ウェーブレットを用いたLSP分析予測残差信号の符号化

正員 清水 忠昭* 非会員 木本 雅也*
 正員 吉村 宏紀* 非会員 井須 尚紀*
 正員 菅田 一博*

Coding Method of LSP Residual Signals Using Wavelets for Speech Synthesis

Tadaaki Shimizu*, Member, Masaya Kimoto*, Non-member, Hiroki Yoshimura*, Member, Naoki Isu*, Non-member,
 Kazuhiro Sugata*, Member

This paper presents a method to use wavelet analysis for speech coding and synthesis by rule. It is a coding system where LSP residual signal is transformed into wavelet coefficients. As wavelet analysis is implemented effectively by filter banks, our method is featured to require less computation than multipulse coding and others where complicated prediction procedures are essential. To achieve good quality speech at low bit rates, we verified to allocate the different bits onto the wavelet coefficients, with more bits in lower frequencies, and less in higher. The synthesized speech by Haar wavelet with 16.538kb/s has nearly same perceptual quality with 6 bits plog PCM (66.15kb/s). We are convinced that coding method of LSP residual signals using wavelet analysis is an effective approach to synthesize speech.

キーワード：規則音声合成，LSP分析，残差信号，ウェーブレット変換

Keywords: speech synthesis by rule, LSP analysis, residual signal, wavelet transform

1. まえがき

規則音声合成の手法として、子音と母音を組み合わせた音声素片を接続することにより任意の音声を生成する単位接続型の合成方式が数多く研究されてきた^{(1)~(9)}。計算機の高性能化に支えられて、合成音声の高品質化を目指した多くの規則音声合成方式は、より多くの記憶容量や計算速度など大規模な計算機資源を必要とするものとなった。一方、少ない情報量で、高品質な音声合成を行うために、麻生らによって音声素片データを記録するパラメータにベクトル量子化を適用する手法が特許出願されている⁽¹⁰⁾。

我々は、小規模な応用に対して高品質な合成音声を与えるために、音声を記録するパラメータとしてLSPパラメータを用い、LSP (Line Spectrum Pair) ベクトルVCV (Vowel Consonant Vowel) 規則音声合成方式の開発研究を続けてきた^{(11)~(14)}。これらの研究においてLSPベクトルVCV規則音声合成方式の有効性を示してきたが、合成音声の自然性は十分とは言えない。現在、数10メガ・バイトから100メガ・

バイト以上の波形辞書を用いて、波形重畳によって自然性に優れた音声合成手法が提案されている。しかし、このような方法は、本研究で目指しているような小規模な分野には適用不可能である。

LSPベクトルVCV規則音声合成方式で合成音声の自然性が損なわれる主な原因は、音声合成にLSP分析合成系を用いていることである。分析合成系を用いる手法でも、疑似音響フィルタの駆動に残差信号を直接用いれば高い自然性を保てるが、情報量は波形を直接用いる方法と同等になる。記憶する情報量の増加を最小限に抑えて符号化し、合成音声の品質の上で十分な精度で復号化できる高能率な残差信号符号化の手法を考案できれば、上記の問題の多くが解決する。本論文では、この問題を解決する一手法として、ウェーブレット変換を用いて残差信号を符号化する方式を検討した。

近年ではウェーブレット変換は音声信号に適用され様々な成果をあげている⁽¹⁵⁾。音声の基礎的な分野では、音声の駆動源の不連続性の検出と特徴づけのためにウェーブレット変換を用いる研究が行われている⁽¹⁶⁾。音声認識の分野では、音声信号の時間一周波数解析にウェーブレット変換を用いる研究⁽¹⁷⁾などが行われている。また、符号化の分野では、ウェーブレット変換を用いて高品質のオーディオ信号

* 鳥取大学工学部
 〒 680-8552 鳥取市湖山町南4丁目101番地
 Faculty of Engineering, Tottori University
 4-101, Koyama-cho Minami, Tottori city, 680-8552

の情報圧縮を行う研究⁽¹⁸⁾や、ウェーブレット変換とベクトル符号化を組み合わせる音声信号の符号化を行う研究⁽¹⁹⁾などが行われている。ウェーブレット変換の音声分野への適用は、基礎的な分野から音声認識や符号化まで広がりをみせている。

ウェーブレット変換はフィルタ・バンクによって行うことができ、複雑な予測等の手続きを必要としないため、本方式の符号化は、マルチパルス符号化⁽²⁰⁾などに比べ少ない計算量で容易に実現できる。本研究では、単純なHaarのウェーブレットを用い、LSP分析における残差信号をウェーブレット係数に分解した場合の各ウェーブレット係数の量子化特性を調べた。その結果、合成音声のスペクトル歪を小さく抑えながら情報圧縮を行うためには、高周波数のウェーブレット係数よりも低周波数のウェーブレット係数に多くのビットを割り当てる方が有利であることが判明した。シミュレーション実験では、残差信号のウェーブレット係数を 11.025 kbit/s および 16.538 kbit/s のビットレートで線形量子化を行って合成音声を作成した。その結果、十分な了解度を持ち、パルス駆動などに比べて自然性の高い合成音声を生成できることが判った。

2. ウェーブレット変換を用いた予測残差信号の符号化

1 段の離散ウェーブレット変換(DWT: Discrete Wavelet Transform)と離散逆ウェーブレット変換(IDWT: Inverse Discrete Wavelet Transform)のブロック図を図1に示す。

図1 離散ウェーブレット変換と逆変換
Fig. 1. One cell DWT and IDWT.

レベルjの信号 c_j は、フィルタ $H(z)$ と $G(z)$ により、高周波成分と低周波成分に分解される。これらの信号成分を1/2にダウン・サンプリングすることにより、レベルj-1のウェーブレット係数である d_{j-1} と c_{j-1} が得られる。本研究では、ウェーブレットの中で最も単純なHaarのウェーブレットを用いた。従って、変換と逆変換に用いるフィルタは式(1)の伝達関数で表される。

$$\begin{aligned} H(z) &= \frac{1-z}{2} & G(z) &= \frac{1+z}{2} \\ Q(z) &= 1-z^{-1} & P(z) &= 1+z^{-1} \end{aligned} \quad \text{..... (1)}$$

本方式では、式(1)のフィルタ群を用いて、LSP分析の残差信号をウェーブレット係数に分解して符号化する。以後、残差信号をウェーブレット変換した係数 d_j 、 c_j をレベルjの残差ウェーブレット係数(RW coefficients: Residual Wavelet Coefficients)と呼ぶ。

本方式による分析合成の手順を図2に示す。本方式の符号化においては、まず音声信号をLSP分析しLSPパラメータと残差信号を算出する。LSPパラメータは適切なビット割り当てを行い伝送するが、この点についてはすでに詳しい研究⁽²¹⁾⁽²²⁾があるので本論文では扱わない。残差信号は、ウェーブレット変換することによって求めた残差ウェーブレット係数の形で符号化し伝送する。音声信号の復号は、残差ウェーブレット係数を逆ウェーブレット変換することによって再生した残差信号で、LSP合成フィルタを駆動することによって行う。以下の章では、本方式による残差ウェーブレット係数の符号化特性について実験的に検証した結果を報告する。

3. 残差ウェーブレット係数の性質

〈3・1〉 音声資料 残差ウェーブレット係数の性質を調べるために、(財)日本情報処理開発協会発行の研究用連続音声データベース中の男性話者の音素バランス文を用いた。以下の実験では、音声資料のサンプリング周波数を

図2 ウェーブレット分析を用いた予測誤差の符号化法
Fig. 2. Block diagram of DWT coder for RW coefficients.

11.025 kHzに変換してから切り出したVCV素片209個、約40秒分を用いた。LSP分析は、12次で分析フレーム長320サンプル、分析インターバル256サンプルで行った。

〈3・2〉 残差ウェーブレット係数のパワー 音声資料から切り出した/sa/の子音部から母音立ち上がり部分の音声信号をLSP分析し、その残差信号を本手法でウェーブレット変換した例を図3に示す。残差ウェーブレット係数は、ダウン・サンプリングの手続きにより、レベルが下がるほど時間軸が粗くなる。図3では波形を見やすくするために縦軸のスケールを適当に変化させている。実際には、残差ウェーブレット係数はレベルが下がるほど平均的な振幅が小さくなっている。無声子音/s/の部分と有声音である母

図3 残差信号のウェーブレット分析例
Fig. 3. Wave form of RW coefficients.

図4 残差ウェーブレット係数のパワー
Fig. 4. Power of RW coefficients.

音/a/の部分とを相対的に比べると、高いレベルのウェーブレット係数ほど無声子音/s/の部分で強調されている。

実験に用いた209個、約40文分のVCV素片から得られた残差信号を有声音部と無声音部に分けて、ウェーブレット変換を行い、残差ウェーブレット係数の平均パワーを求めた。図4は、残差ウェーブレット係数のパワーを元の残差信号のパワーで正規化した結果である。有声音部、無声音部ともにレベルが下がるとウェーブレット係数のパワーは小さくなる。この傾向は無声音部の方がやや強い。

〈3・3〉 残差ウェーブレット係数へのビット割り当て誤差

残差ウェーブレット係数に対するビット割り当ての目安を得るために、線形量子化した場合について、各残差ウェーブレット係数に割り当てるビット数がスペクトル歪に与える影響を調べた。残差ウェーブレット係数 $d_1, d_2, \dots, d_5, c_5$ の各々の1サンプルあたりに10ビットを割り当てた時の合成音声を基準合成音声として、 $d_1, d_2, \dots, d_5, c_5$ のうち一つの残差ウェーブレット係数への割り当てビット数だけを1サンプルあたり8ビットから0ビットまで変えた合成音声の平均スペクトル歪を求めた。平均スペクトル歪は、式(2)で定義する。

$$SD = \sqrt{\frac{1}{N} \sum_n \frac{1}{W} \sum_f \left\{ \hat{S}_n(f) - S_n(f) \right\}^2} \dots\dots\dots (2)$$

ここで、

- $\hat{S}_n(f)$: 評価対象となる合成音声の対数スペクトル
- $S_n(f)$: 基準合成音声の対数スペクトル
- n : 分析フレーム番号
- N : 分析フレーム総数
- f : 周波数
- W : 分析周波数帯域

図5は、各残差ウェーブレット係数の1サンプルあたりのビット割り当て量と平均スペクトル歪の関係を表している。有声音、無声音のどちらの場合も、高いレベルの残差ウェーブレット係数のビット割り当て量を削減した場合の方が低いレベルの残差ウェーブレット係数のビット割り当て量を削減した場合よりも、平均スペクトル歪は大きくなる。この傾向は、無声音の方が強く、特に d_1 のビット割り当て量を削減したときの平均スペクトル歪の増大は顕著である。

一方、ウェーブレット変換の手続きの中でダウン・サンプリングを行っているため、単位時間あたりのサンプル数は低いレベルの残差ウェーブレット係数の方が少ない。具体的な数値をあげると、音声信号32サンプルあたり(約2.9msあたり)の残差ウェーブレット係数のサンプル数は d_1 は16サンプル、 d_2 は8サンプル、 d_3 は4サンプル、 d_4 は2サンプル、 d_5, c_5 にはそれぞれ1サンプルである。このため、高いレベルの残差ウェーブレット係数に対するビット割り当て量を削減する方が、低いレベルの残差ウェーブレット係数に対するビット割り当て量を削減する

よりも、単位時間あたりの情報量削減の効果は大きい。例えば、 d_{-1} の1サンプルあたりのビット割り当てを1ビット減らすと、音声信号32サンプルあたり16ビットの削減になるが、 d_{-5} の場合には音声信号32サンプルあたり1ビットの削減にしかない。

図6は、各残差ウェーブレット係数に対するビット割り当ての変更により音声信号32サンプルあたり削減されるビット数を横軸にして、合成音声の平均スペクトル歪をプ

ロットしたものである。図6から、スペクトル歪を大きくせずに低ビットレート化するためには、高周波数成分を受け持つ高いレベルのウェーブレット係数に少ないビットを割り当てる方が有利であることが判る。この結果は、図5による先の議論に矛盾しているようにも思えるが、その原因は、高いレベルのウェーブレット係数におけるビット削減の情報量削減効果が平均スペクトル歪の増大の効果を上回ったためである。

図5 残差ウェーブレット係数へのビット割り当て量と合成音声の平均スペクトル歪
Fig. 5. Relation between Spectral Distortion and Quantization Bits.

図6 ビット削減量と合成音声の平均スペクトル歪
Fig. 6. Relation between Spectral Distortion and Reductive Bit Rate.

4. 提案手法の評価

〈4・1〉 音声資料 提案した音声符号化手法の性能評価のために、聞き取りによる主観評価実験を行った。主観評価実験には、(財)日本情報処理開発協会発行の研究用連続音声データベースの音素バランス文の中から選んだ10個の短文の読み上げ発話を用いた。短文の読み上げは、男声2名、女性2名の計4名の話者によって行い、合計40発話を主観評価実験の音声資料とした。発話の長さはいずれの短文も2秒から3秒程度であり、音声資料の発話時間の合計は約107秒である。音声資料は、サンプリング周波数を11.025 kHz、量子化数16ビットでA/D変換し、次数12次、分析フレーム長320サンプル、分析インターバル256サンプルでLSP分析を行った。

〈4・2〉 残差ウェーブレット係数へのビット割り当て 残差ウェーブレット係数全体の符号化ビットレートを11.025 kbits/sと16.538 kbits/sとして、各レベルの残差ウェーブレット係数へのビット割り当てを行った。スペクトル歪みを評価基準とする前章の評価実験により、合成音声品質を保ってビット削減を行うには、図6に示される通り高いレベルの残差ウェーブレット係数へのビット割り当てを削減する方が良いといえる。この結果に従って、下記の条件のもとにビット割り当てを行った。

1) 高いレベルの残差ウェーブレット係数へのビット割り当てが、より低いレベルの残差ウェーブレット係数へのビット割り当てを超えないこと。

2) 1つのレベルの残差ウェーブレット係数へのビット割り当てを最大10ビットとすること。

これら2つの条件を満たすビット割り当ての全ての種類を表1で網羅した。表1に示すビット割り当てにより、〈4・1〉節で述べた音声資料を符号化・復号化した際に生じる歪みを、式(2)によるスペクトル歪みSDを用いて評価した結果を表1の最右欄に示す。同じビットレートでは、ス

ペクトル歪みにそれほど大きな差はない。

〈4・3〉 主観評価実験 表1に示す全てのビット割り当てに対して主観評価実験を行う事は、評価対象が多すぎて現実的でない。そこで、研究グループ内で予備的な聞き取り実験を行い、ビットレート別に最も評価が良かったビット割り当てについて主観評価実験を行った。予備実験で選択されたビット割り当ては、11.025 kbits/sでは表1中のC、16.538 kbits/sでは表1中のdである。

予備実験により選択した提案手法C、dと、4ビット、6ビット、8ビットの μ log-PCMの計5種類の符号化法の間で一対比較法による主観評価実験を行った。一対比較の組み合わせは、資料提示順序の前後の入れ替えを含めて $5 \times 4 = 20$ 通りである。〈4・1〉節で述べた40発話の音声資料全てにつ

図7 主観評価実験による品質尺度値
Fig. 7. Perceptual quality of decoded speech.

表1 ビット割り当て
Table 1. Bit allocation for RW coefficients.

Bit Rate (kbits/sec)		Bit Allocation (bits/sample)						SD (dB)
		d_{-1}	d_{-2}	d_{-3}	d_{-4}	d_{-5}	c_{-5}	
11.025	A	0	0	0	6	10	10	11.84
	B	0	0	2	2	10	10	9.65
	C	0	0	0	8	8	8	11.95
	D	0	0	2	4	8	8	9.40
	E	0	0	2	6	6	6	9.41
	F	0	0	4	4	4	4	9.19
16.538	a	0	0	4	6	10	10	9.16
	b	0	0	4	8	8	8	9.16
	c	0	2	2	4	8	8	7.82
	d	0	0	6	6	6	6	9.25
	e	0	2	2	6	6	6	7.82
	f	0	2	4	4	4	4	7.62

図8 本方式による合成音声の波形例
Fig. 8. Synthesized speech by our method.

いて、一対比較の組み合わせ20通りを作成し、合計で40発話×20通り＝800対の比較対を主観評価実験に用いた。

主観評価実験では、被験者の疲労を考慮して、40対の比較対を1セットとし、1セットごとに休憩を挟みながら比較対の提示を行った。1回の実験では、4セット＝1クールを被験者に提示した。1人の被験者に対して5クール＝20セットの実験を行った。

比較対の提示順序の影響を避けるために、被験者への比較対の提示順は次の条件を満たした上でランダムとした。

1) 1セット内に一対比較の組み合わせ20通りが全て出現する。

2) 1セット内で男声と女声が同数である。

3) 5クール内に比較対の重複がなく、1人の被験者が800対の比較対を全て評価する。

被験者には、電子音の合図に続いて、比較対をなす資料を交互に2回ずつ提示し、「いずれの音声の雑音が少ないか。」という基準で評価を行わせた。この際、中間の判断を許さず、必ず比較対のいずれかの資料を選択させた。

主観評価実験は、健康で聴覚に異常のない20代の被験者8名により行った。1セットの比較対の提示にかかる時間は約12分であり、3分間の休憩をはさみながら、1クール約1時間で実験を行った。1つのクールの実験が終了した後、次のクールの実験までには最低2時間の十分な休憩時間を挟んだ。被験者は、全員が合計で5クール、約5時間の評価実験に参加した。

一対比較法による主観評価実験の結果に、Thurstoneの比較判定の法則のケースVを適用して、合成音声の品質尺度値を求めた。図7は、4ビット μ log-PCMによる符号化音声を基準として計算した品質尺度値を示している。提案手法C、dともに、4ビット μ log-PCMに比べて品質尺度値が1.0以上高い。品質尺度値が1.0高いことは、約84%の割合で比較対象よりも品質が優れていると判断されることを示している。一方、提案手法C、dともに、8ビット μ log-PCMに比べて品質尺度値が1.0以上低い。提案手法C、dは、品質尺度値によると、4ビット μ log-PCMと8ビット μ log-PCMの中間の品質を持ち、6ビット μ log-PCM(66.15 kbits/s)に近い品質であると結論できる。16.538 kbits/sのdのビット割り当てを行った場合の合成音声の波形の例を図8に示す。

5. まとめ

LSP分析の残差信号をウェーブレット変換を用いて符号化する手法を検討した。本方式では、単純なHaarウェーブレットを用いた場合でも、高品質な合成音声を得ることができることを一対比較による主観評価実験により示した。

本手法により残差信号を11.025 kbits/sから16.538 kbits/s程度のビットレートで符号化した場合、6ビット μ log-PCM(66.15 kbits/s)に近い高品質な合成音声を得ることができた。この際、合成音声のスペクトル歪を小さく抑えるためには、高いレベルのウェーブレット係数よりも低いレベル

のウェーブレット係数に多くのビットを割り当てる方が有利であることが判明した。

本稿の議論では、各残差ウェーブレット係数に対して有声音、無声音とも同じビット割り当てを行った。しかし、(3・2)節の結果から有声音と無声音とで各残差ウェーブレット係数に対するビット割り当てを変えることで、より高品質な合成音声を得ることができると考えられる。また、適応的なビット割り当てを行うことで、より低ビットレートでの符号化が期待できる。

(平成14年2月8日受付、平成14年12月2日再受付)

文 献

- (1) C. Fruichi, and S. Imai: "Speech Synthesis-by-Rule Based on the Concatenation of CV-Syllable Mel-Cepstral Parameters", *Trans. of IEICE*, Vol.J67-D, No.2, pp.1356-1363 (1984-2) (in Japanese)
古市千枝子・今井 聖: 「CV音節のメルケプストラムパラメータの接続に基づく音声の規則合成」, 信学論D, **J67-D**, 2, pp.1356-1363 (1984-2)
- (2) Y. Arai: "LSP-CV Speech Synthesis by Using the CV-syllable Positioning Rule", *Trans. of IEICE*, Vol.J70-A, No.5, pp.836-843 (1987-5) (in Japanese)
新居康彦: 「CV音節配置規則を用いたLSP-CV規則音声合成」, 信学論(A), **J70-A**, 5, 836-843 (1987-5)
- (3) T. Minowa and Y. Arai: "The Japanese CV-syllable positioning rule for speech synthesis", *Proc. IEEE-IECEJ-ASJ, ICASSP*, 86, pp.2031-2034 (1986)
- (4) 伏木田勝信・三留幸夫・佐伯 猛: 「ホルマントCV-VC方式による規則型音声合成システム」, 情報処理学会第31回全国大会講演論文集, pp.1107-1108 (1985)
- (5) H. Sato: "Speech Synthesis on the Basis of PARCOR-VCV Concatenation Units", *Trans. of IEICE*, Vol.J61-D, No.11, pp.858-865 (1984-11) (in Japanese)
佐藤大和: 「PARCOR-VCV連鎖を用いた音声合成方式」, 信学論D, **J61-D**, 11, pp.858-865 (1978-11)
- (6) H. Sato: "Speech Synthesis Using CVC Concatenation Units and Excitation Waveform Elements", *Trans. of the Committee on Speech Research*, S83-69, pp.541-546 (1984) (in Japanese)
佐藤大和: 「CVCと音源要素に基づく(SYMPLE)音声合成」, 日本音響学会音声研究会資料, S83-69, pp.541-546 (1984)
- (7) M. Ichikawa, K. Iwata, Y. Mitome, and K. Fushikida: "Phoneme Sequence Unit for Speech Synthesis by Rule", *IEICE Technical Report*, SP87-6, pp.41-48 (1987) (in Japanese)
市川昌子・岩田和彦・三留幸夫・伏木田勝信: 「規則合成における単位音声セットの検討」, 信学技報, SP87-6, pp.41-48 (1987)
- (8) K. Takeda, K. Abe, and Y. Sagisaka: "Speech Synthesis by Rule Based on Adaptive Unit Selection", *Trans. of IEICE*, Vol.J73-D-II, No.12, 1945-1951 (1990-12) (in Japanese)
武田一哉・安部勝雄・匂坂芳典: 「選択的に合成単位を用いる規則音声合成」, 信学論(D-II), **J73-D-II**, 12, pp.1945-1951 (1990)
- (9) T. Koyama, and N. Koizumi: "Speech Synthesis By Rule Based On VCV Waveform Synthesis Units", *IEICE Technical Report*, SP96-8, pp.53-60 (1987) (in Japanese)
小山貴夫・小泉宣夫: 「VCVを基本単位とする波形規則合成方式の検討」, 信学技報, **SP96-8**, pp.53-60 (1987)
- (10) 麻生 隆・大洞恭則・藤田 武: 「音声合成方法及びその装置」, 公開特許公報(A), 特開平5-73100 (1993)
- (11) T. Shimizu, H. Yoshimura, H. Nishida, N. Isu, and K. Sugata: "Determination of Dictionary Size and Selection Rule for Speech Synthesis based on LSP Vector VCV Method", *T. IEE Japan*, Vol.119-C, No.8/9, pp.1060-1067 (1999-8/9) (in Japanese)
清水忠昭・吉村宏紀・西田博充・井須尚紀・菅田一博: 「LSPベクトルVCV規則音声合成方式のための合成単位素片数と素片選択法」, 電学論誌C, **119**, 8/9, pp.1060-1067 (1999-8/9)
- (12) T. Shimizu, H. Yoshimura, Y. Sumida, N. Isu, and K. Sugata: "Speech Synthesis by VCV Method based on Vector Quantization of LSP parameter", *T. IEE Japan*, Vol.120-C, No.3, pp.420-427 (2000-3) (in Japanese)
清水忠昭・吉村宏紀・隅田庸市・井須尚紀・菅田一博: 「LSPパラメータにベクトル量子化を適用した小規模応用のためのVCV規

- 則音声合成], 電学論誌 C, **120**, 3, pp.420-427 (2000-3)
- (13) T. Shimizu, H. Yoshimura, M. Kimoto, T. Namiki, N. Isu, and K. Sugata: "Relation between phonemic environmental index and connective distortion index on VCV speech synthesis", *T. IEE Japan*, Vol.121-C, No.3, pp.681-688 (2001-3) (in Japanese)
清水忠昭・吉村宏紀・木本雅也・並木寿枝・井須尚紀・菅田一博: 「VCV規則音声合成における音韻環境指標と接続歪み指標の関係」, 電学論誌 C, **121**, 3, pp.681-688 (2001-3)
- (14) T. Shimizu, H. Yoshimura, T. Namiki, N. Isu, and K. Sugata: "Coding scheme of residual signal to improve the clearness of explosive consonant for LSP-VCV speech synthesis", *T. IEE Japan*, Vol.122-C, No.2, pp.224-231 (2001-2) (in Japanese)
清水忠昭・吉村宏紀・並木寿枝・井須尚紀・菅田一博: 「LSP-VCV規則音声合成における破裂子音の明瞭性向上のための残差信号の符号化」, 電学論誌 C, **122**, 2, pp.224-231 (2002-2)
- (15) T. Irino: "Speech signal processing using wavelet transform", *IEICE Technical Report*, SP92-81, DSP92-66, pp.59-66 (1992) (in Japanese)
入野俊夫: 「Wavelet変換による音声信号処理」, 信学技報, SP92-81, DSP92-66, pp.59-66 (1992)
- (16) H. Kawahara, and T. Irino: "Evaluation of speech excitation events by wavelet transform", *IEICE Technical Report*, SP91-46, pp.25-32 (1991) (in Japanese)
河原英紀・入野俊夫: 「Wavelet変換による音声の駆動源の特徴付けについて」, 信学技報, SP91-46, pp.25-32 (1991)
- (17) Y. Katayama, A. Miyazaki, and H. Sakoe: "Wavelet Analysis of Speech Signals and its Applications", *IEICE Technical Report*, DSP94-12, RCS94-3, pp.17-24 (1994) (in Japanese)
片山喜規・宮崎明雄・迫江博昭: 「ウェーブレットによる音声信号の解析とその応用について」, 信学技報, DSP94-12, RCS94-3, pp.17-24 (1994)
- (18) Deepen Sinha, Ahmed H. Tewfik: "Low Bit Rate Transparent Audio Compression using Adapted Wavelets", *IEEE Transactions on Signal Processing*, Vol.41, No.12, pp.3463-3479 (1993)
- (19) E. Mandridake, M. Najim: "Joint Wavelet Transform and Vector Quantization for Speech Coding", *IEEE Int Symp Circuits Syst*, Vol.1, pp.699-702(1993)
- (20) S.W. Lee, Y. Takahashi, and Y. Kurahashi: "A Multi-Pulse speech coding system using a method for approximate synthesis of transition segment including unvoiced consonants", *Trans. of IEICE*, Vol.J77-A, No.3, pp.293-302 (1994-3) (in Japanese)
李 時雨・高橋 寛・倉橋 裕: 「無声子音を含む遷移区間の近似合成法を用いたマルチパルス音声符号化システム」, 信学論 A, **J77-A**, 3, pp.293-302 (1994-3)
- (21) N. Sugamura, and F. Itakura: "Line spectrum representation of linear predictor coefficients of speech signal and its statistical properties", *Trans. of IEICE*, Vol.J64-A, No.4, 323-330(1981-4) (in Japanese)
管村昇・板倉文忠: 「線形予測係数の線スペクトル表現とその統計的性質」, 信学論 A, **J64-A**, 4, pp.323-330 (1981-4)
- (22) N. Sugamura, and F. Itakura: "Speech data compression by LSP speech analysis-synthesis technique", *Trans. of IEICE*, Vol.J64-A, No.8, pp.599-606 (1981-8) (in Japanese)
管村 昇・板倉文忠: 「線スペクトル対(LSP) 音声分析合成系による音声情報圧縮」, 信学論 A, **J64-A**, 8, 599-606 (1981-8)

清水 忠 昭

(正員) 1963年生。1987年3月 大阪大学基礎工学部生物工学科卒業。同年鳥取大学工学部助手。2002年鳥取大学工学部助教授, 博士(工学)。音声信号処理, ニューラルネットワークの研究に従事。電子情報通信学会, 日本音響学会の各会員。

木 本 雅 也

(非会員) 1977年生。2002年3月 鳥取大学大学院工学研究科博士前期過程修了。2002年 鳥取大学大学院工学研究科博士後期過程在籍。音声信号処理の研究に従事。情報処理学会の会員。

吉 村 宏 紀

(正員) 1970年生。1998年3月鳥取大学大学院工学研究科博士後期課程修了。同年, 九州工業大学情報工学部リサーチアソシエイト。1999年大阪府立大学工学部助手。現在, アイコンヤマト株式会社勤務博士(工学)。ニューラルネットワーク, 音声信号処理, デジタル信号処理の研究に従事。電子情報通信学会の会員。

井 須 尚 紀

(非会員) 1953年生。1978年3月 大阪大学大学院基礎工学研究科前期課程修了。同年航空宇宙技術研究所研究員。1989 福井大・工・電子助教授。1993 鳥取大・工・知能情報助教授。医学博士。動揺病の生理工学・中枢神経系の生理学などの研究に従事。情報処理学会, 神経科学学会, 生理学会, 宇宙航空環境医学会の各会員。

菅 田 一 博

(正員) 1938年生。1966年4月 京都大学大学院工学研究科博士課程修了。同年 同大電気助手。1971年大阪大学基礎工学部助教授。1986年鳥取大学工学部教授。工学博士。しきい値理論, オートマトンと言語理論, 計算の複雑さ, 音声信号処理, ニューラルネットワークの研究に従事。電子情報通信学会, 情報処理学会, 日本音響学会, 人工知能学会, 日本農業気象学会の各会員。