals and 5 mixture signals.



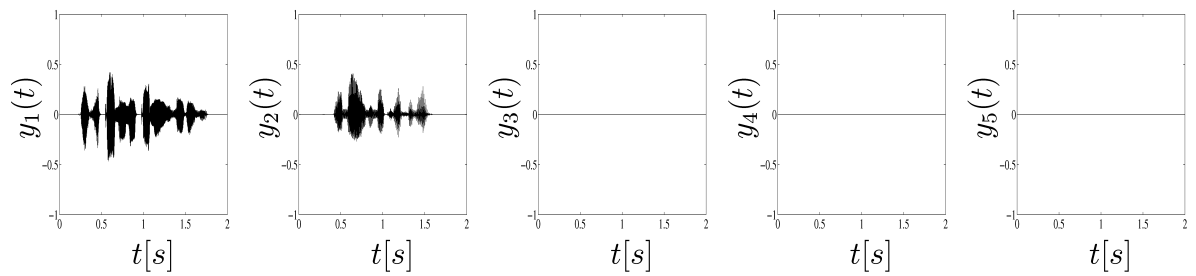
(a) Sound sources



(b) Mixture signals



(c) Separated signals



(d) Restored signals

Fig. 3.4: Estimated signals in the case of 2 source signals and 5 mixture signals.

3.4.3 原信号数が混合信号数より多い場合

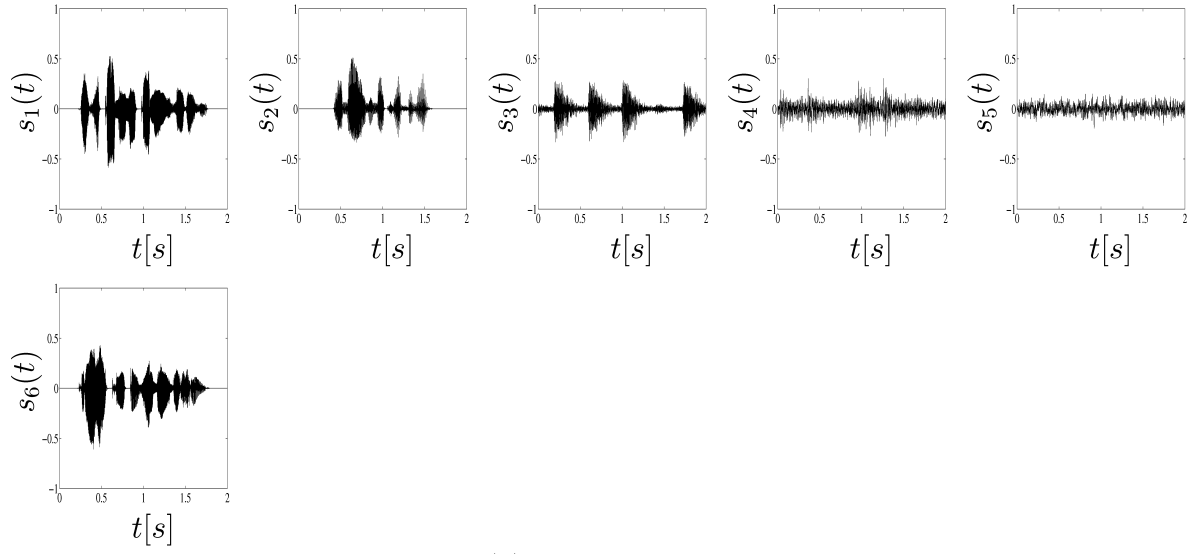
一方，原信号数が観測信号数より多い場合，すなわち， $N = 6$ ， $M = 5$ のときの出力信号について調べた結果を Fig. 3.5 に示す．このときの原信号は，前述の $s_n(t)$ ($n = 1, \dots, 5$) に $s_6(t)$ として女性の発話音声を追加した．Fig. 3.5 から，エンジン音である $s_5(t)$ 以外の音声は，ほぼ正確に推定されていて，それぞれの推定信号に残留雑音が残ることが分かる．特に，人の発話音声である $s_1(t)$ ， $s_2(t)$ ， $s_6(t)$ や突発音である $s_3(t)$ のような非ガウス性の高い原信号については，その特徴をはっきりと捉えていることが分かる．

さらに原信号数を増加させて， $s_7(t)$ に男性の発話， $s_8(t)$ に工場内の突発音， $s_9(t)$ に駅構内の騒音， $s_{10}(t)$ に自動車エンジンの定常雑音を追加した場合，すなわち， $N = 10$ ， $M = 5$ について実験した．このときの結果は，Fig. 3.6 に示すように，原信号数が非常に多い場合は原信号の推定は困難であることが分かる．しかしながら，残留雑音が含まれている中で，話者の発話音声である原信号 $s_1(t)$ は推定信号 $y_1(t)$ に， $s_6(t)$ は $y_3(t)$ に，はっきりと特徴が示されている．また，突発音である $s_3(t)$ は $y_2(t)$ に， $s_8(t)$ は推定信号 $y_4(t)$ に，それぞれの波形の主要な成分が出力されていることが確認できる．この結果は， $N > M$ のときは独立成分分析による分離精度は落ちるが，その分離信号から非ガウス性の高い原信号の特徴抽出が可能であることを意味している．

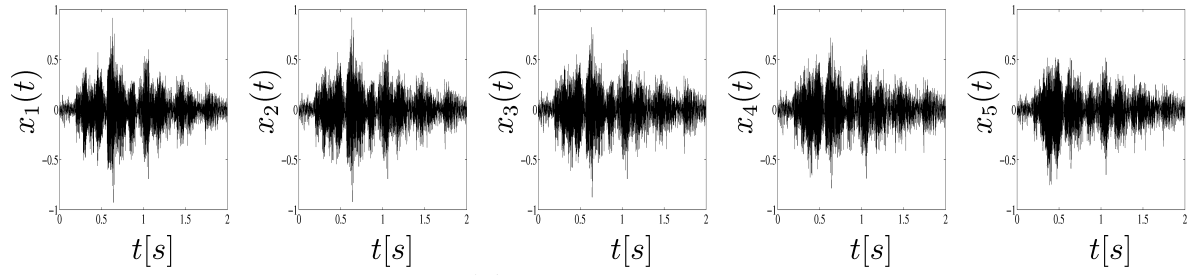
以上の結果をまとめると，原信号数と混合信号数が異なる条件下で，原信号数と観測信号数が等しいと仮定して分離したとき，独立成分分析による分離信号にスケールを回復した後の出力信号は，原信号数が混合信号数より同じか少ないときには原信号に一致し，原信号数が混合信号数より多いときには非ガウス性の高い原信号の特徴を抽出した信号であることが分かった．したがって，原信号数が未知のときでも，独立成分分析によって特徴抽出は可能であり，さらに，原信号数が混合信号数より少ない場合，分割信号を適用することで原信号の数が簡単に推定できる [14]．

3.5 結言

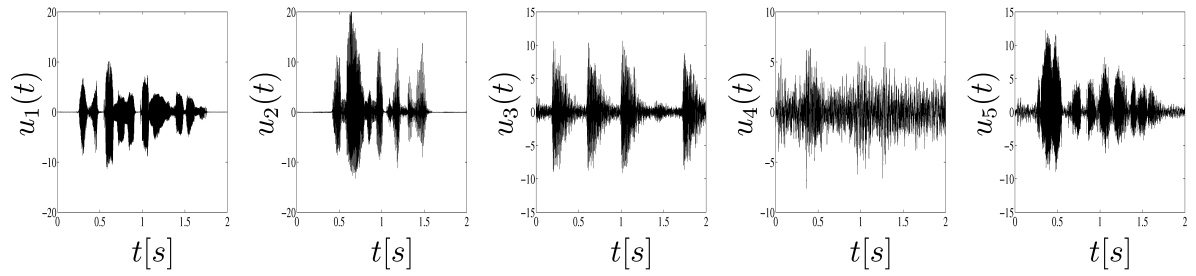
本章では，独立成分分析の不定性問題について述べた．独立成分分析では原信号が互いに統計的に独立という制約の下で解かれるため，非定常信号も容易に分離できる



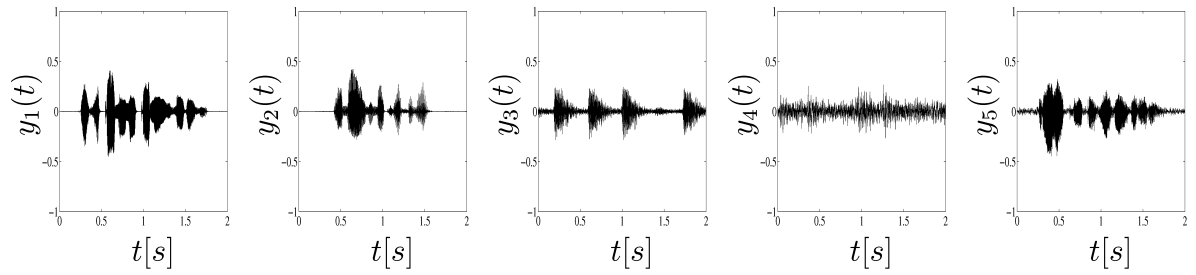
(a) Sound sources



(b) Mixture signals



(c) Separated signals



(d) Restored signals

Fig. 3.5: Estimated signals in the case of 6 source signals and 5 mixture signals.

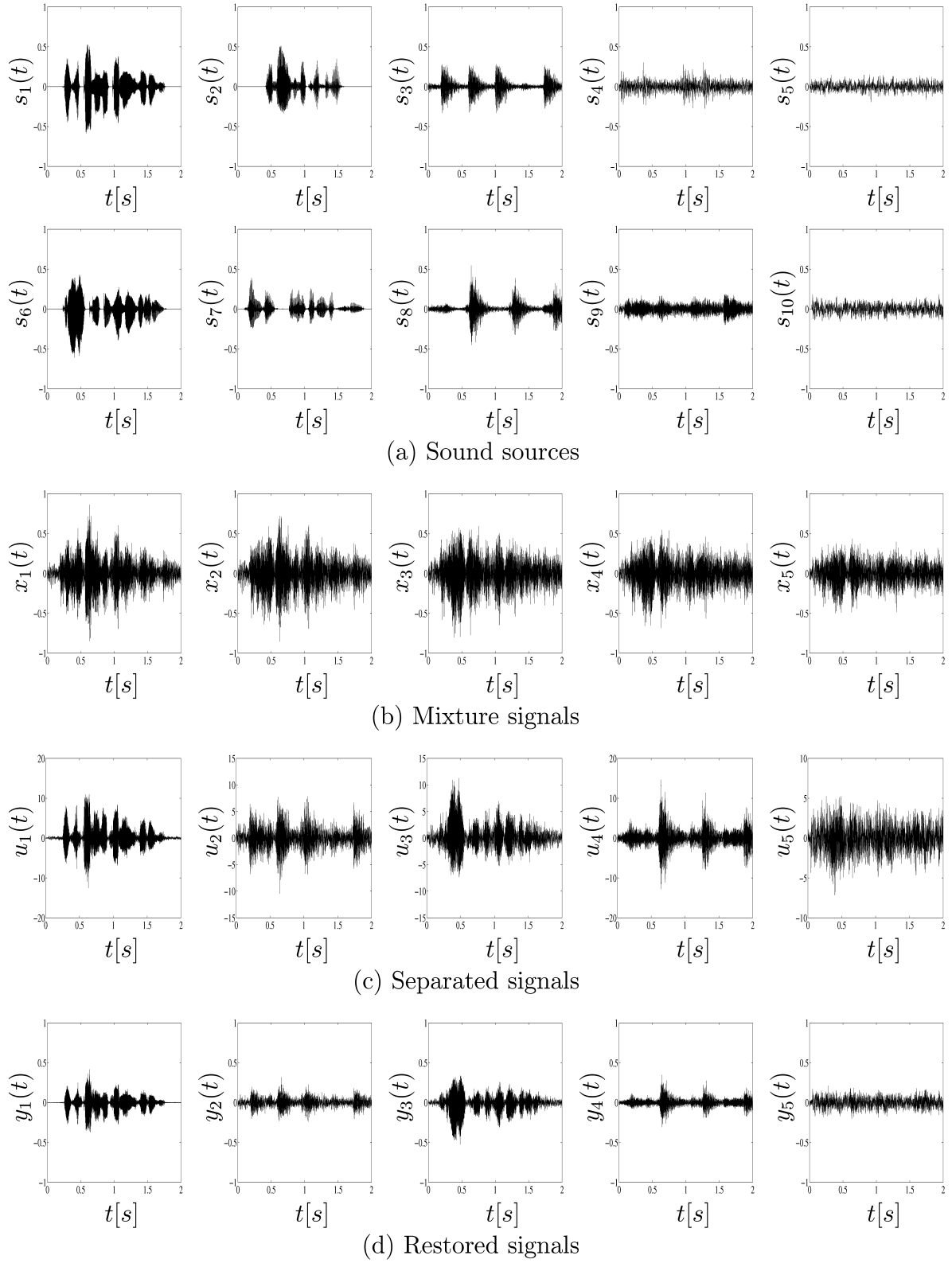


Fig. 3.6: Estimated signals in the case of 10 source signals and 5 mixture signals.

が、出力される分離信号にはスケールの不定性と成分置換の問題が残る。

まず、スケールの不定性問題に対して、分割信号を導入した解消法を述べた。また、分割信号は原信号と伝達関数の積で表現できることを証明した。さらに、分割信号は一つの前信号が一つの伝達関数を通過して観測された信号であるという物理的解釈を与え、分割信号のスケールは伝達関数によって決まるためスケールの不定性は自ずと解消できることを示した。さらに、原信号数が混合信号数より少ない場合、分割信号を適用することで簡単に原信号数の推定が可能であることを述べた。

最後に、瞬時混合での信号分離シミュレーションによって、原信号数と混合信号数が異なる場合の独立成分分析の分離性能を調べた。原信号数と混合信号数が異なる条件下では、分離信号のスケールを回復した推定信号は、原信号数が混合信号数より同じか少ないときには原信号に一致することが分かった。さらに、原信号数が混合信号数より少ない場合、全ての原信号数が推定されて、残りの出力信号は振幅が0になることを確認した。このことから、分割信号を適用することで原信号の数が簡単に推定できることを確認した。一方、原信号数が混合信号数より多い場合、正確に原信号を正確に推定できないものの、原信号の特徴は良く現れていることを確認した。

以上のことから、原信号数が未知である状況下においても、独立成分分析に基づく信号分離や特徴抽出が可能であることを確認した。

第4章 実環境下での音響信号の分離と 目的音声の選択

4.1 緒言

音声認識技術は、パーソナルコンピュータや携帯電話での音声入力だけでなく、広範な分野での利用が期待されている。音声認識の利点は、情報の入力速度がキーボード入力の3～4倍、手書き文字入力の8～10倍と速くなる点や、移動中やほかの作業をしながら入力できる点にある。さらに、マイクロホンや電話機などの既存の機器を入力端末として利用できるため、使用機器の入力操作方法についての訓練を必要としないことや、遠隔地からの入力も可能となることなどが挙げられる。

近年、音声認識技術は著しく進展して、周囲に雑音のない理想的な環境下では認識能力の極めて高い音声認識エンジンが供給できるようになってきている [74]～[78]。しかし、生活音や活動音が存在する一般家庭やオフィスで、期待通りの認識率を得ることは難しい。このような実環境下で音声認識エンジンのもつ本来の認識能力を引き出すためには、雑音混じりの音声から雑音を除去して、話者音声のみを音声認識エンジンに受け渡す前処理が不可欠と考えられる。

実環境下での雑音除去問題について、周波数領域の独立成分分析を適用した手法が数多く提案されている [9], [10], [15]～[22]。しかし、畳み込み混合された観測信号に周波数領域の独立成分分析 (FDICA: Frequency Domain Independent Component Analysis) を適用した場合、個々の周波数でスケールと位相の不定性と成分置換の問題が生じるため、そのまま時間領域に戻しても元の音源は復元されない。

スケールと位相の不定性の問題は、分割スペクトル (Decomposed spectrum) を導入することで解消できる [9], [10]。このとき、分割スペクトルのスケールは伝達関数に支配されることが導出されている [11]～[13]。一方、成分置換の問題は、分離信号の類似度に基づく方法 [9], [10] や、分離行列の推定結果を隣接する周波数の分離行列の初期

値に利用する方法 [23], 混合行列の周波数軸上での連続性を用いた方法 [24] が提案されている. また, Hyvärinen による FastICA アルゴリズム [15], [25], [26] は, 雑音に比べて非ガウス性の高い話者音声を最初に出力する性質があるため, 成分置換問題はある程度緩和される. しかし, これらを実環境下で適用しても実用レベルの結果を得ることは難しい. 他に, マイクロホンアレーとしての指向特性を考慮した方法 [27], [28] が提案されているが, 計算量の問題など改善すべき点が残っている. 話者音声の尖度が雑音に比べて高いことを利用した方法 [29] は, 原理的に, 他人の発話が干渉音となる場合は適用できない. 伝達関数のゲインや位相のそれぞれに基づいた方法 [11], [12], [30], [31] は, 後述するように, ゲインのみや位相のみを用いても成分置換は正しく修正できない.

そこで本章では, 実環境下で有効に機能する雑音除去システムの構築を目的として, 周波数領域の独立成分分析における成分置換問題を解決するとともに, 音源とマイクロホンの相対的な位置情報が得られている場合には目的音声の選択が可能であるアルゴリズムを述べる. まず, 実環境におけるゲインと位相の音響伝達特性を詳細に調べて, その特性が周波数帯域やマイクロホン間隔, 音源からマイクロホンまでの距離に依存するため, ゲインや位相を単独に利用しただけでは成分置換は修正できないことを述べる [32], [33]. 次に, 音響伝達特性のゲインと位相の双方を組み合わせた成分置換の修正法を述べる. 具体的には, ブラインド信号分離では音源からマイクロホンまでの距離は未知であるため, 提案法では実環境の音響特性をもとに, 周波数帯域でゲインと位相を使い分けて成分置換を修正する方法を述べる [34]. 最後に, 音源とマイクロホンの相対的な位置情報が得られない場合に目的音声を簡単に選択する方法について述べる [12].

以降では, 4.2 で周波数領域の独立成分分析と, 周波数領域の FastICA アルゴリズムの概要を述べる. 4.3 では, スケールと位相の不定性と成分置換の問題の解決法を提案する. また, 目的信号の選択法を提案する. 提案法の有効性は, 実環境下での音源分離実験によって検証し, その結果を 4.4 で述べる. 最後に, 本章のまとめを 4.5 で述べる.

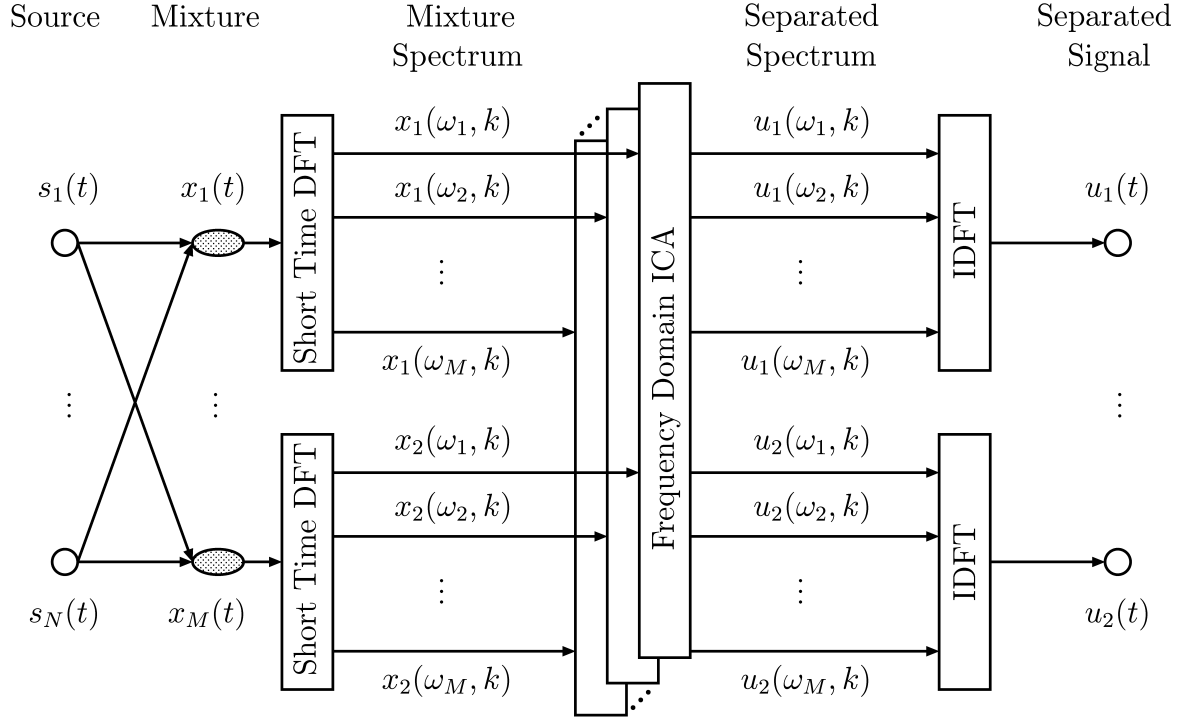


Fig. 4.1: Frequency Domain Independent Component Analysis.

4.2 周波数領域独立成分分析

4.2.1 周波数領域独立成分分析

実環境下で観測される混合信号 $x_m(t)$ は、瞬時混合 (instantaneous mixture) とは異なり、原信号とマイクロホンの中で時間遅れや反射、残響が影響して

$$x_m(t) = \sum_{n=1}^N \sum_t g_{mn}(\tau) s_n(t - \tau) \quad (4.1)$$

のように畳み込まれて観測される．ここに、 $g_{mn}(t)$ は原信号 n からマイクロホン m までの伝達関数である．したがって、周波数領域の独立成分分析 (FDICA: Frequency Domain Independent Component Analysis) のモデルは

$$\mathbf{x}(t) = G(t) * \mathbf{s}(t) \quad (4.2)$$

のように畳み込み混合 (convolutive mixture) または時空間的混合 (spatial-temporal mixture) と呼ばれるモデルで表される．ここに、 $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$, $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$, $G(t)$ は伝達関数 $g_{mn}(t)$ を要素とする混合行列, $*$ は畳み込みを表す．

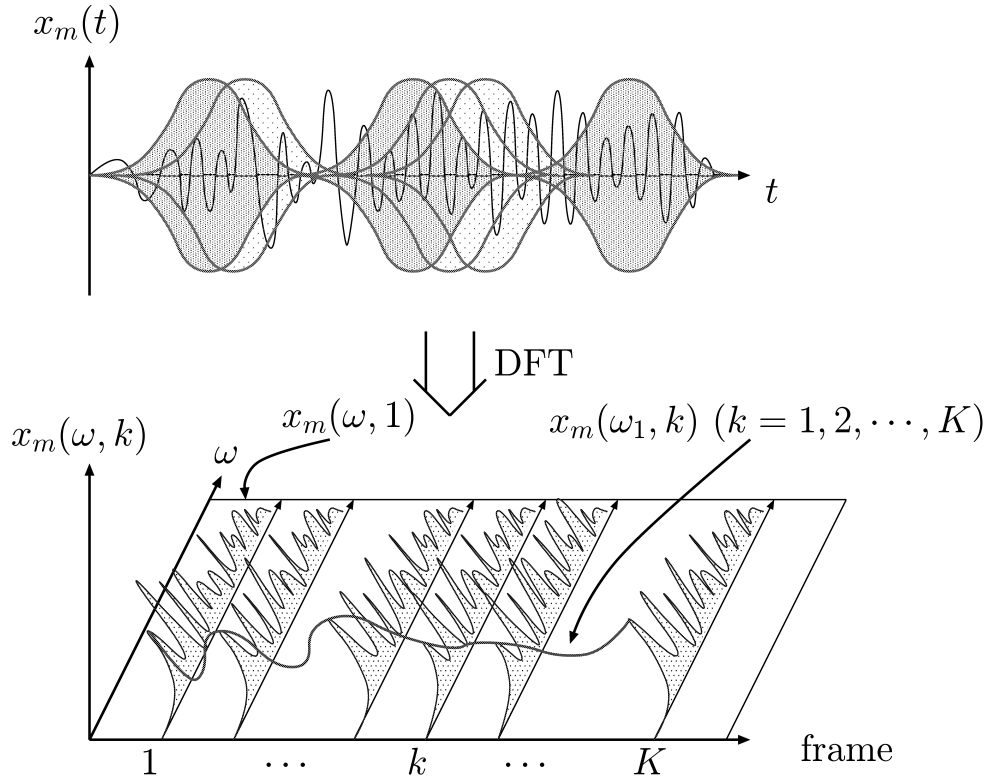


Fig. 4.2: Translation from the time domain to the frequency domain.

混合信号 $\mathbf{x}(t)$ から原信号 $\mathbf{s}(t)$ を周波数領域の独立成分分析で分離する手順を Fig. 4.1 に示す.

まず, 混合信号 $x_m(t)$ を

$$x_m(\omega, k) = \sum_t e^{-j\omega t} x_m(t) w(t - k\tau) \quad (4.3)$$

ように離散フーリエ変換 (DFT: Discrete Fourier Transform) して, その短時間スペクトルを求める. ここに, ω は規格化周波数, k はフレーム番号, τ はフレーム周期, $w(t)$ は窓関数である. この変換の概念を Fig. 4.2 に示す.

混合信号のスペクトル $\mathbf{x}(\omega, k) = [x_1(\omega, k), \dots, x_m(\omega, k), \dots, x_M(\omega, k)]^T$ は, 音源のスペクトル $\mathbf{s}(\omega, k) = [s_1(\omega, k), \dots, s_n(\omega, k), \dots, s_N(\omega, k)]^T$ と伝達関数 $G(\omega)$ の積で

$$\mathbf{x}(\omega, k) = G(\omega) \mathbf{s}(\omega, k) \quad (4.4)$$

のように近似される. Eq. (4.4) に対して, 周波数領域の独立成分分析に基づいて, それぞれの周波数 ω で分離行列 $H(\omega)$ を推定して, 分離信号のスペクトル $\mathbf{u}(\omega, k) =$

$[u_1(\omega, k), \dots, u_n(\omega, k), \dots, u_N(\omega, k)]^T$ は

$$\mathbf{u}(\omega, k) = H(\omega)\mathbf{x}(\omega, k) \quad (4.5)$$

のように求められる。

後述するように、スケールと位相の不定性と成分置換の問題を解決した後に、推定信号 $y_n(\omega, k)$ ($k = 0, 1, \dots, K-1$) を

$$y_n(t) = \frac{1}{2\pi} \frac{1}{W(t)} \sum_k \sum_{\omega} e^{j\omega(t-k\tau)} y_n(\omega, k) \quad (4.6)$$

のように時間領域へ離散フーリエ逆変換 (IDFT: Inverse Discrete Fourier Transform) すれば、 $s_n(t)$ に相当する $y_n(t)$ が復元される。ここに、 $W(t) = \sum_k w(t - k\tau)$ である。

4.2.2 周波数領域 FastICA アルゴリズム

周波数領域の FastICA アルゴリズムは、全ての周波数において、混合信号のスペクトルは平均が 0 で分散が 1 であり、その実部と虚部は無相関で同一の分散を持つという仮定の下で導出されている。したがって、時間領域の場合と同様に、混合信号のスペクトル $\mathbf{x}(\omega, k)$ は

$$\mathbf{x}'(\omega, k) = \mathbf{x}(\omega, k) - E[\mathbf{x}(\omega, k)] \quad (4.7)$$

と中心化した後で

$$\tilde{\mathbf{x}}(\omega, k) = Q(\omega)\mathbf{x}'(\omega, k) \quad (4.8)$$

のように白色化される。ここに、 $Q(\omega)$ は各周波数 ω における白色化行列である。

時間領域の FastICA ならば、Eq. (2.59) の Negentropy の近似式をそのまま適用できるが、周波数領域における分離信号のスペクトル $u_n(\omega, k)$ は複素数値を持つため、Eq. (2.59) が複素数値を扱えるようにする必要がある。そこで、 $u_n(\omega, k)$ の絶対値をとって、Negentropy を

$$J(|u_n(\omega, k)|^2) \propto \{E[F(|u_n(\omega, k)|^2)] - E[F(|\nu|^2)]\}^2 \quad (4.9)$$

のように書き改めて、この $J(|u_n(\omega, k)|^2)$ の最大化を行うことになる [15]。また、式 (2.59) の Negentropy の近似 $J(u_n(t))$ は、 $u_n(t)$ の分散が 1 という制約条件のもとで導

かれたものである．この制約条件は，周波数領域では $u_n(\omega, k)$ の絶対値の分散が1となることに置き換えられる．したがって，制約条件は

$$\begin{aligned} E[|u_n(\omega, k)|^2] &= E\left[(\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k))\overline{(\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k))}^T\right] \\ &= \bar{\mathbf{h}}_n^T(\omega)E[\tilde{\mathbf{x}}(\omega, k)\tilde{\mathbf{x}}^T(\omega, k)]\mathbf{h}_n(\omega) \\ &= \bar{\mathbf{h}}_n^T(\omega)\mathbf{h}_n(\omega) \\ &= \|\mathbf{h}_n(\omega)\|^2 = 1 \end{aligned} \quad (4.10)$$

となって，分離荷重ベクトルのノルムが1となることを要請している．また，式 (4.9) の ν はガウス変数であるので，式 (4.9) 右辺第2項の $E[F(|\nu|^2)]$ は定数として扱える．したがって， $J(|u_n(\omega, k)|^2)$ の最大化をするには，制約条件 $\|\mathbf{h}_n(\omega)\|^2 = 1$ のもとで， $E[F(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2)]$ を最大化すれば良いことになる．

Kuhn-Tucker の条件より，制約条件 $\|\mathbf{h}_n(\omega)\|^2 = 1$ の下で， $E[F(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2)]$ の最大化は

$$E[F(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2)] - \beta E[|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2] = 0 \quad (4.11)$$

の解を求めることに帰着される．ここに， β はラグランジュの未定乗数である．式 (4.11) の解を求めるためニュートン繰り返し近似法を用いると，荷重更新後の分離荷重ベクトル $\mathbf{h}_n^+(\omega)$ は

$$\begin{aligned} \mathbf{h}_n^+(\omega) &= \mathbf{h}_n(\omega) \\ &\quad - \frac{E[\tilde{\mathbf{x}}(\omega, k)\overline{(\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k))}f(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2)] - \beta\mathbf{h}_n(\omega)}{E[f(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2) + |\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2 f'(|\bar{\mathbf{h}}_n^T(\omega)\tilde{\mathbf{x}}(\omega, k)|^2)] - \beta} \end{aligned} \quad (4.12)$$

と逐次的に求まる．ここに， $\mathbf{h}_n^+(\omega)$ は更新後の分離荷重， $f'(\cdot)$ は $f(\cdot)$ の導関数である．式 (4.12) は，両辺に $\beta - E[f(|\bar{\mathbf{h}}_n(\omega)^T \tilde{\mathbf{x}}(\omega, k)|^2) + |\bar{\mathbf{h}}_n(\omega)^T \tilde{\mathbf{x}}(\omega, k)|^2 f'(|\bar{\mathbf{h}}_n(\omega)^T \tilde{\mathbf{x}}(\omega, k)|^2)]$ を掛けることで

$$\begin{aligned} \mathbf{h}_n^+(\omega) &= E[\tilde{\mathbf{x}}(\omega, k)\overline{u_n(\omega, k)}f(|u_n(\omega, k)|^2)] \\ &\quad - E[f(|u_n(\omega, k)|^2) + |u_n(\omega, k)|^2 f'(|u_n(\omega, k)|^2)]\mathbf{h}_n(\omega) \end{aligned} \quad (4.13)$$

と簡略化できる．ここで，エルゴード性が成り立つと考えて，集合平均を時間平均に置き換えれば，荷重更新式は

$$\mathbf{h}_n^+(\omega) = \frac{1}{K} \sum_{k=1}^K \{\tilde{\mathbf{x}}(\omega, k)\overline{u_n(\omega, k)}f(|u_n(\omega, k)|^2)\}$$

$$-[f(|u_n(\omega, k)|^2) + |u_n(\omega, k)|^2 f'(|u_n(\omega, k)|^2)]\mathbf{h}_n(\omega)\} \quad (4.14)$$

と定式化される．さらに， $\mathbf{h}_n^+(\omega)$ は，式 (4.10) の制約条件を満たすように

$$\mathbf{h}_n(\omega) = \frac{\mathbf{h}_n^+(\omega)}{\|\mathbf{h}_n^+(\omega)\|} \quad (4.15)$$

と規格化する．

この更新式は，更新前後の分離荷重 $\mathbf{h}_n(\omega)$ の向きが一致するまで繰り返される．つまり，収束条件

$$\bar{\mathbf{h}}_{n,\text{old}}^T(\omega)\mathbf{h}_{n,\text{new}}(\omega) \simeq 1 \quad (4.16)$$

を満たすまで繰り返すことになる．ここに，添え字の old と new はそれぞれ更新前後を表す．

これまでの処理では，ただ1つの分離荷重ベクトルが得られただけである．2つ以上の分離荷重ベクトルを推定する場合にも同様に，上述の FastICA アルゴリズムを適用することになる．ただし，分離荷重ベクトルが同じ値を推定することを防ぐため，全ての分離荷重ベクトル $\mathbf{h}_n(\omega)$ を直交化して無相関化する必要がある．

直交化を行うには，分離荷重ベクトル $\mathbf{h}_n(\omega)$ が推定された後，Gram-Schmidt の直交化により

$$\mathbf{h}_n(\omega) = \mathbf{h}_n(\omega) - \sum_{i=1}^{n-1} \mathbf{h}_i(\omega) \bar{\mathbf{h}}_i^T(\omega) \mathbf{h}_n(\omega) \quad (4.17)$$

のように全ての分離荷重ベクトルと直交化する．そして，これを式 (4.15) で規格化させたものが，推定する分離荷重ベクトルである．

以上の処理は，一つの周波数において分離荷重ベクトル $\mathbf{h}_n(\omega)$ ($n = 1, \dots, N$) を推定する手順である．したがって，時間領域での分離信号を得るためには，全ての周波数 ω に対して FastICA アルゴリズムを適用して，その各々で分離信号のスペクトル系列を推定することになる．

4.3 周波数領域独立成分分析の不定性問題の解決と目的信号の選択

本節では，周波数領域の独立成分分析におけるスケールと位相の不定性と成分置換の問題を解決する．したがって，分離スペクトルを $u_{\hat{n}}(\omega, k)$ のように書いて，音源の

順番 n と分離信号の順番 \hat{n} を明確にして議論する.

周波数領域の FastICA アルゴリズムによって推定される分離行列 $H(\omega)$ と, 白色化行列 $Q(\omega)$, 混合行列 $G(\omega)$ の関係は

$$H(\omega)Q(\omega)G(\omega) = P(\omega)D(\omega) \quad (4.18)$$

のように表され, 全ての周波数 ω でスケールと位相の不定性と成分置換の問題が生じる. ここに, $P(\omega)$ は各行各列で値が1の要素を一つ持ち, それを除いた全ての要素での値が0で, 周波数 ω ごとに要素の配置が異なる成分置換行列, $D(\omega) = \text{diag}[d_1(\omega), \dots, d_N(\omega)]$ は対角行列で, その要素 $d_{\hat{n}}(\omega)$ は各周波数 ω での振幅の係数を表す. この $d_{\hat{n}}(\omega)$ は白色化処理に起因した任意の複素数値をとる. このことは, 振幅だけでなく位相もが確定できないことを意味する.

したがって, 分離信号のスペクトル $u_{\hat{n}}(\omega, k)$ が $d_{\hat{n}}(\omega)s_{\hat{n}}(\omega, k)$ と推定され, スケール $d_{\hat{n}}(\omega)$ は周波数毎に個々別々の値を取り, $u_{\hat{n}}(\omega, k)$ の順番 \hat{n} と $s_n(\omega, k)$ の順番 n が必ずしも一致しない. そのため, 分離信号のスペクトル $u_{\hat{n}}(\omega, k)$ ($k = 0, 1, \dots, K-1$) を

$$u_n(t) = \frac{1}{2\pi} \frac{1}{W(t)} \sum_k \sum_{\omega} e^{j\omega(t-k\tau)} u_{\hat{n}}(\omega, k) \neq s_n(t) \quad (4.19)$$

のように時間領域に逆変換しても, $s_n(t)$ に相当する $u_n(t)$ は復元されない. したがって以降では, これらの問題の解法について説明する.

4.3.1 スケールと位相の不定性の解決

スケールと位相の不定性の問題に対しては, 分割スペクトル (Decomposed spectrum) $\mathbf{v}_{\hat{n}}(\omega, k) = [v_{\hat{n}1}(\omega, k), \dots, v_{\hat{n}M}(\omega, k)]^T$ を

$$\mathbf{v}_{\hat{n}}(\omega, k) = (H(\omega)Q(\omega))^{-1}[0, \dots, 0, u_{\hat{n}}(\omega, k), 0, \dots, 0]^T \quad (4.20)$$

のように導入して解消できる [9], [10].

この分割スペクトル $\mathbf{v}_{\hat{n}}(\omega, k)$ について, \hat{n} 行 \hat{n} 列の要素に1, 他の要素に0を持つ行列 $Z_{\hat{n}}$ を用いて

$$\begin{aligned} \mathbf{v}_{\hat{n}}(\omega, k) &= (H(\omega)Q(\omega))^{-1} Z_{\hat{n}} \mathbf{u}(\omega, k) \\ &= (H(\omega)Q(\omega))^{-1} Z_{\hat{n}} H(\omega)Q(\omega)G(\omega) \mathbf{s}(\omega, k) \end{aligned}$$

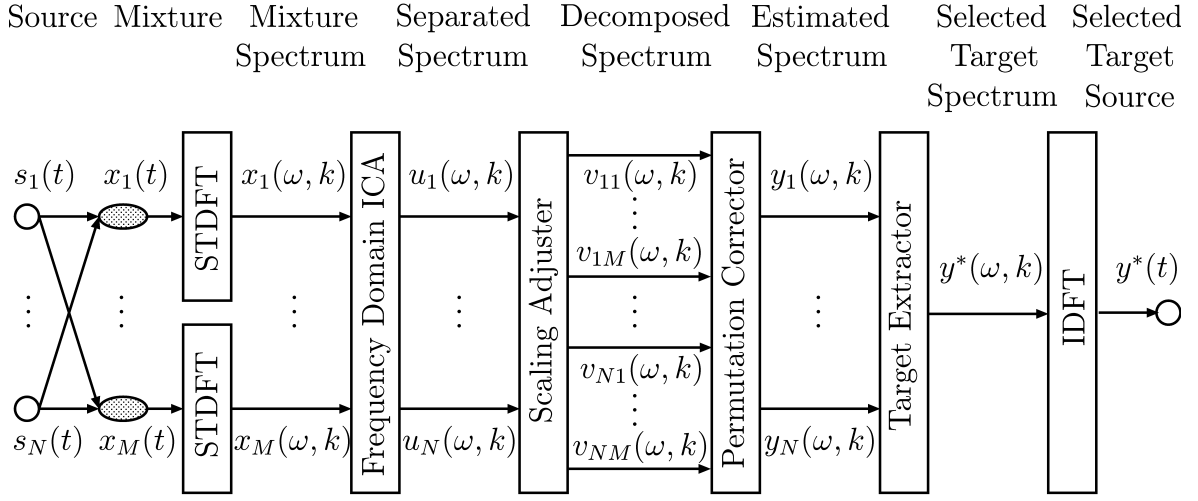


Fig. 4.3: Target source signal estimation based on frequency domain Independent Component Analysis.

$$\begin{aligned}
 &= G(\omega)D^{-1}(\omega)P(\omega)^T Z_{\hat{n}}P(\omega)D(\omega)\mathbf{s}(\omega, k) \\
 &= G(\omega)D^{-1}(\omega)Z_n D(\omega)\mathbf{s}(\omega, k) \\
 &= G(\omega)Z_n \mathbf{s}(\omega, k) \\
 &= [g_{1n}(\omega)s_n(\omega, k), \dots, g_{mn}(\omega)s_n(\omega, k), \dots, g_{MN}(\omega)s_N(\omega, k)]^T \quad (4.21)
 \end{aligned}$$

のように導かれる [12], [13]. すなわち, 分割スペクトル $v_{\hat{n}m}(\omega, k)$ は, n 番目の音源から m 番目のマイクロホンまでの伝達関数 $g_{mn}(\omega)$ と音源のスペクトル $s_n(\omega, k)$ の積で

$$v_{\hat{n}m}(\omega, k) = g_{mn}(\omega)s_n(\omega, k) \quad (4.22)$$

のように一意に決まる. したがって, 分割スペクトルがマイクロホンで観測された原信号の推定値であることは明らかである.

以上のことから, スケールと位相の不定性問題は, 全ての周波数において分離信号のスペクトルに対して分割スペクトルを求めるだけで解消できる. このとき, 分割スペクトルのスケールは, それぞれの周波数での伝達関数 $g_{mn}(\omega)$ そのものの値で一意に決定される. 2 音源 2 マイクロホンの場合の分割スペクトルと音源の対応関係は, 成分置換の有無によって Table 4.1 のように求められる.

Table 4.1: Decomposed spectra in permutation and those in no-permutation.

—	No-permutation ($\hat{n} = n$)	Permutation ($\hat{n} \neq n$)
$v_{11}(\omega, k)$	$g_{11}(\omega)s_1(\omega, k)$	$g_{12}(\omega)s_2(\omega, k)$
$v_{12}(\omega, k)$	$g_{21}(\omega)s_1(\omega, k)$	$g_{22}(\omega)s_2(\omega, k)$
$v_{21}(\omega, k)$	$g_{12}(\omega)s_2(\omega, k)$	$g_{11}(\omega)s_1(\omega, k)$
$v_{22}(\omega, k)$	$g_{22}(\omega)s_2(\omega, k)$	$g_{21}(\omega)s_1(\omega, k)$

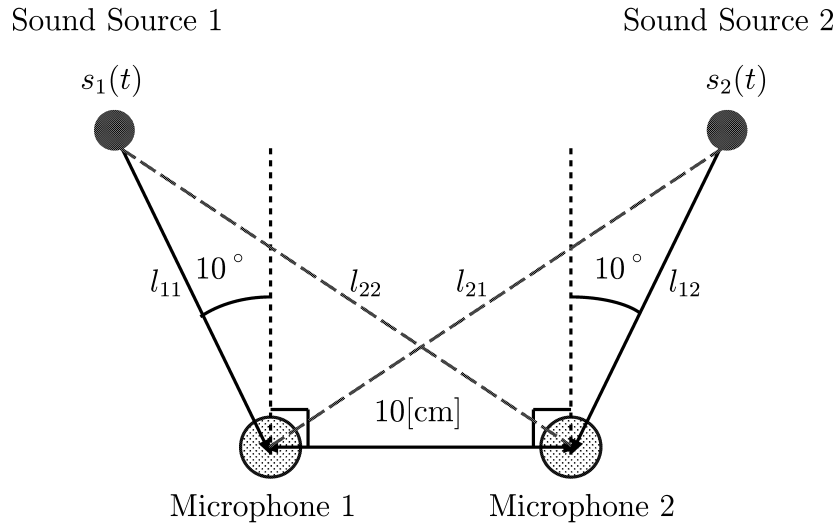


Fig. 4.4: Placement of sound sources and microphones.

4.3.2 伝達関数の推定と成分置換の修正原理

2 音源 2 マイクロホンの環境の下で，成分置換の有無による分割スペクトルと音源の対応関係は Table 4.1 となる．ここで，Fig. 4.4 のように，音源 $s_1(t)$ は相対的にマイクロホン 1 に近く，音源 $s_2(t)$ は相対的にマイクロホン 2 に近いと仮定すれば

$$|g_{11}(\omega)| > |g_{21}(\omega)| \quad , \quad |g_{12}(\omega)| < |g_{22}(\omega)| \quad (4.23)$$

$$\angle g_{11}(\omega) > \angle g_{21}(\omega) \quad , \quad \angle g_{12}(\omega) < \angle g_{22}(\omega) \quad (4.24)$$

なる関係が得られる．以降，Eq. (4.23) をゲイン条件，Eq. (4.24) を位相条件と呼ぶ．この関係を利用すれば成分置換の修正法は以下のように提案できる．

Table 4.2: Decision rule for permutation.

No-permutation ($\hat{n} = n$)	Permutation ($\hat{n} \neq n$)
$ r_1(\omega) > r_2(\omega) $	$ r_1(\omega) < r_2(\omega) $
$\angle r_1(\omega) > \angle r_2(\omega)$	$\angle r_1(\omega) < \angle r_2(\omega)$

まず、一つの音源から二つのマイクロホンまでの伝達関数の比を

$$r_{\hat{n}}(\omega) = \frac{g_{1\hat{n}}(\omega)}{g_{2\hat{n}}(\omega)} \quad (4.25)$$

のように定義すれば、Eq. (4.22) より、一つの分離信号のスペクトル $u_{\hat{n}}(\omega, k)$ から生成される二つの分割スペクトル $v_{\hat{n}1}(\omega, k)$, $v_{\hat{n}2}(\omega, k)$ によって

$$r_{\hat{n}}(\omega) = \frac{1}{K} \sum_{k=0}^{K-1} \frac{v_{\hat{n}1}(\omega, k)}{v_{\hat{n}2}(\omega, k)} \quad (4.26)$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} \frac{g_{1n}(\omega) s_n(\omega, k)}{g_{2n}(\omega) s_n(\omega, k)} \quad (4.27)$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} \frac{g_{1n}(\omega)}{g_{2n}(\omega)} \quad (4.28)$$

のように推定される。このことは、同一音源から二つのマイクロホンまでの伝達関数の比 $r_{\hat{n}}(\omega)$ を K 個の分割スペクトルから推定したものを意味している。したがって、ゲイン条件 Eq. (4.23) と位相条件 Eq. (4.24) を用いれば、伝達関数の比の絶対値 $|r_{\hat{n}}(\omega)|$ と偏角 $\angle r_{\hat{n}}(\omega)$ に基づく成分置換の判定則は Table 4.2 のように導出できる。

Table 4.1 より、2 音源 2 マイクロホンの環境では、一つの音源 $s_n(\omega, k)$ に対する推定値には、二つの分割スペクトル $v_{\hat{n},m=1}(\omega, k) = g_{m=1,n}(\omega) s_n(\omega, k)$ と $v_{\hat{n},m=2}(\omega, k) = g_{m=2,n}(\omega) s_n(\omega, k)$ がある。そこで、より近いマイクロホンで収録された信号は暗騒音の影響が少ないと考え、 $v_{\hat{n},m=n}(\omega, k) = g_{m=n,n}(\omega) s_n(\omega, k)$ を音源 $s_i(\omega, k)$ の推定値として採択する。したがって、ゲインと位相のそれぞれに基づく成分置換の修正則は

Rule1: gain-based method

$$y_n(\omega, k) = \begin{cases} v_{\hat{n}=n,m=n}(\omega, k) & \text{if } |r_1(\omega)| > |r_2(\omega)| \\ v_{\hat{n} \neq n,m=n}(\omega, k) & \text{if } |r_1(\omega)| < |r_2(\omega)| \end{cases} \quad (4.29)$$

Rule2: phase-based method

$$y_n(\omega, k) = \begin{cases} v_{\hat{n}=n, m=n}(\omega, k) & \text{if } \angle r_1(\omega) > \angle r_2(\omega) \\ v_{\hat{n} \neq n, m=n}(\omega, k) & \text{if } \angle r_1(\omega) < \angle r_2(\omega) \end{cases} \quad (4.30)$$

のように提案できる [34].

4.3.3 実環境下の音響特性

前節で提案した成分置換の修正法は、ゲイン条件 Eq. (4.23) と位相条件 Eq. (4.24) を用いて導出されている。したがって、実環境においてもゲイン条件と位相条件を満足することが必要である。そこで本節では、ゲイン条件 Eq. (4.23) と位相条件 Eq. (4.24) が実環境下でどの程度成立するかを調べる。そして、実環境下では条件の成立する割合が周波数によって異なることを述べる。

ここでは、Fig. 4.4 の配置で、 $s_1(t)$ を話者音声、 $s_2(t) = 0$ とした場合を考える。この場合、Eq. (4.4) より、混合信号は $x_1(\omega, k) = g_{11}(\omega, k)s_1(\omega)$ および $x_2(\omega, k) = g_{21}(\omega, k)s_1(\omega)$ と表される。そして、これらの混合信号の比は、 $r_x = x_1(\omega, k)/x_2(\omega, k) = g_{11}(\omega)/g_{21}(\omega)$ となって、音源 1 からマイクロホン 1 および 2 までの伝達関数の比と等しいことが分かる。したがって、この比の絶対値 $|r_x(\omega)|$ と位相 $\angle r_x(\omega)$ からゲイン条件 Eq. (4.23) と位相条件 Eq. (4.24) がどの程度成立するかを調べられる。

残響時間が 500[msec]、暗雑音が 48.0[dB] の実験室 (L7.3[m] × W6.5[m] × H2.9[m]) の中央に、マイクロホンを 2 個並べて実験を行った。音源とマイクロホンの距離 l_{11} は 10[cm], 30[cm] とし、マイクロホン間隔は 10[cm]、サンプリング周波数を 8000[Hz]、分解能を 16[bit] とした。それぞれ 32 組 (8 名 × 4 単語) のデータについて、 $|r_x(\omega)| > 1$, $\angle r_x(\omega) > 0$, つまり、ゲイン条件 Eq. (4.23) と位相条件 Eq. (4.24) が成立した割合を示したのが Table 4.3 である。この結果から、ゲイン条件と位相条件は実環境下でも殆ど成り立つことが読みとれる。

さらに、32 組の全データについて周波数ごとに $|r_x(\omega)|$ と $\angle r_x(\omega)$ の値をプロットしてみると、Fig. 4.5 のように $r_x(\omega) > 1$, $\angle r_x(\omega) > 0$ とならない場合の特徴が明確になる。すなわち、(a)(b) では中域と高域、(c) では低域と高域、(d) では低域に例外が散見される。これらの例外が生じる原因は以下の理由だと考えられる。まず、(a) や (b) については、一般的に低域で音声のパワーが大きく、周波数が高くなるにつれて $|g_{11}(\omega)|$

Table 4.3: Ratios [%] where Eqs. (4.23) and (4.24) hold in real environment.

Distance	10[cm]	30[cm]
gain inequality	99.88	98.50
phase inequality	97.84	99.64

と $|g_{21}(\omega)|$ に有意な差がなくなつて、暗騒音の影響を受けやすくなったためと考えられる。また、(b) の例外の数が (a) に比べて多いのは、伝達関数のゲインの差は音源とマイクロホンの距離が長くなるにつれて小さくなるため、暗騒音の影響が相対的に大きくなったためと考えられる。

一方、位相差は二つのマイクロホンに到達する時間差と波長で計算されるため、高域に対して低域では波長が長くなり、位相差は小さい値になり暗騒音の影響を受けやすい。したがって、(c) や (d) は低域になるほど位相差が小さくなり、約 400[Hz] 以下で例外が生じている。また、(c) の約 3200[Hz] 以上の高域で例外が起きているのは空間的エイリアシング (spatial aliasing) のためである。エイリアシングが起きないようにするには、音速を $c = 340[\text{m/s}]$ 、音源の最高周波数を $f_{max} = 3400[\text{Hz}]$ とすると、音源 1 から二つのマイクロホンまでの距離差は、 $\Delta l_1 = |l_{11} - l_{21}| < c/(2f_{max}) = 5[\text{cm}]$ 未満でなければならない。マイクロホン間隔が $l_0 = 10[\text{cm}]$ の場合、音源 1 からマイクロホン 1 までの距離が $l_{11} = 30[\text{cm}]$ のとき、音源 1 からマイクロホン 2 までの距離は $l_{21} = \sqrt{10^2 + 30^2 - 2 \cdot 10 \cdot 30 \cos 100} = 33.23 [\text{cm}]$ となつて、 $\Delta l_1 = 3.23 < 5[\text{cm}]$ となるため、エイリアシングは起きない。しかし、 $l_{11} = 10[\text{cm}]$ の場合、 $l_{21} = 15.32[\text{cm}]$ となつて、 $\Delta l_1 = 5.32 > 5[\text{cm}]$ となるため、 $f > 3195.5[\text{Hz}]$ でエイリアシングが起きていることが分かる。

以上の結果が、Eq. (4.29) と Eq. (4.30) で提案した成分置換の修正法にどのような影響を与えるかを次で述べる。

4.3.4 周波数による修正能力の違い

Eq. (4.29) や Eq. (4.30) の修正則は、実環境下で用いたときには修正能力が異なるため、ゲインと位相に基づく修正法の周波数に対する修正能力を調べた。Fig. 4.4 の

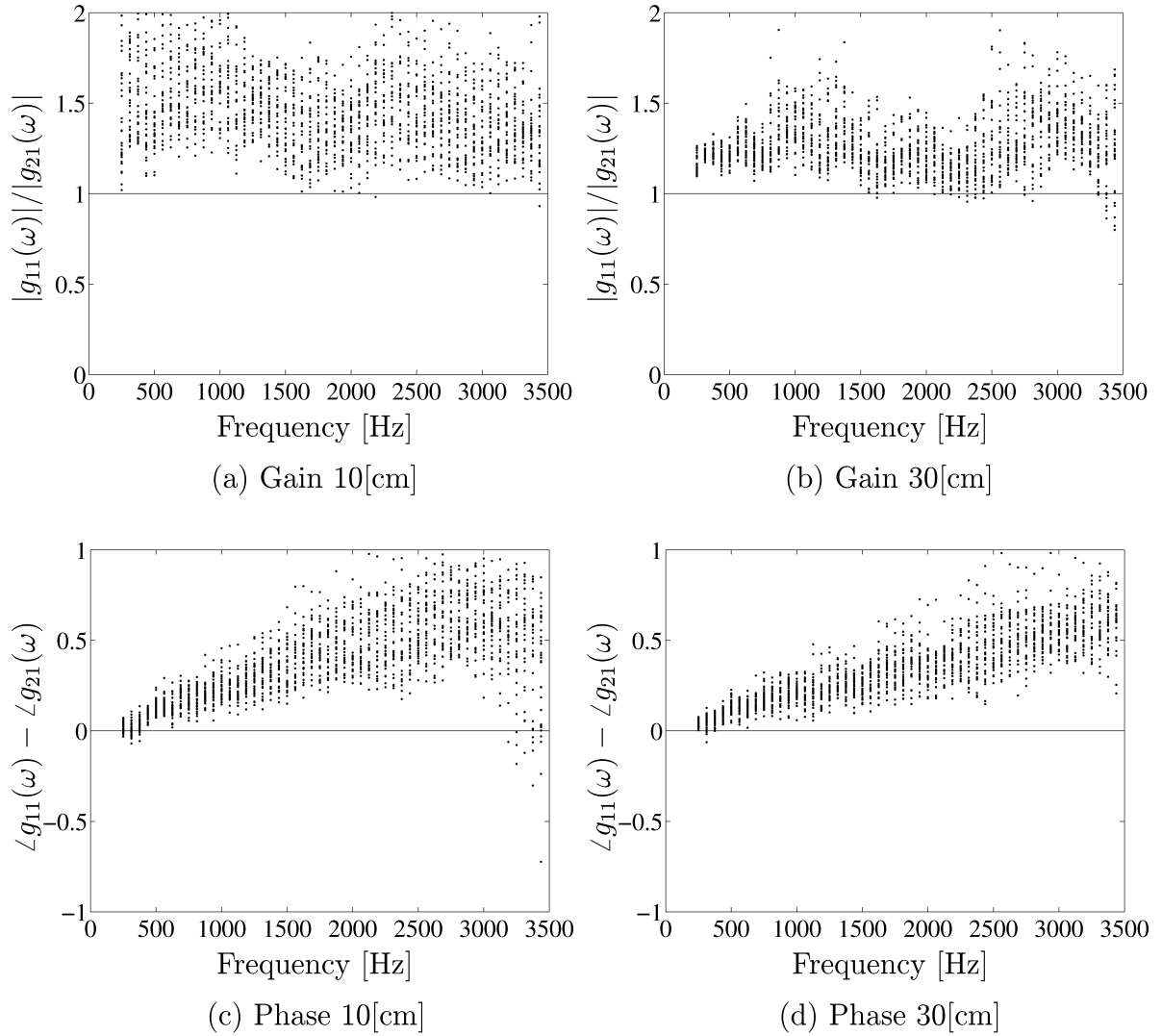


Fig. 4.5: Frequency properties of gain ratio and phase difference.

配置で, $(l_{11}[\text{cm}], l_{22}[\text{cm}]) = (10, 30), (10, 60), (30, 60), (30, 100)$ の組合せで, 音源 1 は男性四名, 女性四名のそれぞれの四単語の発話, 音源 2 はスピーカーから駅構内の騒音 [73] を再生し, 計 128 組の混合信号を収録して実験した.

ゲインと位相に基づく修正法の修正能力の結果を Fig. 4.6 に示す. このときの修正率は [29] に準じて算出した. グラフの横軸は周波数, 縦軸は成分置換の修正率を表す. (a) は $(l_{11}, l_{22}) = (10, 30)$, (b) は $(10, 60)$, (c) は $(30, 60)$, (d) は $(30, 100)$ の実験結果であり, 実線と破線はそれぞれゲインと位相に基づく修正法に対する解決率である.

(a) と (b) の比較から, Rule1 と Rule2 は共に (b) の修正能力が高いことが分かる. これは (b) は (a) よりも雑音源が遠くにあるため, 分離性能が高くなり当然の結果であ

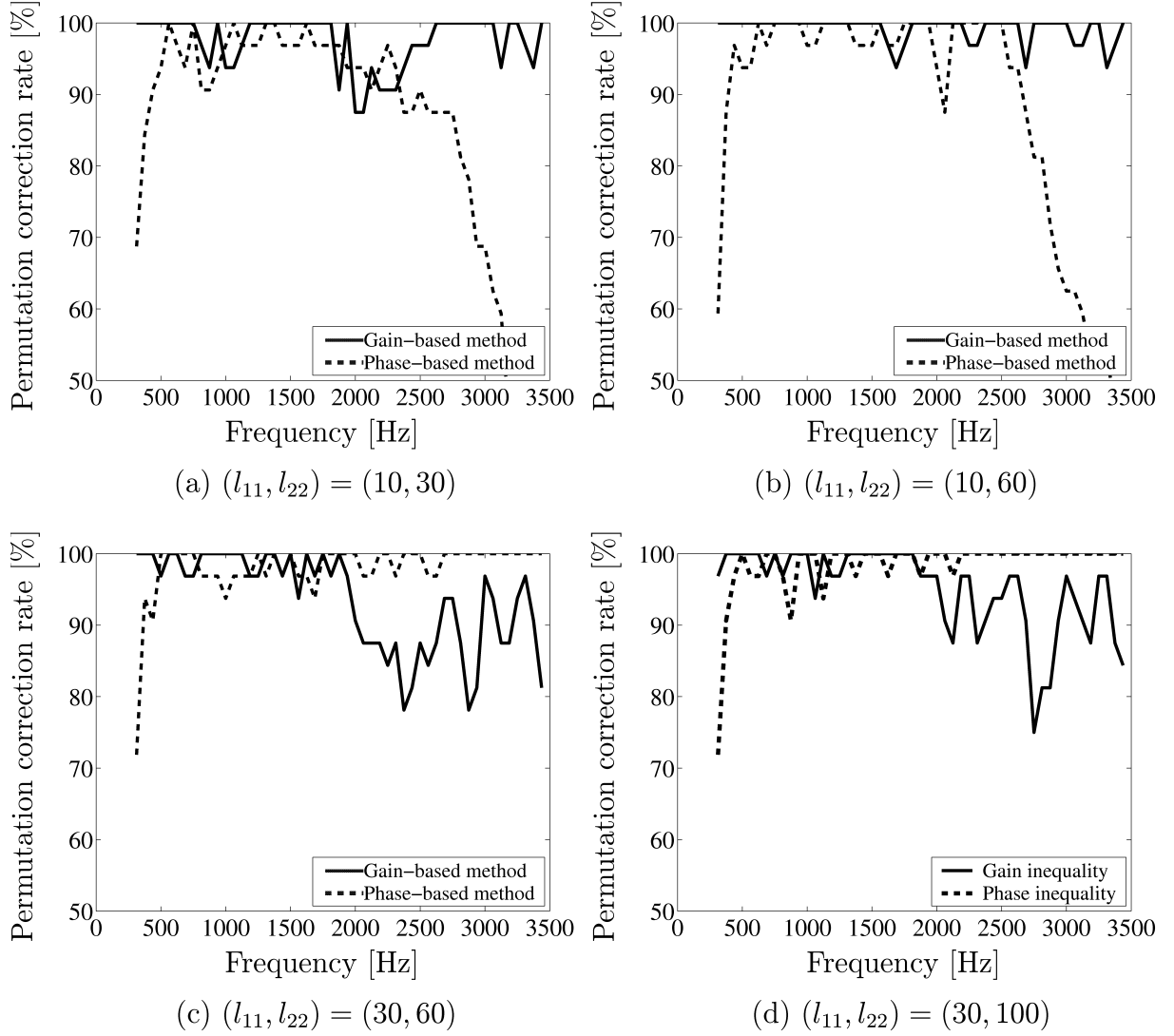


Fig. 4.6: Frequency properties of permutation correction.

る。(c)と(d)の間でも、これと同様のことがいえる。また、全てのグラフにおいて、Rule1による修正率は高域よりも低域において高いことが分かる。これは、音声が高域にパワーを持っているために、高域と比較して周囲の暗騒音の影響を受けにくいためと考えられる。一方、Rule2は、約400[Hz]以下の低域においては全く機能していないことが分かる。この理由は、低域での位相差が顕著に現れないためであり、前述の伝達関数の位相の特徴と一致している。

(a)(b)と(c)(d)を比較してみると、Rule1では $l_{11} = 10$ の場合は全域で修正率が高いのに対して、 $l_{11} = 30$ の場合の修正率は約1800[Hz]以上の高域で修正率の低下が見られる。これは、音源が遠くなって伝達関数のゲイン差が顕著に現れなくなり、音声の

パワーが小さい高域では、暗騒音が大きく影響したためである。一方 Rule2 では、約 2500[Hz] 以上の高域において、 $l_{11} = 10$ の場合の修正率は急激に低下していることが分かる。この理由は空間的エイリアシングが生じたためであり、エイリアシングが生じない $l_{11} = 30$ の修正率はほぼ 100[%] である。

以上の結果から、実環境下で周波数領域の独立成分分析における成分置換を解決するには、Rule1 と Rule2 を適切に使い分ける必要がある。すなわち、音源からマイクロホンまでの距離が近い場合には Rule1 を、遠い場合には Rule2 を用いなければならない。周波数に関しては、低域では Rule1 を、中域では Rule2 を用いることで、高い精度で成分置換を修正できると期待できる。また、高域ではエイリアシングの有無によって、Rule1 と Rule2 を適切に選択しなければならないことが分かる。

4.3.5 統合的成分置換修正法

ゲインと位相のそれぞれに基づく成分置換の修正法の修正能力は、マイクロホン間隔や音源からマイクロホンまでの距離によって異なることが分かった。しかしながら、ブラインド信号分離ではこの距離は未知であるため、ここでは伝達関数の周波数に対する特徴に基づいて、ゲインと位相の情報を統合した成分置換修正法を提案する。

前述の結果から、エイリアシングが無い場合、音源のパワーが大きい低域では Eq. (4.29) のゲインに基づく修正則が有効であり、位相差が大きく現れる高域では Eq. (4.30) の位相に基づく修正則が有効である。したがって、エイリアシングが起きないように考慮すれば、二つの修正則を統合した新たな規則は、以下のように提案できる [34]。

$$f < f_0 \quad \text{Rule1: Gain-based method} \quad (4.31)$$

$$f_0 \leq f \quad \text{Rule2: Phase-based method} \quad (4.32)$$

ここに、 f_0 は低域と高域の周波数帯域の境界を表す周波数であり、本実験環境では Fig. 4.5(b)(d) の結果から、エイリアシングが生じない範囲で、約 1000～2000[Hz] に設定すれば成分置換は良好に解決できると考えられる。このことは、Fig. 4.6(c)(d) の結果からも裏付けされている。

エイリアシングの有無は面音源として近似できる遠距離場のときはマイクロホン間隔のみで決まるが、点音源として近似できる近距離場のときはマイクロホン間隔だけ

でなく音源からマイクロホンまでの距離にも依存する。したがって、実環境におけるエイリアシングの有無は簡単に判別できない。一つの音源と二つのマイクロホンの距離の差が小さい場合、すなわち、距離差 $\Delta l_i = |l_{1i} - l_{2i}|$ が小さく $\Delta l_i < c/(2f)$ のときにはエイリアシングは起きない。このような場合、 $|r_1(\omega)| \simeq |r_2(\omega)| \simeq 1$ になる。したがって、ゲイン比からエイリアシングの有無を推定して、エイリアシングが生じる周波数帯域がある場合は Eq. (4.29) を用いる修正法も考えられる。

提案法をまとめると、 f_0 や目的音声の位置に関する知識を前提に、混合信号を周波数領域に変換して独立成分分析で分離した後、まず分離信号のスケールを回復し、次に成分置換を修正して、最後に目的信号を選んで時間領域に戻す処理の流れとなる。

4.3.6 多音源多マイクロホン環境での成分置換修正法

これまでは、簡単のため、2音源2マイクロホンの条件下で解説した。ここでは、前節までの議論に基づいて、多音源多マイクロホンでの成分置換の修正を述べる。音源 $s_n(t)$ は n 番目のマイクロホンに最も近いと仮定すれば、伝達関数のゲインと位相について

$$|g_{nn}(\omega)| > |g_{mn}(\omega)| \quad (4.33)$$

$$\angle g_{nn}(\omega) > \angle g_{mn}(\omega) \quad (4.34)$$

のように、多音源下でのゲイン条件と位相条件が得られる。

ここで、分割スペクトル $\mathbf{v}_{\hat{n}}(\omega, k)$ の要素について絶対値を求めると、Eq. (4.22) と Eq. (4.33) より、 $|v_{\hat{n}m}(\omega, k)|$ の大小関係は伝達関数にのみ依存して決まる。したがって、 $|v_{\hat{n}m}(\omega, k)|$ が最大値のとき $m = n$ となり、 $v_{\hat{n}, m=n}(\omega, k)$ は音源 m の推定値となる。以上の事実から、ゲインに基づく成分置換の修正則は

$$n = \arg \max_{\hat{n}} |v_{\hat{n}m}(\omega, k)| \quad (4.35)$$

のように定式化できる [31], [79]。同様に、位相に基づく修正則は

$$n = \arg \max_{\hat{n}} \angle v_{\hat{n}m}(\omega, k) \quad (4.36)$$

のように導出できる。

ゲインと位相のそれぞれを利用した修正法を、4.3.3 で検証した実環境下での音響伝達特性に基づいて周波数帯域で使い分ければ、多音源多マイクロホン環境での成分置換は解決できる。

さらに、位相に対して信頼できる周波数帯域を明らかにした結果から、この周波数帯域を利用して音源方向を推定して、その方向に基づいてマイクロホンを配置することで、提案法はさらに高い修正能力が期待できる。

4.3.7 目的音声の選択

スケールと位相の不定性問題と成分置換の問題が解決されて、音源に対する最適な分離信号が得られたとしても、まだ、どの分離信号を目的音声として選択するかというチャンネル選択の問題が残る。実際の応用においても、駅構内や工場内のような騒音環境下で音声認識を用いる場合、話者音声と騒音が混じって観測される混合信号から、話者音声のみを抽出することが要求される。ここでは、観測された混合信号に対して、周波数領域の独立成分分析で分離処理を行い、スケールの不定性と成分置換を解決した後の推定信号から、目的音声を抽出する規則を導出する。

前節で述べた成分置換の修正法によって得られる音源の推定値 $y_n(\omega, k)$ は

$$y_n(\omega, k) = g_{nn}(\omega) s_n(\omega, k) \quad (4.37)$$

のように得られる。このことは、 $y_n(\omega, k)$ は $s_n(\omega, k)$ に対する推定値であることを意味している。もし、目的音声がどのマイクロホンに近いかわかるという情報を先験的に得られるならば、目的音声の抽出問題は自ずと解決される。すなわち、目的音声に最も近いマイクロホン m から

$$y^*(\omega, k) = y_m(\omega, k) = g_{mm}(\omega) s_m(\omega, k) \quad (4.38)$$

のように、目的音声の推定値 $y^*(\omega, k)$ を簡単に選択できる。

このように目的音声の位置情報が既知であれば、単純な選択で目的音声を得ることができる。しかし、この位置情報が未知の場合、音源の位置情報を利用しないで目的音声を抽出する必要がある。したがって、以降では目的音源の位置情報は未知であるとして、騒音環境下で観測された混合信号から話者音声を抽出する方法を提案する。

一般的に、話者音声は騒音に比べて尖度が高い特徴を持つ。この特徴を利用して、Kurtosis を用いた話者音声の抽出方法がある。Kurtosis に基づいて話者音声 $y^*(t)$ を抽出するには、成分置換が解決された推定信号 $y_n(\omega, k)$ をフーリエ逆変換して得られる $y_n(t)$ に対して Kurtosis を計算して、Kurtosis の値が最大となる推定信号を話者音声と判断することになる。しかしながら、このように高次統計量を用いる場合や、調音構造のような話者音声に関する知識を用いる場合には、計算は複雑になり、異常値に対して敏感になるため、話者音声の抽出に用いるには頑健性に問題がある。

そこで、周波数領域の独立成分分析に Hyvärinen の提案した FastICA アルゴリズムを採用し、その特徴を利用することで簡単に目的信号を選択する方法を提案する。一般に話者音声は騒音より尖度が高いため非ガウス性は高くなる。したがって、FastICA アルゴリズムを用いて分離した場合、非ガウス性の高い順話者音声が最初に出力される頻度が高くなることを意味している。言い換えれば、最初に出力される分離信号のスペクトルは話者音声のスペクトルである確率が高い。この事実に基づいて、全ての周波数ビンにおいて成分置換が解決すると同時に、FastICA によって最初に推定された回数をカウントしておけば、その数が最大となる推定信号を話者音声と判断できる。この選択法は、FastICA アルゴリズムによって最初に出力された回数を数えるだけなので、非常に簡単な方法である。また、周波数領域で話者音声を判断できるため、Kurtosis を用いた選択法とは異なり、目的信号だけを時間領域へ逆変換すればよい。計算量が少なくなる特徴を持つ。

4.4 音源分離実験

提案法の有効性を検証するため、実環境下で観測された混合信号に対して実験を行う。ただし、原信号や音源からマイクロホンまでの伝達特性が利用できない実環境下では SNR (Signal to Noise Ratio) や SDR (Signal Distortion Ratio) の正確な算出法が確立していないため、復元音声の品質が評価できない。そこで、まず、原信号と伝達特性が利用できるバーチャルルームにおいて、音源とマイクロホンを空間的エイリアシングが生じないように配置した下で、ゲインと位相のそれぞれに基づく成分置換の修正法に対して、是正能力と復元音声の品質を評価する。その後、実環境下で観測した混合信号を用いて提案法の有効性を検証する。このとき、定常信号に近い駅構内

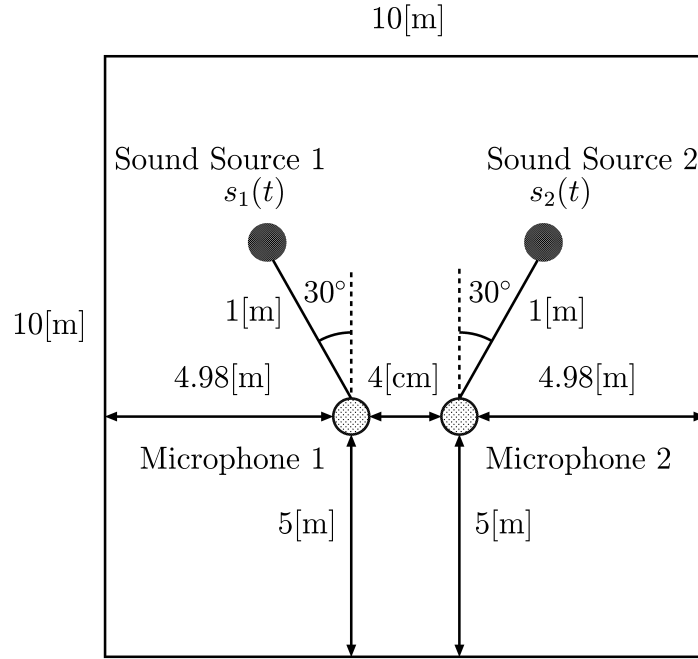


Fig. 4.7: Placement of sources and microphones in the virtual room

の騒音の中で話者が発話した混合信号から雑音除去実験を行い、次に、非定常信号である二人の発話が存在する混合信号から音源分離実験を行う。また、多音源多マイクロホンにおける成分置換の修正結果を述べる。最後に、目的信号の選択実験を行う。

4.4.1 バーチャルルームでの音源分離評価実験

新聞記事読み上げ音声コーパス [72] から 10 パターン（男性 5 パターン，女性 5 パターン）の話者音声と，Ambient Noise Database [73] から 2 パターンの駅構内の騒音を，それぞれ音源 1 と音源 2 として使用した。また， $L10[m] \times W10[m] \times H10[m]$ のバーチャルルームにおいて，音源とマイクロホンを Fig. 4.7 のように配置したときのインパルス応答を求め，これを音源に畳み込んで計 20 組の混合信号を生成した。このとき，音源とマイクロホンは全て床から 1[m] の高さに配置した。

データ取得については，サンプリング周波数を 8000[Hz]，分解能を 16[bit] とした。離散フーリエ変換については，リアルタイム化を考慮して，フレーム長を 16[msec]，フレーム周期を 8[msec]，窓関数はハミング窓を用いた。成分置換の修正と時間領域への復元については，電話音声の品質に相当する 300～3400[Hz] の周波数範囲を処理対象

Table 4.4: Permutation correction rates [%], SNR [dB] and SDR [%] in the virtual room.

	Correction rates	SNR	SDR
FastICA	95.10	6.80	5.11
Similarity	98.23	15.88	14.26
Power	98.86	18.24	16.49
Rule1	98.93	18.41	16.66
Rule2	99.20	18.45	16.68
Proposed	99.99	18.91	17.10

とした．また，FastICA アルゴリズム [15] では，初期荷重をノルムが1の複素乱数，非線形関数を $f(|u_n(\omega, k)|^2) = 1 - 2/(e^{2|u_n(\omega, k)|^2} + 1)$ ，最大繰返し回数を1000，誤差を0.000001， f_0 は1800[Hz] とした．

以上の混合信号をもとに，成分置換の修正率 [29] と SNR [27], [28]，SDR を求めた．ここに，SDR は

$$\text{SDR} = \frac{1}{2} \sum_{n=1}^2 10 \log \frac{\sum_t |g_{nn}(z)s_n(t)|^2}{\sum_t |g_{nn}(z)s_n(t) - b_{nn}(z)s_n(t)|^2} \quad (4.39)$$

と定義される評価量で， $b_{nn}(z)$ は分離フィルタ $H(z)$ と混合フィルタ $G(z)$ の積 $B(z) = H(z)G(z)$ の n 行 n 列の要素である．Table 4.4 は，20 組の混合信号に対する修正率，SNR，SDR のそれぞれの平均値であり，表に記述した FastICA は文献 [15]，Similarity は分離信号の類似度に基づく方法 [10]，Power は分割スペクトルのパワーに基づく方法 [12] を示す．

この結果より，提案法による修正率はほぼ100[%]で，SNR と SDR の値は他のどの手法よりも高いことが分かり，提案法の有効性が確認できる．一方，FastICA アルゴリズムを適用しただけでも，95[%] 程度の修正率が得られることが分かる．このことは，非ガウス性の高い順に原信号を推定する FastICA アルゴリズムの特徴を反映した結果となっている．また，類似度に基づく方法や Power 法，Rule1，Rule2 の修正率は全て98[%]以上であるが，SNR と SDR は Power 法や Rule1，Rule2 の方が高い値を示している．この理由は成分置換是正能力の周波数依存性に関係しており，Power 法や

Rule1 の場合は成分置換が是正できていないのはスペクトルの小さな帯域にほぼ限定されていたが、FastICA アルゴリズムや類似度に基づく方法の場合はスペクトルの大小に関係なく全帯域で散在していたためである。

4.4.2 実環境下での雑音除去実験

残響時間が 500[msec]、暗騒音が 48.0[dB] の実験室 (L7.3[m] × W6.5[m] × H2.9[m]) で実験を行った。このときの騒音レベル L_p [dB] は、人間の聞き取れる最小の音圧である 20[μPa] を基準として

$$L_p = 20 \log_{10} \left(\frac{p}{20} \right) \quad (4.40)$$

で定義される [80]。ここに、 p [μPa] は音の圧力である。

マイクロホンの間隔を 10[cm] として 2 個並べ、これらのマイクロホンを結ぶ線と垂直な方向をマイクロホン正面方向 (Fig. 4.4 の破線) として、正面方向からそれぞれ外側に 10° の方向に、音源 1 をマイクロホン 1 から l_{11} [cm]、音源 2 をマイクロホン 2 から l_{22} [cm] 離して配置した。そして、 $(l_{11}[\text{cm}], l_{22}[\text{cm}]) = (30, 60), (30, 100)$ の組合せで実験を行った。このとき、音源 1 は男性四名、女性四名のそれぞれ四単語の発話で、音源 2 はスピーカーから駅構内の騒音 [73] を再生し、計 64 組の混合信号を収録して実験を行った。混合信号は、短いもので 2.3[sec]、長いもので 9.1[sec] で、 f_0 は 1800[Hz] とした。

Table 4.5 に成分置換の修正率と SNR の近似値 [81] を算出した結果を示す。既存の Power 法による結果は約 9 割程度の解決率を示しているが、音源からマイクロホンまでの距離によって修正率の差異が見られる。これに対して、二つの修正則を統合した提案法による場合、音源の距離の変化に依存することなく、100[%] 近い解決率が得られた。さらに、全ての組み合わせの混合信号に対して、時間領域での推定信号の反転がなかった。すなわち、 $y_1(t)$ に $s_1(t)$ の推定値が出力されていることを確認した。一方、類似度に基づく方法では時間領域での推定信号の順番が反転し、 $y_1(t)$ に $s_2(t)$ の推定値が出力されている割合が 15.63[%] であった。この結果から、提案法は周波数領域での成分置換の修正だけでなく、目的信号の選択までもが可能であることを示している。

Table 4.5: Permutation correction rates [%] and SNR [dB] against station premise noises roared from a loud speaker.

	Correction rates		SNR	
(l_{11}, l_{22})	(30,60)	(30,100)	(30,60)	(30,100)
FastICA	88.61	87.62	6.31	7.11
Similarity	93.87	95.20	15.17	15.45
Power	88.60	90.75	16.42	16.88
Rule 1	94.04	95.51	17.30	17.66
Rule 2	98.13	98.58	17.78	17.84
Proposed	99.98	99.99	18.20	18.29

Fig. 4.8 に、駅構内の雑音が流れる中で、女性話者が発声したときの混合信号と復元信号の波形を示す。(a) はマイクロホン 1 と 2 で観測された混合信号 $x_1(t)$ と $x_2(t)$, (b) は FastICA アルゴリズムによる分離信号波形 $u_1(t)$ と $u_2(t)$, (c) は提案法による復元信号の波形 $y_1(t)$ と $y_2(t)$ である。 $u_1(t)$ と $y_1(t)$ の波形の比較から、本来ゼロとなるべき非発話区間では後者の方が残留雑音のレベルの小さいことが分かる。また、周波数間でのスケールの不定性が解決されていないため、 $u_1(t)$ では $t \simeq 1.2[\text{sec}]$ 付近で突発音のような箇所が出ていることが読み取れるが、提案法ではそのような箇所がないことが分かる。

以上の結果から、提案法は音源からマイクロホンまでの距離に依存せず、実環境下で有効に機能することを確認できた。

4.4.3 実環境下での音源分離実験

次に、二つの音源の両方が話者音声である場合について実験を行った。Fig. 4.4 における話者とマイクロホンの距離 d_1 , d_2 をそれぞれ 10[cm] として、六名（男性三名、女性三名）の話者による発話を実験対象とした。具体的には、二人のペアの一人はマイクロホン 1 側、もう一人はマイクロホン 2 側に位置して、発話区間の短い単語と長い単語の組み合わせ（東京－近畿大学）、（新飯塚－産業技術研究科）、（博多－五反田研

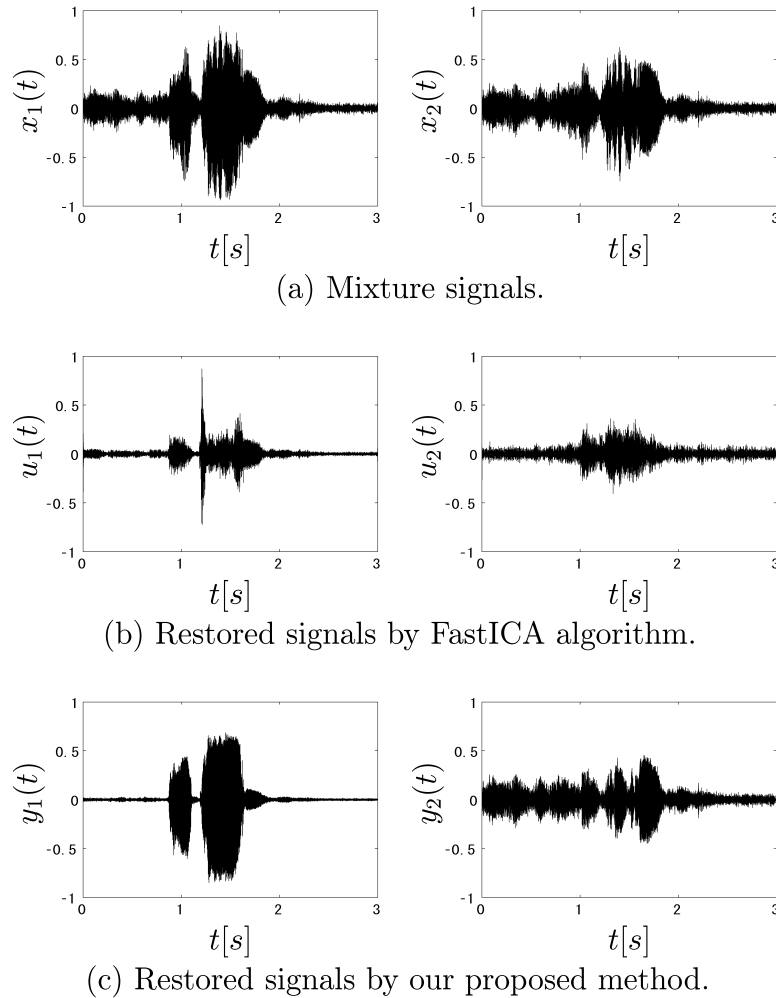


Fig. 4.8: Experimental result when a female speaker uttered under station premise noise.

究室) の3つのパターンを発話を収録し、次に発声単語を入れ替えた3パターンの発話を収録した。さらに二人の位置を入れ替えて上記を繰り返して、計12パターンに対する混合信号を収録した。さらにペアの一人はそのままにして、残りの四人をペア相手として換えながら、上記と同様の発話を収録した。以上により、総計180パターンに対する混合信号のデータを取得した。ここで、発話パターンを発話区間の短い単語と長い単語の組み合わせにしたのは、各周波数における成分置換の目視による判定作業を容易になるようにするためである。これらのデータ長は、短いもので2.3[sec]、長いもので4.1[sec]であった。

Table 4.6: Permutation correction rates [%].

	male	female	average
FastICA	48.43	52.77	50.60
Similarity	91.00	91.12	91.06
Power	98.74	99.43	99.08
Proposed	99.66	99.78	99.72

以上のデータに対する成分置換の解決率を Table 4.6 に示す。Table 4.6 の性別は音源 1 の話者の性別を示している。まず、FastICA による成分置換の解決率は平均で 50.60[%] であった。これは、FastICA が非ガウス性の高い順に分離する手法であり、二つの音源の両方が非ガウス性の高い話者音声であることから当然の結果であり、成分置換の解決には何も寄与していないことが分かる。これに対して、分割スペクトルのゲインと位相に基づく成分置換の修正則を周波数帯域で使い分けた提案法の場合、成分置換の解決率は平均して 99.72[%] と非常に高い解決率であった。これらの結果から、提案法は定常的な雑音のみでなく、非定常雑音に対しても有効であることを確認した。

このときの分離波形を Fig. 4.9 に示す。Fig. 4.9 は、音源 1 として女性話者が「東京」、音源 2 として男性話者が「近畿大学」と発声したときの混合信号と推定信号である。Fig. 4.9(a)(b) はそれぞれマイクロホン 1 とマイクロホン 2 で観測した混合信号、(c)(d) は FastICA による推定信号、(e)(f) は提案法によって成分置換が解決された女性話者音声と男性話者音声の推定信号である。(c)(d) の FastICA による場合、音源は発話区間の短い単語と長い単語であったにもかかわらず、二つの分離信号の音声持続時間がほぼ同じになっていて、成分置換の問題は全く解決されていないことが分かる。これに対して、提案法による場合 (e)(f) には、女性話者と男性話者の音声持続時間が異なることが反映されており、視覚的にも成分置換は正確に修正されていることが読みとれる。また、復元音声を聞いてみたとき、提案法による結果は他の音声が残んど混じることなく元の音声は明瞭に聞こえた。

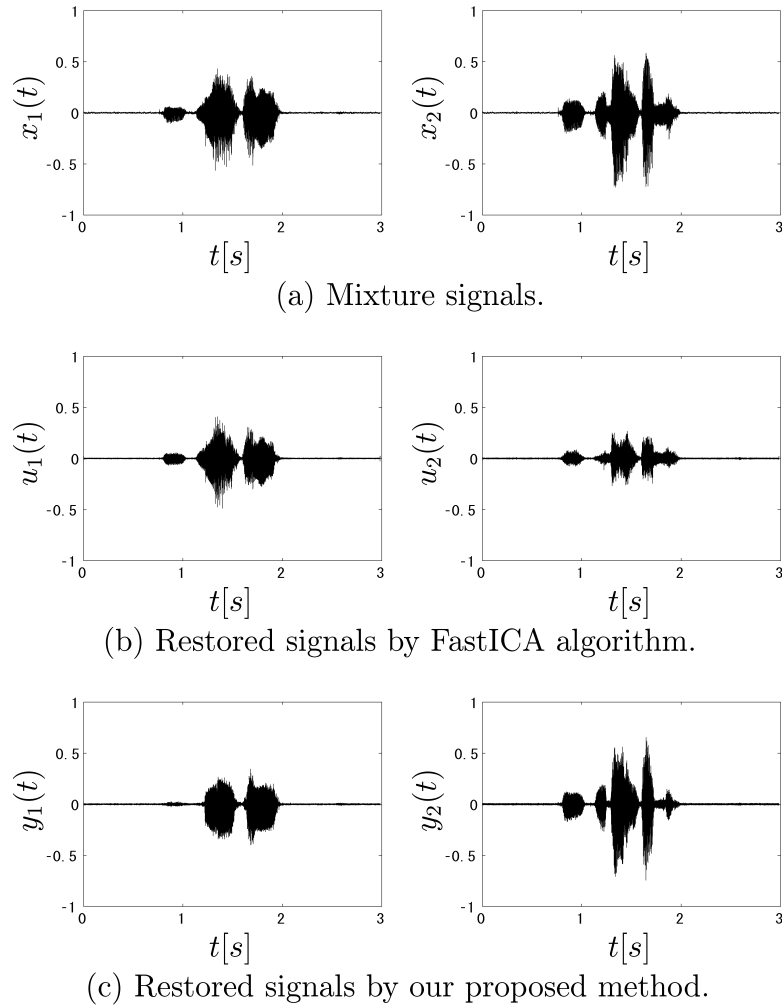


Fig. 4.9: Experimental result when a female and a male uttered.

4.4.4 目的音声選択実験

続いて、騒音環境下で発声された話者音声为正しく抽出されるかを調べるために、前節で取得した、音源が話者音声と騒音であるときのデータを用いて実験を行った。Table 4.7は128組のデータに基づく話者音声の抽出率を示している。話者音声の抽出率 R_s は、話者音声が正確に抽出された数を N_s 、騒音が抽出された数を N_n として

$$R_s = \frac{N_s}{N_s + N_n} \quad (4.41)$$

と定義した。

Table 4.7: Target speech extraction rates [%].

(l_{11}, l_{22})	(30,60)	(30,100)
Kurtosis	65.63	81.25
Proposed	100	100

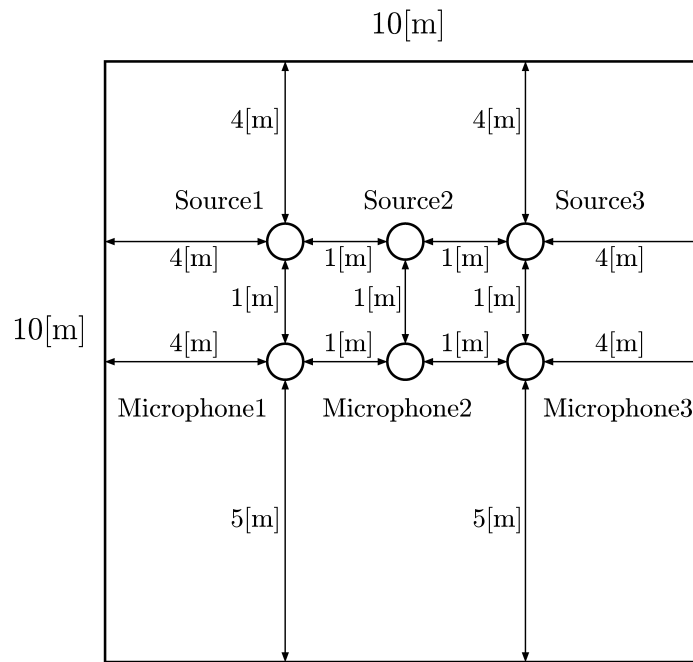


Fig. 4.10: Locations of sources and microphones in the virtual room.

話者音声の非ガウス性が高い特徴を利用して成分置換の頻度を求めて話者音声を抽出する提案法の場合、音源の距離に関係なく、話者音声を 100[%] 正確に抽出できることが分かる。比較のため、推定信号 $y_i(t)$ の Kurtosis を計算し、Kurtosis の大きい信号を話者音声として採用した結果も併せて示す。この Kurtosis に基づいたときの抽出率は、音源の距離によってばらつきがある。以上の結果は、提案法が騒音レベルに対しても頑健であることを示唆している。

4.4.5 多音源多マイクロホン環境での音源分離実験

多音源多マイクロホン環境における提案法の有効性を検証するために、音源1を話者音声、音源2を駅の騒音、音源3を車の騒音として実験を行った。 $(L10[m] \times W10[m] \times H10[m])$ のバーチャルルームにおいて、音源とマイクロホンは床から1[m]の高さでFig. 4.10のように配置してインパルス応答を算出し、音源と畳み込んで合計20組の混合信号を作成した。

Fig. 4.11に実験結果の波形を示す。(a)は音源、(b)は混合信号である。FastICAアルゴリズムのみの場合、(c)のように音源を推定できないことが分かる。一方、(d)は提案法による推定信号の波形である。スケールと位相の不定性と成分置換の問題が解決されて、音源とほぼ等しいことが確認できる。また、FastICAアルゴリズムの特徴に基づく目的音源の選択法は、多音源多マイクロホン環境下でも全データに対して話者音声を選択できることを確認した。

4.5 結言

実環境下でも有効に機能する雑音除去システムの構築を目的として、周波数領域の独立成分分析で大きな問題である成分置換について、音響伝達特性に基づいた解決法を提案した。本章では、まず、周波数領域の独立成分分析を用いた実環境下で観測される音響信号の分離方法を述べた後、非ガウス性の最大化に基づく周波数領域のFastICAアルゴリズムを概説した。また、周波数領域の独立成分分析で得られる分離信号のスペクトルにはスケールと位相の不定性と成分置換の問題が生じ、その値は各周波数で異なるために非常に困難な問題であることを示した。

スケールと位相の不定性に対しては、分割スペクトルを導入することで解消できることを述べた。また、分割スペクトルが音源のスペクトルと伝達関数の積で表せることを導出し、この結果から、成分置換の問題は、各音源からマイクロホンまでの距離によって得られる伝達関数から解決できることを述べた。成分置換に対しては、分割スペクトルから得られる伝達関数の比を推定して、そのゲインと位相に基づいて修正する方法を提案した。さらに、実環境下での音響信号のゲイン比と位相差は、音源からマイクロホンまでの距離でほぼ理論通りに決まることと、周波数帯域やマイクロホ

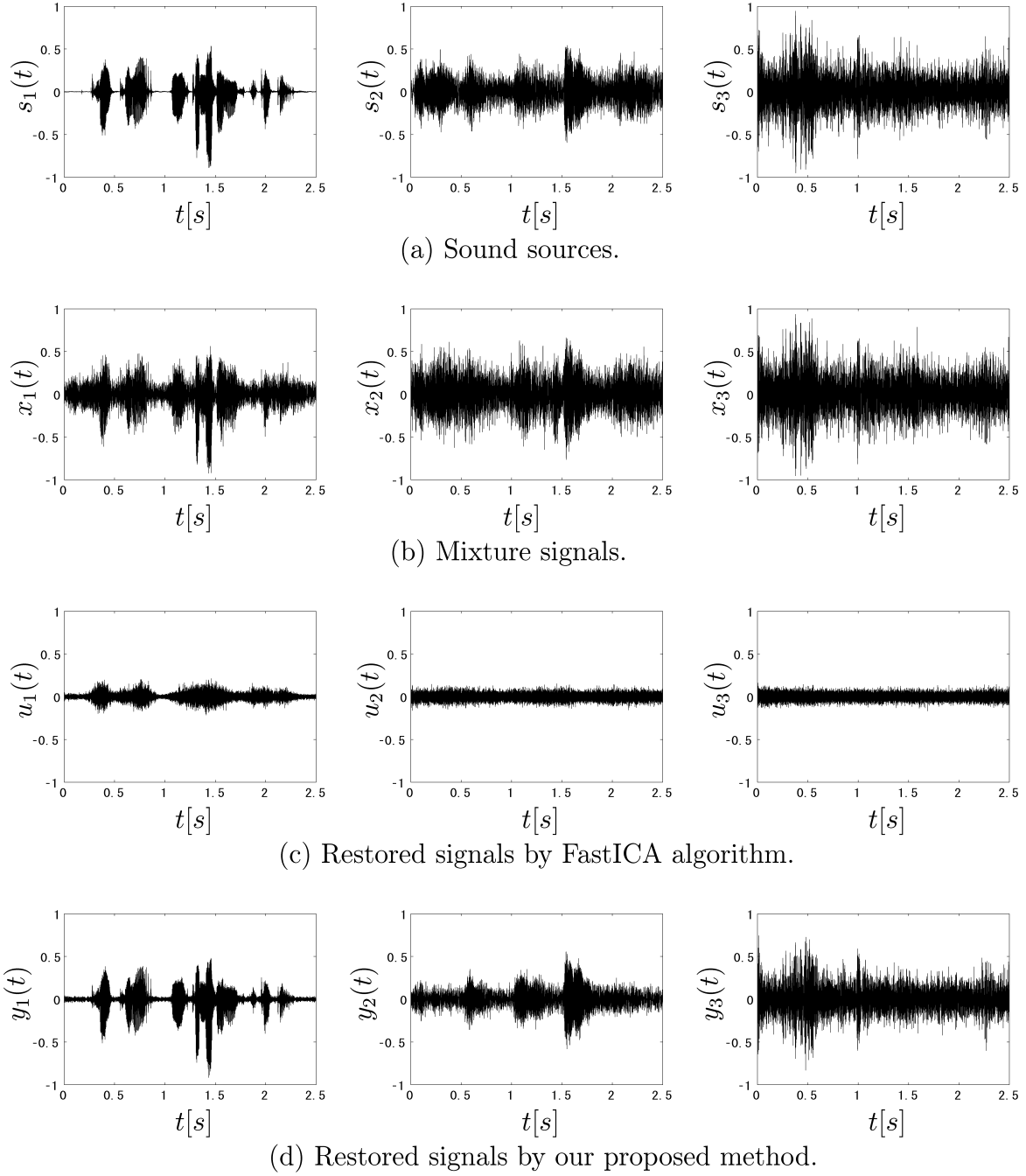


Fig. 4.11: Experimental result under 3 sources and 3 microphone condition.

ン間隔，音源とマイクロホンの距離に依存していることを明確にした．そして，この音響伝達特性に基づいて，伝達関数のゲインと位相の情報を組み合わせて利用する統合的な成分置換の修正法を提案した．

音源分離後の目的音声選択方法については，音源とマイクロホンの相対的な位置関係が分かっている場合には，提案する成分置換の修正法を用いることで目的音声を選択できることを述べた．また，この位置情報が得られない場合，周波数領域のFastICAアルゴリズムの非ガウス性の高い信号を最初に推定する特徴を利用して，簡単に話者音声を選択する方法を提案した．

実環境下での検証実験から，提案した成分置換修正法は音源とマイクロホンの距離や周波数帯域に依存せず，高い修正能力を持つことを確認した．また，騒音環境下で話者音声が存在する場合だけでなく，他人の発話が雑音となる場合においても有効に機能することが分かった．そして，多音源多マイクロホン環境下でも，提案法は成分置換の修正ができることを確認した．さらに，提案した目的音声選択法は，正確に話者音声を選択できることを確認した．

本章では，畳み込み混合信号を分離するときの周波数領域の独立成分分析に，FastICAアルゴリズムを用いた．しかし，ここで提案した成分置換の修正法は他のアルゴリズムに対しても適用できる．また，そのときの音源分離実験の結果も同様に，明瞭な音声として推定されていることを確認した．

第5章 単一試行視覚認識時脳波の特徴 抽出と変動識別

5.1 緒言

脳は、人間の思考、行動を支配するのみならず、その情動面や自律機能をも制御する中枢である。近年、生体計測技術の進歩により、脳波計 (EEG: Electroencephalogram), 脳磁計 (MEG: Magnetoencephalogram), 磁気共鳴画像 (MRI: Magnetic Resonance Imaging), 機能的磁気共鳴画像 (fMRI: Functional Magnetic Resonance Imaging), 陽電子放出断層撮影法 (PET: Positron Emission Tomography) などの装置を用いて、脳や体を痛めずに非侵襲で脳神経活動を計測することが可能になっている。さらに、このような計測装置を利用して高次脳機能に関する研究が盛んになり、臨床診断などにも用いられている [82]。

脳の電気活動を頭皮上から記録した電気信号である脳波は、脳のメカニズムを解明するために、医学の分野でも工学の分野でも多くの研究が行われている。脳波には、大別すると自発性脳波と誘発性脳波がある。自発性脳波は特定の事象の生起に関係なく持続的にゆらいでいる電位である。これに対して、誘発性脳波は何らかの事象生起に関連して一時的に出現する電位である。この誘発性脳波には、感覚刺激などに関連して注意、認識、課題解決、随意運動などの心理的な活動によって変動する事象関連電位 (ERP: Event Related Potential) がある。したがって、脳波から事象関連電位を抽出して識別することができれば、何を認識してどのように意思決定したかを推測できる。

しかし、脳波には、眼電図 (EOG: Electrooculigram) や瞬目 (blink), 筋電図 (EMG: Electromyogram), 心電図 (ECG: electrocardiogram) 等のアーチファクト (artifact) が含まれる。したがって、事象関連電位を見るためには、同一条件下で複数回試行されて計測した脳波を加算平均 (averaging) する処理が必要となる。加算平均するためには、音や光などの外部刺激によって、特定の脳内活動を誘発する必要がある。しかし、

このような脳活動は厳密に同一の時間構造で誘発されているとは限らないため、加算平均によって本来着目しようとしている信号を歪ませる可能性もある。また、外部刺激による誘発を行うことが難しい場合、単一試行波形には雑音によって乱されていて、そのままの波形を解析することは困難である。

これまでに、独立成分分析（ICA: Independent Component Analysis）を適用して、脳内の誘発成分（evoked potential）、自発成分（spontaneous potential）、心拍運動（heart rate）や眼球運動（eye movement）、電源の干渉（interferences）などの成分（artifacts）を計測信号から分解する研究が進められている [35]～[39]。しかし、独立成分分析の脳信号処理への適用に関する研究は、人頭模型による実験（Phantom 実験）や、擬似脳波に対する研究が多い。また、実際に観測された信号を解析する場合、多試行データに加算平均処理をした後に独立成分分析を適用して分離する報告が多く [40]～[42]、実際に観測された脳波の変動識別の研究に関しては見当たらない。

そこで本研究では、実際に観測された脳波に対して独立成分分析を適用して、単一試行データから特徴を抽出することと、その変動を識別することを目的とする。具体的には、視覚認識時の応答脳波に対して特徴を抽出するために、視覚性オドボール課題実験によって得られる観測信号に独立成分分析を適用する。このようにして、事象関連電位や閃光刺激に対する応答と考えられる成分を抽出して、その特徴から、脳波の変動を判断する方法を提案する。また、脳波解析では原信号数が完全に未知であるため、原信号数を推定したもとで独立成分分析を行う特徴抽出法を提案する。この特徴抽出法を前頭部で観測された脳波に用いて思考に関連した特徴を抽出し、その特徴に基づいた脳波変動識別法を提案する。

本章の構成は以下の通りである。脳波信号は音声信号と異なり一般になじみがない。そこでまず、**5.2** で脳波研究の歴史と脳波計測方法、視覚誘発電位、事象関連電位について述べる。次に、**5.3** では、脳波信号に独立成分分析を適用した特徴抽出法を提案する。さらに、原信号数を推定することで、より明確な特徴を抽出する方法を述べる。単一試行視覚認識時脳波に対して特徴抽出して脳波変動を識別する実験により提案法の有効性を確認し、その結果を **5.4** で述べる。最後に、**5.5** で本章のまとめを述べる。

5.2 脳波研究の歴史と脳波の概要

5.2.1 脳波研究の歴史

脳は、人間の思考、行動を支配するのみならず、その情動面や自律機能をも制御する中枢である。しかしながら、時々刻々と進行する脳機能の変化を客観的に捉える方法は、脳波の発見以前にはほとんど存在しなかった。すなわち、それまでは、中枢の機能状態を知るためには、主に刺激に対する反応を末梢で観察する方法が用いられるにすぎなかった。ところが、1791年に Galvani が筋の収縮時に電流が発生することを観察して以来、脳が活動するときにも同様に電氣的变化が生じると考えられるようになった。

最初に脳の電氣活動を記録したのは、1875年に Caton であるといわれる。Caton は、ウサギおよびサルの大脳皮質から直流電流と思われる電氣活動を記録し、それが脳機能に関係があると考えた。次いで、1890年にイヌの視覚領域皮質が光刺激を与えると大きな電氣変動を示し、光が与えられないと小さな電位変動しか生じないことが発表された。また、同時期には、これらの電位変動は、硬膜上あるいは頭蓋骨上からも記録できることが報告されている。その後、動物の大脳皮質の電氣活動や、それが末梢刺激に対して示す反応などについて発表されているが、当時の技術では目を見張るべき成果を挙げるにはいたらなかった。この後、弦電流計や真空管増幅器が実用化されてから、心電図の研究や脳の電氣活動の研究が急速に進歩した。

1924年、Hans Berger はヒトの脳の電氣活動を最初に記録し正確に記載した。すなわち、2本の白金針電極を頭蓋骨欠損部から大脳皮質に向けて挿入し、ヒトの脳から規則正しい電氣活動を記録することに成功した。また、このような電氣活動が、電極を挿入しなくても、頭皮上に装着した電極からも同様に記録できることを確かめた。さらに、健常者の安静閉眼時に、後頭部や頭頂部にみられる周波数が10[Hz]、振幅が数[μ V]前後の規則正しい波をアルファ波 (α wave) と命名した。また、開眼して注視するとアルファ波が消失し、これに代わって18~20[Hz]で20~30[μ V]の波が出現することを確認し、これをベータ波 (β wave) と命名した。そして、これらの脳の電氣活動を総称して、脳電図、または脳波 (Elektrenkephalogram) と命名した。

現在では、脳波計、脳磁計、機能的磁氣共鳴器などの生体計測技術の進歩により、脳

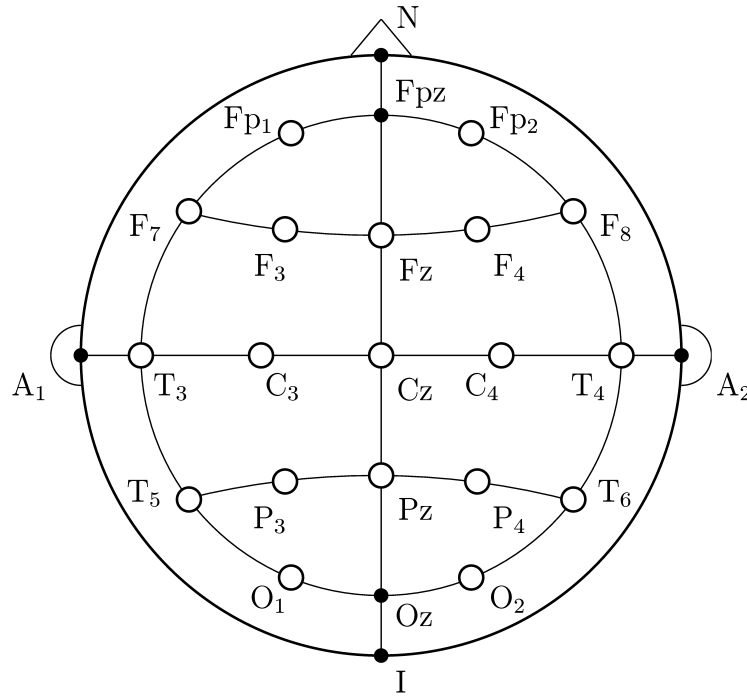


Fig. 5.1: International ten-twenty electrode system.

や体を痛めずに脳神経活動を計測することが可能になっている。さらに，高次脳機能に関する研究が盛んになり，このような計測装置を利用した臨床診断の研究が進められている。

5.2.2 国際脳波学会連合標準電極配置法

国際脳波学会連合標準電極配置法（International ten-twenty electrode system）による電極の位置は Fig. 5.1 に示す 19 点であり，それぞれ以下のように定められている。それぞれの電極は，まず，正中線上に鼻根（N: nasion），頭蓋頂（V: vertex），後頭極（I: inion）を通る線を考え，その線上に五つの点を決める。すなわち，前頭極正中中部（Fpz: frontal pole midline point）を鼻根から後頭極までの距離の 10[%] 鼻根から後ろに配置し，前頭正中中部（Fz: frontal midline point）をそれから 20[%] 後方に配置して，さらにその後 20[%] 後方に中心正中中部（Cz: central midline point）がありこれは頭蓋頂に一致する。中心正中中部から 20[%] 後方には頭頂正中中部（Pz: parietal midline point），その 20[%] 後に後頭中央部（Oz: occipital midline point）がありこれは後頭

極の10[%]前に位置する．このように，10[%]，20[%]と配置していくので，国際10-20 system と呼ばれている．

外側部の配置は，中心部の冠状線を基準にする．左右の耳介前点を定め，左右耳介前点間距離の10[%]の位置に左側頭中部（ T_3 : left mid-temporal）と右側頭中部（ T_4 : right mid-temporal），それから20[%]の位置に左中心部（ C_3 : left central）と右中心部（ C_4 : right central）を配置する．

左前頭極（ Fp_1 : left frontal pole）と右前頭極（ Fp_2 : right frontal pole）は，前頭極正中中部から左右側頭中部と後頭中央部を通る周線をそれぞれ100[%]として，前頭極正中中部の10[%]後方に配置する．左側頭前部（ F_7 : left anterior temporal）と右側頭前部（ F_8 : right anterior temporal）はその20[%]後方に配置する．左側頭後部（ T_5 : left posterior temporal）と右側頭後部（ T_6 : right posterior temporal）は側頭中部から20[%]後方に，左後頭部（ O_1 : left occipital）と右後頭部（ O_2 : right occipital）は後頭中央部から10[%]前方に配置する．

最後に，前頭正中中部と左右側頭前部を通る冠状線上で，それぞれ等間隔の位置に，左前頭部（ F_3 : left mid-frontal）と右前頭部（ F_4 : right mid-frontal）を配置して，頭頂正中中部と左右側頭後部を通る冠状線上で，それぞれ等間隔の位置に，左頭頂部（ P_3 : left mid-parietal）と右頭頂部（ P_4 : right mid-parietal）を配置する．

5.2.3 脳波の計測

頭皮上脳波を記録するためには，まず，被験者の頭皮に電極を装着する．ここでは，直径約1[cm]の銀一塩化銀皿型電極を用いる．皿型電極には電導性ペーストを塗り，電極と頭皮との接触抵抗を低くする．このようにして，頭皮に電極を装着すると，二つの電極の間にわずかな電位差が生じる．実際の電位差は数十[μ V]である．

脳波を記録する場合，脳の電気活動を受けにくい部位に貼り付けた電極を基準として，頭皮上の探索電極（exploring electrode）から記録するのが一般的である．基準電極（reference electrode）は，左右の耳朵や鼻尖，頭外部に置くことが多い．また，接地電極（ground electrode）は額や手首，足首などの頭外部に置く．

電極で観測された脳波は電極箱に接続され，そこから遮蔽した導線によって脳波計に導かれる．脳波計では差動増幅器によって数万倍に増幅される．このとき，それぞ

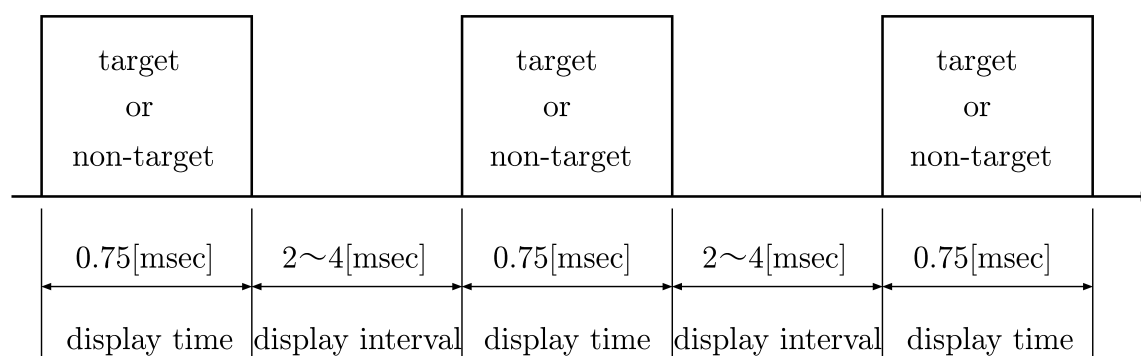


Fig. 5.2: Oddball task experiment.

れの探查電極から基準電極の差分が脳波として観測されることになる。

5.2.4 視覚性オドボール課題実験

オドボール (oddball) 課題とは、識別可能な二種類の感覚刺激をランダムで呈示し、一方の刺激の呈示頻度を他方よりも少なくし、呈示頻度の低い刺激が呈示されたときに、それに応じて、低頻度刺激の数をカウントさせたり、スイッチを押させたりする反応を行わせる課題をいう。低頻度刺激は標的刺激 (target stimuli) と呼ばれ、高頻度刺激は標準刺激 (standard stimuli) や非標的刺激 (non-target stimuli) と呼ばれる。

本論文におけるオドボール課題実験では、視覚認識時の脳波を対象に、標的刺激に赤色、標準刺激に緑色の光刺激とした。その出現割合は、赤色 : 緑色 = 3 : 7 とした。また、標準刺激が呈示されたときには何も考えないようにして、標的刺激が呈示されたときに標的刺激の出現数を頭の中でカウントさせる課題を与えた。各刺激の呈示時間は 0.75[msec]、刺激の順序はランダム、刺激提示間隔は 2~4[sec] をランダムで変更し、標的刺激が 30 回呈示されるまで続けた。また、脳波計のサンプリング周波数は 1000[Hz] とした。この実験の流れを Fig. 5.2 に示す。被験者は 20 歳代前半の男性 1 名、女性 1 名の合計 2 名で、上述の実験を 1 日に 3 回行い、それを 3 日間行った。

このようなオドボール課題を付加して、標的刺激、標準刺激によって誘発される電位をそれぞれ個別に頭皮上で記録して 20~50 回で加算平均すると、いくつかの頂点を持つ電位がみられる。このうちのいくつかは誘発電位、すなわち、外因性電位であ

り、標的刺激と標準刺激の両方によって同様に誘発される。これに対して、いくつかのピークは標的刺激に対してのみ出現する。この電位は事象関連電位と呼ばれ、以降で詳しく述べる。

5.2.5 視覚誘発電位

視覚誘発電位 (VEP: visual evoked potential) の研究の発達は、他の誘発電位と同じく平均加算法が導入されてから発展し、1960～1970年代には閃光刺激による視覚誘発電位の正常値や臨床応用についての研究が進められた。また、同時期にパターン反転刺激 (pattern reversal stimulation) という新しい視覚刺激方法が開発され、単眼の半側視野刺激が可能なこともあって急速に普及してきた。1970年代末頃からは短潜時視覚誘発電位の研究が行われている。

閃光刺激は放電型ストロボスコプを用いて、開眼で眼前 30～45[cm] の距離から刺激を与える。閃光時間は 10～40[μ sec]、刺激感覚は 1～3[sec] で、周期的あるいはランダムとする。加算平均回数は 30～50 回である。このように閃光刺激を与えたとき、視覚誘発電位には 1[Hz] 前後の頻度の刺激によって生じる transient 型誘発電位と、頻度の速い反復刺激によって生じる steady-state 型誘発電位がある。

閃光刺激による transient 型誘発電位の正常波形は、刺激から 250[msec] 以内に 5～7 個の頂点を持つ。パターン反復刺激による transient 型誘発電位は、全視野刺激の場合、正中後頭電極を中心に三相性波形が現れる。半側視野刺激の場合、正中後頭電極および刺激と同側半球側に三相性波形が現れる。刺激視野と反対側の視覚領に出現するはずの誘発電位が刺激視野と同側の頭皮上から記録される逆説的側性化 (paradoxical lateralization) は、ヒトの視覚野が後頭用に内側面に存在するため、そこで生じた誘発電位のベクトルが刺激視覚と同側の方向を向いているためと説明されている。

閃光刺激による steady-state 型誘発電位は、高頻度の反復閃光刺激による正弦波様の視覚誘発電位である。パターン反復刺激による steady-state 型誘発電位は、全視野刺激、半側視野刺激の場合とも正中後頭部で振幅が最大で、外側後頭部、外側側頭部の順に振幅は低下し、左右はほぼ対称である。

5.2.6 事象関連電位

事象関連電位（ERP: Event Related Potential）は感覚刺激などに関連した注意，認識，課題解決，随意運動などの心理的な活動によって変動する脳波であり，同一条件下で観測された脳波を加算平均することで，P300，随伴陰性電位（CNV: Contingent Negative Variation），運動関連電位（MRCP: Movement Related Cortical Potentials）などが見られる [82]。P300 は後期陽性複合体（LPC: Late Positive Complex）とも呼ばれ，刺激の不確実さに関連して，約 300[msec] で誘発される大きな陽性電位である。随伴陰性電位は，刺激に対して反応させるときに，反応に先行してゆっくりと立ち上がる陰性の電位である。運動関連電位は，随意運動に伴って，運動開始に先行する電位変化である。

P300 を含む事象関連電位を構成する各成分を分解して出現時間の順序に並べた模式図は，Fig. 5.3 のようになる。P300 は情報呈示による被験者の心理的不確実さの解決に関連して出現し，刺激の情報内容により変動する内因性の電位であると考えられていて，いくつかの成分によって構成されている。まず，P3a は 220～280[msec] で出現し，低頻度の刺激に対して誘発され，前頭部と中心部優位である。これに対して，P3b は 310～380[msec] で，オドボール課題の標的刺激に対して出現する。したがって，P3b は P300 の主成分であり，頭頂部優位である [83]。

P3a は期待に基づいた刺激の感知，あるいは期待された刺激とのミスマッチを感知する信号の登録に関連するのに対して，P3b は刺激感知後に，それを同定するためにさらに評価・分析が行われるときに出現し，認知過程における文脈の更新に関連するものと考えられている。また，標的でない低頻度の新奇（novel）な刺激に対しても，P300 様の成分が出現し，novelty P300 と呼ばれている [84]。

P300 の振幅は頭皮上では正中部で高く，頭頂部で 10～20[μ V] 最大で，中央部，前頭部の順に低下する。振幅には個人差が大きく，小児期には成長とともに増大し，15 歳頃が最大で，以後は加齢とともに低下する傾向を示すことが分かっている。また，P300 の潜時は 250～500[msec] で，年齢や標的刺激識別の難易度によって異なる。年齢との関係を見ると，小児期には長く，成長とともに短縮して 15 歳頃に最短になり，以後加齢とともに延長する。

さらに，P300 の各種条件下での変動について調べてみると，P300 は聴覚，視覚，体

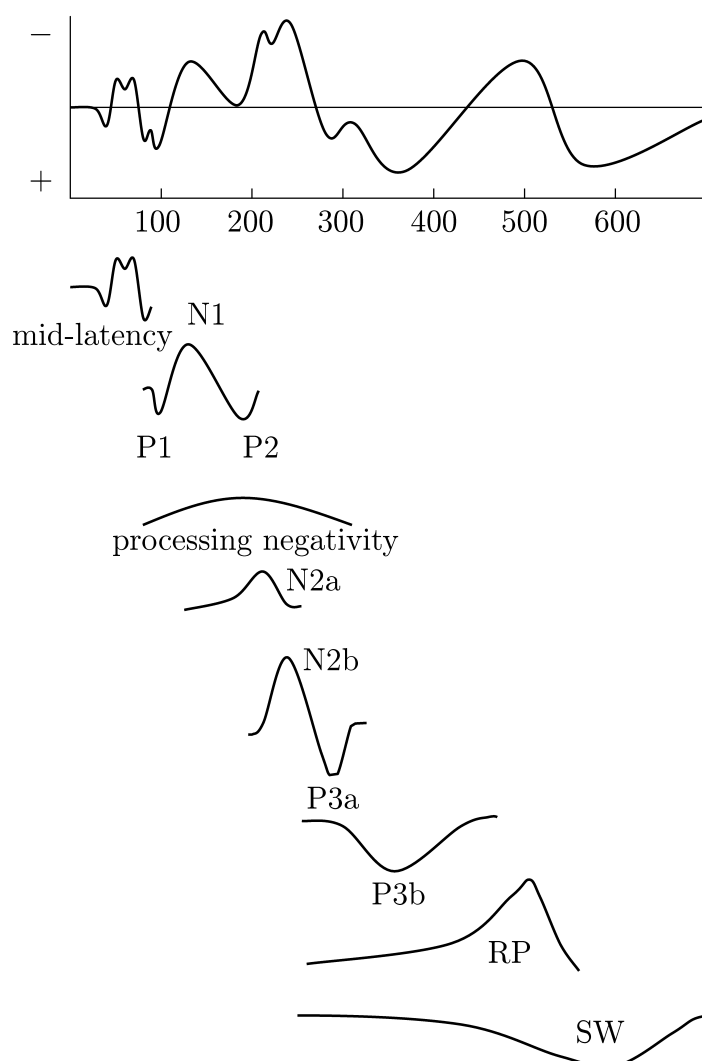


Fig. 5.3: Event Related Potential

性感覚などの感覚刺激の種類には関係なく、標的刺激に対して出現する。これに対して、早期に出現する外因性電位は、感覚の種類によっても異なり、N100, P200 は聴覚刺激に依存して出現する。

P300 が生じるためには、課題遂行中に被験者が標的刺激に注意している必要があり、検査中に注意をそらすと P300 は出現しなくなる [85]。また、標準刺激の呈示頻度に比べて標的刺激の呈示頻度が低いほど、P300 が明瞭に出現する [86]。さらに、標的刺激と標準刺激の呈示のしかたが規則的であるほど、すなわち、被験者が標的刺激の出現を予測しやすいほど、P300 は不明瞭になる。そのうえ、標的刺激と標準刺激の差が少

なく、標的刺激が標準刺激の系列の中から識別しにくいときには、P300の潜時が長くなる。

5.3 独立成分分析を用いた脳波の特徴抽出

本章では、独立成分分析に基づく単一試行視覚認識時脳波の特徴を抽出して、その変動を識別するシステムの構築を目的とする。このように視覚認識時脳波の変動を識別することは、人間の考えを直接コンピュータに伝えるマンマシンインターフェース (man-machine interface) の実現が可能になる。また、コンピュータに伝えた情報が正しいか否かを人間が視覚によって判断したときの脳波情報をフィードバックすることで、正確な情報伝達が期待できる。

ここではまず、視覚認識時脳波に対して、独立成分分析の適用方法を述べる。その後、独立成分分析に基づいて単一試行視覚認識時脳波の特徴を抽出する方法を提案する。このとき、独立成分分析による分離信号にフィルタ処理を施すことで、事象関連電位や閃光刺激に対する応答と考えられる特徴が抽出できることを示す。また、分離信号のスケールを回復した分割信号に基づいて原信号数を推定することで、オドボール課題実験による思考に対する応答と考えられる特徴が抽出できることを述べる。さらに、以上の特徴に基づいた脳波の変動識別法を提案する。

5.3.1 独立成分分析の脳波解析への適用

国際脳波学会連合標準電極配置法 (International ten-twenty electrode system) による19箇所に加えてOzを加えた20箇所を観測された信号を $\mathbf{x}(t) = [x_1(t), \dots, x_{20}(t)]^T$ のように表す。一方、脳内の原信号は完全に未知であるため、 $\mathbf{s}(t) = [s_1(t), \dots, s_{20}(t)]^T$ のように表現できて、互いに統計的に独立であると仮定する。また、脳内の電気信号は非常に短い時間で伝達すると考えられるので、瞬時混合モデルを用いて

$$\mathbf{x}(t) = A\mathbf{s}(t) \quad (5.1)$$

のように表せるものとする。

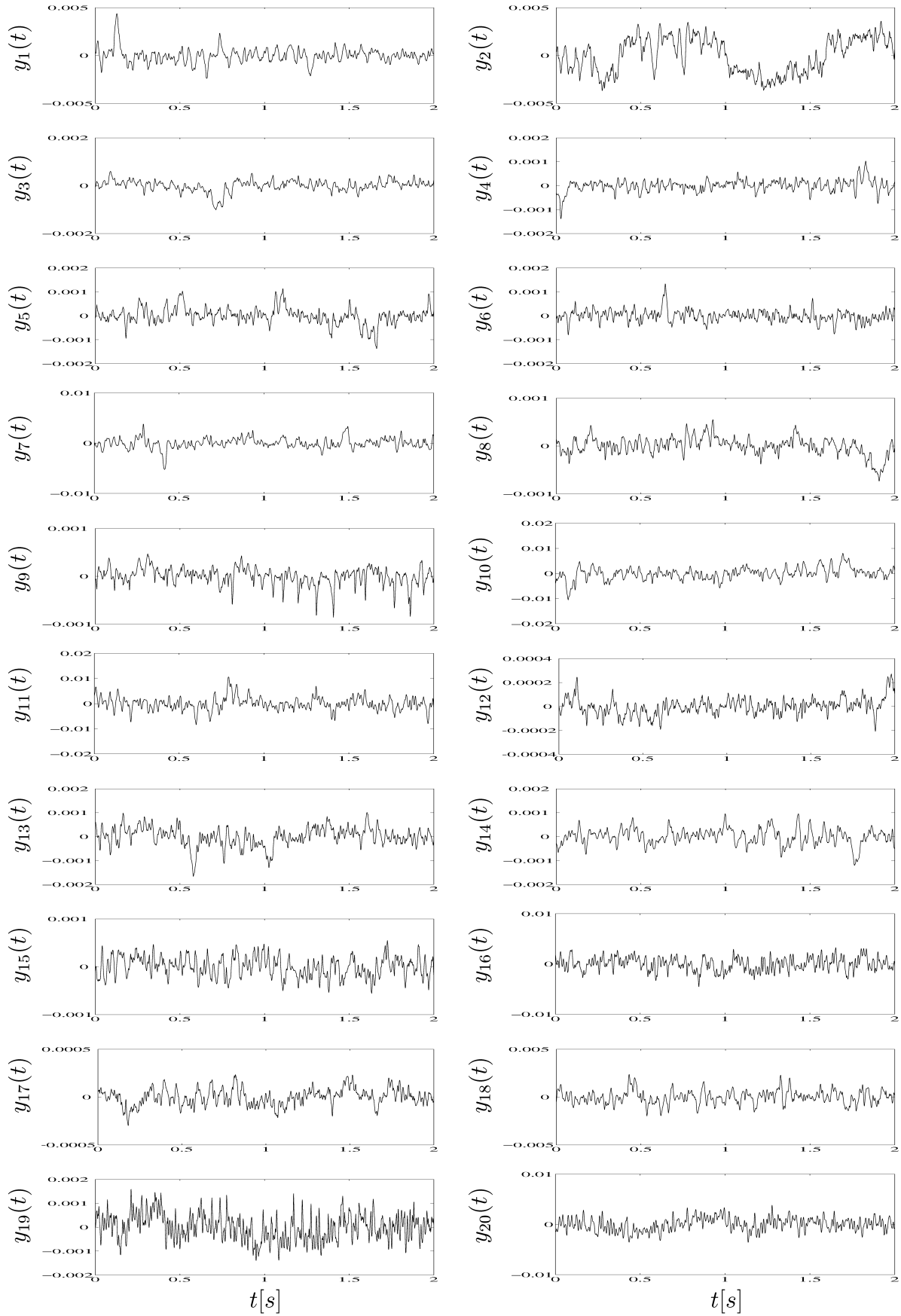


Fig. 5.4: Separated signals by ICA in the case of the target stimuli.

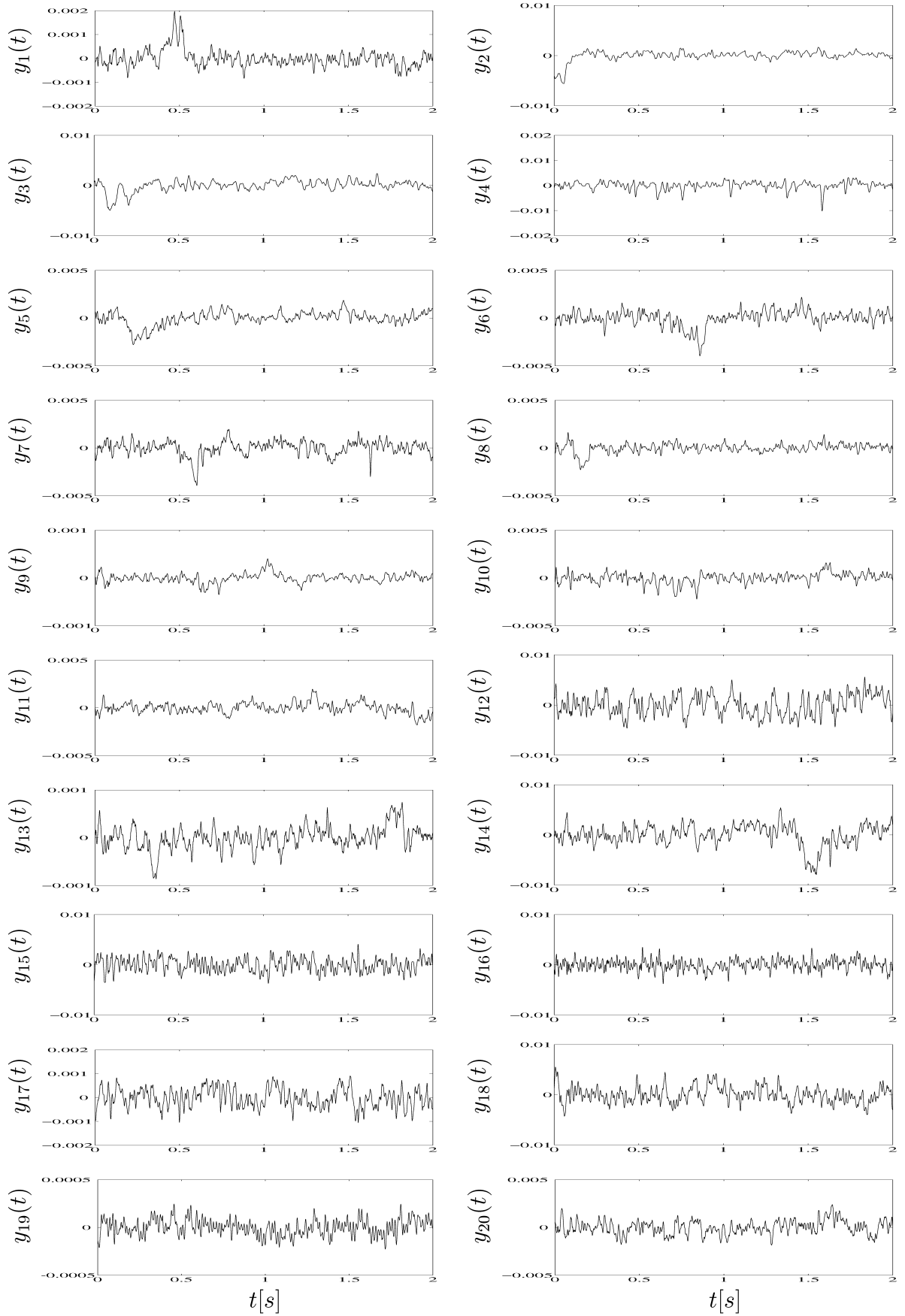


Fig. 5.5: Separated signals by ICA in the case of the non-target stimuli.

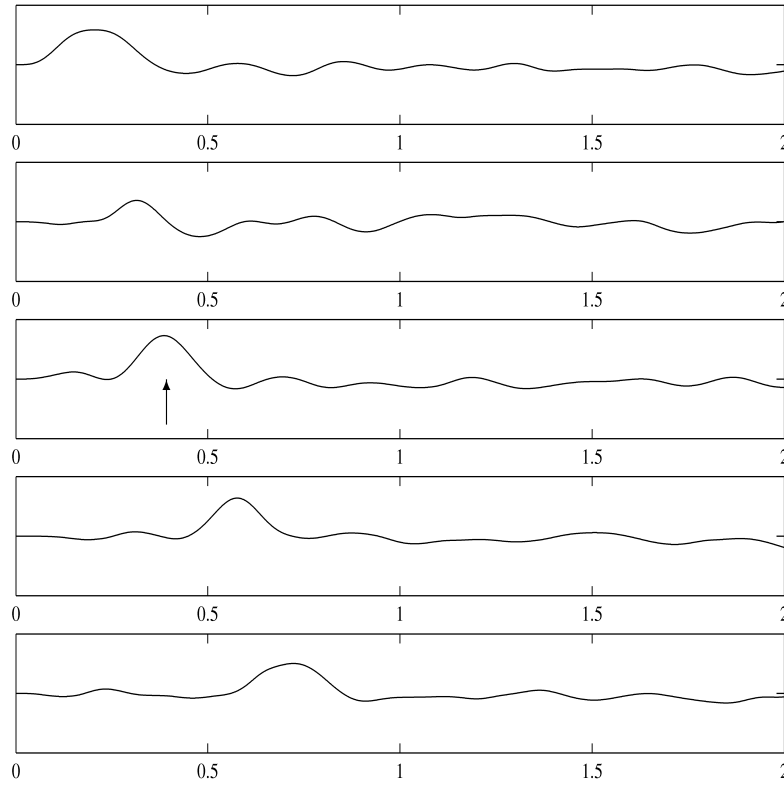


Fig. 5.6: Separated signals after the LPF in the case of the target stimuli.

この瞬時混合モデルを用いることで、独立成分分析を利用した脳波解析では

$$\mathbf{u}(t) = W\mathbf{x}(t) \quad (5.2)$$

のように分離信号 $\mathbf{u} = [u_1(t), \dots, u_{20}(t)]^T$ を推定することになる．このときの分離信号の波形は Fig. 5.4 と Fig. 5.5 である．Fig. 5.4 は標的呈示時の応答脳波に対する分離信号であり，Fig. 5.5 は標準刺激呈示時の応答脳波に対する分離信号である．

以降では，まず，これらの分離信号にフィルタ処理を施した脳波の特徴抽出法を提案する．また，原信号数を推定して独立成分分析を適用することで，さらに明瞭な特徴を抽出できる方法を提案する．さらに，それぞれの特徴に基づいた脳波変動識別法を提案する．

5.3.2 事象関連電位と閃光刺激に関連した成分の抽出

ここでは，事象関連電位 P300 と同時刻に出現する信号のピークを見つけるために，まず，分離信号にローパスフィルタを掛けて，その出現時刻の早い順に分離信号を並び

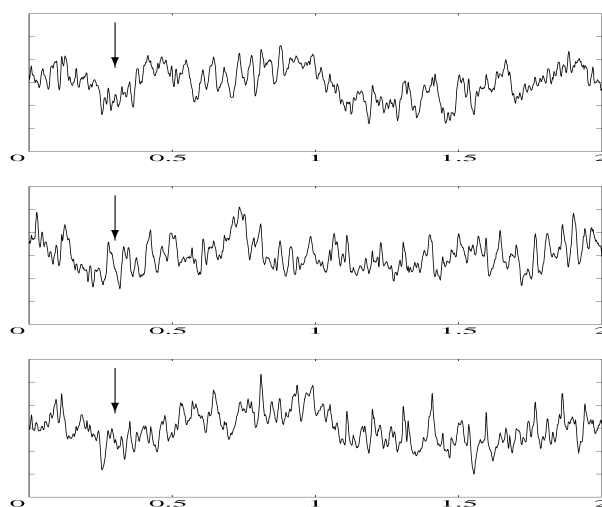


Fig. 5.7: ERP by averaging in the case of the target stimuli.

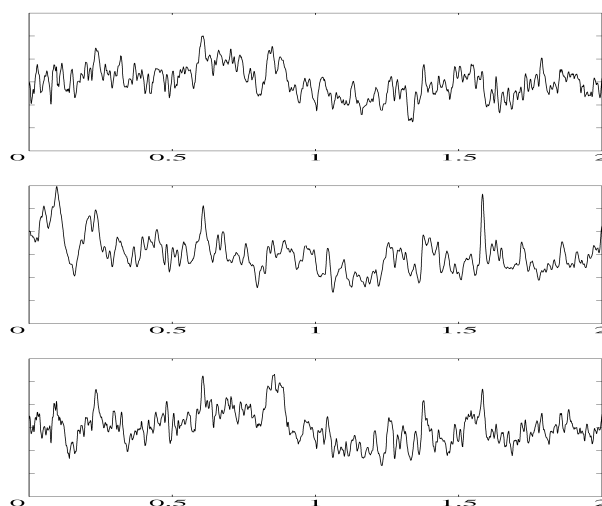


Fig. 5.8: Signal by averaging in the case of the non-target stimuli.

替えることで、分離信号のピークの出現時刻を明確にする．この処理によって得られる標的刺激呈示時の分離信号の波形を Fig. 5.6 に示す．ここでは、得られた 20 個の分離信号のうち、ピークが出現する順に、5 成分のみを示している．Fig. 5.6 の第 3 成分に見られるように、標的刺激を呈示した場合、分離信号の一つに刺激呈示後約 380[msec] においてピークが出現することが確認できる．このピークの出現は、標的刺激を呈示した場合のほとんどの分離信号で確認することができたが、標準刺激に対する分離信号では確認できなかった．このことは、独立成分分析によって、標的刺激に対する単一試行応答脳波から事象関連電位 P300 の成分が抽出されたと考えられる．

このようにして得られた独立成分分析の分離信号が事象関連電位を抽出しているこ

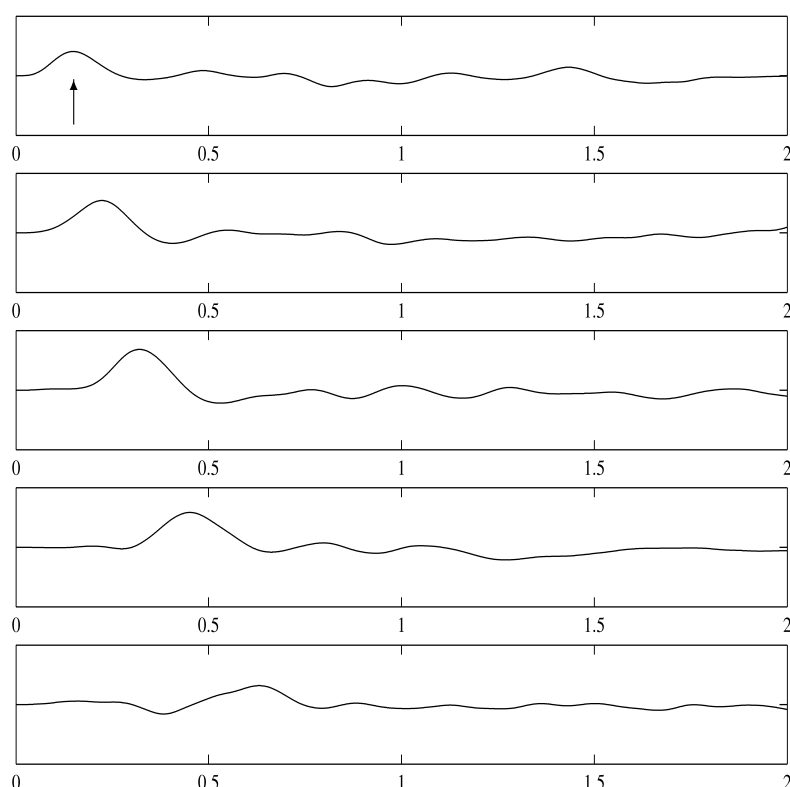


Fig. 5.9: Separated signals after the LPF in the case of the non-target stimuli.

とを確認するため、標的刺激に対する脳波データから得られる加算平均波形を Fig. 5.7 に示す。Fig. 5.7 から、刺激呈示後約 380[msec] において、大きな変動が確認できて、これは事象関連電位 P300 と考えられる。一方、標準刺激に対する脳波データを加算平均した結果は Fig. 5.8 で、事象関連電位が存在していないことが分かる。以上のことから、独立成分分析によって得られる分離信号の一つは、加算平均処理によって得られる事象関連電位と同時刻のピークを持つことが分かる。また、独立成分分析は単一試行データから事象関連電位と考えられる成分を抽出できることが分かる。

同様に、標準刺激を呈示したときの観測信号に対する分離波形を Fig. 5.9 に示す。標準刺激を呈示した場合の分離信号には、標的刺激を呈示したときとは異なり、刺激呈示後約 380[msec] におけるピークの出現はほとんど見られないことが確認された。しかし、標準刺激を呈示した場合、第1成分のピークは、約 180[msec] でほぼ同時刻に出現するという規則が見られた。一方、標的刺激を呈示した場合、最初に出現するピークは約 200[msec] 以降であることが分かった。このことは、標準刺激の緑色の光が、標的刺激の赤色の光より明るいため、被験者が感じる眩しさに影響した結果と考えら

れる.

以上の結果は、独立成分分析によって、単一試行視覚認識時脳波から、有意な特徴が抽出されることを意味している. さらに、この分離信号のピークが出現する時刻に基づいて、簡単に脳波の変動を識別できることを意味している.

上述の結果に基づいて、独立成分分析を適用した後の分離信号波形について、約 380[msec] でピークが出現する分離信号が存在する場合を標的刺激呈示後の応答脳波、存在しない場合を標準刺激呈示後の応答脳波と判断する脳波変動識別法が提案できる. 同様に、第1成分のピークが約 180[msec] に出現する場合を標準刺激呈示後の応答脳波、出現しない場合を標的刺激呈示後の応答脳波と判断する脳波変動識別法が提案できる [87].

5.3.3 原信号数の推定と思考に関連した成分の抽出

脳波に独立成分分析を適用する場合、脳内の原信号数は完全に未知であるため、原信号数は観測信号数と等しいと仮定して解かれることが多い. ここで、原信号数を知ることができれば、独立成分分析による分離信号は、さらに原信号を明確に推定できると考えられる.

実際に本論文で用いている多センサで観測された脳波信号に対して独立成分分析を用いた解析を行った結果、分離信号のいくつかは非常に似ている波形となって出力されていた. また、それらの波形は分割信号

$$\mathbf{v}_n(t) = W^{-1}[0, \dots, 0, u_n(t), 0, \dots, 0]^T \quad (5.3)$$

を利用してスケールを回復させたときに、振幅が非常に小さくなって出力される信号があることを確認した. この事実から、原信号数はセンサ数よりも少ないと予想できる.

そこで、原信号数を推定するために、まず、分割信号によってスケールが回復した推定信号から、振幅が大きな推定信号数を原信号数として推定する. さらに、推定信号から特徴的な信号数を推定する. 分離信号の非ガウス性は

$$\kappa_3(y_n(t)) = E[y_n^3(t)] \quad (5.4)$$

$$\kappa_4(y_n(t)) = E[y_n^4(t)] - 3E[y_n^2(t)]^2 \quad (5.5)$$

$$J(y_n(t)) \approx \frac{1}{12}\kappa_3(y_n(t))^2 + \frac{1}{48}\kappa_4(y_n(t))^2 \quad (5.6)$$

のように、歪度 (Skewness) $\kappa_3(y_n(t))$, 尖度 (Kurtosis) $\kappa_4(y_n(t))$, Negentropy $J(y_n(t))$ を用いて計算できる. 特徴的な信号は非ガウス性の分布となることが多いといわれているので, $\kappa_3(y_n(t))$, $\kappa_4(y_n(t))$, $J(y_n(t))$ の値が大きい信号数が, 最終的に特徴的な分離信号数 \hat{N} として推定できる.

特徴的な原信号数 \hat{N} を推定した後, 再度, FastICA アルゴリズムによって混合信号 $x_n(t)$ ($n = 1, \dots, \hat{N}$) から FastICA アルゴリズムによって分離荷重 w_n ($n = 1, \dots, \hat{N}$) を推定し, 分離信号 $y_n(t)$ ($n = 1, \dots, \hat{N}$) が求められる. 以上の処理によって, 特徴的な分離信号を得ることができる [88], [89].

実際に, 観測信号に対して独立成分分析によって分離した後, スケールを回復した推定信号は, 前述の Fig. 5.4 と Fig. 5.5 であり, これらの振幅を統一した波形が Fig. 5.10 と Fig. 5.11 である. Fig. 5.10 から, 振幅が大きく, 非ガウス性が高い推定信号は, $y_2(t)$, $y_7(t)$, $y_{10}(t)$, $y_{11}(t)$, $y_{16}(t)$, $y_{20}(t)$ の 6 個であることが分かる. 同様に, Fig. 5.11 では, $y_2(t)$, $y_3(t)$, $y_4(t)$, $y_{12}(t)$, $y_{14}(t)$, $y_{18}(t)$ の 6 個の推定信号について, 振幅が大きく, 非ガウス性が高いことが分かる. このようにして, 特徴的な原信号数を推定した結果, 本論文で観測した脳波においては, どの場合でも 6 個以下になることを確認した.

したがって, 微弱な信号は多数あるものの, 視覚刺激や課題に対して活発に反応する信号数は 6 個以下であると推定できて, 6 個の観測信号に対して独立成分分析を適用すれば, 推定信号に明確な特徴が抽出されと考えられる. このとき, 観測信号の組み合わせについては, 詳しく調べる必要がある.

そこでまず, 前頭部に配置した Fp1, Fp2, F7, F3, Fz, F4 を用いて実験した. 前頭部の前頭前野では思考や意思決定を行っているため, オドボール課題実験における課題に関して, 前頭部で観測された脳波は特徴が得られやすいと考えられる.

観測信号を分離した結果, 標的刺激を提示した場合には Fig. 5.12 のような分離信号が得られた. また, 標準刺激を呈示した場合には Fig. 5.13 の分離信号が得られた. これらの波形から, 推定した原信号数に基づいて分離した波形には, 矢印で示す時刻に特徴的なピークが現れていることが確認できる. すなわち, Fig. 5.12 の標的刺激では, $y_3(t)$ の 100[msec], $y_6(t)$ の 700[msec], $y_1(t)$ の 1800[msec] である. 同様に, Fig. 5.13 の標準刺激には, $y_5(t)$ の 100[msec], $y_4(t)$ の 800[msec], $y_1(t)$ の 1600[msec] に特徴的なピークが確認できる.

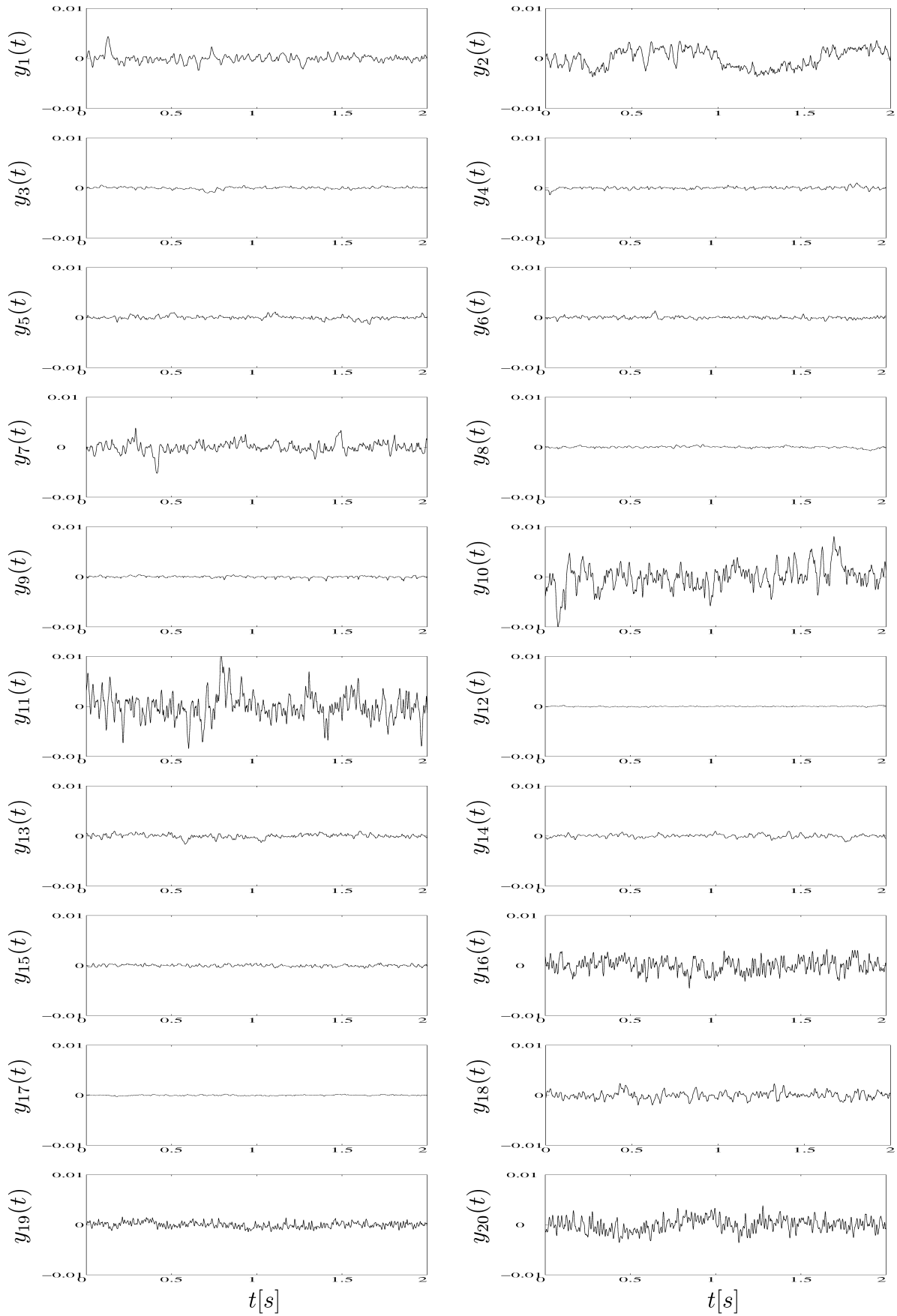


Fig. 5.10: Scaling adjusted signals in the case of the target stimuli.

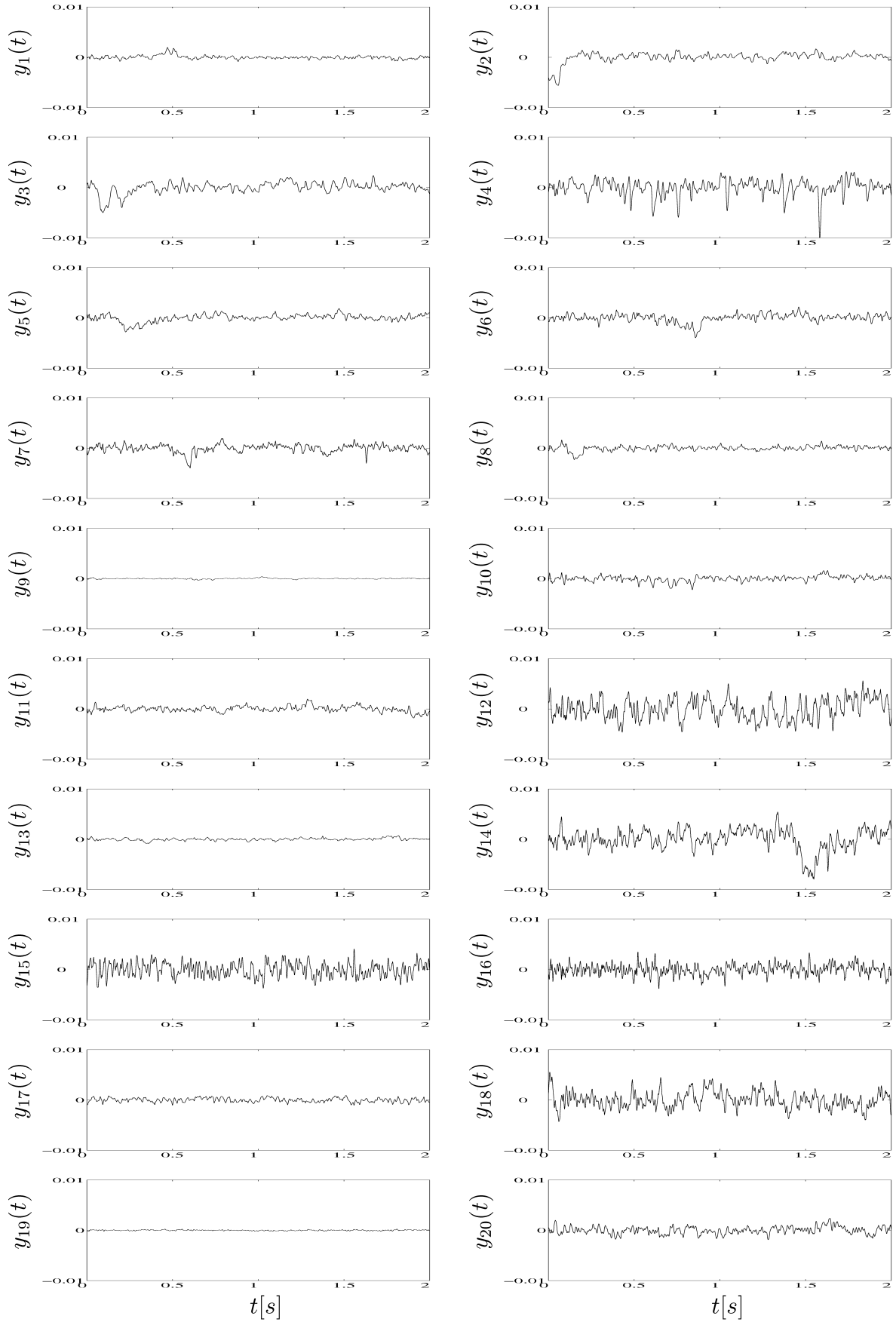


Fig. 5.11: Scaling adjusted signals in the case of the non-target stimuli.

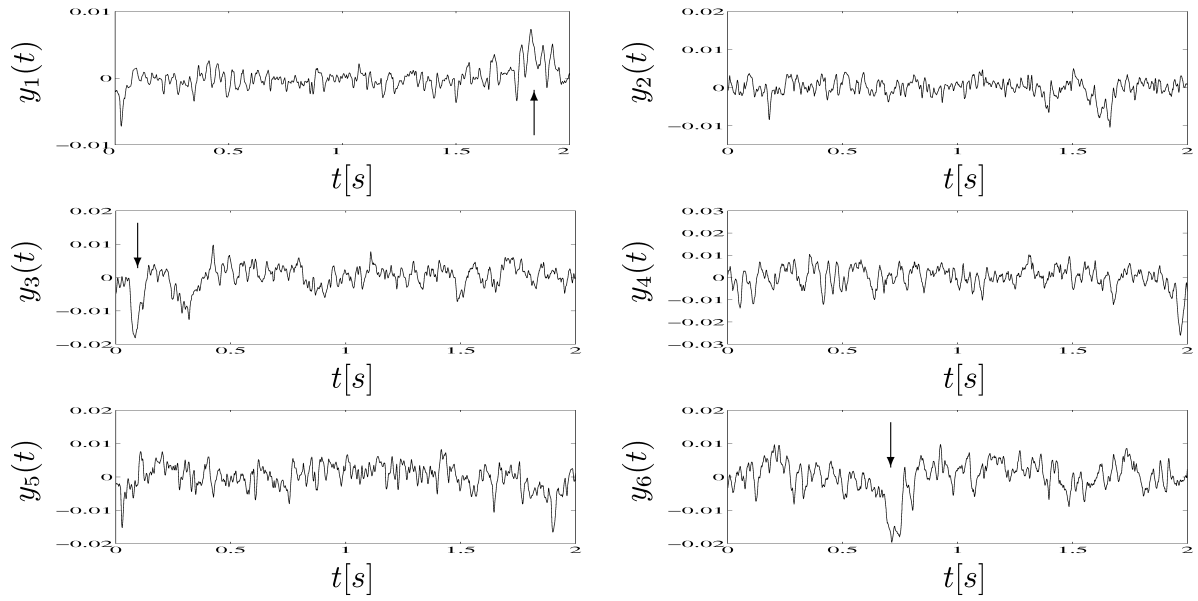


Fig. 5.12: Separated signals of frontal region in the case of the target stimuli.

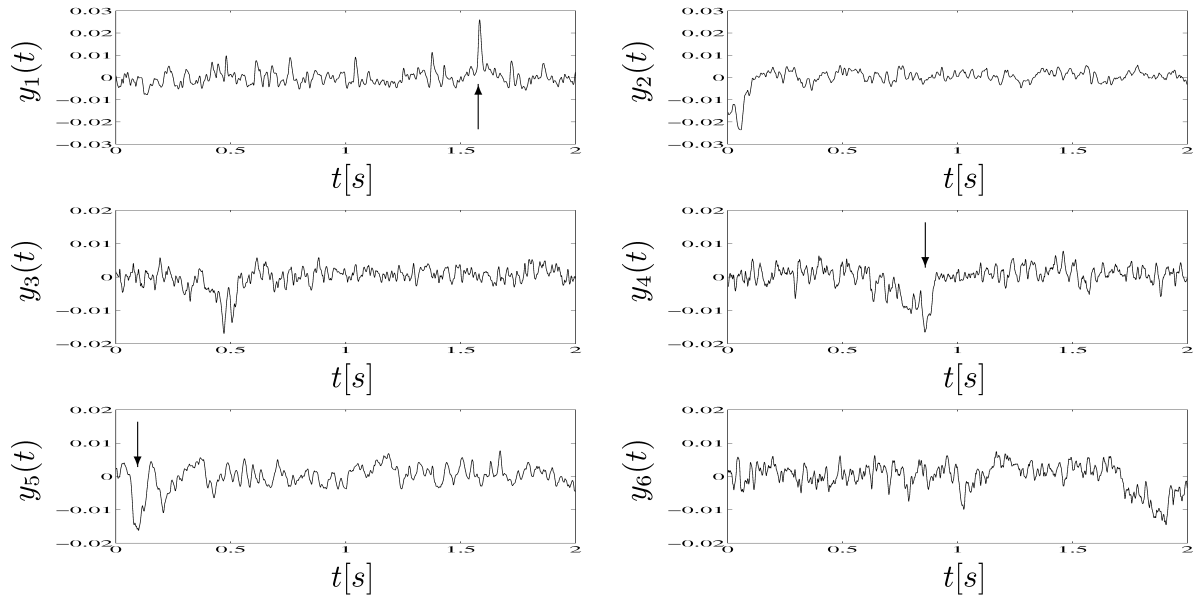


Fig. 5.13: Separated signals of frontal region in the case of the non-target stimuli.

さらに、これらの特徴的なピークが同時刻に出現する頻度を Table 5.1 に示す。この結果から、100[msec] のピークは、標的刺激と標準刺激の両方で高い出現頻度を示していることから、閃光刺激に対する応答であると考えられる。700[msec] での特徴的な

Table 5.1: Frequency of characteristic peaks [%] from separated signals of frontal region.

characteristic peaks		target	non-target
100	[msec]	73.3	71.5
700	[msec]	93.3	0.0
800	[msec]	10.0	71.5
1600	[msec]	16.0	82.1
1800	[msec]	80.0	63.3

ピークは、標的刺激を呈示した場合には93[%]以上出現しており、標準刺激を呈示した場合には全く出現していないことが分かる。このことから、オドボール課題実験の課題に対する思考に関連した信号ではないかと推察できる。また、800, 1600[msec]のピークは、標準刺激を呈示した場合に高い頻度で出現しており、課題のカウントが無かったことに関連した思考であると考えられる。最後の1800[msec]でのピークは、標的刺激と標準刺激の両方の出現頻度が高いことから、次の刺激に対する準備のための思考による信号と考えられる。以上のことから、前頭部で観測された脳波を用いた場合、そのピークの出現時刻に一貫した規則性があり、この規則に基づいて脳波の変動を識別することができることが分かる。

次に、後頭部に配置した T_3 , T_4 , P_3 , P_z , P_4 , O_z を用いて実験した。後頭部には視覚野が存在しているため、後頭部の観測脳波には視覚刺激に関連した特徴が得られやすいと考えられる。

観測信号を分離した結果、標的刺激を呈示した場合にはFig. 5.14のように、標準刺激を呈示した場合にはFig. 5.15のように分離信号が得られた。これらの波形から、推定信号の波形には、矢印で示す時刻に特徴的なピークが現れていることが確認できる。すなわち、Fig. 5.14の標的刺激では、 $y_1(t)$ の400[msec]、 $y_4(t)$ の600[msec]である。同様に、Fig. 5.15の標準刺激では、 $y_6(t)$ の800[msec]、 $y_1(t)$ の1000[msec]、 $y_3(t)$ の1300[msec]に特徴的なピークが確認できる。

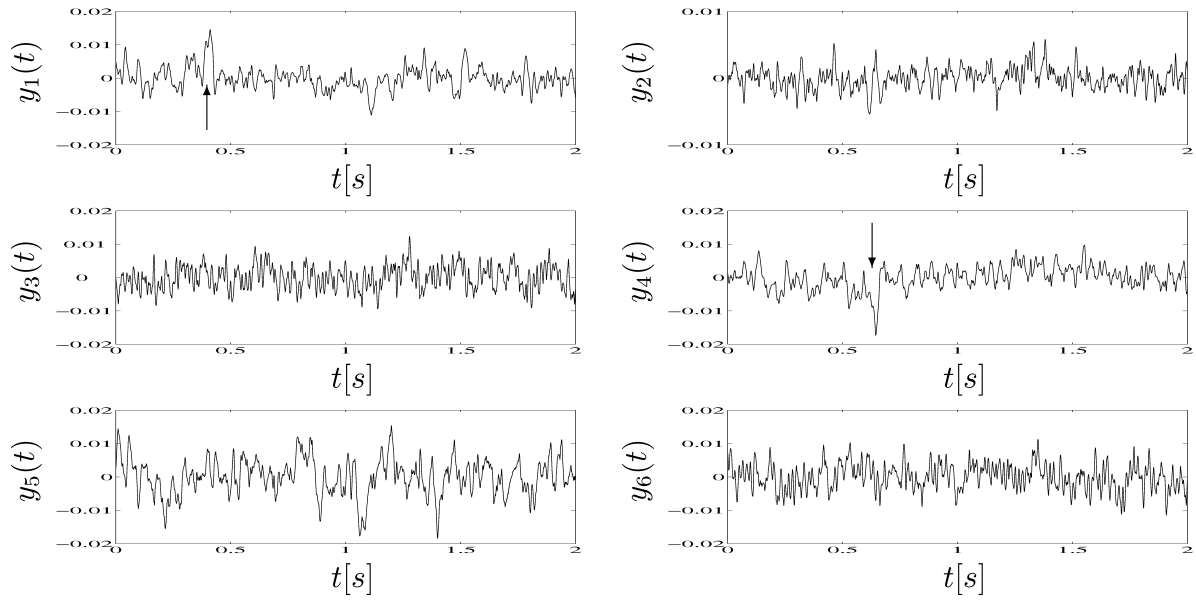


Fig. 5.14: Separated signals of occipital region in the case of the target stimuli.

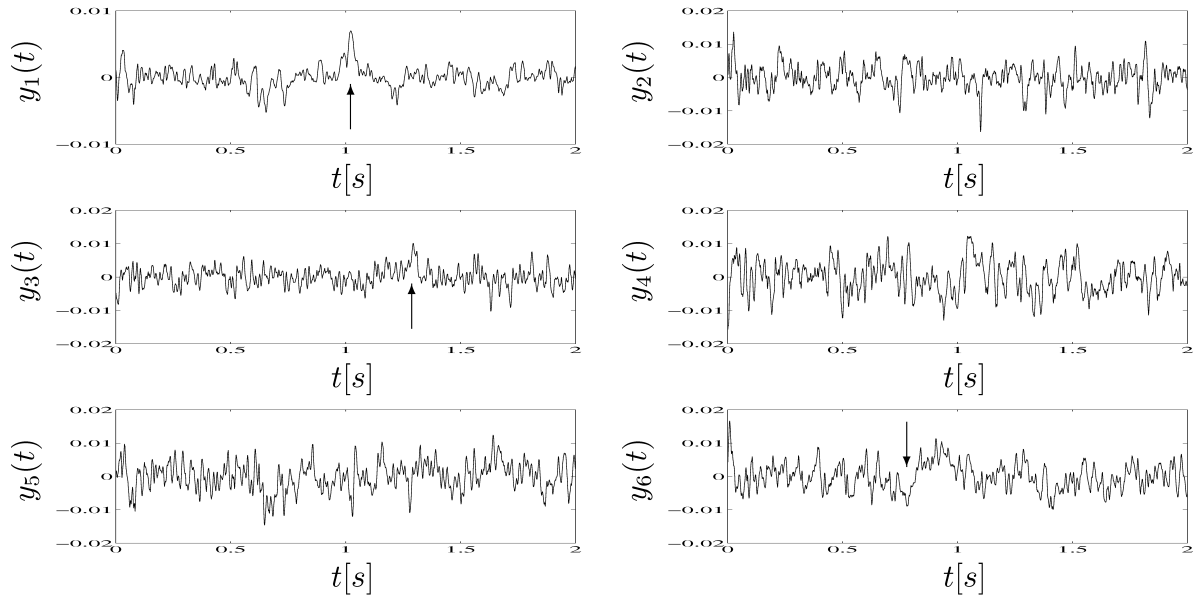


Fig. 5.15: Separated signals of occipital region in the case of the non-target stimuli.

さらに、これらの特徴的なピークが同時刻に出現する頻度を Table 5.2 に示す。この結果から、400, 600[msec] の初期に存在するピークは標的刺激で高い出現頻度を示していて、1000, 1300[msec] の後期に存在するピークは標準刺激で比較的高い頻度を示

Table 5.2: Frequency of characteristic peaks [%] from separated signals of occipital region.

characteristic peaks		target	non-target
400	[msec]	43.3	3.2
600	[msec]	40.0	3.2
800	[msec]	0.0	15.9
1000	[msec]	0.0	38.1
1300	[msec]	0.0	50.8

している。しかしながら、それらの出現頻度は4割程度で、脳波変動を識別できる特徴とは言い難いことが分かる。このことから、後頭部で観測された脳波を用いた場合、そのピークの出現時刻に規則性を見出すことが困難であることが分かる。

以上の実験結果から、前頭部で観測された混合信号を用いて独立成分分析による推定信号から、700[msec]のピークの存在の有無による脳波変動識別を提案できる。このとき、この特徴的なピークの形状は、ほぼ同一の形状となるため、単純なピークの出現時刻だけでなく、パターンマッチングを用いた識別が可能である。さらに、800, 1600[msec]のピークを補助的に用いることで、非常に正確な識別法が提案できる。

5.4 単一試行視覚認識時脳波の変動識別実験

5.4.1 事象関連電位と閃光刺激に関連した成分に基づく脳波変動識別

独立成分分析を適用した後の分離信号波形について、約380[msec]でピークが出現する分離信号が存在する場合を標的刺激呈示後の応答脳波と判断して、存在しない場合を標準刺激呈示後の応答脳波と判断したときの脳波変動識別率をTable 5.3に示す。Table 5.3は男性被験者と女性被験者に対して、それぞれの実験回数ごとの識別率とそれらを平均した識別率を示している。男性被験者に対する脳波変動識別率は平均して95.28[%]、女性被験者に対しては93.12[%]と非常に高い識別率を得られることが確認できる。

Table 5.3: Electroencephalogram fluctuation recognition rates [%] using the feature on the ERP.

	1	2	3	4	5	6	7	8	9	average
male	98.89	80.00	95.45	94.44	97.65	100	97.78	94.44	98.89	95.28
female	71.11	95.56	100	95.56	79.22	98.89	100	97.73	100	93.12

Table 5.4: Electroencephalogram fluctuation recognition rates [%] using the feature on the photic stimuli.

	1	2	3	4	5	6	7	8	9	average
male	97.78	100	92.05	92.22	98.82	100	98.89	100	76.67	95.16
female	98.89	100	76.67	63.64	84.44	90.91	97.78	92.05	91.11	88.39

同様に、第1成分のピークが約180[msec]に出現する場合を標準刺激呈示後の応答脳波と判断して、出現しない場合を標的刺激呈示後の応答脳波と判断したときの変動識別率を Table 5.4 に示す。Table 5.4 から、男性被験者に対して 95.16[%]、女性被験者に対して 88.39[%] と高い識別率が得られることが分かる。

比較のため、自己回帰（AR: Auto-Regressive）モデルに基づく認識法 [90] を適用した結果は、男性被験者に対して 51.37[%]、女性被験者に対して 52.32[%] であった。この結果からも、提案する認識手法の有効性が確認できる。

5.4.2 思考に関連した成分に基づく脳波変動識別

推定した原信号数に基づいて、前頭部での観測信号から独立成分分析によって得られる分離信号波形を用いて脳波変動識別を行った。このとき、標的刺激を呈示したときに現れる 700[msec] と 1800[msec] のピークの形状に着目した。標的刺激が呈示されたときの応答脳波から 5 パターンを用いてテンプレートを作成し、パターンマッチング法によって識別した。識別方法については、いずれかの波形が分離信号内に存在す

Table 5.5: Electroencephalogram fluctuation recognition rates [%] based on the AR model.

	1	2	3	4	5	6	7	8	9	average
male	55.73	52.00	54.13	53.27	50.20	49.32	48.97	49.12	49.62	51.37
female	53.52	51.83	45.93	48.12	67.33	48.12	52.50	53.32	50.23	52.32

Table 5.6: Electroencephalogram fluctuation recognition rates [%] using the feature on the frontal region.

	1	2	3	4	5	6	7	8	9	average
male	100	100	100	98.96	100	98.90	99.25	99.13	96.84	99.23
female	98.92	100	100	100	98.92	98.70	98.11	100	100	99.41

る場合を標的刺激呈示後の応答脳波と判断し、どちらも存在しない場合を標準刺激呈示後の応答脳波と判断した。このときの脳波変動識別率を Table 5.6 に示す。Table 5.3 から、男性被験者に対する脳波変動識別率は平均して 99.23[%]、女性被験者に対しては 99.41[%] と非常に高い識別率を得られることが確認できる。また、このとき標準刺激に対する識別率は全ての実験で 100[%] を示していた。

これらの結果は、Table 5.3 の事象関連電位に基づいた場合や Table 5.4 の閃光刺激に関連した成分に基づく場合と比較しても、高い識別率を示していることが分かる。このことから、視覚性オドボール課題実験において、前頭部で観測された脳波を用いて独立成分分析によって特徴抽出することは、脳波の変動を識別について非常に有効であることが分かった。

以上の実験結果から、独立成分分析を用いて単一試行視覚認識時脳波から事象関連電位や閃光刺激に対する応答、思考に関連する応答に関する特徴の抽出ができて、その特徴から脳波の変動を識別できることが分かる。このとき、提案法は分離信号のピークの出現時刻を調べて、その存在の有無によって脳波の変動が識別できる非常に簡単なアルゴリズムである。このようにして単一試行脳波から抽出した特徴に関しては、被

験者や実験日の変化によって出現する時刻に変動が見られた。しかし、個々の実験においては一貫して出現していて、高い精度で変動を識別できた。したがって、脳波変動識別において非常に有効な情報が独立成分分析によって抽出できたと考えられる。

5.5 結言

本章では、単一試行で得られる視覚認識時脳波から、脳波の変動を検出することを目的として、視覚性オドボール課題実験による応答脳波の特徴抽出を行った。

まず、独立成分分析によって得られる分離信号にフィルタ処理を施して特徴を抽出する方法を述べた。この処理によって、事象関連電位や閃光刺激に応答と考えられる成分が抽出できた。抽出した事象関連電位は標的刺激を呈示したときに高頻度で出現し、標準刺激を呈示したときにはほとんど出現しないことを確認した。さらに、この成分は380[msec]でほぼ同時刻に出現することから、380[msec]でピークが出現する分離信号が存在する場合を標的刺激呈示後の応答脳波と判断して、存在しない場合を標準刺激呈示後の応答脳波と判断する脳波変動識別法を提案し、93[%]以上の識別率を得た。

次に、標準刺激呈示時に高頻度で出現する閃光刺激に関連した特徴に基づいて、180[msec]のピークが出現する分離信号が存在する場合を標準刺激呈示後の応答脳波と判断して、存在しない場合を標的刺激呈示後の応答脳波と判断する脳波変動識別法を提案し、約90[%]の識別率を得た。以上のことから、独立成分分析に基づく脳波の特徴抽出法と変動識別法は非常に有効な手法であることが検証できた。

さらに、分離信号のスケールを回復した推定信号の振幅と非ガウス性に基づいて、原信号数を推定する方法を提案した。この推定信号数を用いて、再度、独立成分分析を行うことで、単純に分離した場合よりも特徴が見出し易いことを確認した。推定原信号数に基づいた推定信号のピークの出現時刻に着目したところ、前頭部で観測された脳波を用いた場合、その推定信号のピークには一貫して同時刻に出現する規則があることを確認した。すなわち、700[msec]に現れるピークは、標的刺激を呈示したときには9割以上の出現頻度を示し、標準刺激を呈示したときには全く出現しないことを確認した。また、800, 1600[msec]に現れるピークは、標的刺激呈示時には1割程度の出現頻度であるが、標準刺激呈示時には約8割の高頻度で出現することが分かった。し

たがって、前頭部で観測された混合信号を利用して独立成分分析による推定信号から、これらのピークに基づきパターンマッチング法によって脳波変動識別を行ったところ、99[%]以上の非常に高い精度で識別できることが確認できた。以上のことから、提案した原信号数の推定法と、その推定信号数を用いた特徴抽出法は非常に有効であることが分かった。

以上の手法によって、単一試行脳波から人間の脳波の変動を観察できるようになるので、人間の状態をリアルタイムに推定できて、今後、多くの臨床への応用が期待される。このとき、本研究において抽出された特徴は、生体学では厳密に明確な意味付けができないかもしれないが、この特徴によって脳波の変動識別が極めて高い精度で可能になったことから、非常に意味のある情報を取り出しているといえる。

第6章 結論

本論文では、原信号と伝達関数が未知の条件下で、観測された混合信号のみを利用して原信号を推定するブラインド信号分離に対して、実環境下で実際に観測された混合信号を用いた解析方法を述べた。

第2章では、まず、ブラインド信号分離問題を解くための代表的な手法について概説し、その中で、独立成分分析は原信号が互いに統計的に独立であるという仮定の下で、混合信号から原信号を推定する統計的手法であり、定常信号だけでなく非定常信号も分離できる特徴を持っていることを述べた。次に、独立成分分析を実行するため、Natural Gradient アルゴリズムと FastICA アルゴリズムについて述べた。Natural Gradient アルゴリズムは、独立性の尺度に Kullback-Leibler 情報量を用いて、それが最小になるように分離する逐次更新アルゴリズムで、分離信号を全て同時に出力する特徴を持っている。また、FastICA アルゴリズムは、非ガウス性の最大化に基づく手法で、その尺度に Negentropy を用いて、Negentropy が最大となるように分離するアルゴリズムであり、分離信号を一つずつ推定するデフレーション法であり、非ガウス性の高い信号から推定されやすい特徴を持っていることを述べた。その後、瞬時混合信号によるシミュレーションによって、Natural Gradient アルゴリズムと FastICA アルゴリズムの分離性能を確認した。シミュレーションでは、スケールの不定性と成分置換の問題が残るが、原信号を正確に復元できることを確認した。さらに、成分置換の問題はアルゴリズムの初期パラメータに依存しているが、FastICA アルゴリズムでは非ガウス性の高い信号から出力される頻度が高いことを確認した。

第3章では、独立成分分析の不定性問題とその解決法について述べた。まず、独立成分分析では原信号が互いに統計的に独立という制約の下で解かれるため、分離信号にはスケールの不定性と成分置換の問題が残ることを述べた。次に、スケールの不定性問題に対して、分割信号を導入した解消法を述べた。このとき、分割信号は原信号と伝達関数の積で表現できることを証明した。さらに、分割信号は一つの実信号が一

つの伝達関数を通過して観測された信号であるという物理的解釈を与え、分割信号のスケールは伝達関数によって決まるためスケールの不定性は自ずと解消できることを示した。さらに、原信号数が混合信号数より少ない場合、分割信号を適用することで簡単に原信号数の推定が可能であることを述べた。その後、瞬時混合での信号分離シミュレーションによって、原信号数と混合信号数が異なる場合の独立成分分析の分離性能を調べた。分離信号のスケールを回復した推定信号は、原信号数が混合信号数より同じか少ないときには原信号に一致することが分かった。さらに、原信号数が混合信号数より少ない場合、全ての原信号数が推定されて、残りの出力信号は振幅が0になることを確認した。このことから、分割信号を適用することで原信号の数が簡単に推定できることを確認した。一方、原信号数が混合信号数より多い場合、正確に原信号を正確に推定できないものの、原信号の特徴は良く現れていることを確認した。

第4章では、実環境下でも有効に機能する雑音除去システムの構築を目的として、周波数領域の独立成分分析で大きな問題である成分置換について、音響伝達特性に基づいた解決法を提案した。まず、周波数領域の独立成分分析を用いた実環境下で観測される音響信号の分離方法を述べた後、非ガウス性の最大化に基づく周波数領域のFastICAアルゴリズムを概説した。また、周波数領域の独立成分分析で得られる分離信号のスペクトルにはスケールと位相の不定性と成分置換の問題が生じ、その値は各周波数で異なるために非常に困難な問題であることを示した。次に、スケールと位相の不定性に対しては、分割スペクトルを導入することで解消できることを述べた。また、分割スペクトルが音源のスペクトルと伝達関数の積で表せることを導出し、この結果から、成分置換の問題は、各音源からマイクロホンまでの距離によって得られる伝達関数から解決できる。そこで、伝達関数のゲインと位相の大小関係に基づく成分置換の修正法を提案した。さらに、実環境下での音響信号のゲイン比と位相差は、音源からマイクロホンまでの距離でほぼ理論通りに決まることと、周波数帯域やマイクロホン間隔、音源とマイクロホンの距離に依存していることを明確にした。そして、この音響伝達特性に基づいて、伝達関数のゲインと位相の情報を組み合わせて利用する統合的な成分置換の修正法を提案した。最後に、音源分離後の目的音声選択方法については、音源とマイクロホンの相対的な位置情報が得られている場合には、提案する成分置換の修正法を用いることで目的音声を選択できることを述べた。また、この位置情報が得

られない場合、周波数領域の FastICA アルゴリズムの特徴を利用して、簡単に話者音声を選択する方法を提案した。実環境下での検証実験では、提案した成分置換修正法は音源とマイクロホンの距離や周波数帯域に依存せず、高い修正能力を持つことを確認した。また、騒音環境下で話者音声が存在する場合だけでなく、他人の発話が雑音となる場合においても有効に機能することが分かった。そして、多音源多マイクロホン環境下でも、提案法は成分置換の修正ができることを確認した。さらに、提案した目的音声選択法は、正確に話者音声を選択できることを確認した。

第5章では、単一試行で得られる視覚認識時脳波から、脳波の変動を検出することを目的として、視覚性オドボール課題実験による応答脳波の特徴抽出を行った。まず、独立成分分析によって得られる分離信号にフィルタ処理を行うことで、簡単に事象関連電位や閃光刺激に応答した成分が抽出できることを示した。次に、推定信号の振幅と非ガウス性に基づいて原信号数を推定した。この推定信号数を用いて、再度、独立成分分析を行うことで、単純に分離した場合よりも特徴が見出し易いことを確認した。さらに、前頭部で観測された脳波から思考に関連した成分が抽出できた。これらの特徴的なピークの出現時刻に着目して、脳波変動識別を行った結果、高い精度の変動識別結果が得られ、提案した特徴抽出法と脳波変動識別法が、単一試行脳波解析に非常に有効であることを確認した。

以上の研究によって、独立成分分析に基づく実環境下での音源分離に対する成分置換問題の解決法と目的信号の選択法と、実際に頭皮で観測された単一試行視覚認識時脳波解析に対する特徴抽出法と脳波変動識別法の有効性を確認した。

今後、実環境下での音源分離問題に関しては、以下の問題の解決までを含めた雑音除去システムの構築を考えている。まず、実際の環境では、音源数は未知で時々刻々と変動するため、音環境の変動検出と原信号数の推定が必要となる [91], [92]。また、これに伴って、原信号数が未知の場合でも機能する独立成分分析のアルゴリズムについても考慮する必要がある [93]。さらに、残響が強い場合や暗騒音が多い場合でも、原信号を推定できるアルゴリズムについて考えている [94]~[97]。その後、音声認識までを組み込んだシステムの構築について研究する予定である。

単一試行脳波解析に関しては、独立成分分析に基づいて抽出した特徴は被験者や実験日の変化によって出現する時刻に若干の変動が見られる。しかし、個々の実験におい

ては一貫した特徴として出現しているため、個々の実験において脳波の特徴抽出と同時に出現時刻を推定する方法について研究する予定である。さらに、原信号数の推定に関しては、混合信号の分布に基づく原信号数の推定法 [91] や、独立成分分析による分離信号同士の相関を用いた原信号数の推定法 [92] が提案されている。したがって、これらの手法を併用して明確な原信号数を推定した後に特徴抽出する方法も考えられる。

謝辞

本論文は、2004年に九州工業大学大学院情報工学研究科に在籍して、本研究に着手して以来、今日に至るまでの研究をまとめたものである。本論文を構成する研究のほとんどは、多くの方々との共同研究として遂行されたものである。

九州工業大学情報工学部の井上勝裕教授には、このような研究の機会を与えて頂いたことに対して深く感謝する。また、研究における様々な御指導、御討論頂いたことに感謝の言葉を述べる。近畿大学産業理工学部の五反田博教授には、1998年に近畿大学に入学してから現在まで、研究の基礎から論文の書き方までもを含めて、終始懇切丁寧に根気強く、時には厳しく御指導頂いた。ここに、深く感謝する。九州工業大学の熊丸耕介名誉教授には、研究での御指導および御支援頂いたことに謝意を表す。本論文をまとめるにあたり、岡崎悦明教授、江島俊朗教授、尾知博教授、前田誠講師には貴重な御意見や有益な御助言を頂いたことに感謝する。また、近畿大学産業理工学部の久良修郭教授、園田敏勝教授、藤原英徳教授、原谷直実教授、江崎秀教授、白土浩講師には様々な御指導および御支援頂いたことに対し、深く感謝する。研究の初期段階において、信和幸博士には、貴重な御意見や御討論頂いた。本研究の動機付けは、氏との議論によって与えられたといっても過言ではない。ここに、感謝の言葉を述べる。さらに、本研究を一緒に行った古屋武志博士、金田圭市博士、中河史成氏に感謝の意を表す。最後に、著者に対して崇高な教育を施してくれた両親と、著者が論文執筆に専念できるように協力してくれた妻に感謝する。

参考文献

- [1] A. Cichocki and S. Amari, “Adaptive blind signal and image processing: learning algorithm and applications,” John Wiley & Sons, Ltd, 2002.
- [2] A. Hyvärinen, J. Karhunen and E. Oja, “Independent component analysis,” John Wiley & Sons, Ltd, 2001.
- [3] T. W. Lee, “Independent component analysis,” Kluwer Academic Publishers, 1998.
- [4] A. Hyvarinen, “Survey on independent component analysis,” Neural Computing Surveys, Vol. 2, pp. 94-128, 1999.
- [5] 甘利俊一, 狩野裕, 佐藤俊哉, 松山裕, 竹内啓, 石黒真木夫, “多変量解析の展開 隠れた構造と因果を推理する,” 株式会社岩波書店, 2002.
- [6] 甘利俊一, 村田昇, “独立成分分析 - 多変量データ解析の新しい方法,” 株式会社サイエンス社, 2002.
- [7] 村田昇, “入門 独立成分分析,” 東京電機大学出版局, 2004.
- [8] 松岡清利, “独立成分分析の基礎,” システム制御情報学会誌, Vol. 48, No. 10, pp. 395-400, 2004.
- [9] S. Ikeda and N. Murata, “A method of ICA in time-frequency domain,” In Proceedings of the First International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99), pp.365-371, 1999.
- [10] N. Murata, S. Ikeda and A. Ziehe, “An approach to blind source separation based on temporal structure of speech signals,” Neurocomputing, Vol.41, Issue 1-4, pp.1-24, 2001.

-
- [11] K. Nobu, T. Koya, K. Kaneda, N. Haratani and H. Gotanda, "Noise reduction using locational information on target sound source," *Journal of Robotics and Mechatronics*, Vol. 15, No. 1, pp. 15-23, 2003.
 - [12] H. Gotanda, K. Nobu, T. Koya, K. Kaneda, T. Ishibashi and N. Haratani, "Permutation correction and speech extraction based on split spectrum through FastICA," *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, pp. 379-384, 2003.
 - [13] T. Ishibashi, K. Inoue, T. Koya, H. Gotanda and K. Kumamaru, "Blind source separation without permutation and scaling indeterminacy," *SICE Annual Conference 2004*, pp. 2298-2301, 2004.
 - [14] T. Ishibashi, S. Tamatsuka, M. Sugahara, K. Inoue, H. Gotanda and K. Kumamaru, "Separation performance of ICA algorithms dependent on the source number," *Proceedings of the 38th ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, 2006.
 - [15] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis for complex valued signals," *Int. J. of Neural Systems*, Vol.10, No.1, pp.1-8, 2000.
 - [16] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, Vol. 22, pp. 21-34, 1998.
 - [17] 河本満, A. K. Barros, A. Mansour, 松岡清利, 大西昇, "時空間的に混合した非定常信号のブラインド信号分離," *電子情報通信学会論文誌 A*, Vol. J82-A, No. 8, pp. 1320-1328, 1999.
 - [18] K. Matsuoka, Y. Ohba, Y. Toyota and S. Nakashima, "Blind separation for convolutive mixture of many voices," *International Workshop on Acoustic Echo and Noise Control (IWAENC2003)*, pp. 279-282, 2003.
 - [19] F. Asano, S. Ikeda, M. Ogawa, H. Asoh and N. Kitawaki, "A combined approach of array processing and independent component analysis for blind separation of

- acoustic signals,” IEEE transaction on Audio and Speech Processing, Vol.11, No.3, pp.204-215, 2003.
- [20] 浅野太, “反射のある音場における信号分離,” 電子情報通信学会技術研究報告, EA2001-8, pp. 57-64, 2001.
- [21] H. Saruwatari, T. Takatani, H. Yamajo, T. Nishikawa and K. Shikano, “Blind separation and deconvolution for real convolutive mixture of temporally correlated acoustic signals using SIMO-model-based ICA,” 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), pp. 549-554, 2003.
- [22] 牧野昭二, 荒木章子, 向井良, 澤田宏, “畳込み混合のブラインド音源分離,” システム制御情報学会誌, Vol. 48, No. 10, pp. 401-408, 2004.
- [23] 丁数学, 大塚将史, 芹沢正樹, 新妻照夫, 須貝和義, “時間-周波数領域のICAを用いる実環境における音響信号のブラインド分離処理,” 電子情報通信学会技術研究報告, EA2001-1, pp. 1-8, 2001.
- [24] 浅野太, 麻生英樹, “周波数領域のブラインド信号分離における permutation について,” 電子情報通信学会技術研究報告, EA2001-19, pp. 9-16, 2001.
- [25] A. Hyvärinen, “Fast and robust fixed-point algorithms for independent component analysis,” IEEE Transactions on Neural Networks, Vol.10, No.3, pp. 626-634, 1996.
- [26] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” Neural Networks, Vol.13, No.4-5, pp.411-430, 2000.
- [27] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda and F. Itakura, “Evaluation of blind signal separation using directivity pattern under reverberant conditions,” IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2000), pp. 3140-3143, 2000.
- [28] H. Sawada, R. Mukai, S. Araki and S. Makino, “A robust and precise method for solving the permutation problem of frequency-domain blind source separa-

- tion,” 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), pp. 505-510, 2003.
- [29] 金田圭市, 古屋武志, 五反田博, “分割スペクトルのエントロピーに基づく成分置換解消法,” 電子情報通信学会論文誌 A, Vol. J87-A, No. 7, pp. 1065-1069, 2004.
- [30] H. Sawada, S. Araki, R. Mukai and S. Makino, “Blind extraction of a dominant source signal from mixtures of many sources,” IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005), vol. III, pp. 61-64, 2005.
- [31] T. Ishibashi, K. Inoue, H. Gotanda and K. Kumamaru, “A solution of permutation problem inherent in frequency domain ICA,” Proceedings of the 36th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, pp. 259-264, 2004.
- [32] T. Ishibashi, K. Inoue, H. Gotanda and K. Kumamaru, “Studies on permutation problem in FDICA,” SICE Annual Conference 2005, pp. 1754-1759, 2005.
- [33] T. Ishibashi, K. Inoue, H. Gotanda and K. Kumamaru, “Properties of acoustic transfer function in real environments and their application to ICA,” Proceedings of the 36th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, pp. 246-251, 2005.
- [34] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介, “実環境下での伝達特性を利用した周波数領域 ICA の成分置換問題の解決,” システム制御情報学会論文誌, Vol. 19, No. 12, pp. 471-478, 2006.
- [35] S. Ikeda and K. Toyama, “Independent component analysis for noisy data: MEG data analysis,” Neural Networks, Vol. 13, No. 10, pp. 1063-1074, 2000.
- [36] 池田思朗, “独立成分解析の信号処理への応用,” 計測自動制御学会誌, Vol. 38 No. 7, pp. 461-467, 1999.
- [37] 池田思朗, 村田昇, “Independent Component Analysis を用いた MEG データの解析,” 電子情報通信学会技術研究報告, NC98-28, pp. 29-36, 1998.

-
- [38] J. Cao, N. Murata, S. Amari, A. Cichocki, and T. Takeda, "Independent component analysis for unaveraged single-trial MEG data decomposition and single-dipole source localization," *Neurocomputing* Vol. 49, pp. 255-277, 2002.
- [39] 曹建庭, "単試行脳信号の独立成分解析による脳内活動情報可視化の試み," システム制御情報学会誌, Vol. 48, No. 10, pp. 409-414, 2004.
- [40] J. Cao, N. Murata, S. Amari, A. Cichocki, T. Takeda, H. Endo and N. Harada, "Single-trial magnetoencephalographic data decomposition and localization based on independent component analysis approaches," *IEICE Transactions on Fundamentals of Electronics, Communications Computer Sciences*, Vol. E83-A, No. 9, 2000.
- [41] 川勝真喜, 石橋秀則, 安達雅春, 内川義則, 小谷誠, "独立成分分析を用いた脳磁界計測のノイズ軽減に関する研究," 電気学会論文誌 C, Vol. 124-C, No. 9, pp. 1685-1691, 2004.
- [42] 小澤拓哉, 野澤昭雄, 井出英人, "ICA によるアーティファクト成分を含む脳波の成分分離," 電気学会論文誌 C, Vol. 124-C, No. 1, pp. 211-212, 2004.
- [43] スチュアート・ローゼン, ピーター・ハウエル, "音声・聴覚のための信号とシステム," 海文堂出版株式会社, 1998.
- [44] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech", *Proceedings of the IEEE*, Vol. 67, No.12, pp. 1586-1604, 1979.
- [45] J. D. Gibson, B. Koo and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Transactions on Signal Processing*, Vol. 39, No. 8, pp. 1732-1741, 1991.
- [46] Flanagan J. L. Flanagan, D. A. Berkley, G. W. Elko, J. E. West and M. M. Sondhi, "Autodirective microphone systems," *Acustica*, Vol. 73, No. 2, pp. 58-71, 1991.
- [47] 西隆司, 三上淳一, 井上友幸, 古川宣一, 清水寧, 川上福司, "多方向同時集音装置," 電子情報通信学会技術研究報告, EA88-65, pp. 33-39, 1988.

-
- [48] 金森丈郎, 茨木悟, 古川博基, 直野博之, 斎藤浩, 西川清, “2次元ディジタルフィルタを用いた超指向性マイクロホン,” 電子情報通信学会技術研究報告, EA91-84, pp. 31-36, 1992.
- [49] 東貞男, 打越聰, 城戸健一, “交差形円弧状配列マイクロホンシステムの実用化に関する実験的検討,” 日本音響学会誌, Vol. 46, No. 2, pp. 111-120, 1990.
- [50] 金田豊, “アダプティブマイクロホンアレー,” 電子情報通信学会論文誌 BII, Vol. J75-B2, No. 11, pp. 742-748, 1992.
- [51] L. J. Griffiths and C. W. Jim, “An alternative approach to linearly constrained adaptive beamforming,” IEEE Transactions on Antennas and Propagation, Vol. AP-30, No. 1, pp. 27-34, 1982.
- [52] 大賀寿郎, 山崎芳男, 金田豊, “音響システムとディジタル処理,” 社団法人電子情報通信学会, 1995.
- [53] S. F. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, No. 2, pp. 113-120, 1979.
- [54] 北岡教英, 赤堀一郎, 中川聖一, “スペクトルサブトラクションと時間方向スムージングを用いた雑音環境下音声認識,” 電子情報通信学会論文誌 D, Vol. J83-D2, No. 2, pp. 500-508, 2000.
- [55] 水町光徳, 赤木正人, “マイクロホン対を用いたスペクトルサブトラクションによる雑音除去法,” 電子情報通信学会論文誌 A, Vol. J82-A, No. 4, pp. 503-512, 1999.
- [56] 永田仁史, 藤岡豊太, 安部正人, “二つの指向性マイクロホンを用いた目的音検出に関する検討,” 電子情報通信学会論文誌 A, Vol. J83-A, No. 12, pp. 1445-1454, 2000.
- [57] M. Aoki, M. Okamoto, S. Aoki, H. Matsui, T. Sakurai and Y. Kaneda, “Sound source segregation based on estimating incident angle of each frequency component of input signals acquired by multiple microphones,” Acoustical Science and Technology, Vol. 22, No. 2, pp. 149-157, 2001.

-
- [58] 青木真理子, 古屋賢一, “騒音下音声強調における空間情報の利用について,” 電子情報通信学会技術研究報告, SP2002-11, pp. 23-30, 2002.
- [59] 青木真理子, 古屋賢一, 片岡章俊, “残響下における音源分離方式 SAFIA の性能改善,” 電子情報通信学会論文誌 A, Vol. J87-A, No. 9, pp. 1171-1186, 2004.
- [60] 青木真理子, 山口義和, 古家賢一, 片岡章俊, “音源分離方式 SAFIA を用いた高騒音下における近接音源の分離抽出,” 電子情報通信学会論文誌 A, Vol. J88-A, No. 4, pp. 468-479, 2005.
- [61] T. W. Lee, M. Girolami and T. J. Sejnowski, “Independent component analysis using an extended informax algorithm for mixed subgaussian and supergaussian sources,” Neural Computation, Vol.11, pp.417-441, 1999.
- [62] A. J. Bell and T. J. Sejnowski, “An information maximization approach to blind separation and blind deconvolution,” Neural Computation, Vol. 7, No. 6, pp. 1129-1159, 1995.
- [63] S. Amari, “Natural gradient works efficiently in learning,” Neural Computation, Vol.10, pp.251-276, 1998.
- [64] K. Matsuoka, M. Ohya and M. Kawamoto, “A neural net for blind separation of nonstationary signals,” Neural Networks, Vol. 8, No. 3, pp. 411-419, 1995.
- [65] R. Boscolo, H. Pan, V. P. Roychowdhury, “Independent component analysis based on nonparametric density estimation,” IEEE Transactions on Neural Networks, Vol. 15, No. 1, pp. 55-65, 2004.
- [66] M. R. Álvarez, F. Rojas, C. G. Puntonet, J. Ortega, F. Theis and E. W. Lang, “A geometric ICA procedure based on a lattice of the observation space,” 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), pp. 1101-1106, 2003.
- [67] H. Shindo and Y Hirai, “Blind source separation by a geometrical method,” Proceedings of 2002 International Joint Conference on Neural Networks (IJCNN2002), pp.1109-1114, 2002.

-
- [68] S. Haykin, "Neural Networks: A Comprehensive Foundation," Macmillan Publishing Company, 1994.
- [69] M. Jones and R. Sibson, "What is projection pursuit?," Journal of the Royal Statistical Society, Series A, No. 150, pp. 1-36, 1987.
- [70] A. Hyvärinen, "New approximations of differential entropy for independent component analysis and projection pursuit," In Advances in Neural Information Processing Systems, Vol.10, MIT Press, pp.273-279, 1999.
- [71] G. J. Borse, "MATLAB 数值解析," 株式会社オーム社, 1998.
- [72] Acoustical Society of Japan, "ASJ continuous speech corpus japanese newspaper article sentences," JNAS Vols.1-16, 1997.
- [73] NTT Advanced Technology Corporation, "Ambient noise database for telephonometry 1996," 1996.
- [74] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," Proceedings of the IEEE, Vol.77, No.2, pp257-286, 1989.
- [75] 古井貞熙, "音声情報処理," 森北出版株式会社, 1998.
- [76] 古井貞熙, "音響・音声工学," 株式会社近代科学社, 1999.
- [77] 鹿野清宏, 中村哲, 伊勢史郎, "音声・音情報のデジタル信号処理," 株式会社昭晃堂, 1997.
- [78] 鹿野清宏, 伊藤克亘, 河原達也, 武田一哉, 山本幹雄, "音声認識システム," 株式会社オーム社, 2001.
- [79] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介, "周波数領域ICAに基づく多音源下での音源分離," 第1回人工頭脳工学シンポジウム講演論文集, pp. 54-55, 2005.
- [80] 日本音響学会編, "音のなんでも小事典," 株式会社講談社, 1996.
- [81] 古屋武志, 金田圭市, 五反田博, "ブラインド信号分離による雑音除去法のSN比改善量の近似的評価法," 電子情報通信学会論文誌 A, Vol. J87-A, No. 7, pp. 1054-1058, 2004.
- [82] 大熊輝雄, "臨床脳波学," 株式会社医学書院, 1999.

-
- [83] N. K. Squires, K. C. Squires and S. A. Hillyard, "Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man," *Electroencephalography and clinical neurophysiology*, Vol. 38, No. 4, pp. 387-401, 1975.
- [84] E. Courchesne, S. A. Hillyard and R. Galambos, "Stimulus novelty, task relevance and the visual evoked potential in man," *Electroencephalography and clinical neurophysiology*, Vol. 39, No. 2, pp. 131-143, 1975.
- [85] T. W. Picton and S. A. Hillyard, "Human auditory evoked potentials. II. Effects of attention," *Electroencephalography and clinical neurophysiology*, Vol. 36, No. 2, pp. 191-199, 1974.
- [86] C. C. Dunean-Johnson and E. Donchin, "On quantifying surprise: the variation of event-related potentials with subjective probability," *Psychophysiology*, Vol. 14, No. 5, pp. 456-467, 1977.
- [87] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介, 松岡成明, "独立成分分析に基づく視覚認識時脳波の変動識別," 第2回人工頭脳工学シンポジウム講演論文集, pp. 46-47, 2006.
- [88] T. Ishibashi, M. Sugahara, S. Tamatsuka, K. Inoue, H. Gotanda and K. Kumamaru, "Estimation of the number of unknown source signals and its application to EEG analysis," *Proceedings of the 38th ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, 2006.
- [89] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介, 松岡成明, "独立成分分析に基づく視覚刺激応答脳波の解析," 第3回人工頭脳工学シンポジウム講演論文集, 2007.
- [90] K. Inoue, K. Kumamaru, R. Yano, A. Hayashi, A. Yamauchi and S. Matsuoka, "Feature Extraction Method of EEG Waves under Visual Recognition Task," *SICE Annual Conference 2002*, pp. 892-894, 2002.
- [91] T. Ishibashi, K. Inoue, H. Gotanda and K. Kumamaru, "Studies on estimation of the sources number in blind source separation problems," *SICE-ICASE International Joint Conference 2006*, pp. 5169-5174, 2006.

- [92] H. Sawada, R. Mukai, S. Araki and S. Makino, “Estimating the number of sources using independent component analysis,” *Acoustical Science and Technology, the Acoustical Society of Japan*, Vol. 26, No. 5, 2005.
- [93] S. Amari, T. P. Chen and A. Cichocki, “Nonholonomic orthogonal learning algorithms for blind source separation,” *Neural Computation*, Vol. 12, No. 6, pp. 1463-1484, 2000.
- [94] C. Servière, “Separation of speech signals with segmentation of the impulse responses under reverberant conditions,” *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, pp. 511-516, 2003.
- [95] 古屋武志, 金田圭市, 五反田博, “独立成分分析に基づく耐高残響音源分離に関する研究,” *電子情報通信学会技術研究報告*, NC2005-21, pp. 7-12, 2005.
- [96] H. Attias, “Independent factor analysis,” *Neural Computation*, Vol. 11, No. 4, pp. 803-851, 1999.
- [97] A. Hyvärinen, “Independent component analysis in the presence of gaussian noise by maximizing joint likelihood,” *Neurocomputing*, Vol. 22, pp. 49-67, 1998.