Wavelet による音声信号の特徴抽出

学生員 森合 正悟* 正員 花崎 泉*

Characteristic Extraction of Speech Signal Using Wavelet

Shogo Moriai*, Student Member, Izumi Hanazaki*, Member

In the analysis-synthesis coding of speech signals, realization of the high quality in the low bit rate coding depends on the extraction of its characteristic parameters in the pre-processing. The precise extraction of the fundamental frequency, one of the parameters of the source information, guarantees the quality in the speech synthesis. But its extraction is difficult because of the influence of the consonant, non-periodicity of vocal cords vibration, wide range of the fundamental frequency, etc..

In this paper, we will propose a new fundamental frequency extraction of the speech signals using the Wavelet transform with the criterion based on its harmonics structure.

キーワード：基本周波数, 分析合成方式, 離散ウェーブレット変換, 調波構造

Keywords: Fundamental frequency, Analysis-synthesis method, Discrete wavelet transform, Harmonics structure

1. はじめに

音声符号化方式の一つである分析合成方式において高品質な音声を実現するには、その前処理における音声信号の特徴パラメータの正確な抽出が必要となる。

分析合成方式は、音声信号に対し音声生成機構の線形分離等価回路モデルに基づいて分析を行い、特徴パラメータを抽出する。特徴パラメータは、音源モデルのパラメータとして有声音/無声音の判別情報、有声音の基本周波数、音源の振幅がある。さらに、声道モデルのパラメータとして線形予測係数がある⁽¹⁾⁽²⁾。分析合成方式はこれらの特徴パラメータを伝送するので、低ビットレートの通信が可能となる反面、音声品質は低くなってしまうという問題がある。音声品質の低下の原因として、音源モデルが極めて簡略化されている、さらには背景雑音による音源情報のパラメータの抽出ミスなどが挙げられる。

そこで我々は、分析合成方式において高品質な音声を実現するために、音源モデルの改善として有声音の基本周波数の抽出に注目した。

音声信号の特徴パラメータの一つである有声音の基本周波数抽出は、音声合成の際の品質に大きな影響を与えるので、いかに正確に抽出するかが重要である。有声音の基本

周波数の抽出は、これまでに比較的単純なアルゴリズムで実現できるケプストラム法、自己相関法、基本波フィルタリング法など数多くの提案がなされてきているが、それぞれ課題が残されており、今なお決定的な抽出法はない⁽³⁾⁽⁴⁾。基本周波数の抽出の困難な原因として、音声信号の非定常性、背景雑音による影響、子音さらにはフォルマントの影響を受け易いといったことが挙げられる。

音声信号の非定常性に対処するためには、音声信号を時間領域と周波数領域の両面から同時にとらえる時間周波数解析が必要となる。時間周波数解析の手法の一つにウェーブレット変換がある。これまでに、有声音の基本周波数の抽出にウェーブレット変換を適用した例が報告されているが、これらの手法はアルゴリズムが複雑になるなど、リアルタイム性を必要とする音声符号化においては実現が困難となってしまう⁽⁵⁾⁽⁶⁾。

そこで本稿では、比較的簡単なアルゴリズムで実現できる、ウェーブレット変換を用いた有声音の基本周波数を抽出方法を提案する。我々はこれまでに、音声信号にウェーブレット変換を適用し、得られた近似関数に対し基本周波数を抽出するという方法を提案している⁽⁷⁾。この方法は、音声信号のスペクトルは周波数領域において基本周波数のところでピークを有するという考えに基づき、ウェーブレット変換をして得られた近似関数に対し、スペクトルの最大ピークを検出して基本周波数を抽出するという方法である。実験の結果、この方法が加法的な雑音にロバストであることを確認している。しかし、検出した最大ピークが基本周波数かどうかの判定基準として、単にパワースペクトルに対

*東京電機大学大学院理工学研究科システム工学専攻

〒350-0394 埼玉県比企郡鳩山町大字石坂

Department of Computers and Systems Engineering, Faculty of Science and Engineering, Tokyo Denki University, Ishizaka, Hatoyama, Hiki-gun, Saitama 350-0394

して閾値を設けただけであったので、対象となる音声信号によっては抽出誤差を生じてしまうという欠点があった。そこで、本稿では新たに音声信号の調波構造による判定基準を採用した有声音の基本周波数の抽出法について述べる。

2. 音声信号処理システムの構成

提案する基本周波数の抽出は、図1に示す分析合成方式の一つのLPCボコーダにおける送信側の一部である。LPCボコーダは、送信側では入力された音声信号に対しフレーム分割を行い、各フレーム区間に対して、

- (1) LPC分析により線形予測係数
- (2) 音源の振幅(フィルタのゲイン)
- (3) フレーム区間の音声の有声音か無声音かを表す有声音/無声音の判別情報
- (4) フレーム区間が有声音であれば基本周波数の推定を行う。

音声信号の特徴パラメータはこれらの4つのパラメータである。また、線形予測係数は、量子化に必要なビット数を低減するために、PARCOR係数などに交換された後に伝送される。受信側では伝送された特徴パラメータから、その区間が有声音ならピッチ周期のインパルス列、無声音なら固定の雑音系列で音源をモデル化した励振ベクトルで駆動し、これを声道フィルタに通すことにより合成音を生成する。

図1 音声符号化システム

Fig. 1. Speech coding system.

図2 基本周波数抽出の流れ

Fig. 2. Flow of extraction of fundamental frequency.

LPCボコーダは図1に示すように、音源モデルが極めて簡素化されており、有声音の基本周波数、有声音/無声音の判別を誤ると極めて音声品質が劣化してしまう。

3. 提案する基本周波数の抽出法

図2に提案するウェーブレット変換と音声信号の調波構造を利用した基本周波数抽出の手順を示す。本手法は、次の4段階から成る。

- (1) 離散ウェーブレット変換による多重解像度解析
- (2) 短時間フーリエ変換
- (3) ピーク検出
- (4) 基本周波数の判定

〈3・1〉 離散ウェーブレット変換による多重解像度解析
ここでは、音声信号に対し離散ウェーブレット変換を行い近似関数を求める。

〈3・1・1〉 多重解像度解析と分解レベルの決定
離散ウェーブレット変換は、連続ウェーブレット変換におけるスケールパラメータ a と時間軸パラメータ b をそれぞれ $2^{-j}, 2^{-jk}$ と離散化すると、信号 $f(t) \in L^2(R)$ を以下のように展開する⁽⁵⁾。

$$f(t) = \sum_{j=1}^L \sum_{k=-\infty}^{\infty} d(j,k) \psi(2^{-j}t - k) + \sum_{k=-\infty}^{\infty} a(L,k) \phi(2^{-L}t - k) \dots\dots\dots (1)$$

$\psi(t)$ はマザーウェーブレット、 $\phi(t)$ はスケーリング関数である。また、 $a(L,k)$ はレベル L の近似係数、 $d(j,k)$ はレベル j の詳細係数であり、それぞれ以下のように定義される。

$$a(L,k) = \frac{1}{\sqrt{2^L}} \int_{-\infty}^{\infty} f(t) \phi(2^{-L}t - k) dt \dots\dots (2)$$

$$d(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t) \psi(2^{-j}t - k) dt \dots\dots (3)$$

また、レベル j の近似関数 $f_j(t)$ は近似係数 $a(j,k)$ を用いて表すと以下のように定義される。

$$f_j(t) = \sum_{k=-\infty}^{\infty} a(j,k) \phi(2^{-j}t - k) \dots\dots\dots (4)$$

さらに、レベル j の詳細部分 $g_j(t)$ はレベル $j-1$ とレベル j との差であり、以下のように示すことができる。

$$g_j(t) = f_{j-1}(t) - f_j(t) = \sum_{k=-\infty}^{\infty} d(j,k) \psi(2^{-j}t - k) \dots\dots\dots (5)$$

よって、レベル j の近似関数は式(6)に示すようにレベル $j+1$ の近似関数 $f_{j+1}(t)$ と詳細部分 $g_{j+1}(t)$ との和で再帰的に示すことができる。

図3 多重解像度解析

Fig. 3. Multiresolution analysis.

$$f_j(t) = f_{j+1}(t) + g_{j+1}(t) \dots\dots\dots (6)$$

ここで、各レベルに対する分割数を解像度とみなせば、レベルが一つ増すごとにその解像度は半分になり、図3のような一連の階層構造が得られる。図3を多重解像度解析という。また、音声信号のサンプリング周波数 F_s をとすると、各レベルのサブサンプリング周波数 F_{sj} は、

$$F_{sj} = F_s / 2^j \dots\dots\dots (7)$$

と表せる。

本研究では、音声信号のサンプリング周波数を 8[kHz] としているので、音声信号の基本周波数の男性平均値 125[Hz] と基本周波数の存在範囲 60~250[Hz] を考慮すると、式(7)から基本周波数はレベル4に存在することになる。よって、基本周波数抽出のレベルは4とする。

〈3・1・2〉 近似関数の計算 離散ウェーブレット変換によりレベル4の近似関数を算出する。また、このとき基本周波数の判定に用いるためにレベル3の近似関数も算出する。

本研究で用いたマザーウェーブレットは、直交基底を作る連続かつサポート・コンパクトな Daubechies 10 である。

〈3・2〉 短時間フーリエ変換 得られたレベル3, 4の近似関数に対し、短時間フーリエ変換を計算する。このとき、ハミング窓で信号の切りだしを行い、以下の仕様に従って短時間ごとのパワースペクトルを求める。

分析窓長 : 37.5[ms]

分析窓周期: 22.5[ms]

〈3・3〉 最大ピークの検出 以下の手順に従って各帯域から図4に示すような最大ピークを検出する。

- (1) レベル3, 4のパワースペクトルの帯域を二分割
レベル3: 250~375[Hz], 376~500[Hz]
レベル4: 60~125[Hz], 126~250[Hz]
- (2) 60~125[Hz]の帯域に対し、スペクトルの最大ピークを検出(第1ピーク F_{4L1} と第2ピーク F_{4L2})
- (3) 126~250[Hz]の帯域に対し、スペクトルの最大ピークを検出(第1ピーク F_{4H1} と第2ピーク F_{4H2})
- (4) 250~375[Hz]の帯域に対し、スペクトルの最大ピークを検出(第1ピーク F_{3L1} と第2ピーク F_{3L2})
- (5) 376~500[Hz]の帯域に対し、スペクトルの最大ピークを検出(第1ピーク F_{3H1} と第2ピーク F_{3H2})

図4 ピーク検出

Fig. 4. Peak detection.

〈3・4〉 基本周波数の判定 検出したレベル4の各帯域のそれぞれのピークに対し、それが基本周波数であるかどうかの判定を行う。

4. 音声信号の調波構造を利用した判定

〈4・1〉 基本周波数抽出に関する問題 前節に示した方法において、音声信号は図5(a)に示すように、周波数領域で基本周波数 F_0 のところでピークを有するはずである。しかし雑音成分が混入した場合、図5(b)に示すように、本来の基本周波数と同等もしくはそれより大きいピークがしばしば生じることがある。この信号に対し、スペクトルの閾値により基本周波数を判定する方法を適用すると、閾値による判定を満足した最初のピークを基本周波数とみなしてしまうので、誤った基本周波数を抽出してしまうことになる⁽⁷⁾。そこで、これらの2つのピークに対して、どちらが基本周波数であるかどうかの判定を行う必要がある。

〈4・2〉 調波構造を利用した判定 前項の判定に関する問題の解決策として、基本周波数の判定に音声信号の調波構造を導入する。音声信号は、基本周波数に対しその2

(a) Without noise

(b) Including additive noise

図5 音声信号のパワースペクトル

Fig. 5. Power spectrum of speech signal.

倍に第2高調波が存在するという原理である⁽⁸⁾。よって、〈3・3〉節で検出したレベル4の各帯域の2つのピークに対して、それが第2高調波を持つかどうかの判定を行うことで、誤った基本周波数の抽出を回避することが可能となる。

〈4・2・1〉 基本周波数判定の手順 レベル4の各帯域のピークに対する音声信号の調波構造を利用した基本周波数の判定は、以下の通りである。

- (1) 60～125[Hz]の帯域に基本周波数 F_0 が存在すれば、126～250[Hz]の帯域において第2高調波が存在するはずである。
- (2) 126～250[Hz]の帯域に基本周波数 F_0 が存在すれば、251～375[Hz]または376～500[Hz]の帯域において第2高調波が存在するはずである。

検出したレベル4の各ピークの判定は、60～125[Hz]の帯域の最初的基本周波数の候補である F_{4L1} から行う。 F_{4L1} はまずその第2高調波の候補 F_{4H1} を用いて判定がかけられる。判定式は以下のように示される。

$$|F_{4H1} - 2 \times F_{4L1}| \leq k_1 \quad \dots\dots\dots (8)$$

ただし、 k_1 は判断誤差である。もし、これらの条件が満たされれば基本周波数 F_0 を以下のように決定する。

$$F_0 = F_{4L1} \quad \dots\dots\dots (9)$$

逆に条件が満たされなければ、図6に示すように以下同様の判定を繰り返す。もし、すべての条件が満たされなければ基本周波数 F_0 は0となる。

〈4・2・2〉 判断誤差の決定 ここで、判断誤差 k_1 の決定について述べる。本研究では、サンプリング周波数8[kHz]で2048ポイントのFFTを行っているので、周波数領域において隣合うスペクトルのピークの間隔は約3.9[Hz]となる。よって、基本周波数の候補の2倍とその第2高調波の候補の2倍との誤差が ± 3.9 [Hz]以内であれば、それらが基

図6 基本周波数の判定
Fig. 6. Criterion of fundamental frequency.

本周波数と高調波の関係にあると判断できる。従って、本研究では判断誤差 k_1 を3.9[Hz]と設定している。

5. ウェーブレットの導入の必要性

ここでは、基本周波数の抽出におけるウェーブレットの導入の必要性について示す。

従来、音声信号は非定常性信号であり、背景雑音による影響、子音さらにはフォルマントの影響により基本周波数の抽出は困難となってしまう。これらの理由から、ウェーブレットの導入を考える。ウェーブレットを導入することの利点として、以下のことが挙げられる。

- (1) 原信号に含まれる雑音をある程度除去できる
- (2) 多重解像度解析により各レベルはオクターブバンドに分解されるので、基本周波数の存在する範囲で解析できる
- (3) ホルマントの影響を取り除いて、声帯音源信号を直接取り出せる

〈5・1〉 (1)の詳細 雑音の付加した信号に対し、離散ウェーブレット変換を行い、分解・再構成してみる。図7(a)に原信号を、図7(b)に離散ウェーブレット変換による分解・再構成の結果を示す。

図7から、ウェーブレット変換により、雑音の除去がある程度可能であることが分かる。

〈5・2〉 (2)の詳細 本システムでは、多重解像度解析で得られた近似関数の基本周波数抽出対象レベルを4と設定している。つまり、〈3・1〉節の式(7)から、レベル4の

(a)Original signal

(b)Result of decomposition and reconstructed

図7 雑音の除去の様子

Fig. 7. State of the removal noise.

図8 原信号とそのレベル4の近似関数

Fig. 8. Original signal and approximation function at level 4.

近似関数のサブサンプリング周波数は500[Hz]となり、基本周波数の存在範囲である60~250[Hz]の範囲で解析ができるので、より正確な基本周波数の抽出が可能となる。

〈5・3〉 (3)の詳細 図8(a)はある区間の音声信号、図8(b)はレベル4の近似関数を示している。このように、離散ウェーブレットの多重解像度解析により分解・再構成されたレベル4の近似関数は、音声信号の周期性が明確になり、声帯音源信号を直接取り出すことが可能である。

6. 実験

本章では、以下に示した方法による抽出結果を示す。

- (1) 基本周波数の判定基準をスペクトルの閾値のみとした場合の提案法とケプストラム法との抽出結果の比較
- (2) 基本周波数の判定基準をスペクトルの閾値のみとした場合と調波構造による判定基準を採用した場合の、判定方法の違いによる抽出結果の比較
- (3) 調波構造による判定基準を採用した場合の提案法とケプストラム法との抽出結果の比較

〈6・1〉 基本周波数の判定基準をスペクトルの閾値のみとした場合 ここでは、以前に我々が提案した方法、つまり基本周波数の判定基準をスペクトルの閾値のみとした場合の提案法とケプストラム法との基本周波数の抽出結果の比較を示す。

〈6・1・1〉 音声資料 使用した音声試料は表1、図9に示す音声信号である。また、雑音の影響を見るために、この信号にS/N比それぞれ5[dB], 0[dB]の白色雑音を付加した音声も資料に加える。

〈6・1・2〉 基本周波数抽出結果 雑音なしの場合の結果である図10では、両者とも基本周波数を忠実に抽出している。また、S/N比が5[dB]の場合の結果である図11では、提案法は図10の場合とほぼ同じ結果であるのに対

表1 音声試料

Table 1. Speech sample

Sampling frequency [kHz]	8
Speech time [s]	1.0
Speaker	Male

図9 音声試料

Fig. 9. Speech sample.

(a) Proposed method

(b) Cepstrum method

図10 基本周波数抽出結果 (雑音なし)

Fig. 10. Result of extraction of fundamental frequency (without noise)

し、ケプストラム法では、基本周波数に乱れが生じており、無声音区間まで誤った基本周波数を抽出している。さらに、図 12 では、ケプストラム法では雑音の影響によりさらに基本周波数に乱れが増し、誤った基本周波数を抽出している。しかし、提案法では若干基本周波数の抽出に乱れがあ

図 11 基本周波数抽出結果 (S/N 比 5[dB])
Fig. 11. Result of extraction of fundamental frequency (S/N rate 5[dB])

図 12 基本周波数抽出結果 (S/N 比 0[dB])
Fig. 12. Result of extraction of fundamental frequency (S/N rate 0[dB])

るものの、ケプストラム法と比べ、精度良く基本周波数を抽出している。

〈6・2〉 基本周波数の判定基準の違いによる影響
ここでは、基本周波数の判定基準をスペクトルの閾値のみとした場合と、調波構造による判定基準を採用した場合の、判定方法の違いによる抽出結果の比較を行う。

〈6・2・1〉 音声資料
使用した音声試料は表 2、図 13 に示す音声信号である。

〈6・2・2〉 基本周波数抽出結果
図 14 (a) は、基本周波数の判定基準がスペクトルの閾値のみの場合で、音声信号によっては誤った基本周波数を抽出してしまうという例である。これは、子音の影響によって図 14 (a) に見られるように、本来の基本周波数部分以外で最大ピークを検出してしまったためである。一方、レベル 4 の帯域からいくつかのピーク、つまり基本周波数の候補を検出し、そのピークに対して調波構造を利用した基本周波数の判定を行った図 14 (b) は、基本周波数の判定が強化されているために、基本周波数を忠実に抽出していることがわかる。

〈6・3〉 基本周波数の判定基準に調波構造を採用した場合
ここでは、基本周波数の判定基準に調波構造を利用した提案法に対し、雑音の影響を確かめるために、従来法であるケプストラム法との比較を行いながら基本周波数抽出の結果を示す。

〈6・3・1〉 音声資料
使用した音声試料は表 3、図 15(a)、図 15(b) に示す音声信号である。

〈6・3・2〉 基本周波数抽出結果
提案法による基本周波数抽出結果は、図 16、図 18 の (a)、(c) となった。提案法は、S/N 比が低くても雑音にロバストであることが確認できる。一方、ケプストラム法の結果である図 17、図 19 の (a)、(c) から、ケプストラム法は S/N 比が低くなるにつれ雑音の影響を受け、本来無音区間の部分にも誤った基本周波数を抽出していることがわかる。

表 2 音声試料
Table 2. Speech sample

Sampling frequency [kHz]	8
Speech time [s]	0.6
Speaker	Male

図 13 音声試料
Fig. 13. Speech sample

図 14 基本周波数抽出結果

Fig.14. Result of extraction of fundamental frequency

表 3 音声試料

Table 3. Speech sample

No	1	2
Sampling frequency [Hz]	8	8
Speech time [s]	0.6	0.7
Speaker	Male 1	Male 2

図 15 音声試料

Fig.15. Speech sample

図 16 提案法 (音声 1)

Fig.16. Proposed method (Speech 1)

さらに、これらの基本周波数のパラメータを用いて LPC ボコーダで音声合成した結果、提案法は図 16、図 18 の (b)、(d) という結果になった。実際にこの音声聞いたところ、すべての音声において音声の意味は十分聞き取れることが確認された。一方、ケプストラム法の結果は図 17、図 19

(a) Fundamental frequency (without noise)

(b) Synthesized speech (without noise)

(c) Fundamental frequency (S/N rate 5[dB])

(d) Synthesized speech (S/N rate 5[dB])

図 17 ケプストラム法 (音声 1)

Fig. 17. Cepstrum method (Speech 1)

の (b), (d) となった。雑音なしの場合の結果である図 17, 図 19 の (b) は、音声の意味は聞き取れることが確認されたが、S/N 比 5[dB] の結果である図 17, 図 19 の (d) と聞いたところ音声の意味を聞き取ることが困難であった。

符号化の音声品質の評価には、実際に合成音を聴取して

(a) Fundamental frequency (without noise)

(b) Synthesized speech (without noise)

(c) Fundamental frequency (S/N rate 5[dB])

(d) Synthesized speech (S/N rate 5[dB])

図 18 提案法 (音声 2)

Fig. 18. Proposed method (Speech 2)

評価を行う主観評価が有効であるが、参考までに客観評価の一つである S/N 比による評価を表 4 に示す。

また、LPC ボコーダの遅延時間は約 20[ms] である。通常、LPC ボコーダの基本周波数の決定は AMDF (average magnitude difference function) が用いられているが、本

図 19 ケプストラム法 (音声 2)

Fig. 19. Cepstrum method (Speech 2)

表 4 合成音の S/N 比

Table 4. S/N rate of synthesized speech

Speech No	Proposed method [dB]	Cepstrum method [dB]
1 (without noise)	12.5255	10.5613
1 (S/N rate 5[dB])	7.5144	2.7558
2 (without noise)	8.7497	8.2457
2 (S/N rate 5[dB])	5.4534	2.2865

研究では基本周波数の抽出に提案法を用いている。これは、比較的簡単なアルゴリズムであるので、遅延時間が 20[ms] 以上となることはない。

7. おわりに

本稿では、比較的簡単なアルゴリズムで実現できるウェーブレット変換を用いた有声音の基本周波数の抽出法を提案した。実験より、ウェーブレット変換を用いた抽出法はケプストラム法よりも加法的な雑音の影響を受けにくいことを確認した。

しかし、ウェーブレット変換を用いた抽出法⁽⁷⁾では、閾値のみによる基本周波数の判定であるため、対象となる音声信号によっては誤った基本周波数を抽出してしまうという問題を指摘し、新しく音声信号の調波構造による判定基準を採用した方法を提案した。実験の結果から、調波構造による判定基準を採用することにより、問題を解決できることを確認した。

さらに、抽出した基本周波数を用いて LPC ボコーダにより音声合成した結果、提案法は雑音が混入しても音声の意味が聴き取れるほどの性能を有していることを確認した。

今後は、基本周波数抽出に付随し有声音/無声音の判別が行えるよう改良を加え、分析合成方式における更なる音源部分の改善を行う必要がある。

(平成 13 年 12 月 4 日受付, 同 14 年 6 月 17 日再受付)

文 献

- (1) 古井貞照: デジタル音声処理, 東海大学出版会, (1985)
- (2) 森谷健弘: 音声符号化, 電子情報通信学会, (1998)
- (3) H. Ohmura and K. Tanaka: "Fine pitch contour extraction by voice fundamental wave filtering method", *ASJ*, Vol. 51, No. 7, pp. 509-518 (1995-7) (in Japanese)
大村 浩・田中和世: 「基本波フィルタリング法による精細ピッチパターンの抽出」, 日本音響学会誌, 51, 7, pp. 509-518 (1995-7)
- (4) N. Kunieda, T. Shimamura, and J. Suzuki: "Pitch Extraction by Using Autocorrelation Function on the Log Spectrum", *IEICE Trans.*, Vol. J80-A, No. 3, pp. 434-443 (1997-3) (in Japanese)
國枝伸行・島村徹也・鈴木誠史: 「対数スペクトルの自己相関関数を利用したピッチ抽出法」, 信学論誌, J80-A, 3, pp. 435-443 (1997-3)
- (5) S. Kadambe and G. F. Boudreaux-Bartels: "Application of the wavelet transform for pitch detection of speech", *IEEE Trans. Inf. Theory*, Vol. 38, No. 2, pp. 917-924 (1992-3)
- (6) A. Satou and S. Nakamura: "A Pitch Extraction Method Using Wavelet Transform", *IEICE Trans.*, Vol. J80-A, No. 11, pp. 1848-1856 (1997-11) (in Japanese)
左宗 晃・中村尚五: 「ウェーブレット変換を用いたピッチ抽出の方法」, 信学論誌, J80-A, 11, pp. 1848-1856 (1997-11)
- (7) S. Moriai and I. Hanazaki: "Pitch Detection of Speech Signals Using Wavelet Transform", 2001 National Convention Record, IEE Japan (in Japanese)
森合正悟・花崎 泉: 「Wavelet 変換を用いた基本周波数抽出について」, 平成 13 年電気学会全国大会
- (8) J. Li, T. Shimamura, and S. Suzuki: "Fundamental Frequency Extraction of Degraded Speech Using AR Spectral Estimation Method", *IEICE Technical Report*, SP94-86, pp. 7-14 (1995) (in Japanese)
李 金中・島村徹也・鈴木誠史: 「AR スペクトル推定法を用いた低 SN 比音声の基本周波数の抽出」, 信学技報, SP94-86, pp. 7-14, (1995)

森 合 正 悟 (学生員) 2002 年 3 月東京電機大学大学院理工学研究科システム工学専攻修士課程修了。同年 4 月, 三菱電機 (株) に入社, 電力産業システム事業所に配属, 現在に至る。在学中は主として, 音声信号処理の研究に従事。

花 崎 泉 (正員) 1986 年 3 月早稲田大学大学院博士課程修了。早稲田大学助手, 小山高専講師などを経て, 1991 年東京電機大学講師、現在助教授。主として, 確率システム理論, 信号処理などの研究に従事。計測自動制御学会, システム制御情報学会, ファジイ学会 IEEE の各会員。