

Title	経験的モード分解を用いた音信号分析に関する基礎研究
Author(s)	澤口, 知希
Citation	
Issue Date	2010-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8946
Rights	
Description	Supervisor: 鶴木祐史, 情報科学研究科, 修士

修 士 論 文

経験的モード分解を用いた
音信号分析に関する基礎研究

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

澤口 知希

2010 年 3 月

修 士 論 文

経験的モード分解を用いた 音信号分析に関する基礎研究

指導教官 鵜木 祐史 准教授

審査委員主査 鵜木 祐史 准教授
審査委員 赤木 正人 教授
審査委員 徳田 功 准教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

0810030 澤口 知希

提出年月: 2010 年 2 月

概要

短時間フーリエ変換やウェーブレット変換は、信号音の時間-周波数分析法として非常によく利用される手法である。これらは、被解析信号を定常信号と仮定することで、時間-周波数領域において信号の変化を分析することができる。しかし、音声信号などの現実的な信号は非定常信号である。従って、定常性を仮定しなければならないこれらの手法では、信号の瞬時振幅・瞬時位相の非定常な変化を正確に分析することができない。

近年、非定常信号を分析する手法として、経験的モード分解 (EMD) が利用されている。これは、Huang らによって提案されたものである。最近、音声信号処理の分野で利用されはじめている。例えば、EMD を利用した雑音除去法が提案されている。Taufiq は、雑音除去後に残された音声信号上のミュージカルノイズを除去するために、Molla & Hirose は、雑音にロバストな音声信号の有声無声判別をするために、EMD を利用した雑音除去法が提案されている。

音声信号は一般に非定常信号であるため、EMD を利用した音声信号の情報表現は、従来の分析法で得られたものと比較して、正確に音声の非定常な変化を表現しているが、音声と雑音の混合信号 (雑音音声) が EMD によってどのように表現されるか不明である。Taufiq や Molla & Hirose の雑音除去法では、雑音のエネルギー - 分布だけに着目して、EMD で表現された成分上で雑音除去を行っているため、音声信号の重要な情報 (非定常な性質) まで除去している可能性も否定できない。

本研究では、EMD による信号分析の特質と、固有モード関数 (IMF) の特徴を調査し、EMD を用いた信号表現について明らかにすることを目的とする。この分析の特徴を基に、EMD の信号分析の特徴を最大限に活かした分析法を検討する。

EMD の信号分解のプロセスから被解析信号をどのように分解するか、また、IMF の制約条件よりどのような特性をもつ信号であるかを調査する。

EMD の分解過程より、EMD に基づく音分析の本質は、共通なエンベロープ分解に基づいた信号表現を行う方法であることが分かった。この本質に基づく、振幅包絡が一定な定常信号と、ある区間にのみ振幅包絡を持つ非定常信号とに分けられると考えることができる。EMD の分析を最大限活かせる音信号処理として、雑音除去法を提案した。振幅包絡が一定な定常雑音と、ある区間にのみ振幅包絡を持つ非定常な音声信号とに分ける方法について検討し、提案手法の有効性を検証するために、雑音除去の評価実験を行った。付加する雑音は、SNR を変化させた合計 6 種類の白色雑音とした。用いる評価尺度は、SNR の改善量、LSD の改善量、PESQ の 3 つの評価尺度を用いた。結果、SNR、LSD 共に改善された。PESQ については、同程度の音声品質となった。

EMD の特質を活かした分析法として、本研究では雑音除去法を提案した。その結果、雑音音声を共通なエンベロープ分解に基づいた信号表現を行うことで、エンベロープが一定な雑音と、音声区間にのみ振幅包絡を持つ非定常な音声とに分解することを実現した。

目次

第1章	序論	1
1.1	はじめに	1
1.2	研究の背景	1
1.3	本研究で取り組む問題	4
1.4	本研究の目的	4
1.5	本論文の構成	4
1.6	記号の定義	5
第2章	経験的モード分解 (EMD)	8
2.1	EMD とは	8
2.2	固有モード関数 (Intrinsic Mode Function: IMF)	8
2.3	EMD の信号分解のプロセス	9
2.3.1	Hilbert 変換	12
2.3.2	Hilbert–Huang スペクトラムの解析例	14
2.4	EMD を用いた音信号処理	16
2.4.1	Taufiq らによって提案された雑音除去法	16
2.4.2	Molla & Hirose によって提案された雑音除去法	17
2.4.3	Matsuda らによって提案された基本周波数推定	24
第3章	経験的モード分解に基づく音声信号の分析	25
3.1	EMD の特徴	25
3.1.1	分析合成系の精度	25
3.1.2	信号分解の特徴	26
3.2	EMD を用いた信号の分析例	27
3.3	EMD を用いた信号分析の本質	29
3.4	EMD を用いた音信号処理の可能性について	29
第4章	経験的モード分解を用いた雑音除去法の提案	30
4.1	EMD を用いた雑音除去法	30
4.2	音声と雑音の切り分け方法	32
4.3	定帯域通過フィルタ	32
4.4	音声と雑音の変調度	34

第5章	評価シミュレーション	36
5.1	シミュレーションの条件	36
5.2	結果	38
5.3	一音声サンプルにおける分析結果	42
5.3.1	分析例 1	42
5.3.2	分析例 2	42
5.4	考察	46
第6章	結論	47
6.1	本研究で明らかになったこと	47
6.2	残された課題	48
6.3	今後の展望	48

目 次

2.1	EMD のアルゴリズムの PAD	9
2.2	IMF の制約条件を満たした信号の例	10
2.3	信号の分解過程 ($x(t)$ に対してエンベロープの平均値を求めたもの (上図), $x(t) - a_1(t)$ に対してエンベロープの平均値を求めたもの (下図))	11
2.4	固有モード関数 (IMF)	12
2.5	$x(t)$ の信号波形	14
2.6	Hilbert–Huang スペクトラム	15
2.7	(a) ミュージカルノイズと (b) 各 IMF におけるミュージカルノイズのエネルギー分布 (文献 [22] より引用)	16
2.8	短時間における音声信号 $x(t)$ とその振幅スペクトル $ X(\omega) $	18
2.9	短時間における雑音音声 $y(t)$ とその振幅スペクトル $ Y(\omega) $	19
2.10	音声信号 $x(t)$, 白色雑音 $n(t)$ とその合成信号 $y(t)$	20
2.11	雑音音声 $y(t)$ を EMD により分解した結果 (6 個の IMF)	21
2.12	IMF の瞬時周波数 (6 個の IMF)	22
2.13	Molla & Hirose の雑音除去法による回復信号 (上図) と回復信号の振幅スペクトル (下図)	23
3.1	非定常信号 $x(t)$ と定常信号 $n(t)$ の和で構成される合成信号 $y(t)$	27
3.2	合成信号 $y(t)$ の分解結果 (4 個の IMF $c_k(t)$)	28
3.3	$x(t)$ と $n(t)$ の再合成信号	28
4.1	音声信号 $x(t)$ と $\text{SNR} = 0$ [dB] のガウス雑音 $n(t)$ の和で構成される雑音音声 $y(t)$	30
4.2	雑音音声 $y(t)$ の分解結果 (最初の 6 個の IMF)	31
4.3	提案法の処理体系 (上図) と IMF のチャンネル選択の処理体系 (下図)	33
4.4	音声 IMF と雑音 IMF の変調周波数特性	35
5.1	雑音除去の結果: SNR の改善量	39
5.2	雑音除去の結果: LSD の改善量	40
5.3	雑音除去の結果: PESQ	41
5.4	回復信号 $\hat{x}(t)$	43
5.5	各信号の短時間スペクトル	43

5.6	音声信号 $x_2(t)$ と $\text{SNR} = \infty$ [dB] のガウス雑音 $n_2(t)$ の和で構成される雑音 音声 $y_2(t)$	44
5.7	雑音音声 $y_2(t)$ の分解結果 (最初の 6 個の IMF)	44
5.8	回復信号 $\hat{x}_2(t)$	45
5.9	各信号の短時間スペクトル	45

表 目 次

1.1	記号の定義	6
2.1	振幅周波数変調の各パラメータ	14
5.1	シミュレーションに使用した音声信号	37

第1章 序論

1.1 はじめに

我々の5感（視覚，聴覚，触覚，味覚，嗅覚）の一つである聴覚は，外界で発生している音を感知するための機能として知られている．我々は，様々な音が混在する環境で生活しており，外界で発生している音を取り入れ，音の大きさや高さ，音色，到来方向を分析することで，判断を行い，身の危険を守ってきた．また，我々は，複数ある音の中から聞きたい音（目的音）だけを取り出す優れた能力をもっている．もし，この優れた聴覚の機能を計算機上で模擬することが可能であれば，複数の音で構成された音から目的音を分離抽出することや，言語情報を機械が認識をする音声認識への応用などが考えられる．

ヒトの聴覚機能を計算機上で実現させるためには，音声信号などの非定常な信号を計算機上でどのように表現することが適切であるか議論をしなければいけない．音信号の性質を表現する一般的な方法として，時間解析や周波数解析がある．時間解析は，振幅強度の時間的变化を表す方法であり，周波数解析は，信号に含まれている周波数成分を表す方法である．この二つの方法を用いることで定常な信号を正確に表現することは可能であるが，現実的な信号は一般的に非定常な信号であり，信号に含まれる周波数成分が時間的に変化するために，これらの方法では，正確に信号を表現することが出来ない．そこで，周波数の時間的な変化を表現する方法として，時間周波数解析がある．これにより，信号のエネルギーを時間と周波数で表現することが可能となった．しかし，一般的に用いられる方法は，信号の微小な時間を定常な信号と仮定することで，この表現を実現しているため，信号を厳密に表現できていないわけではない．従って，信号の定常性を仮定しない分析手法について調査を行い，非定常信号を適切に表現する方法について検討をする．

本研究では，人間の優れた音の分析能力を計算機上で実現するために，音信号を計算機上でどのように表現することが適切であるか検討を行う．それにより，精度の高い音分析の手法を実現することを狙いとする．

1.2 研究の背景

短時間フーリエ変換（Short Time Fourier Transform: STFT）やウェーブレット変換（Wavelet Transform: WT）[1] は，信号音の時間-周波数分析法として非常によく利用される手法である．これらは，被解析信号を定常信号と仮定することで，時間-周波数領域において信

号の変化を分析することができる．これらの手法は，脳波や地震信号，音信号等の時系列信号の解析，画像処理等に広く用いられている．

以下に、Fourier 変換対 [2] を示す．

$$X(\omega) = \frac{1}{\sqrt{2\pi}} \int x(t) e^{-j\omega t} dt \quad (1.1)$$

$$x(t) = \frac{1}{\sqrt{2\pi}} \int X(\omega) e^{j\omega t} d\omega \quad (1.2)$$

原信号を $x(t)$ ，スペクトルを $X(\omega)$ とする．このとき，基底関数は，複素正弦波となることが分かる．また，Fourier 変換を適用するためには，上で述べたように，被解析信号は定常信号でなければならない．STFT は，被解析信号の微小な時間を定常信号と仮定した上で，Fourier 変換を行い時間周波数解析を行う手法である．

以下に，ウェーブレット変換対 [3] を示す．

$$W(a, b) = \frac{1}{\sqrt{a}} \int x(t) \Psi^*\left(\frac{t-b}{a}\right) dt \quad (1.3)$$

$$x(t) = \frac{1}{D_\Psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(a, b) \Psi\left(\frac{t-b}{a}\right) \frac{da db}{a^2} \quad (1.4)$$

ここで， $x(t)$ は被解析信号， $W(a, b)$ はウェーブレット係数， a はスケールパラメータで， b はシフトパラメータとする． Ψ は，アナライジングウェーブレット（ウェーブレット変換における基底関数）とし， Ψ^* は Ψ の複素共役とする．以下に Ψ の式を示す．

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (1.5)$$

ただし，ウェーブレット逆変換の条件として，下記の二つを満たさなければならない．

- 直流成分を持たないこと

$$\int_{-\infty}^{\infty} \Psi(x) dx = 0 \quad (1.6)$$

- 逆変換が存在すること

$$\int_{-\infty}^{\infty} \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega = \frac{1}{2} D_\Psi < \infty \quad (1.7)$$

ここで， $\hat{\Psi}(\omega)$ は $\Psi(\omega)$ の Fourier 変換である．

ウェーブレット変換は，ウェーブレット関数とよばれる基底関数を用いた分析手法である．この手法もまた，Fourier 変換と同様に線形性を仮定した上で，分析をしなければいけない．

しかし、現実的な信号（例えば、脳波や地震波、音声信号など）は非定常信号である。従って、非定常信号を STFT 等で分析することは、分析窓幅を周期とした信号の分析を行うこととなり、窓の起点と終点を結ぶ定常信号と仮定することになる。それゆえに、これらの手法では、信号（たとえばインパルス信号など）の瞬時振幅・瞬時位相の非定常な変化を正確に分析することができない。

近年、非定常信号を分析する手法として、経験的モード分解（Empirical mode decomposition: EMD）[4] が利用されている。これは、Huang らによって提案されたものである。主に脳波解析 [5, 6] や心電図（ECG）波形解析 [7]、地震の反射波解析 [8]、天文学 [9]、画像工学 [10, 11, 12, 13, 14]、画像認識 [15]、コンクリート工学 [16]、信号処理 [17, 18]、金融時系列解析 [19]、情報ハイディング [20]、手ぶれ検出 [21] などの研究分野で利用されている。これらの分析はどれも信号の非定常な変化をみるために用いられている。文献 [5] では、様々な脳波信号が混在する中から α 波を抽出し、脳死状態と昏睡状態を判定する方法を提案している。文献 [8] では、EMD を用いた衝撃弾性波法のテストのための信号処理法を提案している。観測信号を適応的に分析し、雑音、弾性波、表面波に分離を行う方法を提案している。文献 [10] では、EMD を用いた指動作映像の認識を行う方法を提案し、位相の同期を保証した分析を実現している。文献 [11] では、EMD を用いた電子透かし法を提案し、一般的に用いられる画像処理歪にロバストな方法を実現した。文献 [17] では、雑音除去法が Tsung らによって提案されている。これは、EMD の適用的なフィルタリング処理の特徴を活かした雑音除去法となっている。

最近、音声信号処理の分野でも利用されはじめており、例えば、EMD を利用した雑音除去法が提案されている [22, 23]。文献 [22] では、雑音除去後に残された音声信号上のミュージカルノイズを除去するために、文献 [23] では、雑音にロバストな音声信号の有声無声判別をするために、EMD を利用した雑音除去法が提案されている。また、基本周波数推定 [24, 25] が提案されている。文献 [24] は、雑音にロバストな基本周波数推定を行うために EMD が利用されている。文献 [25] は、病的音声のための頑健な基本周波数抽出法を提案した。この他にも、Navin & John による音声強調 [26] や、Molla らによる音源分離 [27]、音声信号の分析 [28]、音源到来方向推定法 [29] などがある。

音声信号は一般に非定常信号であるため、EMD を利用した音声信号の情報表現は、従来の分析法で得られたものと比較して、正確に音声の重要な性質（非定常な変化）を表現している。しかし、音声と雑音の混合信号（雑音音声）が EMD によってどのように表現されるか、またその表現上でそれぞれがどのように分離可能な状態にあるのか、不明である。しかし、先の研究 [22, 23] では、雑音のエネルギー - 分布だけに着目して、EMD で表現された成分上で雑音除去を行っているため、音声信号の重要な情報（非定常な性質）まで除去している可能性も否定できない。また、文献 [24] においても、信号のエネルギーが特定の帯域ごとに集中する遅れ時間空間にのみ EMD を適用させられる基本周波数推定法で、EMD を利用した音声信号の情報表現を明らかにすることで、この手法は改善の余地を残している。

1.3 本研究で取り組む問題

EMD は、信号の非定常な変化を正確に表現することが可能ではあるが、入力信号に依存した分析手法（被解析信号を適応的に分解する手法）であるために、被解析信号がどのように表現されるかが分からない。この信号表現について議論を行った上で、EMD の分析の特性を活かした分析法について検討する必要がある。本研究では、EMD を用いた信号表現を問題として捉え、音信号における EMD の信号分析の特質と、分解された信号である固有モード関数の特性について検討する。

1.4 本研究の目的

本研究の目的は、EMD による信号分析の特徴を調査し、EMD によって分解された固有モード関数の信号表現を検討することである。この分析の特徴を基に、音信号に適用させ、その分析の特徴について明らかにする。

波及効果として、本研究が明らかにされることで、音信号を分析する基礎技術として確立することが期待できる。これにより、音声信号、音楽信号の非定常性を考慮した分析（特に、これまで音信号処理に用いられてきた線形処理では、見ることのできなかつた非定常性の分析）が可能になると考えられる。例えば、非定常な信号である音声信号（特に、基本周波数の推定やフォルマントの推定）を高精度に（非定常な変化を厳密に）分析することが期待できる。また、他の分野においても、本研究で明らかにした EMD の分析の特徴を活かすことで、非定常信号の厳密な分析をするための基礎技術として、確立することが期待できる。

1.5 本論文の構成

本論文は 6 章で構成される。

第 1 章では、本研究で対象とする研究分野の背景、問題点を述べる。そして、本研究で取り扱う問題とその目的を述べる。音を分析する上で重要な性質である非定常な変化を分析するのに有効である経験的モード分解に着目し、この手法を用いた音信号分析の特徴について調査したこと、明らかにしたことを以後の章で述べる。

第 2 章では、Huang らによって提案された経験的モード分解の理論を示す。ここでは、信号分解の過程を示し、実際に、作成した信号を分解することで、その振る舞いについて紹介する。また、すでに提案されている EMD を用いた音信号の分析をここで紹介する。

第 3 章では、2 章で示した理論を基に、経験的モード分解を用いた分析の特徴について示す。経験的モード分解の信号を分解する過程から、どのような信号表現を可能にするか検討をする。この検討結果を基に、作成した信号を分解し、どのように信号表現するか、その振る舞いについて紹介する。

第4章では、3章に示した分析の特徴から、どのような分析に適しているのかを考え、この分析手法を活かすことのできる音信号分析を考える。ここでは、経験的モード分解を活かすことのできるアプリケーションとして、雑音除去を取り上げ、これについて解く。音声信号と雑音信号を EMD の分析の特徴を活かし、分離を試みる。その分離の方法について述べる。

第5章では、提案法が有効であるか評価を行う。用いる雑音の定義を行い、その雑音を音声信号に付加（音声対雑音比 SNR を調整）して、その信号を提案法により音声を抽出し、どれだけ改善することが出来たか評価を行う。その後で、EMD を用いた雑音除去について考察を行う。

6章では、本研究で明らかにしたこと、本研究の今後の課題と展望について述べる。

1.6 記号の定義

ここでは、本論文で用いる記号を定義し、それを表 1.1 に示す。

表 1.1: 記号の定義

記号	定義
t	時刻
$x(t)$	原信号
ω	角周波数
$X(\omega)$	$x(t)$ のスペクトル
a	スケールパラメータ
b	シフトパラメータ
$W(a, b)$	ウェーブレット係数
Ψ	アナライジングウェーブレット
Ψ^*	アナライジングウェーブレットの複素共役をとったもの
$\hat{\Psi}$	Ψ のフーリエ変換
D_Ψ	許容条件の有限性を表す尺度
$c(t)$	固有モード関数
k	固有モード関数の番号
K	固有モード関数のチャンネル数
$r(t)$	残差
$h(t)$	固有モード関数の候補関数
$u(t)$	上側のエンベロープ
$l(t)$	下側のエンベロープ
$a(t)$	上側と下側のエンベロープの平均値
P	コーシーの主値
t'	ある時刻
$z(t)$	複素信号
$y(t)$	$z(t)$ の虚数成分
$\theta(t)$	瞬時位相
$\alpha(t)$	瞬時振幅
$H(t, \omega)$	Hilbert–Huang スペクトラム
A	振幅
f_c	中心周波数
m_{AM}	AM における振幅変調度
f_{AM}	AM における変調周波数
m_{FM}	FM における変調度
f_{FM}	FM における変調周波数

記号	定義
ω_c	中心角周波数
ω_f	周波数変調における角周波数
ω_a	振幅変調における各周波数
$\lambda_{d,i}$	i 個目における固有モード関数の分散
σ_d^2	ミュージカルノイズの分散
$x(t)$	音声信号 (2.4.2 節以降)
$n(t)$	雑音信号
$y(t)$	雑音音声 (2.4.2 節以降)
$Y(\omega)$	雑音音声のスペクトル (2.4.2 節以降)
$IF_k(t)$	瞬時周波数特性
$\hat{x}(t)$	回復音声
$\hat{X}(\omega)$	回復音声のスペクトル
$ X(\omega) $	音声信号の振幅スペクトル
$ Y(\omega) $	雑音音声の振幅スペクトル
$R_{yy}(\tau)$	自己相関関数
T	総時間
SNR_{c+r}	信号対雑音比 (固有モード関数と残差全てを含む)
SNR_c	信号対雑音比 (固有モード関数のみに対して)
LSD_{c+r}	対数スペクトル距離 (固有モード関数と残差全てを含む)
LSD_c	対数スペクトル距離 (固有モード関数のみに対して)
$e_{k,m}(t)$	固有モード関数の変調周波数特性
M	定帯域通過フィルタのチャンネル数
$M_{k,m}$	変調度
I_{SNR}	SNR の改善度
I_{LSD}	LSD の改善度

第2章 経験的モード分解 (EMD)

2.1 EMD とは

非定常信号を分析する手法として、経験的モード分解 (Empirical mode decomposition: EMD) [4] が Huang らによって提案された。EMD は、次式に示すように、信号 $x(t)$ を固有モード関数 (Intrinsic Mode Function: IMF) の和と残差 $r(t)$ に分解する手法である。

$$x(t) = \sum_{k=1}^K c_k(t) + r(t) \quad (2.1)$$

ただし、 K は IMF のチャンネル数を表す。

IMF を Hilbert 変換することで時間-周波数空間の信号として表現が可能 [30, 31] である。Hilbert 変換をすべての IMF に対して適用すると、Hilbert-Huang スペクトラムが得られる。これは、STFT などのように、ある区間を定常信号と仮定することなく、非定常信号を時間周波数空間に変換することが可能である。

2.2 固有モード関数 (Intrinsic Mode Function: IMF)

EMD によって分解される IMF は次に示す二つの制約条件を満たさなければならない。

1. 信号の極値の数と信号のゼロ交差数は同じか一つ違いであること。
2. 信号の任意の点において、極大点と極小点から作られるエンベロープ (上側エンベロープと下側エンベロープ) の平均値がゼロであること。

ここで、チャンネル数 K は、被解析信号の性質に依存して決まるため、すべての信号に対して同じ IMF の個数をもつことはない。

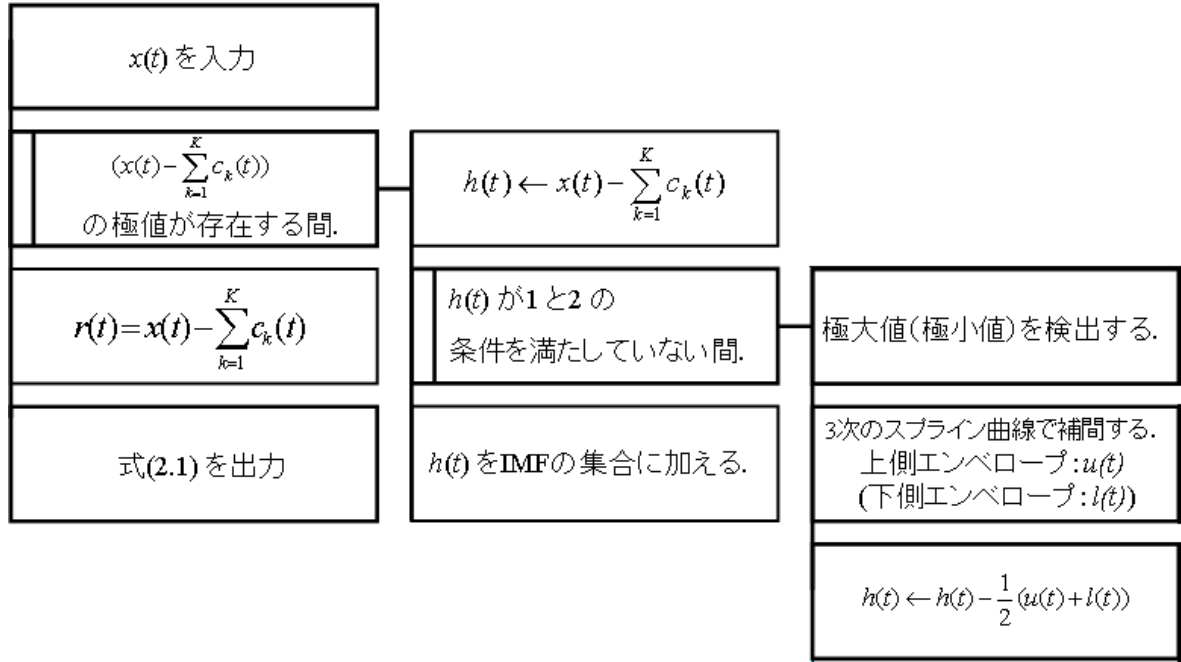


図 2.1: EMD のアルゴリズムの PAD

2.3 EMD の信号分解のプロセス

EMD の分解手順 [4, 33] について以下に示す．図 2.1 に EMD のアルゴリズム (PAD) を示す．EMD の信号分解の手順は，まず， $x(t)$ に極値があるかを確認し，極値がある場合，原信号 $x(t)$ を $h(t)$ とおき， $h(t)$ が二つの制約条件を満たしているか確認する．満たしていなければ， $h(t)$ の極大点（極小点）を 3 次スプライン曲線で補間を行い，上側のエンベロープを $u(t)$ ，下側のエンベロープを $l(t)$ を求める．上側エンベロープと下側エンベロープから平均値 $a(t)$ を求める．その式を以下に示す．

$$a(t) = \frac{1}{2}(u(t) + l(t)) \quad (2.2)$$

エンベロープの平均値がゼロになるまで， $h(t)$ からエンベロープの平均値を減算する．IMF の制約条件が成立するまで，これらの減算を続けることで原信号を IMF に分解する．IMF を得たら，原信号から IMF を減算し，これを次の $h(t)$ とする．その式を以下に示す．

$$h(t) = x(t) - \sum_{k=1}^K c_k(t) \quad (2.3)$$

その信号に対して同様の操作を行い，極値がなくなるまで IMF に分解し続ける．最後に，原信号から全ての IMF を減算した信号が残差 $r(t)$ となる．

図 2.2 に上記の二つの制約条件を満たした信号 (IMF) の例を示す．この信号は制約条件通りに，極値とゼロ交差を交互に繰り返す信号であり，且つ，極大点同士から作られる上

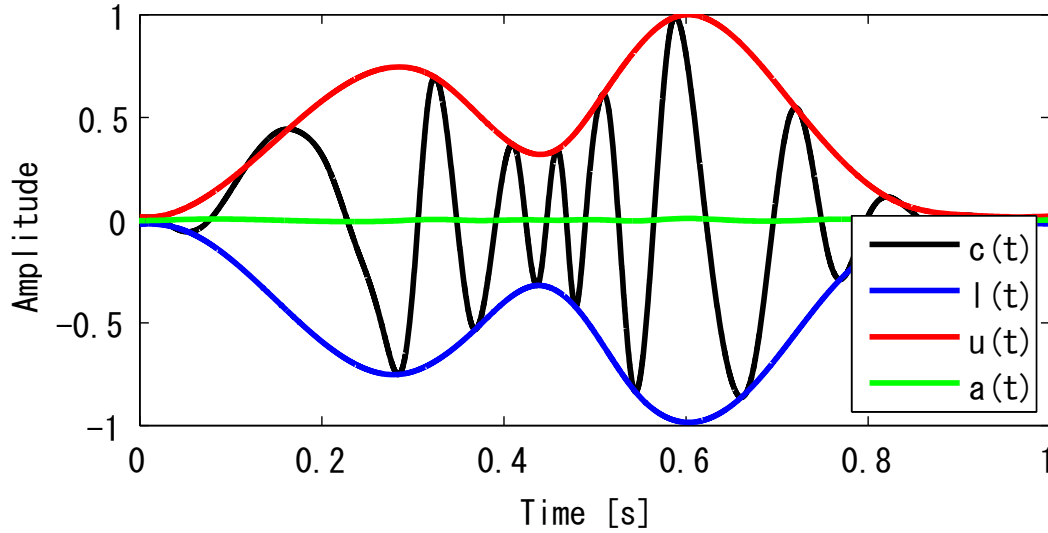


図 2.2: IMF の制約条件を満たした信号の例

側エンベロープ $u(t)$ と極小点同士から作られる下側エンベロープ $l(t)$ の平均値 $a(t)$ がゼロとなる信号であることが分かる．

ある信号 $(0.5 \times \sin(2\pi \times 100 \times t) + 0.5 \times \sin(2\pi \times 1 \times t))$ の分解過程を図 2.3 に示す．図に示すような原信号 $x(t)$ を用いて，この信号を EMD により分解する過程を示す．まず，上図で， $x(t)$ の極大値，極小値をそれぞれ，3 次のスプライン曲線で補間し，上側のエンベロープ $u_1(t)$ と下側のエンベロープ $l_1(t)$ を得る．次に， $u_1(t)$ ， $l_1(t)$ の平均値 $a_1(t)$ を求める．ここで，IMF の制約条件 2 より，上下のエンベロープの平均値がゼロであるかを判定する．しかし， $x(t)$ は，その条件を満たさないため，原信号 $x(t)$ からエンベロープの平均値 $a_1(t)$ を引く． $x(t) - a_1(t)$ の信号を図 2.3 下図に示す．この信号に対しても同様に，エンベロープの平均値を求め，IMF の制約条件 2 が満たされるか確認する．この信号の場合，制約条件 2 が満たされていることが分かる．さらに，制約条件 1（信号の極値の数と信号のゼロ交差数は同じか一つ違い）も満たしていることが分かる．二つの制約条件が満たされているので，この信号（ $x(t) - a_1(t)$ ）を $IMF(c_1(t))$ とする．次に，原信号から IMF の差分をとった信号 $(x(t) - \sum_{k=1}^K c_k(t))$ ，ただし，この場合 $K=1$ となる．) を求める．以下， $x(t) - \sum_{k=1}^K c_k(t)$ に対して，同様の操作を行う．図 2.4 に EMD による信号分解の結果を示す．結果より，作成した信号と同じ信号に分解されたことが分かる．

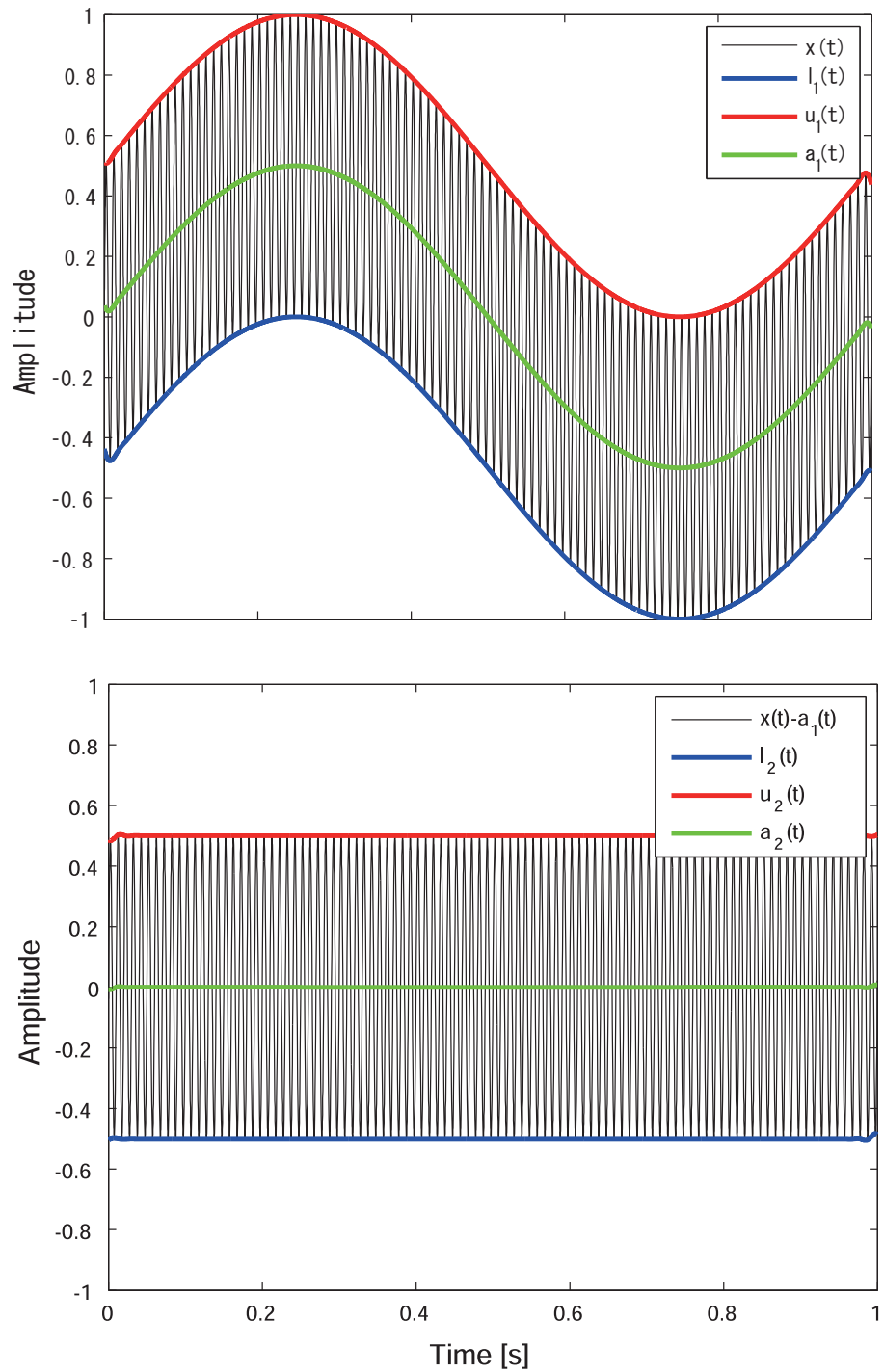


図 2.3: 信号の分解過程 ($x(t)$ に対してエンベロープの平均値を求めたもの (上図), $x(t) - a_1(t)$ に対してエンベロープの平均値を求めたもの (下図))

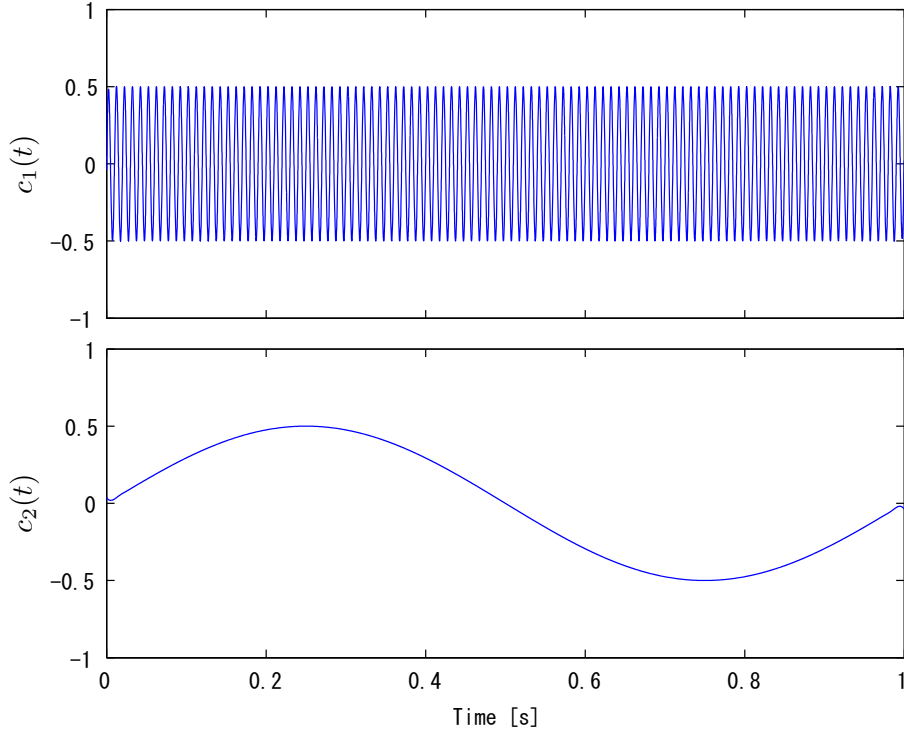


図 2.4: 固有モード関数 (IMF)

2.3.1 Hilbert 変換

単一の周波数しか持たない信号に対して, 振幅と位相を表現する手法を提案した [30, 31] . ヒルベルト変換 $y(t)$ は以下の式で表される .

$$y(t) = \frac{1}{\pi} P \int_{-\pi}^{\pi} \frac{x(t')}{t - t'} dt' \quad (2.4)$$

P はコーシーの主値である . このとき , 複素信号 $z(t)$ は以下の式で表される .

$$z(t) = x(t) + iy(t) = \alpha(t)e^{i\theta(t)} \quad (2.5)$$

上記の式より , 瞬時振幅 $\alpha(t)$ と瞬時位相 $\theta(t)$ は以下の式で表される .

$$\alpha(t) = \sqrt{x^2(t) + y^2(t)} \quad (2.6)$$

$$\theta(t) = \arctan \frac{y(t)}{x(t)} \quad (2.7)$$

瞬時位相の式より , 瞬時周波数は以下の式で求められる .

$$\omega(t') = \left. \frac{d\theta(t)}{dt} \right|_{t=t'} \quad (2.8)$$

ここで， $c_k(t)$ に対応する瞬時振幅を $\alpha_k(t)$ ，瞬時周波数を $\omega_k(t)$ とおくと， $x(t)$ は次のように表される．

$$\begin{aligned} x(t) &= \sum_{k=1}^K c_k(t) + r(t) \\ &= \operatorname{Re}\left(\sum_{k=1}^K \alpha_k(t) \exp(i \int \omega_k(t) dt)\right) + r(t) \end{aligned} \quad (2.9)$$

すべての IMF に対して，時間–周波数空間で表現した信号は Hilbert–Huang スペクトラム [4, 32] と呼ばれる．Hilbert–Huang スペクトラムは以下の式で表される．

$$H(t, \omega) = \begin{cases} \alpha_k(t) & (\omega_k(t) = \omega) \\ 0 & (\text{otherwise}) \end{cases} \quad (2.10)$$

これは，短時間フーリエ変換や，ウェーブレット変換とは違い，周波数分解能を高める（低くする）と，時間分解能が低く（高く）なる性質（不確定性）を持たない．

表 2.1: 振幅周波数変調の各パラメータ

搬送波		振幅変調		周波数変調	
振幅 A	1.0	振幅変調度 m_{AM}	0.9	周波数変調度 m_{FM}	1.0
周波数 f_c	50 [Hz]	変調周波数 f_{AM}	2 [Hz]	変調周波数 f_{FM}	10 [Hz]

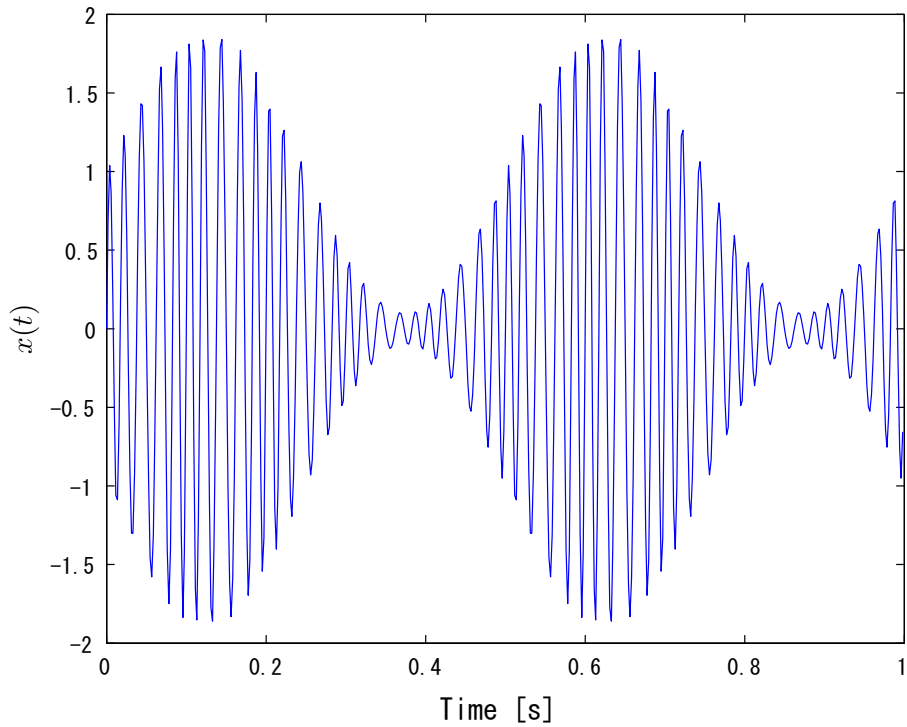


図 2.5: $x(t)$ の信号波形

2.3.2 Hilbert–Huang スペクトラムの解析例

こちら側で作成した信号の Hilbert–Huang スペクトラムを解析した例を紹介する．信号は，単一周波数を持つ正弦波を振幅周波数変調した信号とする．次式にそれを示す．

$$x(t) = A \sin(\omega_c t + m_{FM} \sin(\omega_f t)) \times (1 + m_{AM} \sin(\omega_a t)) \quad (2.11)$$

ω は角周波数で各パラメータを表 2.1 のように設定した．なお， $\omega = 2\pi f$ とする．図 2.5 に作成した信号を示し，その Hilbert–Huang スペクトラムを解析した結果を図 2.6 に示す．図 2.6 の色の違いは，スペクトルの強度を示している．赤ほど強度は強く，青ほど強度は弱い．Hilbert–Huang スペクトラムを見ると，瞬時周波数が時間的に変動しつつ，振幅が変化してる様子が分かる．

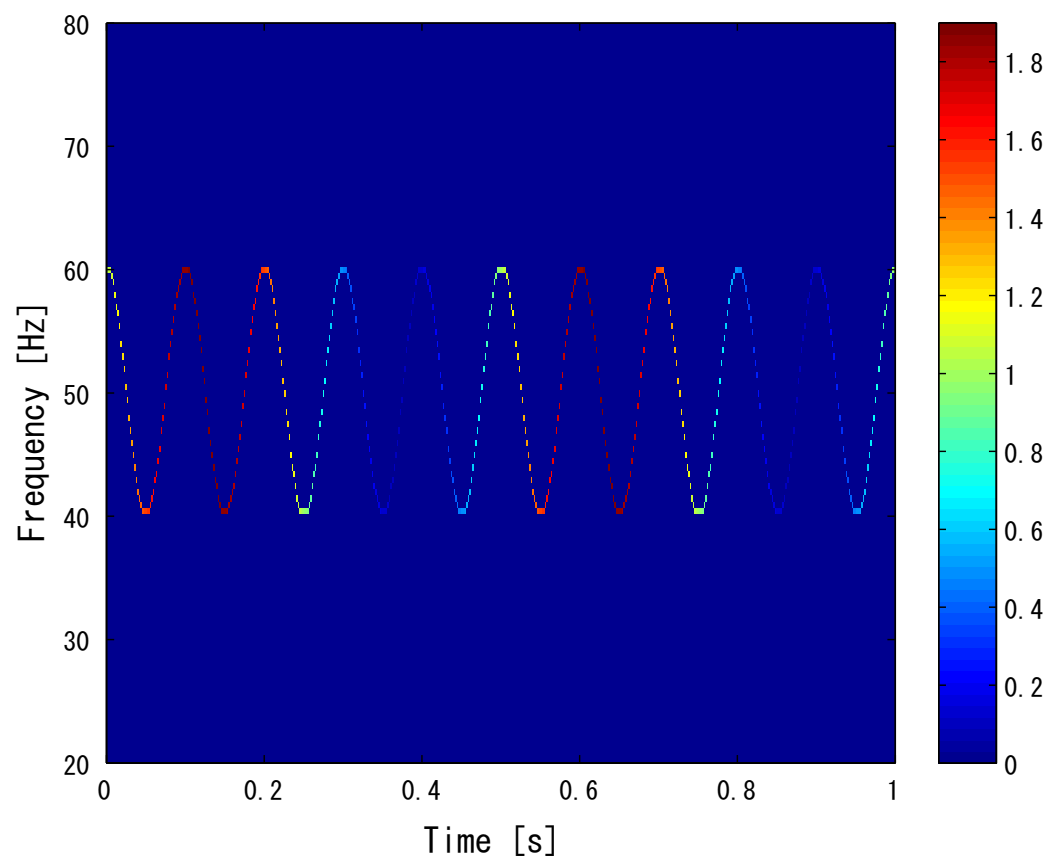


図 2.6: Hilbert-Huang スペクトラム

2.4 EMD を用いた音信号処理

2.4.1 Taufiq らによって提案された雑音除去法

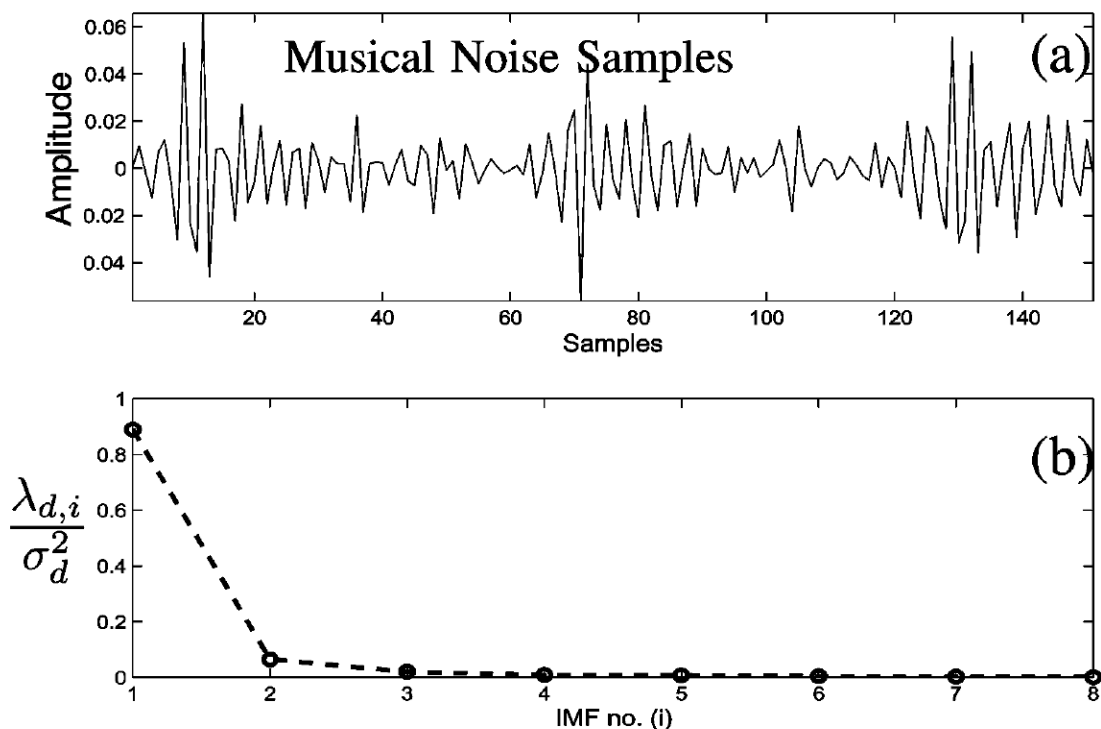


図 2.7: (a) ミュージカルノイズと (b) 各 IMF におけるミュージカルノイズのエネルギー分布（文献 [22] より引用）

Taufiq らは、EMD を用いたミュージカルノイズの抑圧法 [22] を提案している。Taufiq らは、スペクトルサブトラクション等で雑音除去をした結果発生するミュージカルノイズのエネルギーの大半は、ミュージカルノイズが付加された音声信号を EMD により分解して得られるはじめのいくつかの IMF に集中すると述べている。図 2.7 にミュージカルノイズのサンプル (a) と、各 IMF におけるミュージカルノイズのエネルギー分布を示す（文献 [22] より引用）。 $\lambda_{d,i}$ は i 個の IMF の分散で、 σ_d^2 はミュージカルノイズが付加された音声信号の分散である。図 2.7 より、最初の 1 個の IMF に 88.94 [%] のミュージカルノイズのエネルギーが集中していることが分かる。これにより、ミュージカルノイズが付加された音声信号を EMD により分解し、最初の 1 個の IMF を取り除くことで、Taufiq らはミュージカルノイズの抑制の方法を実現した。

2.4.2 Molla & Hirose によって提案された雑音除去法

有声・無声の分類の前処理として，Molla & Hirose は EMD を用いた雑音除去法 [23] を提案している．Molla & Hirose の雑音除去法は，音声と雑音をもつスペクトルの違いに着目し，EMD を用いて雑音音声进行分解した後，高周波成分を含む IMF を取り除くことで，雑音除去を行っている．

図 2.8 に短時間における音声信号 $x(t)$ とそのスペクトル $X(\omega)$ を，図 2.9 に短時間における雑音音声 $y(t)$ (SNR = 5 [dB] の白色雑音を付加) とそのスペクトル $Y(\omega)$ を示す．スペクトルの横軸はナイキスト周波数で正規化した．図 2.8 より，音声信号のスペクトルは，低域に集中しているのに対して，図 2.9 より，雑音音声のスペクトルは，全帯域に分布していることが分かる．Molla & Hirose は，この二つの図からも分かるように，高周波成分を含む IMF を取り除くことで，雑音除去を行う方法を提案した．次に実際に雑音除去の手順を示す．図 2.10 に音声信号 $x(t)$ ，雑音信号 (SNR = 5 [dB] の白色雑音) とその合成信号 $y(t)$ を，図 2.11 に $y(t)$ を EMD により分解した結果 (最初の 6 本の IMF) を示す．図 2.11 より，高周波雑音の IMF である $c_1(t)$ と $c_2(t)$ と，音声信号を含む IMF $c_3(t)$ ， $c_4(t)$ ， $c_5(t)$ ， $c_6(t)$ に分けられていることが分かる．Molla & Hirose は，最初の何個の IMF が高周波雑音であるか判別を行うために，各 IMF の瞬時周波数を調べることで実現した．各 IMF の瞬時周波数は式 2.8 により求める．図 2.12 に IMF の瞬時周波数 (最初の 6 本) を示す． $IF_1(t)$ と $IF_2(t)$ の瞬時周波数をみると，音声スペクトルには含まれない瞬時周波数であることが分かる．Molla & Hirose は付加されている雑音が白色雑音の場合，最初の 2 本の IMF を取り除くことで，高周波雑音の除去を行っている．以下に付加された雑音が白色雑音の場合における回復信号の式を示す．

$$\hat{x}(t) = \sum_{k=3}^K c_k(t) + r(t) \quad (2.12)$$

図 2.13 に回復信号 $\hat{x}(t)$ とそのスペクトル $\hat{X}(\omega)$ を示す．この図より，高周波雑音が除去されていることが分かる．

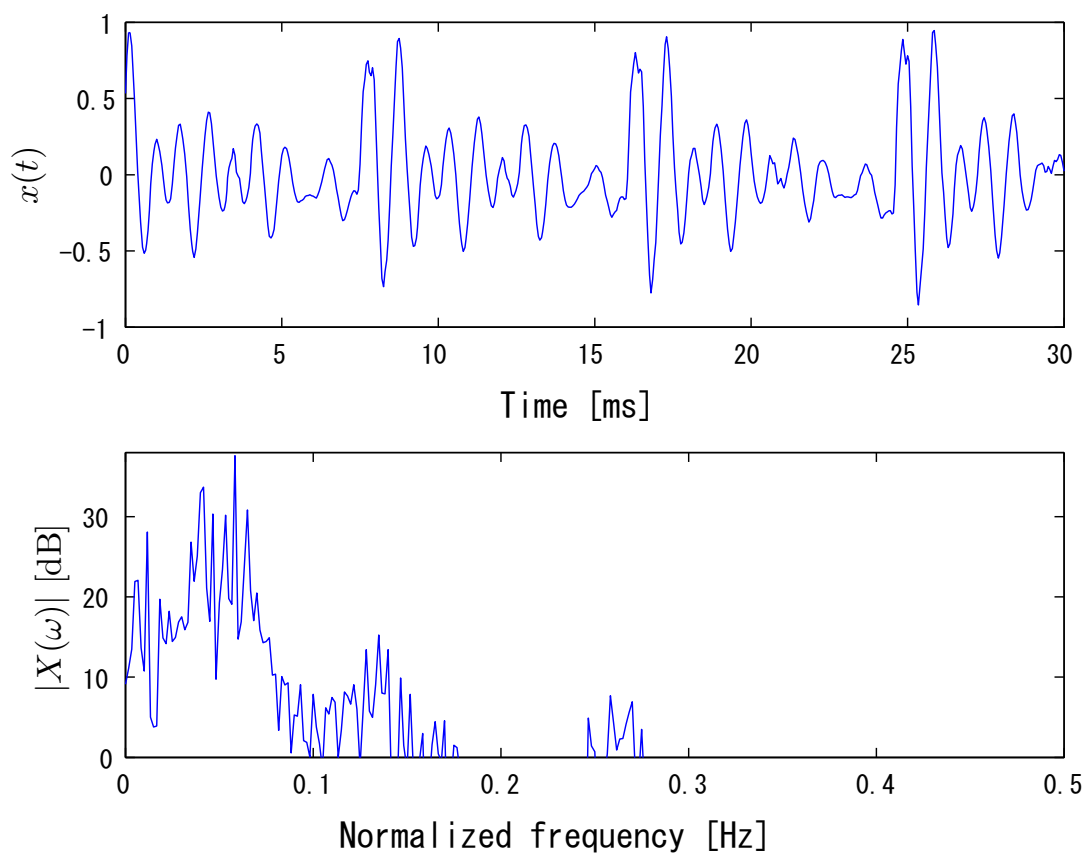


図 2.8: 短時間における音声信号 $x(t)$ とその振幅スペクトル $|X(\omega)|$

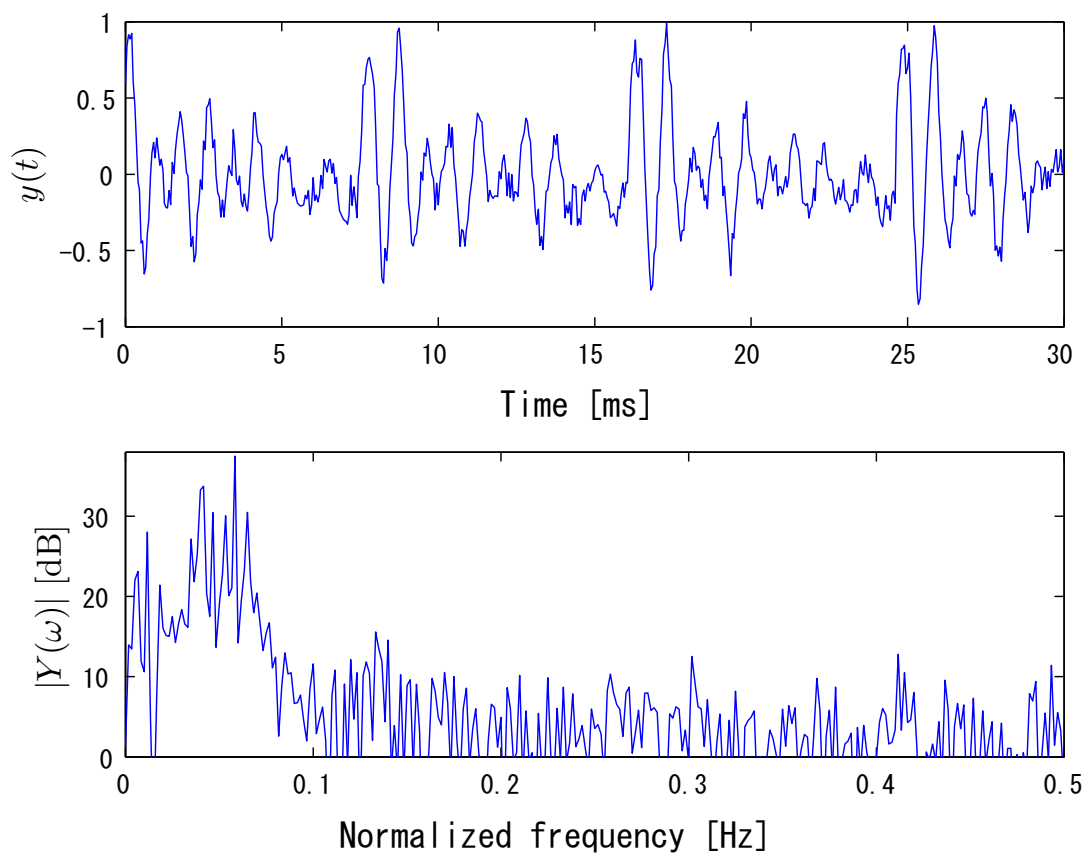


図 2.9: 短時間における雑音音声 $y(t)$ とその振幅スペクトル $|Y(\omega)|$

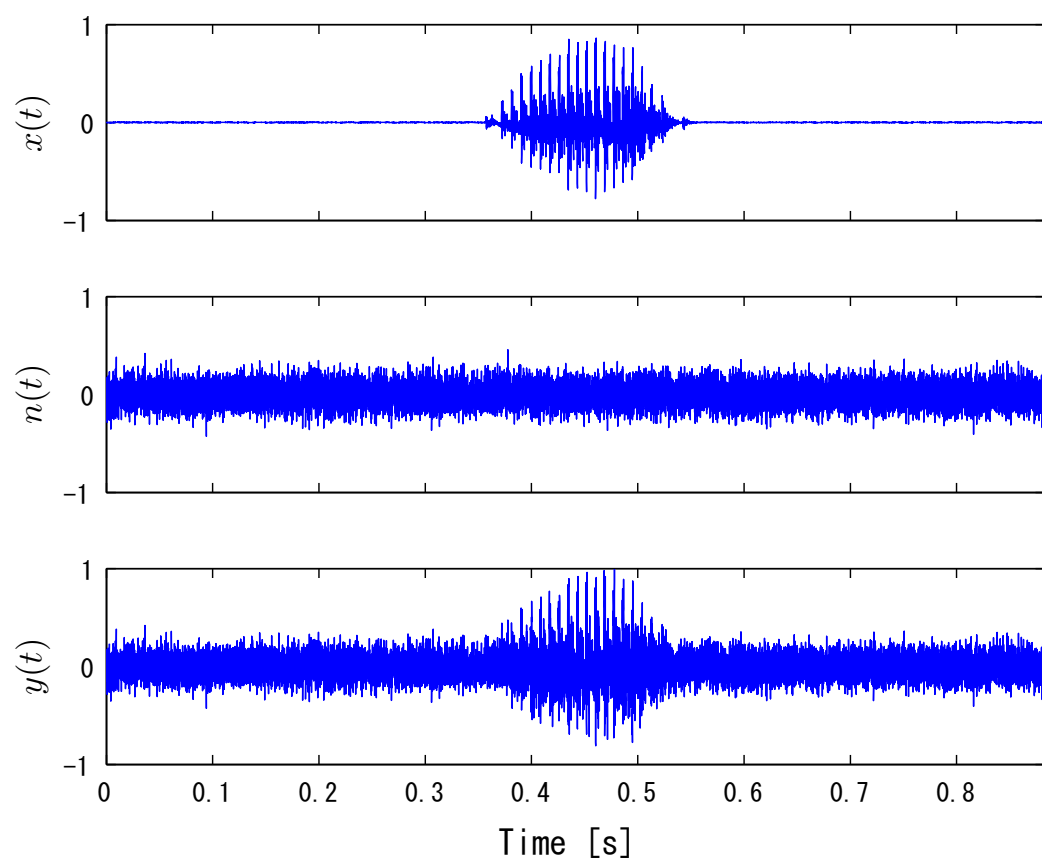


図 2.10: 音声信号 $x(t)$, 白色雑音 $n(t)$ とその合成信号 $y(t)$

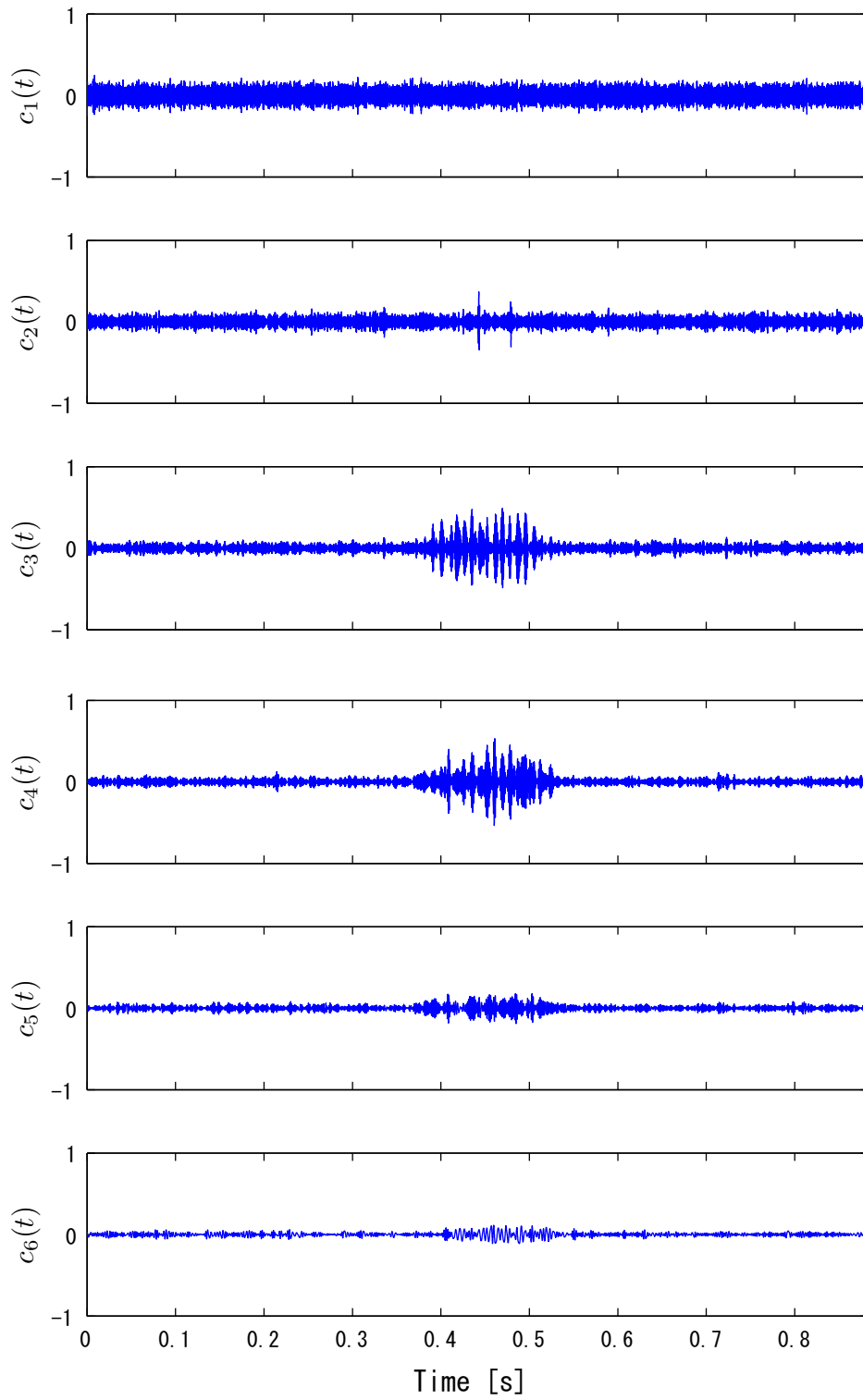


図 2.11: 雑音音声 $y(t)$ を EMD により分解した結果 (6 個の IMF)

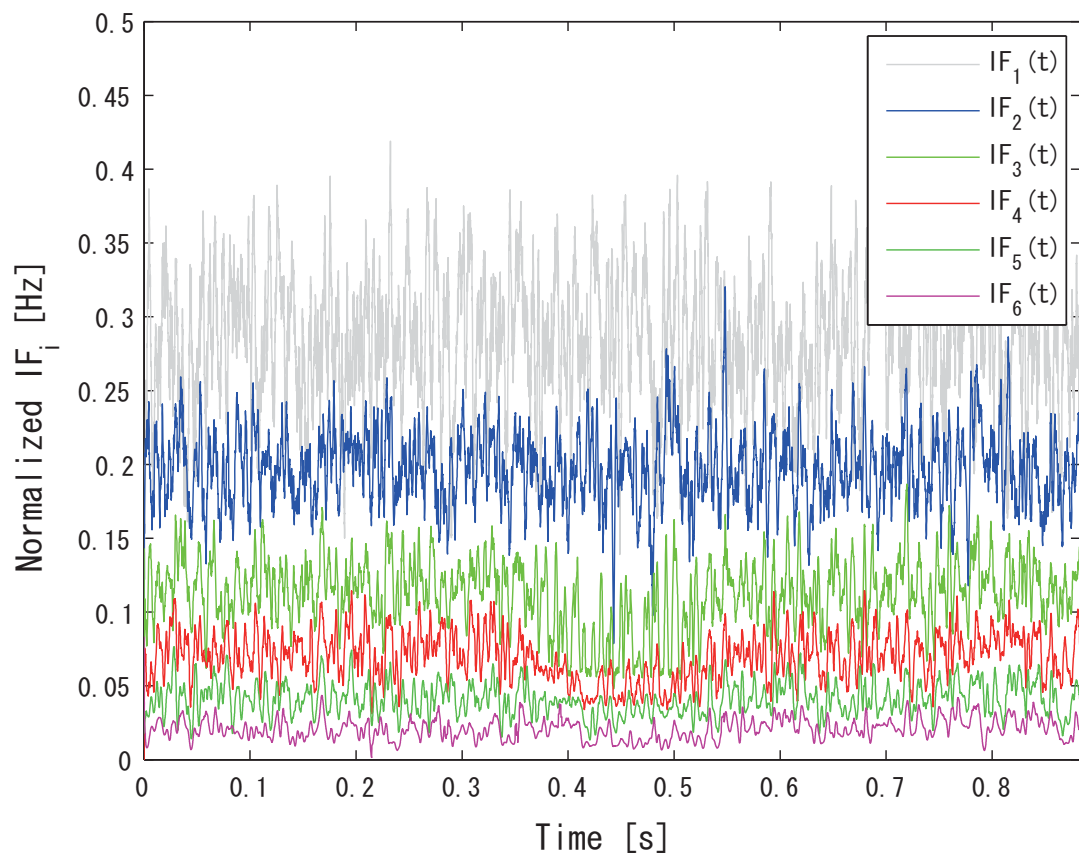


図 2.12: IMF の瞬時周波数 (6 個の IMF)

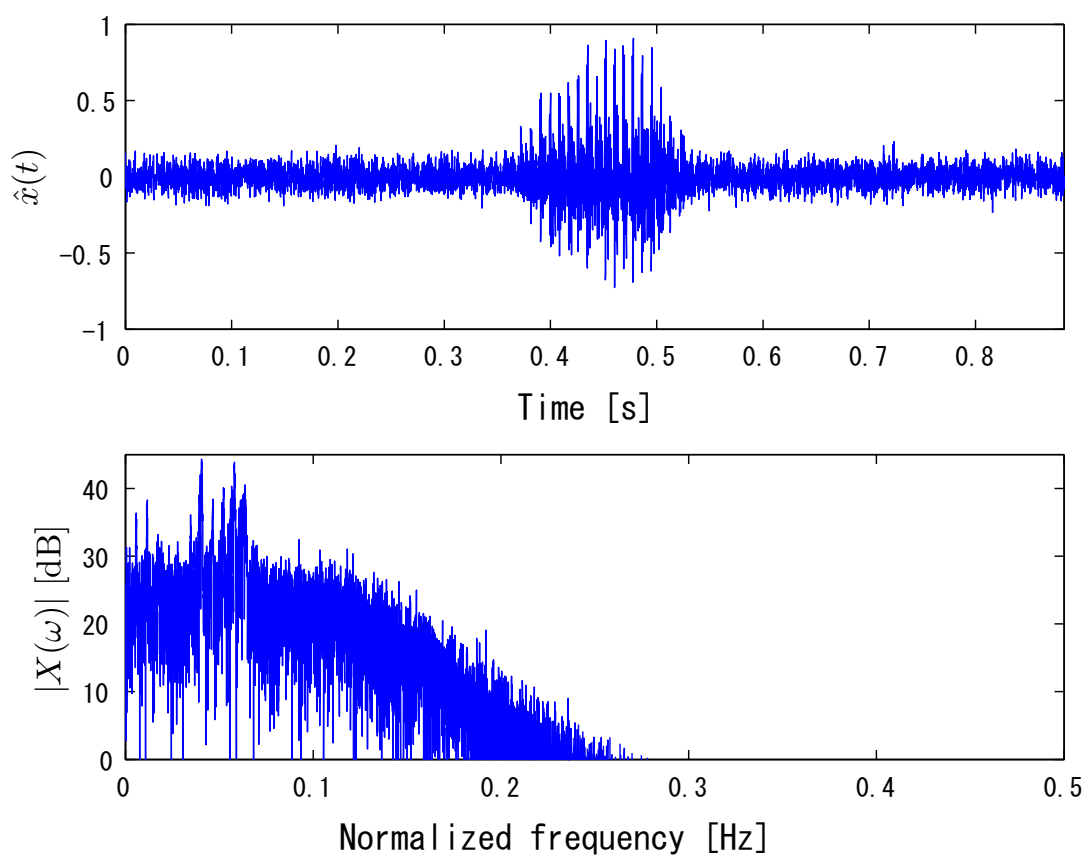


図 2.13: Molla & Hirose の雑音除去法による回復信号（上図）と回復信号の振幅スペクトル（下図）

2.4.3 Matsuda らによって提案された基本周波数推定

Matsuda らは，EMD を用いた基本周波数推定法 [24] を提案している．非定常な信号を分析するのに適した EMD を用いることで，基本周波数の非定常な変化を正確に捉えた分析を実現した．EMD は入力信号に依存した分析手法であるために，基本周波数に関する信号成分が複数の IMF に分解される問題を抱えていた．この問題に対して，Matsuda らは，遅れ時間空間において基本周波数のエネルギーが特定の帯域に集中することに着目し，雑音にロバストな基本周波数推定法を実現した．この手法は，観測信号を EMD によって分析するのではなく，自己相関関数に対して間接的に EMD を用いることで，基本周波数に関わるエネルギーを集中させ，推定精度の向上を図った．音声信号を $x(t)$ ，白色雑音を $n(t)$ ，雑音音声を $y(t)$ としたときに，自己相関関数 $R_{yy}(\tau)$ は以下の式で表現される．

$$R_{yy}(\tau) = \frac{1}{T} \int_0^T (y(t))(y(t + |\tau|)) dt \quad (2.13)$$

$$\begin{aligned} &= \frac{1}{T} \int_0^T (x(t) + n(t))(x(t + |\tau|) + n(t + |\tau|)) dt \\ &= \frac{1}{T} \int_0^T (x(t)x(t + |\tau|) + n(t)n(t + |\tau|) + x(t)n(t + |\tau|) + n(t)x(t + |\tau|)) dt \end{aligned} \quad (2.14)$$

式 2.14 より， $x(t)x(t + |\tau|)$ から音声区間に音声の周期性を見ることが出来， $n(t)n(t + |\tau|)$ は $\tau = 0$ のときのみ相関が 1 でそれ以外は 0 となり， $x(t)n(t + |\tau|)$ ， $n(t)x(t + |\tau|)$ は音声区間に音声と雑音の間の相関が見ることが出来る．Matsuda らは，この操作により音声の周期性を強調し，基本波のエネルギーを集中させた．その信号に対して，EMD を用いて分析することにより，基本周波数推定法を実現した．

第3章 経験的モード分解に基づく音声信号の分析

3.1 EMD の特徴

3.1.1 分析合成系の精度

EMD を用いて分析および合成をおこなう際、分析合成系の精度を確認する必要がある．一般的に、フィルタバンク等の分析合成系の精度は 30 [dB] 程度と知られている．ここで、EMD の分析合成系の精度を調べる．経験的モード分解の分析合成系に対する検討を行う．分析合成系の精度を確かめるため、分解して得られた固有モード関数と残差を再合成することで、入力信号と再合成された信号で損失した信号との比 (SNR_{c+r}) を求める．以下にその式を示す．

$$\text{SNR}_{c+r} = 10 \log_{10} \frac{\int_{-\infty}^{\infty} x^2(t) dt}{\left(\int_{-\infty}^{\infty} x(t) dt - \left(\sum_{k=1}^K \int_{-\infty}^{\infty} c_k(t) dt + \int_{-\infty}^{\infty} r(t) dt \right) \right)^2} \quad (3.1)$$

次に、入力信号 $x(t)$ と固有モード関数 $c_k(t)$ のみを再合成した信号（残差 $r(t)$ 以外を再合成した信号）との比 (SNR_c) を求める．以下にその式を示す．

$$\text{SNR}_c = 10 \log_{10} \frac{\int_{-\infty}^{\infty} x^2(t) dt}{\left(\int_{-\infty}^{\infty} x(t) dt - \left(\sum_{k=1}^K \int_{-\infty}^{\infty} c_k(t) dt \right) \right)^2} \quad (3.2)$$

SNR の条件と同様に対数スペクトル距離 (Log Spectral Distance: LSD) LSD_{c+r} , LSD_c を求める． $S(\omega)$ を入力信号のスペクトル、 $\hat{S}(\omega)$ を再合成した信号のスペクトルとする．以下に式を示す．

$$\text{LSD} = \sqrt{\frac{1}{L} \sum_{l=1}^L 20 \log_{10} \frac{S(\omega_l)}{\hat{S}(\omega_l)}} \quad (3.3)$$

分析に用いた信号は ATR 音声データベースの a-set[34] から 30 種類（男性 5 名、女性 5 名の各 3 音声ずつ）の音声を用いた．

分析合成系の精度を評価した結果、入力信号と、固有モード関数ならびに残差を再合成した信号に対する SNR_{c+r} は平均値が 314.5 [dB]、分散が 0.12 となった． LSD_{c+r} の平均値

は 2.44×10^{-12} [dB], 分散が 7.70×10^{-24} となった。入力信号と、固有モード関数のみを再合成した信号に対する SNR_c は平均値が 33.1 [dB], 分散が 52.6 となった。 LSD_c の平均値は 1.04 [dB], 分散が 0.18 となった。これらの結果からも分かるように、分析合成系の精度としては非常に高いことが確認された。

3.1.2 信号分解の特徴

ここでは、EMD の信号分解の特徴と IMF の性質について検討を行う。先行研究として、文献 [35] では、非整数ガウス雑音を EMD により分解し、分解された IMF をみることで、EMD と同様のフィルタバンクを設計し、どのような信号表現であるかを議論をしている。しかし、EMD の分析は、入力信号に依存する分析手法であるために、文献 [35] の信号表現は、どの信号に対しても Flandrin が示した信号表現になるとは限らない。従って、2.3 節で示した EMD の分解過程から分解される信号について検討を行うことで、どのような信号表現を実現するか検討をしなければいけない。この信号表現について検討を行った上で、各分野の分析に応用することで、EMD の特徴を最大限活かした分析を可能に出来ると考えられる。音信号分析に EMD を適用させる上でも、EMD の信号表現について検討することは重要である。次に、EMD の信号分解の過程から EMD の信号分解の特徴について検討を行う。

原信号を IMF に分解するためには、まず、原信号の極大点と極小点を 3 次のスプライン曲線で補間し、上側と下側のエンベロープの平均値を求める。次に、エンベロープの平均値がゼロになるまで、原信号からエンベロープの平均値を減算する処理を繰り返し、IMF の制約条件を満たすまで以上の処理を続ける。従って、IMF はエンベロープベースで分解されていることがわかる。固有モード関数が得られたら、原信号から、ここまで得られた全ての IMF を減算し、残った信号に対して、同様に IMF の制約条件を満たすまでエンベロープの平均値を減算し続け、次の IMF を求める。このとき、IMF は動きの速いもの（振動の速いもの）から動きの遅いもの（振動の遅いもの）に分解されるものと解釈できる。

次に、EMD により得られた IMF の性質について検討を行う。一つ目の制約条件より、IMF は必ず極値とゼロ交差を繰り返す信号でなければならないことがわかる。これは、極値とゼロ交差の頻度に対して制限がないため、IMF は帯域幅に制限の無い FM 波であると解釈することができる。従って、三角波やノコギリ波のような全周波数帯域を含む信号であっても、1 本の IMF として取り出される。二つ目の制約条件より、IMF は時間軸を対称としたエンベロープを持つ信号でなければならないことがわかる。そのため、この条件を満たす信号は AM 波であると解釈することができる。以上のことから、IMF は、共変調された AM-FM 信号ごとに分解されるものであり、振幅包絡が一定である定常な波から、ある区間にのみ振幅を持つ非定常な波へと順番に分解されるものであると解釈できる。

3.2 EMD を用いた信号の分析例

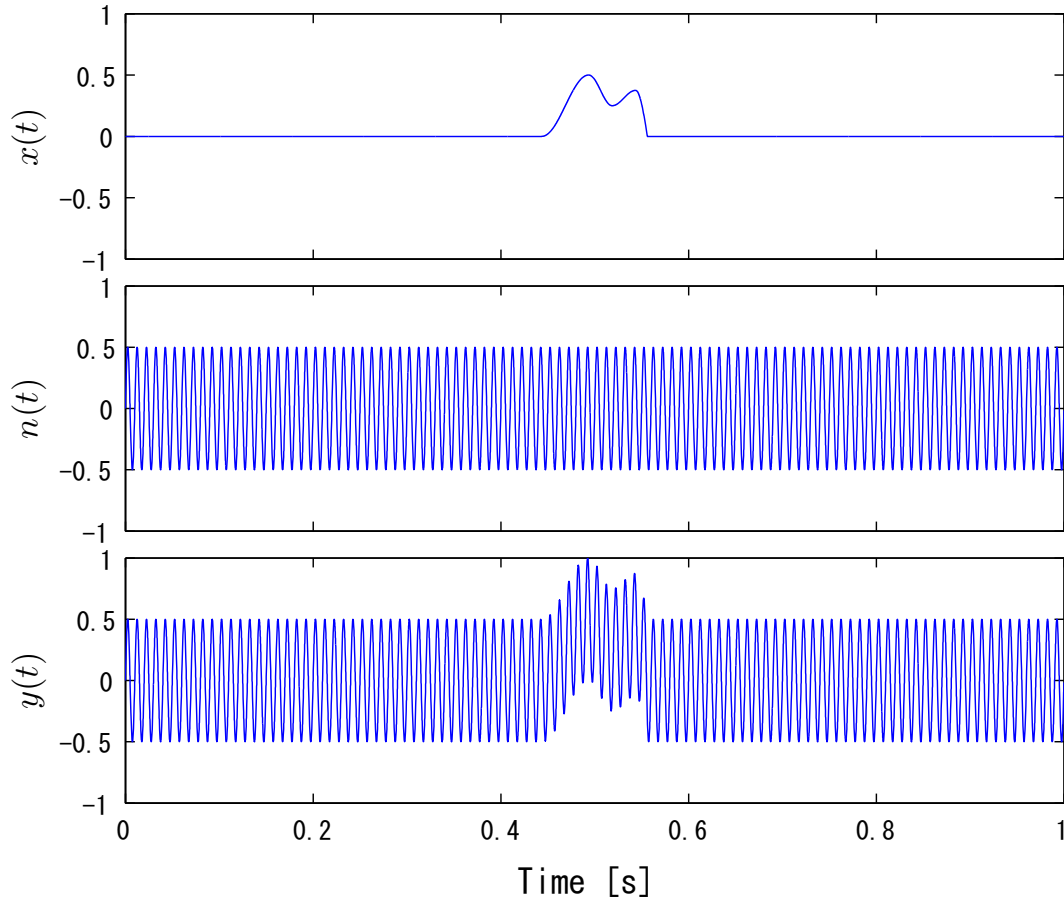


図 3.1: 非定常信号 $x(t)$ と定常信号 $n(t)$ の和で構成される合成信号 $y(t)$

ここで，EMD を用いた信号分析の一例をみる．ある区間にのみ振幅を持つ非定常信号 $x(t)$ と定常信号 $n(t)$ の和で構成される合成信号 $y(t)$ を考える．図 3.1 にこれらの信号波形を，図 3.2 に EMD によって分解された $y(t)$ の IMF を示す．上記で述べた IMF の性質を逆手にとれば，一定なエンベロープをもつ IMF $c_1(t)$ は定常信号であり，それ以外の IMF ($c_2(t), c_3(t), c_4(t)$) は非定常信号であると解釈できる．そのため，前者の IMF ($c_1(t)$) を $\hat{n}(t)$ ，後者の IMF の総和 ($c_2(t) + c_3(t) + c_4(t)$) を $\hat{x}(t)$ として，再合成すると図 3.3 のようになる．この結果から，EMD を用いた音信号の分析では，合成信号 $y(t)$ を振幅包絡が一定な信号 $n(t)$ とある区間にのみ振幅をもつ信号 $x(t)$ に分解可能であるといえる．

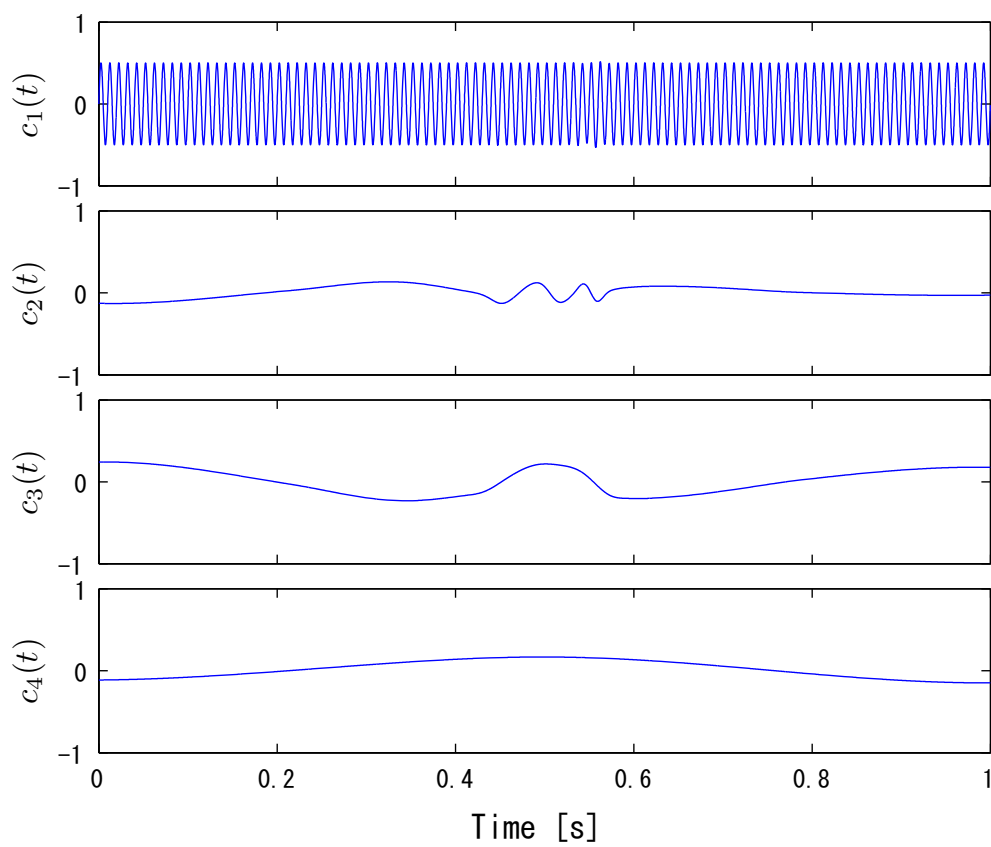


図 3.2: 合成信号 $y(t)$ の分解結果 (4 個の IMF $c_k(t)$)

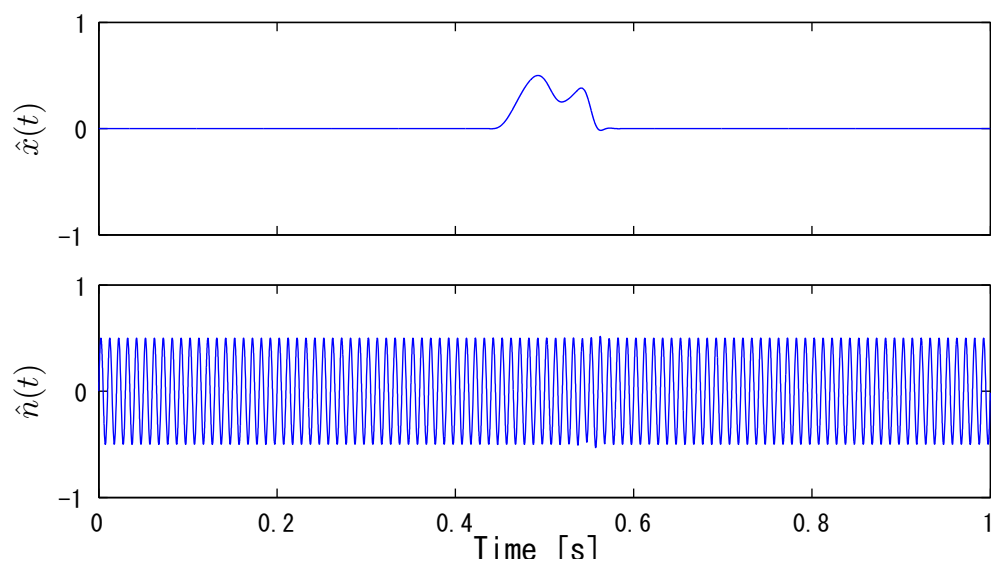


図 3.3: $x(t)$ と $n(t)$ の再合成信号

3.3 EMD を用いた信号分析の本質

これまで、EMD を用いた信号分析について調査してきた。EMD の分解過程より、音信号を、エンベロープが一定な定常信号と、ある区間にのみエンベロープを持つ非定常信号に分解可能であることを示した。この結果より、EMD に基づく音分析の本質は、共通なエンベロープ分解に基づいた信号表現を行う方法であることがわかった。従って、被解析信号を共変調された信号ごとに分解を行うことが可能であると考えることができる。

3.4 EMD を用いた音信号処理の可能性について

EMD を用いた信号分析の本質に基づき、この分析の性質を最大限に活かすことのできるアプリケーションを検討する。EMD を用いることで、音信号を、エンベロープが一定な定常信号と、ある区間にのみエンベロープを持つ非定常信号に分解可能であることを示した。この特質を活かし、定常信号である白色雑音と非定常信号である音声信号の分離可能性を検討する。次章では、EMD を用いた雑音除去の可能性について検討する。

第4章 経験的モード分解を用いた雑音除去法の提案

4.1 EMD を用いた雑音除去法

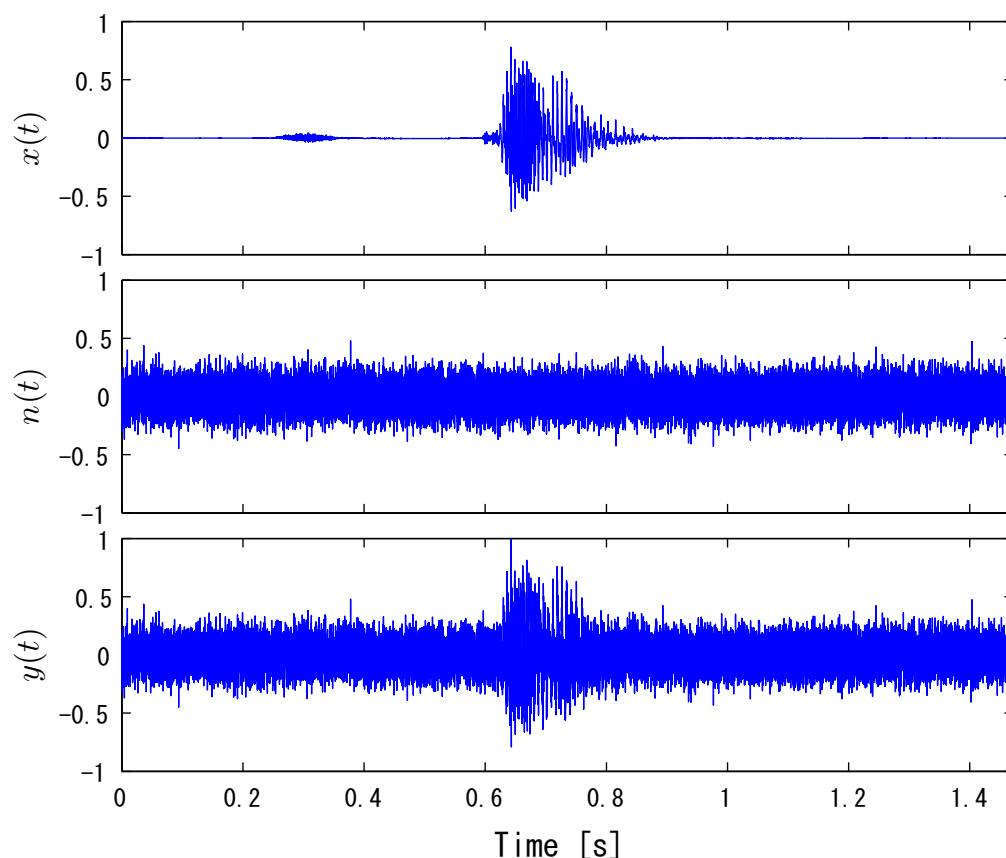


図 4.1: 音声信号 $x(t)$ と $\text{SNR} = 0$ [dB] のガウス雑音 $n(t)$ の和で構成される雑音音声 $y(t)$

すでに，EMD を用いた雑音抑圧法は，Molla & Hirose [23] によって提案されている．しかし，この手法では，雑音のエネルギー分布だけに着目し，特定の IMF のみを限定的に取り除くため，結果的には雑音の高周波成分だけが取り除かれることになり，雑音抑圧が完全であるとは考え難い．

ここで，EMD を用いて音声を分解した一例をあげる．図 4.1 に信号対雑音比 (SNR) が

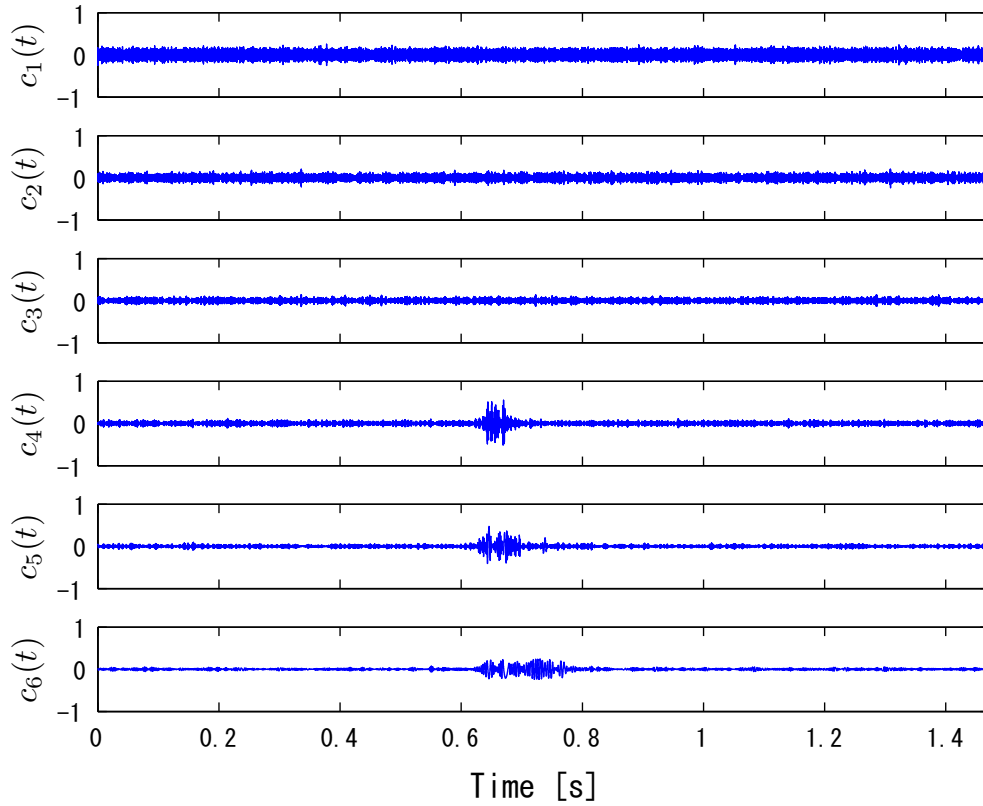


図 4.2: 雑音音声 $y(t)$ の分解結果 (最初の 6 個の IMF)

0 [dB] の白色雑音を付加した音声信号 $y(t)$ を示す．次に，この信号を，EMD を用いて分解した結果 (最初の 6 個の IMF のみ) を図 4.2 に示す．これらの結果は，前節の結果と同様に，分解された IMF は，振幅包絡が一定な白色雑音と，ある区間에만振幅包絡を持つ音声信号とに分けられていることがわかる．これは，EMD の特徴である共通なエンベロープ分解に基づく信号表現によって，雑音音声と共変調された雑音信号と共変調された音声信号とに分離することによって，このような結果になったものと考えられる．これに対し，Molla & Hirose の雑音抑圧法では，雑音が白色雑音 $n(t)$ のとき，EMD を用いて雑音音声と分解して得られた最初の 2 個の IMF ($c_1(t)$ と $c_2(t)$) を強制的に取り除くことで雑音抑圧を行っている．ここで，最初の 2 個の IMF の選定は，雑音のエネルギー分布を考慮した結果によるものである．しかし，図 4.2 から分かるように，必ずしも最初の 2 個の IMF だけが雑音成分を表しているとは限らない．

本研究では，音声信号の IMF と雑音成分の IMF の違いを考慮し，Molla & Hirose の雑音抑圧では取り除くことが出来なかった他の雑音成分の IMF を取り除くことで，EMD の特質を活用した雑音除去の可能性を検討する．そのため，音声の特徴を持つ IMF のみを次式のように再合成することで，雑音音声から非定常信号である音声信号のみを抽出する

ことにする．

$$\hat{x}(t) = \sum_{k \in S} c_k(t) \quad (4.1)$$

ただし， S は，音声の IMF を表すチャンネル番号の集合である．

4.2 音声と雑音の切り分け方法

雑音音声を EMD により分解し，音声 IMF と雑音 IMF に分類する方法を考える．本稿では，音声と雑音の変調度の違いに着目し，これらの分類を行う．音声信号のエンベロープで見られる変調周波数特性は，約 2 - 8 [Hz] の変調周波数にピークを持つことが知られている [36, 37, 38]．この特徴に基づき，各 IMF の変調周波数特性を求め，雑音 IMF と音声 IMF を判別する．

提案手法の処理体系のモデルと音声 IMF と雑音 IMF のチャンネル選択の処理体系を図 4.3 に示す．雑音音声を EMD により分解し，IMF を求める．次に，各 IMF から，IMF のエンベロープを抽出する．音声の変調度は 20 [Hz] までに重要な成分が含まれているため，エンベロープをカットオフ周波数 20 [Hz] のローパスフィルタに通し，ダウンサンプリングをする．ダウンサンプリングしたエンベロープを変調フィルタバンク（ここでは，定帯域通過フィルタを利用）にて再度，周波数分析を行い， $e_{k,m}(t)$ を得る．次節に，定帯域通過フィルタの仕様を記載する．

4.3 定帯域通過フィルタ

各 IMF の変調度を計算するために用いたバンドパスフィルタは，定帯域通過フィルタとする．変調スペクトルのピークは 2 - 8 [Hz] の帯域に存在することから，ピークを検出するためには，周波数分解能が最低 1 [Hz] 程度でなければいけないため，提案手法では，0.5 [Hz] 刻みで分析を行えるよう，チャンネル数は 40 とした．

定帯域バンドパスフィルタは，ローパスフィルタとハイパスフィルタを組み合わせたもので設計し，ローパスフィルタとハイパスフィルタは，それぞれ 3 次のチェビシェフ多項式を用いた．また，チェビシェフフィルタを採用した理由として，各 IMF のエンベロープの帯域が 0 - 20 [Hz] と狭い帯域であるためである．

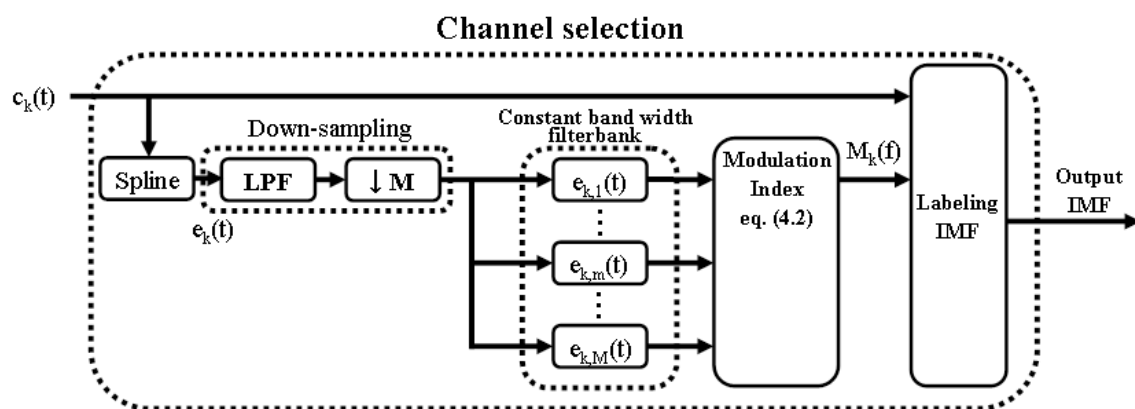


図 4.3: 提案法の処理体系（上図）と IMF のチャンネル選択の処理体系（下図）

4.4 音声と雑音の変調度

前節で設計された定帯域バンドパスフィルタを用いて，各 IMF のエンベロープを分析し，各帯域における変調度を次式により求める．

$$\begin{aligned} M_{k,m} &= \frac{(\max(e_{k,m}(t)) + \overline{e_{k,m}(t)}) - (\min(e_{k,m}(t)) + \overline{e_{k,m}(t)})}{\max(e_{k,m}(t)) + \overline{e_{k,m}(t)}} \\ &= \frac{\max(e_{k,m}(t)) - \min(e_{k,m}(t))}{\max(e_{k,m}(t)) + \overline{e_{k,m}(t)}} \end{aligned} \quad (4.2)$$

この式に基づき，図 4.2 における音声 IMF と雑音 IMF の変調度を求めた結果の一例を図 4.4 に示す．音声 IMF の変調周波数特性は，5 [Hz] にピークが存在し，かつその変調度が高い傾向にある．これに対して，雑音 IMF の変調周波数特性は音声 IMF のものに対して低い傾向にある．提案法では，2～8 [Hz] に変調周波数のピークが存在し，かつその変調度が 0.25 以上の IMF を音声 IMF として判断する．最後に，音声 IMF と雑音 IMF を判別した後，音声 IMF のみを 4.1 を利用して再合成する．

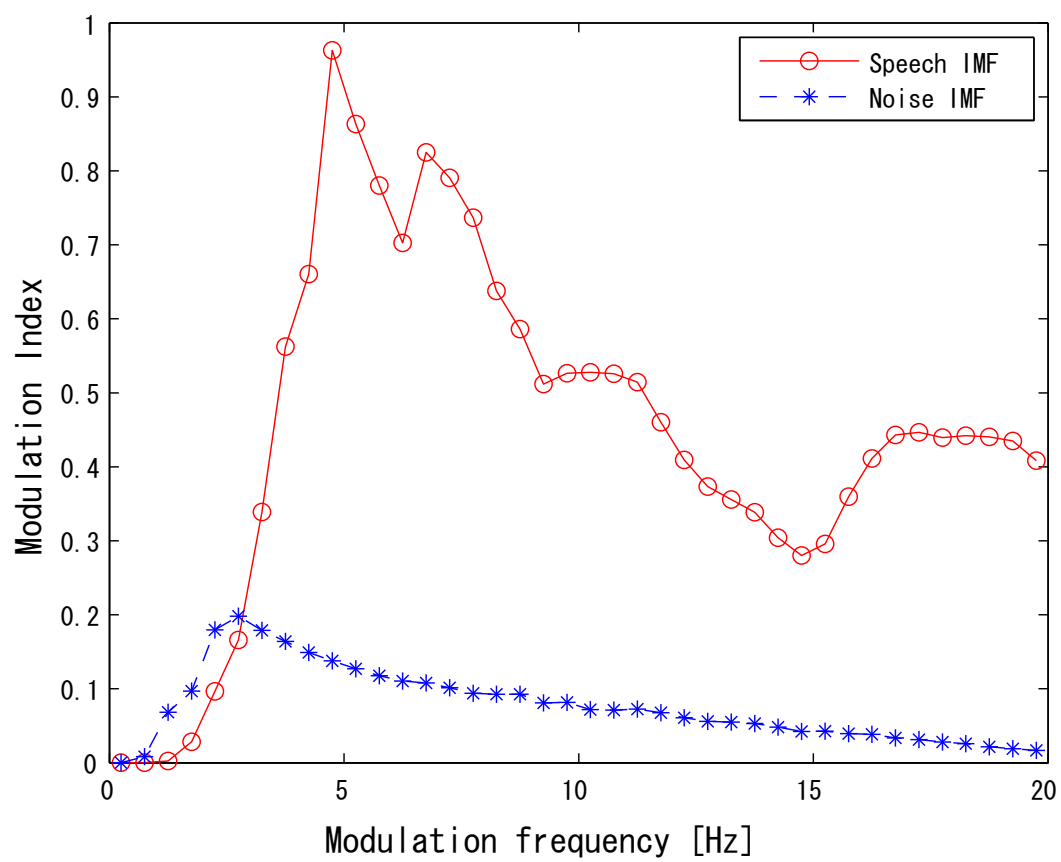


図 4.4: 音声 IMF と雑音 IMF の変調周波数特性

第5章 評価シミュレーション

5.1 シミュレーションの条件

提案手法の有効性を検証するために、雑音除去の評価実験を行った。使用するデータは、ATR 音声データベース a-set [34] 中の 30 種類の音声（男性 5 名、女性 5 名の各 3 音声ずつ）とした。表 5.1 にその音声を示す。表には、各音声サンプルにおける音韻とラベル（頭文字の "m" は男性を、"f" は女性を表している）を記載する。

音声信号に付加する雑音は、音声区間における信号対雑音比（SNR）が $-5, 0, 5, 10, 15, 20$ [dB] の合計 6 種類の白色雑音とした。

提案法が有効であるかを評価する尺度として、SNR、LSD（Log Spectral Distance）、PESQ（Perceptual Evaluation of Speech Quality）の 3 つを用いた。

本研究では、対象信号の非定常性を考慮しているため、信号の振幅情報だけでなく位相情報も加味して評価可能な SNR を採用した。次式に示すように、雑音音声と回復音声の SNR の改善量 I_{SNR} [dB] とした。

$$I_{\text{SNR}} = 10 \log_{10} \frac{\int \hat{x}^2(t) dt}{\int \hat{n}^2(t) dt} - 10 \log_{10} \frac{\int x^2(t) dt}{\int n^2(t) dt} \quad (5.1)$$

ただし、 $\hat{x}(t)$ を推定した音声信号、 $\hat{n}(t)$ を推定した雑音信号とする。

LSD は、回復信号と原信号のスペクトル距離を測るために利用された。LSD の改善量 I_{LSD} [dB] により評価される。次式に LSD の計算式を示す。

$$\text{LSD} = \sqrt{\frac{1}{\Phi} \sum_{\phi=1}^{\Phi} \left(10 \log_{10}(S(\omega_{\phi})) - 10 \log_{10}(\hat{S}(\omega_{\phi})) \right)^2} \quad (5.2)$$

ただし、 S を音声信号のスペクトル、 $\hat{S}$ を回復音声のスペクトルとする。ここで、雑音音声の LSD を LSD_n として、回復音声の LSD を LSD_r とする。その時の LSD を改善量 I_{LSD} を以下に示す。

$$I_{\text{LSD}} = \text{LSD}_r - \text{LSD}_n \quad (5.3)$$

最後に、音声品質を評価する客観的手法の一つとして、PESQ を用いて評価する。これは、被験者が直接耳で聞いた評価データを統計する MOS（Mean opinion score）を客観的に評価するものであり、人間の感覚に近い評価尺度として利用された。PESQ と MOS に

表 5.1: シミュレーションに使用した音声信号

Sample #	Phoneme	Label
No. 1	/aimai/	mau10014
No. 2	/aNnai/	mau10176
No. 3	/oboeru/	mau10640
No. 4	/gaineN/	mht11005
No. 5	/kudamono/	mht11295
No. 6	/kuchibiru/	mht11300
No. 7	/shikisai/	mms12001
No. 8	/jouyaku/	mms12500
No. 9	/sunawachi/	mms12600
No. 10	/tatoeba/	mmy13003
No. 11	/teNkai/	mmy13401
No. 12	/tojimari/	mmy13510
No. 13	/haNkei/	mn14006
No. 14	/hikkomu/	mn14113
No. 15	/fuutou/	mn14203
No. 16	/akarui/	faf10028
No. 17	/afureru/	faf10113
No. 18	/ijimeru/	faf10223
No. 19	/gaNsho/	ffs11030
No. 20	/kinyuu/	ffs11107
No. 21	/kiraku/	ffs11189
No. 22	/shikumi/	fkm12012
No. 23	/shichigatsu/	fkm12057
No. 24	/shuukai/	fkm12146
No. 25	/youteN/	fkn15017
No. 26	/ryougae/	fkn15120
No. 27	/wagamama/	fkn15207
No. 28	/haNdoru/	fks14018
No. 29	/bariki/	fks14052
No. 30	/hyoumeN/	fks14155

は高い相関があることが知られている．0 から 5 までの値で評価を行い，0 であるほど音質が悪く，5 に近づくほど音質が非常によい．

提案法の比較対象として，Molla & Hirose の雑音除去法 [23] を用いた．

5.2 結果

30 種類の音声サンプルに対して，提案法と Molla & Hirose の雑音除去法を用いて雑音除去を行い，SNR の改善量を求めた．この結果を図 5.1 に示す．二つの手法共に，SNR が改善された．提案法は，低 SNR 時に Molla & Hirose の雑音除去法に比べて，SNR が改善されている．また，Molla & Hirose の雑音除去法では，高 SNR 時に音声信号を削りすぎてしまったため，回復信号を歪ませてしまう．これに対して，提案手法では，音声 IMF と雑音 IMF を判断することで，高 SNR 時において，回復信号を歪ませることなく，雑音除去可能であることが分かった．

次に，LSD の改善量を求めた結果を図 5.2 に示す．二つの手法共に，LSD が改善された．しかし，提案法は，Molla & Hirose の雑音除去法に比べ，LSD 改善量が低い結果となった．

最後に，PESQ の解析結果を図 5.3 に示す．提案手法で回復された音声信号は，雑音音声と同等の音質であることがわかった．これに対して，Molla & Hirose の雑音除去法で回復された音声信号は，高 SNR 時において，回復信号を歪ませてしまい，音質が低下してしまったものと考えられる．

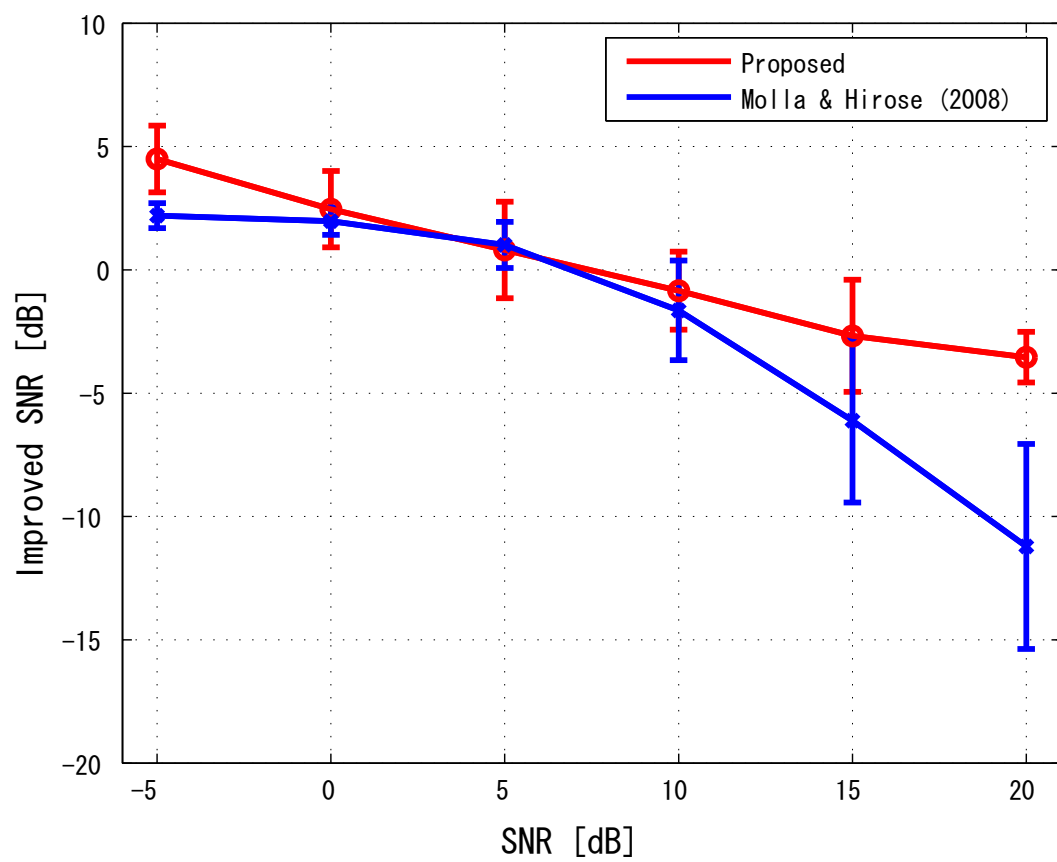


図 5.1: 雑音除去の結果：SNR の改善量

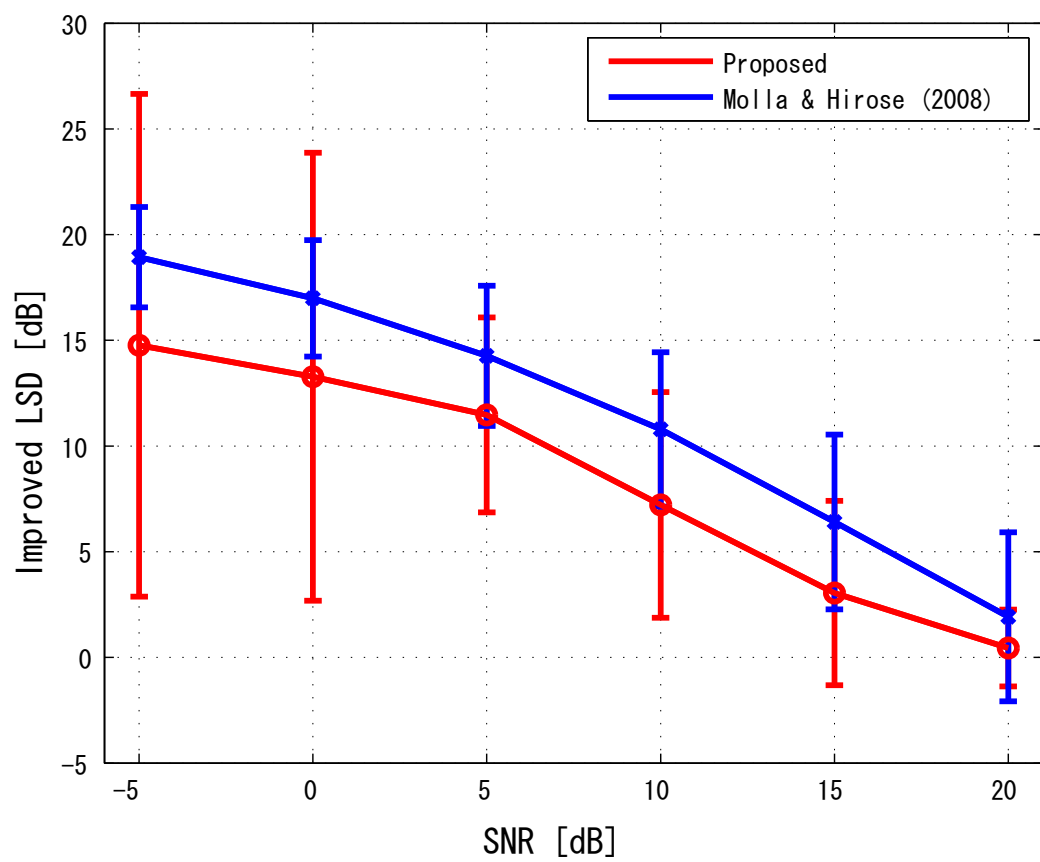


図 5.2: 雑音除去の結果：LSD の改善量

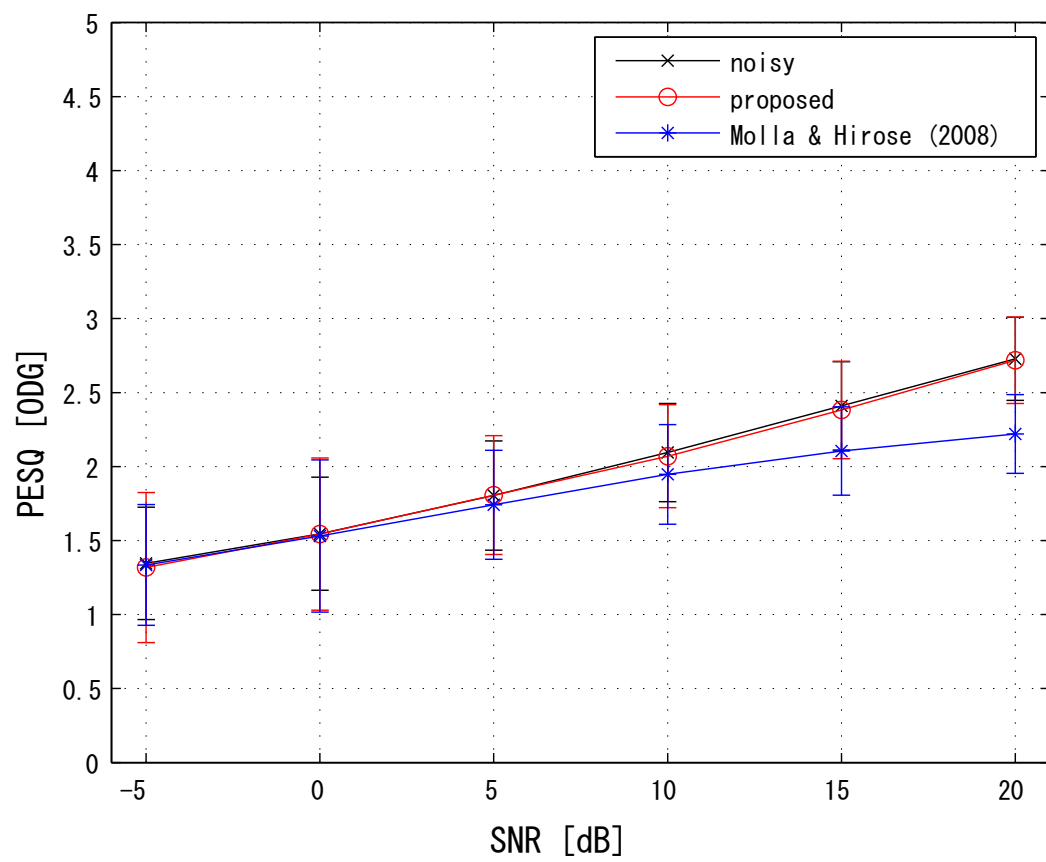


図 5.3: 雑音除去の結果：PESQ

5.3 一音声サンプルにおける分析結果

5.3.1 分析例 1

図 4.1 に示す雑音音声に対し，提案法を利用して雑音除去したときの結果を図 5.4 に示す．また，図 5.5 に，原音声，雑音音声，提案手法で回復された音声，Molla & Hirose の雑音除去法で回復された音声の各短時間スペクトルを示す．雑音音声に比べ，提案法と Molla & Hirose の雑音除去法で回復された音声は原音声のスペクトルに近づいていることが分かった．また，提案法は，Molla & Hirose の雑音除去法で取り除くことが出来なかった中域（2.0～6.0 [kHz]）の雑音成分を強力に取り除くことが出来ていることが分かった．

5.3.2 分析例 2

もう一つの音声サンプルとして，音声信号 $x_2(t)$ に $\text{SNR} = \infty$ [dB] のガウス雑音 $n_2(t)$ の和で構成される雑音音声（この場合， $\text{SNR} = \infty$ [dB] なので，クリーン音声と同等）に対して，同様に解析した結果を示す．図 5.6 に $x_2(t)$ ， $n_2(t)$ ， $y_2(t)$ の信号波形を，図 5.7 に $y_2(t)$ を EMD により分解した結果（最初の 6 本）を示す．図 5.8 に提案法と Molla & Hirose の方法により回復した信号を，図 5.9 にクリーン音声，提案法により回復した信号，Molla & Hirose の雑音除去法で回復した音声の各短時間スペクトルを示す．図 5.7 より，分解した信号は音声信号だけ（ $\text{SNR} = \infty$ なので）であるため，IMF はすべて非定常信号であることが分かる．図 5.8，5.9 より，提案法で回復した信号は，音声の特徴を捉えた分析により，波形を歪ませないことが分かった．それに対して，Molla & Hirose の雑音除去法では，音声の特徴を持つ IMF を減算してしまっているため，音声波形を歪ませていることが分かった．以上の結果から，EMD の信号表現について深く検討した上で，音信号の分析に応用したことで，提案法は音声の性質と，雑音の性質を上手に捉えた分析を実現することができたと言える．

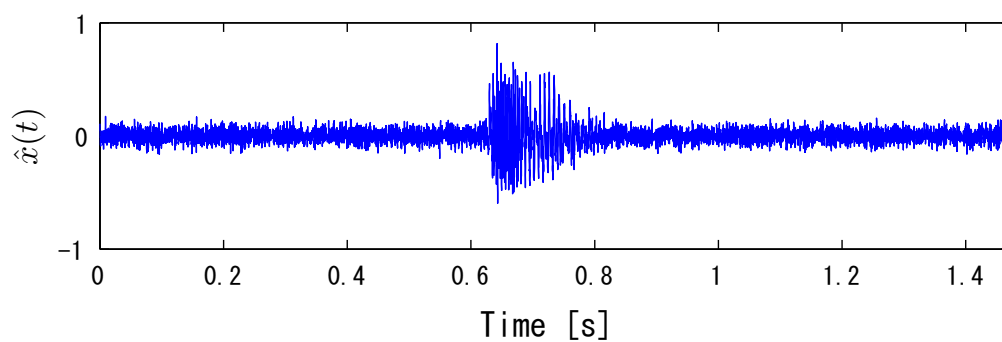


図 5.4: 回復信号 $\hat{x}(t)$

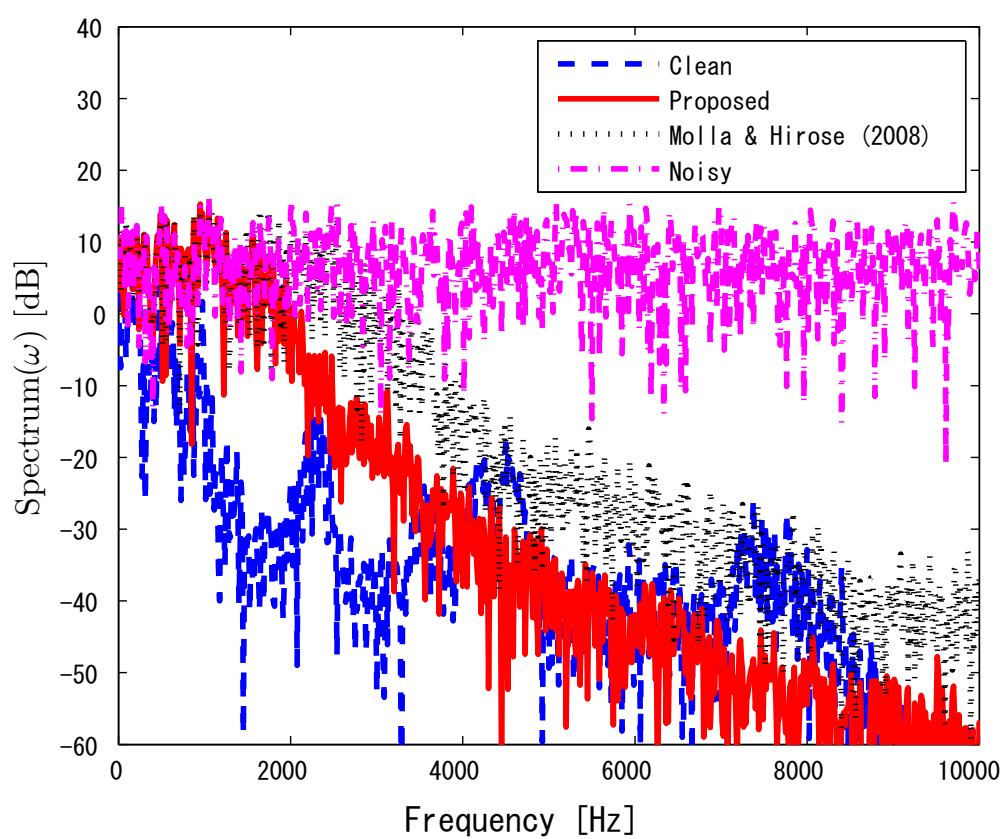


図 5.5: 各信号の短時間スペクトル

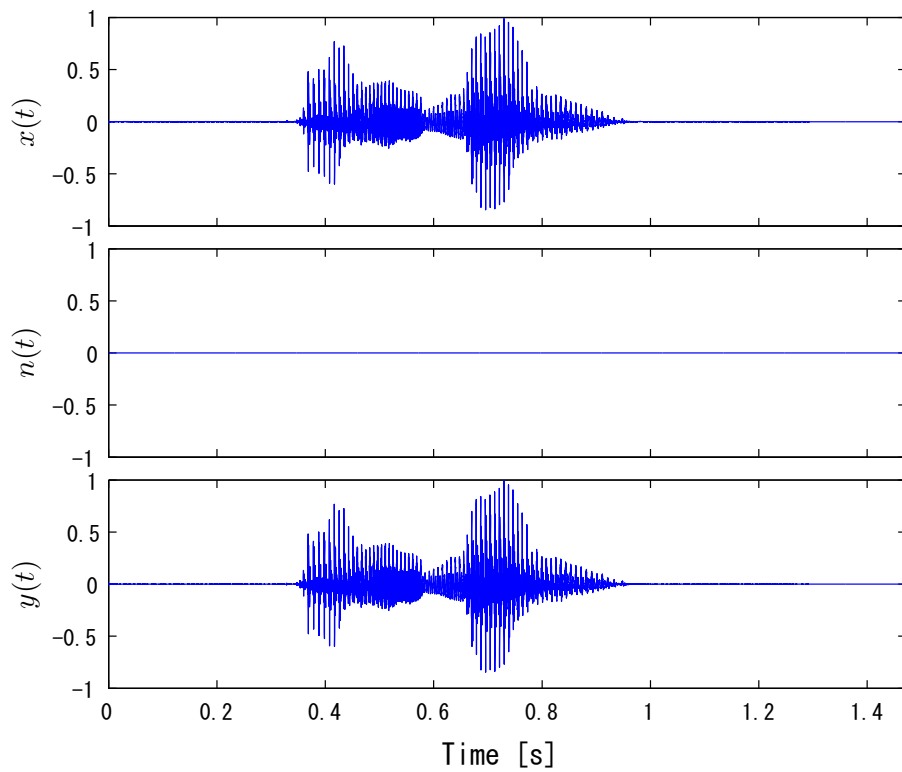


図 5.6: 音声信号 $x_2(t)$ と $\text{SNR} = \infty$ [dB] のガウス雑音 $n_2(t)$ の和で構成される雑音音声 $y_2(t)$

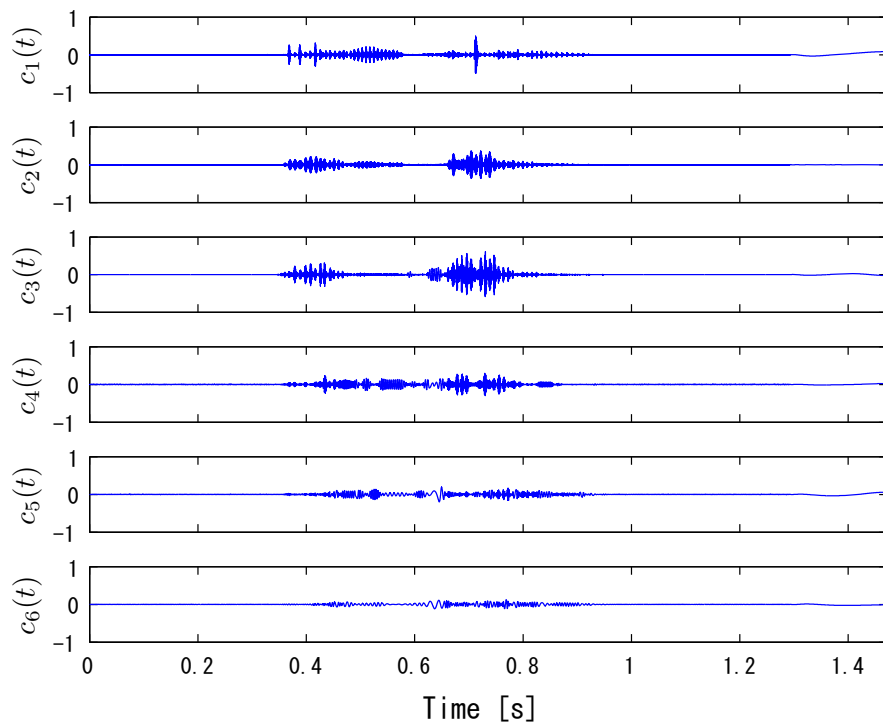


図 5.7: 雑音音声 $y_2(t)$ の分解結果 (最初の 6 個の IMF)

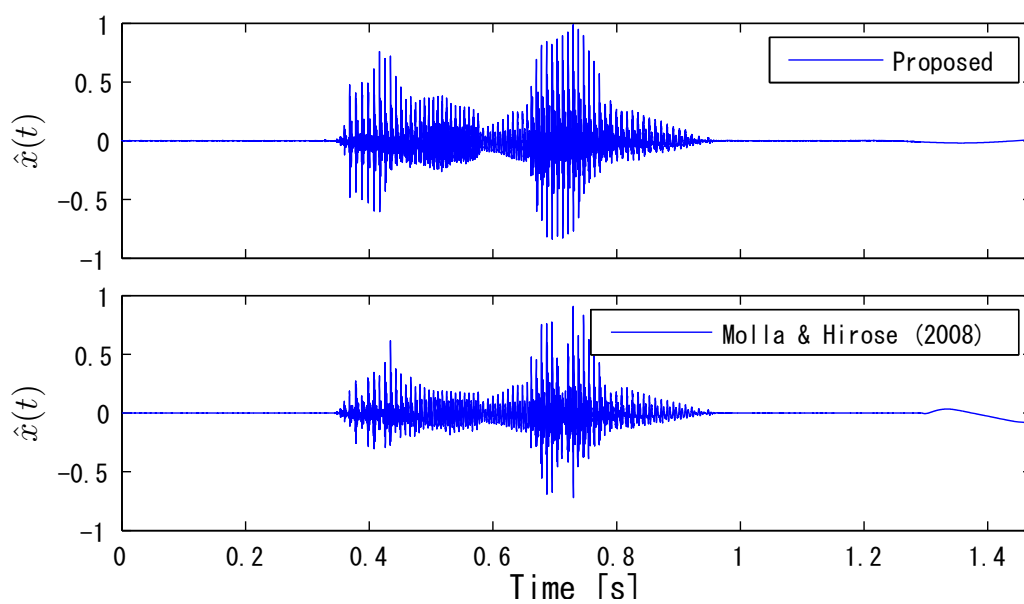


図 5.8: 回復信号 $\hat{x}_2(t)$

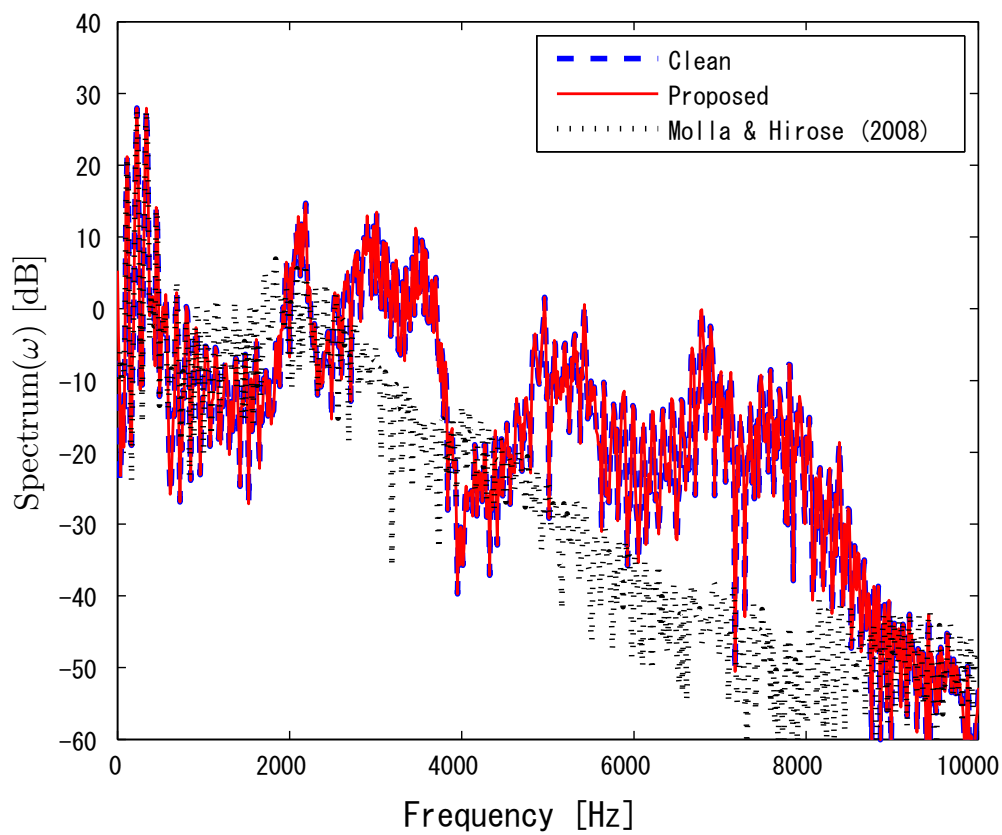


図 5.9: 各信号の短時間スペクトル

5.4 考察

これまで提案されてきた雑音除去法と提案法を含めた形で，EMD を用いた雑音除去法について考察を行う．ここで再び，これまで提案されてきた EMD を用いた雑音除去法について簡単に紹介する．

Taufiq らによって提案されたミュージカルノイズの抑圧法も Molla & Hirose の雑音除去法と同様に，EMD で分解された IMF の初めの数法がミュージカルノイズのエネルギー分布を多く含むことに着目し，この IMF を取り除くことで，抑圧を行っている．

Molla & Hirose によって提案された雑音除去法は，音声のエネルギー分布と雑音のエネルギー分布の違いに着目し，高周波成分を含む IMF を取り除くことで，雑音除去を行っている．

本研究で明らかにした，EMD の信号分析の本質（共通なエンベロープ分解に基づく信号表現を行う方法）に基づくと，Taufiq らによって提案されたミュージカルノイズの抑圧法は，エンベロープが一定であるミュージカルノイズと，音声区間にのみ振幅包絡を持つ非定常な音声信号とに分離することで，雑音除去を実現したと考えることができる．また，Molla & Hirose の雑音除去法，提案法に対しても，同様にエンベロープが一定なホワイトノイズと，音声区間にのみ振幅包絡を持つ非定常な音声とに分離することで，雑音除去を実現したと考えることができる．

これより，これまで紹介してきた EMD を用いた雑音除去はどれも，音声の非定常な変化を上手に捉えた分析を可能にしたものと考えられる．

第6章 結論

6.1 本研究で明らかになったこと

本研究では，非定常信号分析に用いられる経験的モード分解（EMD）の特質について調査した結果及び，EMDによって分解される固有モード関数（IMF）の特徴を調査した結果，以下のことを明らかにした．

- EMDの特質は，EMDの信号分解のプロセスから，被解析信号を振幅包絡が一定な定常信号と，ある区間にのみ振幅包絡を持つ非定常信号とに分けられると考えることができる．
- IMFの特徴は，IMFを定義する二つの制約条件から，帯域幅に制限の無いAM-FM波であることが分かった．

以上のことを明らかにした上で，EMDに基づく音分析の本質は，共通なエンベロープ分解に基づいた信号表現を行う方法であることが分かった．

EMDの分析を最大限活かせる音信号処理として，雑音除去法や，基本周波数推定，フォルマントの推定などが考えられるが，本研究では，雑音除去法の可能性について検討をした．

振幅包絡が一定な定常雑音と，ある区間にのみ振幅包絡を持つ非定常な音声信号とに分ける方法について検討し，提案法が有効であるか検証するために，SNRの改善量，LSDの改善量，PESQの3つの評価尺度を用いた．

結果，SNR，LSD共に改善された．PESQについては，雑音音声に対して，改善されることがなかった．EMDを用いた雑音除去法として，Molla & Hiroseが提案した方法と比較を行うと，Molla & Hiroseの雑音除去法では，高SNR時に，信号を歪ませることがあったが，提案法では，音声の特徴を捉えた分析を行うことで，高SNR時は信号を歪ませず，低SNRの時は有効に雑音を除去することが可能となり，Molla & Hiroseの雑音除去法より，有効性が高くなったことを明らかにした．

これまでに提案されたEMDを用いた雑音除去法と，本研究で明らかにしたEMDの信号表現から，考察を行った．結果，EMDを用いた雑音抑圧は，共通なエンベロープ分解に基づいた信号表現を行うことで，エンベロープが一定な雑音と，音声区間にのみ振幅包絡を持つ非定常な音声とに分解する方法であると解釈することが出来た．

6.2 残された課題

これまで非定常信号を分析する手法として EMD による分析の特徴について調査してきた。EMD による信号分析は一般的に信号を分析する手法（信号の定常性を仮定した分析手法）として用いられている短時間フーリエ変換による分析や、フィルタバンクによる分析、ウェーブレット変換による分析などと、どの程度有効であるかを示す必要がある。これは、瞬時パラメータ（瞬時振幅、瞬時周波数）を推定するシミュレーションにより、分析精度を求めることで確認をする。

6.3 今後の展望

本研究で示した EMD の本質に関する裏付けをとる必要がある。これは、音信号処理の分野だけではなく、EMD を用いた信号分析の全般的な知識として、分析の特徴を示さなければならない。これによって、今後ますます、EMD を用いた信号分析の有用性が高まると考えられる。本質に関する裏付けをとるためには、これまでに提案されてきた EMD を用いた分析法を説明付けることで示さなければならない。

EMD の本質に関する裏付けがとれたあとで、EMD の特質を活かした信号処理について検討する必要がある。本研究で示した EMD の本質（共通なエンベロープ分解に基づいた信号表現）及び、EMD による雑音除去の結果を基に、雑音にロバストな基本周波数の推定の可能性について検討する必要がある。雑音環境下における基本波は EMD によってどのような信号として表現されるか調査することで、基本波の非定常な変化を分析可能にできると考えられる。もう一つのアプローチとして、EMD を用いた基本周波数推定の現状の問題点として基本波付近の帯域幅の不安定になる問題がある。これを、EMD の帯域幅制御 [39, 40] を用い、基本波付近の帯域を安定にさせることで基本周波数推定法の可能性について検討する必要がある。

また、本研究では、EMD の特質が被解析信号を共通なエンベロープごとに分解する信号表現であることを明らかにし、この特質を活かすことで、音信号分析への応用について検討を行ってきた。この特質の他に、音信号を表現する上でどのような信号表現であれば高精度な分析を可能にするかを検討し、それに対する EMD の改良を行った上で、その特質を活かした音信号分析をする方法についても検討する余地があると考えられる。

最後に、従来法である STFT や WT に比べて、EMD は信号の非定常な変化を正確に捉えることができるが、STFT や WT に比べて、計算コストが高く、リアルタイム化を実現するには至っていない。また、計算にかかる時間も、被解析信号に依存するために、一般的に計算時間を示すことが難しい問題がある。この問題についても今後取り組む必要があると考える。

謝辞

本研究を進めるにあたり，多大な助言と懇切丁寧かつ，熱心な御指導をして頂きました鵜木祐史准教授に心から感謝致します．本研究を進めるにあたり，多大な助言と熱心な指導をして頂きました赤木正人教授に心から感謝致します．本研究に関して多大な助言をして頂いた李軍峰助教，宮内良太助教，博士後期課程の羽二生篤氏，木谷俊介氏に心より感謝致します．有意義な討論，助言を賜った赤木・鵜木研究室の皆様方に心から感謝いたします．

参考文献

- [1] 山崎 芳男, “音響信号の時間周波数分析,” 日本音響学会誌, Vol.53, No. 2, pp. 147–153, 1997.
- [2] A. パポーリス, 町田 東一訳, 村田 忠夫訳, アナログとデジタルの信号解析, 現代工学社, 1986.
- [3] 金井 浩, 音・振動のスペクトル解析, コロナ社, 1999.
- [4] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, Henry H. Liu., “The Empirical Mode Decomposition and the Hilbert Spectrum for nonlinear and nonstationary time series analysis,” *Proc. the Royal Society: Math., Physi. & Eng. Sci.*, **A454**, pp. 903–995, 1998.
- [5] 斉藤 佑樹, 田中 聡久, 曹 建庭, “経験的モード分解を用いた準脳死患者の脳波解析,” 電子情報通信学会総合大会論文集, A-4-7, Mar. 2007.
- [6] Hualou Liang, Steven L. Bressler, Robert Desimone, Pascal Fries, “Empirical mode decomposition: a method for analyzing neural data”, *Computational Neuroscience: Trends in Research 2005.*, Vol. 65–66, pp. 801–807, 2005.
- [7] Yan Lu, Jingyu Yan, and Yeung Yam, “Model-based ECG Denoising Using Empirical Mode Decomposition,” *2009 IEEE International Conference on Bioinformatics and Biomedicine*, pp. 191–196, Nov. 2009.
- [8] Bradley Matthew Battista, Camelia Knapp, Tom McGee, and Vaughn Goebel, “Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data,” *Geophysics.*, vol. 72, No. 2, pp. H29–H37, 2007.
- [9] Katie Coughlin and Ka Kit Tung, “11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method,” *Advances in Space Research*, Vol. 34, Issue 2, pp. 323–329, Feb. 2003.
- [10] 石田 皓之, 高橋 友和, 井手 一郎, 村瀬 洋, “経験的モード分解を用いた Hilbert ワーブ法による指動作映像の認識,” 画像の認識・理解シンポジウム (MIRU2008), pp.1162–1168, July 2008.

- [11] Youngseock Lee, Jihah Nah, Jongweon Kim, "Digital Image Watermarking Using Bidimensional Empirical Mode Decomposition in Wavelet Domain," *Multimedia, ISM '09. 11th IEEE International Symposium*, pp. 583–588, Dec. 2009.
- [12] Peter C. Tay, "AM–FM Image Analysis using the Hilbert–Huang Transform," *Image Analysis and Interpretation.*, pp. 13–16, Mar. 2008.
- [13] 深井 寛修 , 滝本 裕則 , 満倉 靖恵 , 田中 聡久 , "経験的モード分解を用いた見た目年齢推定," 信学技報 , vol. 107, no. 413, pp.127–130, Jan. 2008.
- [14] Shaohui Chen, Hongbo Su, Renhua Zhang, Jing Tian, and Lihu Yang, "Improving Empirical Mode Decomposition Using Support Vector Machines for Multifocus Image Fusion," *Sensors 2008.*, Vol. 8, Issue 4, pp. 2500–2508, Apr. 2008.
- [15] Jen-Chun Lee, Ping S. Huang, Chung-Shi Chiang, T-M Tu, and Chien-Ping Chang, "An Empirical Mode Decomposition Approach Iris Recognition," *Image Processing, 2006 IEEE International Conference*, pp. 289–292, Oct. 2006.
- [16] Chia-Ching Lin, Pei-Ling Liu, Po-Liang Yeh, "Application of empirical mode decomposition in the impact-echo test," *NDT & E International*, Vol. 42, No. 7, pp. 589–598, Oct. 2009.
- [17] Tsung-Ying Sun, Chan-Cheng Liu, Jyun-Hong Jheng, Tsung-Ying Tsai, "An Efficient Noise Reduction Algorithm Using Empirical Mode Decomposition and Correlation Measurement," *Intelligent Signal Processing and Communications Systems, 2008. ISPACS 2008. International Symposium.*, pp. 1–4, Feb.2009.
- [18] Vinayak Kamath, Ying-Cheng Lai, Liqiang Zhu, "Empirical Mode Decomposition and Blind Source Separation Methods for Antijamming with GPS Signals," *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, pp. 335–341, Apr. 2006.
- [19] Norden E. Huang, Man-Li Wu, Wendong Qu, Steven R. Long, Samuel S. P. Shen and Jin E. Zhang, "Applications of Hilbert-Huang transform to non-stationary financial time series analysis," *Applied Stochastic Models in Business and Industry*, Vol. 19–361, 2003.
- [20] LingFei Liang, and ZiLiang Ping, "Information Hiding Based On Empirical Mode Decomposition," *Computational Intelligence and Industrial Application, 2008. PACIIA '08. Pacific-Asia Workshop*, Vol. 1, pp. 561–565, Dec. 2008.
- [21] James Z. Zhang, Brant T. Price, Robert D. Adams, Kenneth Burbank, and Theodore J. Knaga, "Hand Tremor Detection via Adaptive Empirical Mode Decomposition and Hilbert–Huang Transform," *The Technology Interface Journal/Winter Special Issue*, Vol. 10, No. 2, ISSN 1523–9926, 2009.

- [22] Hasan, T. and Hasan, M. K., "Suppression of Residual Noise From Speech Signals Using Empirical Mode Decomposition," *IEEE signal process. letters.*, Vol. 16, Issue 1, pp. 2–5, 2008.
- [23] Molla, M. K. I. and Hirose, K., "Robust Voiced/Unvoiced Classification of Speech Signal Using Hilbert-Huang Transformation," *Signal processing*, Vol. 12, No. 6, 473–482, 2008.
- [24] 松田 徹也, 広瀬 啓吉, 峯松 信明, "経験的モード分解による主構造抽出を介した雑音環境下における音声信号の基本周波数推定," 電子情報通信学会技術研究報告, vol. 109, No. 57, pp. 49–54, May 2009.
- [25] Gaston Schlotthauer, Maria Eugenia Torres and Hugo L. Rufiner, "A new algorithm for instantaneous f0 speech extraction based on ensemble empirical mode decomposition," *Proc. 17th European Signal Processing Conference (EUSIPCO 2009)*, Aug. 2009(CDROM).
- [26] Navin Chatlani and John J. Soraghan, "Adaptive Empirical Mode Decomposition for Signal Enhancement with Application to Speech," *15th International Conference on Systems, Signals and Image Processing, IWSSIP 2008.*, pp. 101–104, June 2008.
- [27] Md. Khademul Islam Molla, Keikichi Hirose, Nobuaki Minematsu, "Separation of Mixed Audio Signal by Decomposing Hilbert Spectrum with Modified EMD," *IEICE transactions on fundamentals of electronics, communications and computer science*, Vol. 89, No. 3, pp. 723–734, Mar. 2006.
- [28] Bouzid, Aicha, Ellouze, Noureddine, "Empirical mode decomposition of voiced speech signal," *Communications and Signal Processing., First International*, pp. 603–606. 2004.
- [29] Özgür Gültekin, I. Erer, M. Kaplan, "Empirical mode decomposition based denoising for high resolution direction of arrival estimation," *Proc. 17th European Signal Processing Conference (EUSIPCO 2009)*, Aug. 2009(CDROM).
- [30] L. コーエン, 吉川 昭訳, 佐藤 俊輔訳, 時間–周波数解析, 朝倉書店, 1998.
- [31] Boualem Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I–Fundamentals. II–Algorithms and applications", *IEEE Proceedings.*, Vol. 80, pp. 520–538, 1992.
- [32] Jun-Wei Huang and Bernd Milkereit, "Empirical Mode Decomposition Based Instantaneous Spectral Analysis and its Applications to Heterogeneous Petrophysical Mode Construction," *Frontiers + Innovation - 2009 CSPG CSEG CWLS Convention*, pp. 205–210, 2009.

- [33] Haiyong Zhang and Qiang Gai, “Research on properties of empirical mode decomposition method,” *Intelligent Control and Automation.*, Vol. 2, pp. 10001–10004, June 2006.
- [34] Takeda, K., Sagisaka, Y., Katagiri, S. and Kuwabara, H., “A Japanese speech database for various kinds of research purposes,” *J. Acoust. Soc. Jpn.*, Vol. 44 Issue 10, 747–754, 1988.
- [35] Patrick Flandrin, Gabriel Rilling, Paulo Goncalves, “Empirical Mode Decomposition as a Filter Bank,” *Signal Processing Letters, IEEE.*, Vol. 11, Issue 2, Part 1, pp.112–114, Feb. 2004.
- [36] Rob Drullman, Joost M. Festen, and Reinier Plomp, “Effect of temporal envelope smearing on speech reception,” *J. Acoust. Soc. Am.*, Vol. 95, No. 2, pp. 1053–1064. Feb. 1994.
- [37] Greenberg, S., Carvey, H., Hitchcock, L. and Chang, S., “Temporal properties of spontaneous speech-a syllable-centric perspective,” *Phonetics.*, Vol. 31 Issue 3–4, 465–485, 2003.
- [38] 荒井 隆行 , 金寺 登 , 吉井 順子 , “変調スペクトルによる雑音下における自動音声区
間検出 –音声周波数帯域及び変調周波数帯域の検討–,” 日本音響学会講演論文集,
1–R–8, Sep. 2009.
- [39] Bo Xuan, Qiwei Xie, and Silong Peng, “EMD Sifting Based on Bandwidth,” *IEEE Signal
Processing Letters.*, Vol. 14, No. 8, Aug. 2007.
- [40] 鷲沢 嘉一 , 田中 聡久 , Danilo Mandic, Andrzej Cichocki, “経験的モード分解における
帯域幅制御,” 第 21 回信号処理シンポジウム講演論文集, A9-4 , Nov. 2006.

学会発表リスト

- T. Sawaguchi and M. Unoki, “A study on a method for analyzing speech signal using the empirical mode decomposition and its applications,” *Proc. 2010 RISP International Workshop on Nonlinear Circuits and Signal Processing*, pp. 345–348, Mar. 2010.
- 澤口知希, 鵜木祐史, “経験的モード分解を用いた音信号分析とその応用の検討”, 日本音響学会 2010 年春季研究発表会, 2-P-18, Mar. 2010.