

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

國枝 伸行[†] 島村 徹也[†] 鈴木 誠史[†]

Pitch Extraction by Using Autocorrelation Function on the Log Spectrum

Nobuyuki KUNIEDA[†], Tetsuya SHIMAMURA[†], and Jouji SUZUKI[†]

あらまし 音声の基本周波数の抽出（ピッチ抽出）は、音声処理の幅広い分野で必要とされているが、決定的な手法はいまだないのが現状である。本論文は、対数スペクトルの自己相関関数を利用した新しいピッチ抽出法（ACLOS）を提案する。その原理は、音声の対数スペクトルの自己相関関数を計算することによって、スペクトル上における高調波の周期（基本周波数）を抽出するものである。音声の SN 比を変えてピッチ抽出を行った結果、本手法は従来法に比べピッチ抽出誤りである gross pitch error を改善することができた。特に SN 比が -5 dB 以下の低 SN 比音声に対しては、自己相関関数法よりも gross pitch error を 3.4~7.2% 改善することができ、本方式が雑音に強いことを示す。

キーワード ピッチ抽出, 基本周波数, 自己相関関数, 対数スペクトル

1. ま え が き

音声の基本周期/基本周波数の抽出（ピッチ抽出）は、分析合成系や音声強調、音声認識、話者認識のほか、音声言語、歌声、唝声の分析などの幅広い分野で必要とされている。これまでにさまざまな抽出法が検討されてきたが、その精密さ、耐雑音性、信頼性等のすべての条件を満たすものはいまだないと言える [1], [2]。近年、抽出精度の向上や時間応答の向上、耐雑音性、高速化を目指した研究が多く見られる [3]~[13]。しかし、雑音によって品質劣化した音声に対するピッチ抽出法では、フィルタバンク [6]~[9] やパターンマッチング [9]~[11] などの複雑な処理を必要とするものが多い。

従来のピッチ抽出法は、時間波形に対する処理、相関関数による処理、スペクトル領域での処理に大別される。時間波形に対する零交差数や波形ピークを利用する手法 [15], [16] は、処理量が少なく実時間処理が可能だが、品質劣化した音声に対してはフィルタバンクを用いる [8] などしないと抽出が期待できない。また、相関処理を利用する手法 [5], [17], [18] は比較的単純なアルゴリズムで位相の変化や雑音に対しても強い

というメリットがあるが、ホルマントの影響を受けやすく、周期雑音に弱いという欠点が指摘されていた。

一方、スペクトル領域で処理を行うピッチ抽出法としては、高調波の公約数のヒストグラムを作成する手法 [19]~[21] およびケプストラム法 [11]~[14] が代表的である。これらの手法は、スペクトル上で波形を観測すると高調波のピークが周期的に現れる（この周期が基本周波数に一致する）ことに注目したものである。ケプストラム法は、自己相関関数とは逆にホルマントの影響に強いが、雑音に弱いという性質をもつ。ケプストラム法が雑音に弱い原因として、対数スペクトルの周期を逆 FFT を利用して求めるために基本周期に対応するピークがインパルス状であることが挙げられる。

ところで、このスペクトル上における高調波の周期を得るためには、自己相関関数を利用して検出することも可能である。自己相関関数においては、波形の周期が周期関数の極大値として観測されると同時に波形自身の周期性も強調されるため [22]、雑音の影響が少なくなると考えられる。そこで本論文は、音声の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（ACLOS: AutoCorrelation function on the Log Spectrum）を提案し、その性能について評価する。本方式は、ケプストラム法における逆 FFT の代わりに自己相関関数を利用する方式とみなすことが

[†] 埼玉大学工学部情報システム工学科, 浦和市
Faculty of Engineering, Saitama University, Urawa-shi, 338
Japan

できる。すなわち、ケプストラム法と自己相関法を組み合わせてことによって、雑音により robust なピッチ抽出を実現するのが本方式である。

スペクトル領域に自己相関関数を導入したピッチ抽出法としては、Lahat ら [6] が試みている。この手法では音声のパワースペクトルの自己相関関数を利用している。しかしながら、この方法では振幅差やホルマントの影響を取り除くためにフィルタバンクを利用しており、処理が煩雑になっていた。本手法では、周期性がより明確に現れる対数スペクトルに対して相関処理を行うため、比較的単純な処理によってピッチ抽出が実現できる。

本論文では、まず 2. で提案法の原理について述べる。3. で実音声についてピッチ抽出を行い、従来のケプストラム法および自己相関関数法と比較評価する。更に、音声の SN 比を 20 ~ -10 dB に変化させてピッチ抽出を行い、その抽出誤りおよび抽出誤差についての検討を行う。最後に 4. で提案法の有効性を確認し、結論をまとめる。また、本方式の可能性についても述べる。

2. 提案法の原理

2.1 抽出原理

図 1(a) に示すように、有声音の対数スペクトルには調波構造が現れる。すなわち、対数スペクトルを波形としてみると周期波形が観測できる。この周期性は、リフタ処理によってホルマントの影響を取り除くことによって明確になる（図 1 で使用したリフタの特性については 3.2 で述べる）。周波数軸上におけるこの周期が基本周波数に一致し、この周期を検出するのがスペクトル領域におけるピッチ抽出法の基本原理である。

ケプストラム法は、対数スペクトルの周期を逆 FFT によって求める手法である。ところが、ケプストラム法では、この周期を示す波形ピークがインパルス状に現れるため、雑音に弱いという欠点があった。本手法では、ケプストラム法における逆 FFT の代わりに自己相関関数を利用してこの周期を求める。自己相関関数による周期検出は雑音に robust であるから、ケプストラム法よりも雑音に強いピッチ抽出が実現できる。

この対数スペクトル $S(i)$ に対する自己相関関数は次式で定義される。

$$r(j) = \frac{1}{N} \sum_{i=0}^{N-1} S(i)S(i+j) \quad (1)$$

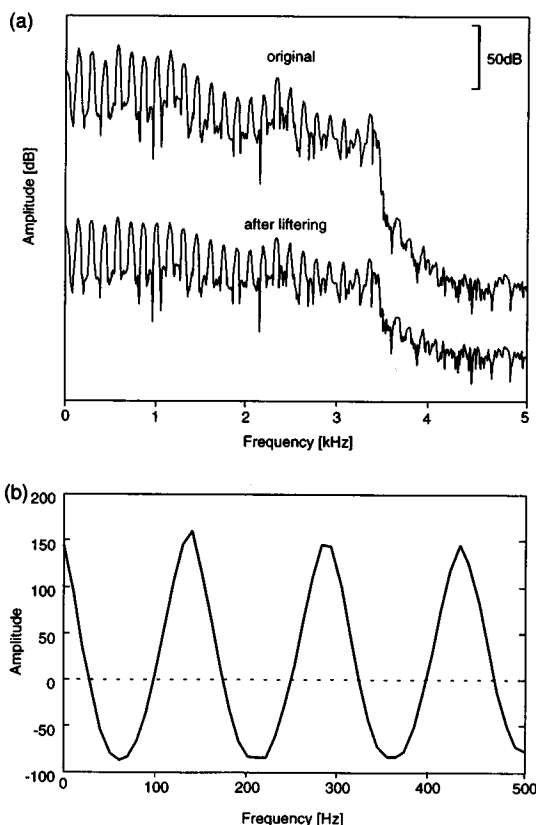


図 1 音声の対数スペクトルとその自己相関関数 (a) リフタ処理前後の対数スペクトル (FIR リフタ使用), (b) 対数スペクトルの自己相関関数

Fig. 1 Log spectrum of speech and its autocorrelation function. (a) Log spectrum of before and after lifting (FIR lifter). (b) Autocorrelation function of log spectrum.

$$j = 0, 1, \dots, M$$

ここで $S(i)$ と $r(j)$ は、ともに周波数についての離散信号を表す。 N は、自己相関関数の積和計算を行う周波数範囲の上限を決定する定数である。よく知られているように、音声の調波構造は高い周波数領域ほどくずれる性質をもつ。従って相関関数を計算する範囲 N は、調波構造が明確な周波数範囲 (0~2.5 kHz 程度) を考慮して設定する。また M は、自己相関関数 $r(j)$ を計算する j の上限を決定する定数である。この値は、基本周波数の存在する範囲 (50~400 Hz 程度) を示す $r(j)$ を計算するように設定する。

実際にはリフタ処理によってホルマントの影響を取り除いた後に自己相関関数を求める。図 1 (b) に音声の対数スペクトルの自己相関関数を示す。このとき $r(j)$

は、 j が基本周波数 F_0 の整数倍となるときに極大値をとる。従って、この極大値をとる最小の j を検出すれば、基本周波数を求めることができる。これが提案法の原理である。

ところで、音声信号のケプストラムを $c(k)$ としたとき、対数スペクトル $S(i)$ の自己相関関数 $r(j)$ は、ケプストラムの2乗 $c^2(k)$ のフーリエ変換であると考えることができる。従って、本方式とケプストラム法は原理的に一見等価であると考えられよう。しかし、本方式では、以下の特徴によってケプストラム法よりも robust なピッチ抽出を実現する。

(1) 対数スペクトル $S(i)$ の自己相関関数 $r(j)$ を計算する範囲は、調波構造の明確な部分のみを使用し、調波構造の乱れる高周波数部分 (2.5 kHz 以上) は使用しない。

(2) 自己相関関数 $r(j)$ は、調波構造を乱す対数スペクトル上のランダム成分を $r(0)$ 付近に集中させ、対数スペクトルの周期性を強調することができる [22]。

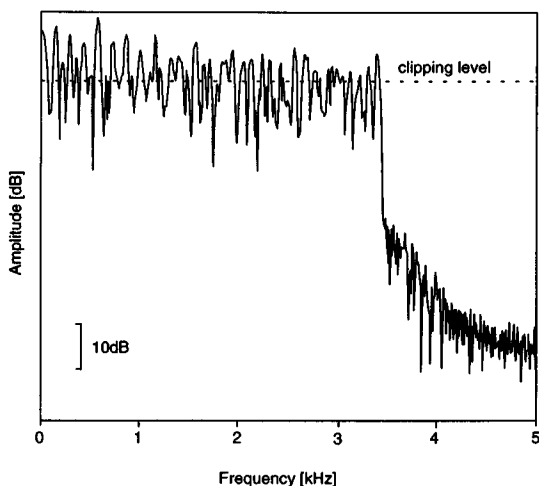


図2 雑音重畳音声のリフタ処理後の対数スペクトル
Fig. 2 Log spectrum of noisy speech after liftering.

(3) ケプストラム法では、周期がインパルス状のピークによって現れるが、提案法ではピッチが周期関数の極大値として現れ、内挿に適している。

すなわち、本方式で使用する対数スペクトルの自己相関関数は、ケプストラムの2乗のフーリエ変換と等価的なものではなく、より周期性が明確に表れるものであることに注目すべきである。

2.2 雑音に対する改良

図2に示すように低 SN 比音声の対数スペクトルを観測すると、雑音によるひずみは調波構造の谷の部分に顕著に現れる。このような場合には、スペクトルの山の部分だけの相関を計算した方が良い結果が得られると考えられる。すなわち、振幅の一定値以下を0とするクリッピングが有効である。

そこで、より雑音に robust なピッチ抽出を実現するために、対数スペクトルのリフタ処理後にクリッピングを行うことにする。本論文では、0~2.5 kHz の区間の平均値をクリッピングレベルとして使用する。

2.3 流れ 図

提案法の流れ図を図3に示す。まず、音声波形 (10 kHz で標準化) をハニング窓により切り出す (512 点)。分析精度を高めるために0を付加してFFTを行い (1024 点)、対数スペクトルを求める。この対数スペクトルは声道の影響を受けているため、そのまま自己相関関数を計算すると抽出誤りが増える。そのためリフタ処理を行うことによってスペクトルを平坦化する。更に雑音の影響を少なくするためにクリッピングを行う。こうして得られた対数スペクトル $S(i)$ について、式 (1) で定義される自己相関関数を計算する。

この周波数軸上における自己相関関数 $r(j)$ の値は、 $10000/1024=9.7$ Hz ごとに求められている。このままでは周波数分析精度が低いので、計算された自己相関関数のピークから補間によって基本周波数 F_0 を求める。

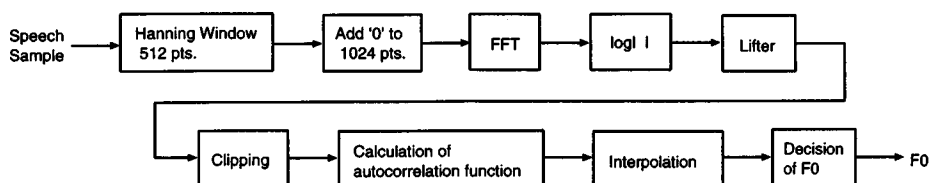


図3 提案法の流れ図
Fig. 3 Flowchart of proposed method.

3. 実験結果

本章では、従来のピッチ抽出法との比較をシミュレーション実験により行う。音声の SN 比を変えてピッチ抽出を行い、その抽出精度について検討する。

3.1 音声試料

使用した音声試料は、NTT アドバンステクノロジー(株)の「20ヶ国語音声データベース」に収録されている男女各 4 名の発声した約 10 秒の短文である。また、付加雑音は白色雑音であり、その振幅を変えて音声に加えることにより SN 比を調整した。これらは 3.4 kHz に帯域制限され、標準化周波数 10 kHz で標準化されている。SN 比は以下の式によって計算する。

$$SNR = 10 \log(P_S/P_N) \quad (2)$$

ここで P_S および P_N は、それぞれ音声信号と雑音の平均電力を示す。音声信号の平均電力を計算する際、音声区間を判別するために信号を 5 ms のフレームに区切った。そして、それぞれのフレームにおける信号の最大振幅が、信号全体の最大振幅の -30 dB 以上であるとき、音声区間として平均電力を求めた。

本研究では、あらかじめ視察によって有聲/無聲の弁別を行い、有聲音のみピッチ抽出を行うことにした。有聲音区間の判別および基準とするピッチ周波数（基準ピッチ周波数）の設定法としては、SAPD[23]等さまざまな手法が考えられるが、ここではより精密に設定するために以下のようにして行った。まず、音声信号を分析窓長 512 点で切り出し、波形および対数スペクトルから周期性を視察によって確認して有聲音区間を決定する。このとき、対数スペクトルの第 n 次高調波の周波数を n で割ることによって、基準ピッチ周波数を設定する。この際に使用した高調波は、 $n=1\sim 15$ の範囲で波形ピークとして明確に観測できるものを選択し、補間によって基準ピッチ周波数を決定した。

3.2 比較した従来法

提案法の性能を調べるために従来のピッチ抽出法と比較を行った。比較の対象としたピッチ抽出法は、これまでの文献[2],[24]においても比較の対象として広く用いられているケプストラム法 (CEPST) と自己相関法 (AUTOC) である。なお Hess の文献においては、ケプストラム法など四つの手法を比較し、自己相関法が最も雑音に robust であることが示されている[25]。また提案法 (ACLOS) におけるリフタ処理は、25 次の FIR ショートパスリフタ (5 ms) を使用し

表 1 実験の諸定数

Table 1 Constants for analysis.

標準化周波数	10kHz
分析窓	ハニング窓, 51.2 ms
分析周期	10 ms
FFT ポイント数	1024 (0 を付加)
自己相関関数の定数	$N=250, M=300$

表 2 ピーク検出のしきい値を決定する定数 α Table 2 Threshold level α for deciding peak.

Speaker	Method		
	CEPST	AUTOC	ACLOS
Male	0.85	0.95	0.70
Female	1.00	0.70	1.00

た。各処理の定数を表 1 に示す。

一般に波形ピークを検出することによって実現されるピッチ抽出法では、複数のピークから基本周期/基本周波数に対応するものを選択することは難しく、倍ピッチや半ピッチ等のピッチ抽出誤りの原因となっている。本論文では、こうしたピッチ抽出誤りを減らすために各処理ごとにしきい値を設定し、しきい値以上のピークのうち最小の周波数を基本周波数として抽出する。しきい値は、基本周波数の存在範囲内 (50~400 Hz) における最大ピークを P_{\max} とし、 αP_{\max} によって設定する。ここで、 α の値は予備実験の結果をもとに表 2 のように設定した。また、 $\alpha=1.0$ のときは最大ピークをそのまま利用することを意味する。

提案法では、各処理における分析精度は 9.7 Hz、自己相関関数およびケプストラム法では 0.1 ms となっている。従って、精度を向上させるために補間を行う方が良い結果が得られる。本論文では、極大値前後の 3 点から 2 次関数に近似する補間によって基本周波数を抽出した。

3.3 抽出精度の評価法

抽出精度の評価は、Rabiner ら[24]の手法をもとに以下のように設定した。まず、ピッチ抽出時間 i における基準ピッチ周期（基準ピッチ周波数の逆数）を $p(i)$ 、各手法によって抽出した基本周期を $\hat{p}(i)$ として、抽出誤差を

$$e(i) = \hat{p}(i) - p(i) \quad (3)$$

によって求める。そして、 $|e(i)| \geq 1 \text{ ms}$ のとき gross pitch error とし、ピッチ抽出誤りとする。また、 $|e(i)| < 1 \text{ ms}$ のとき fine pitch error とし、その標準偏差と平均誤差を求めた。

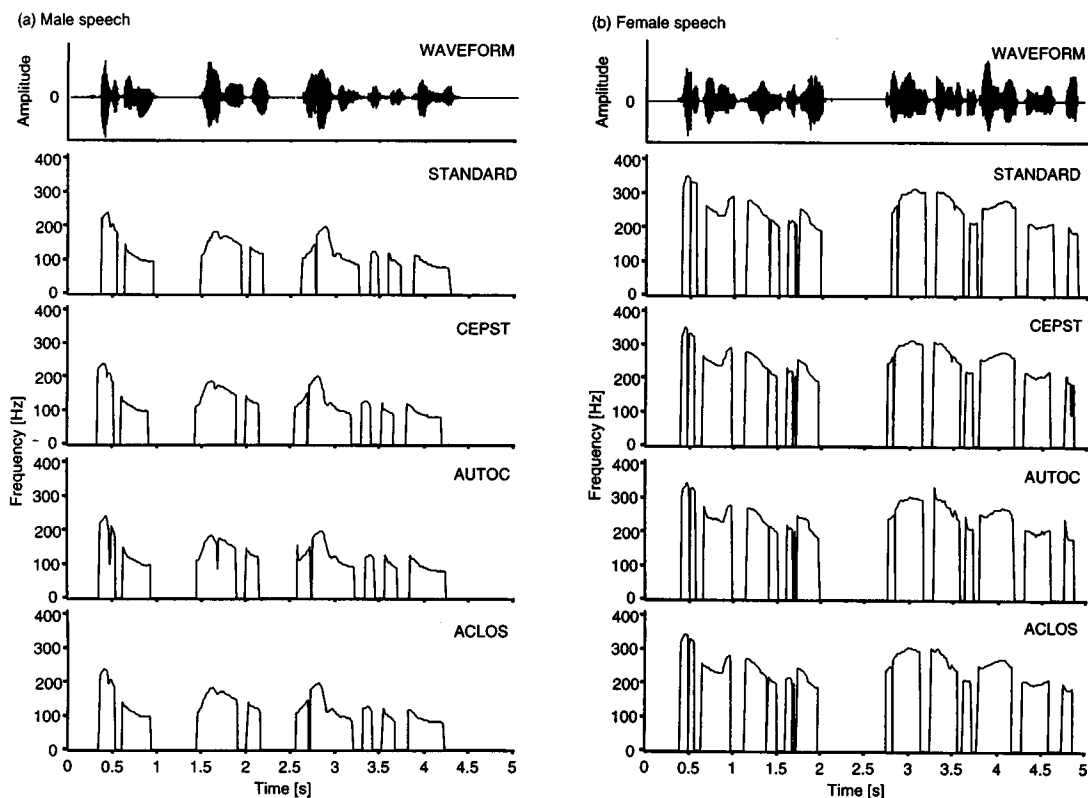


図4 種々のピッチ抽出法による基本周波数抽出結果 (a) 男声 (発話内容:「人々の屏風絵と如来像に対する興味は」), (b) 女声 (発話内容:「遙かな宇宙への旅を可能にしたのだと考えていました」)

Fig. 4 Examples of pitch contour of standard pitch by observation (STANDARD) and of the extracted by several methods. (a) Male speech:/hitobitono bjoobueto njoraizonitaisuru kjoomiwa/. (b) Female speech:/harukana uchuueno tabiwo kanoonifitanodato kaNgaeteimafita/.

3.4 抽出結果

連続音声に対する抽出結果の例を図4, 5に, 抽出精度の結果を表3~5に示す。まず図4を見ると, 提案法によるピッチ抽出の結果, 文頭や文末においてもピッチ抽出誤りがほとんど起きていないことがわかる。また, 図5のSN比が0dBの音声に対する処理を見ると, 従来のケプストラム法や自己相関法では倍ピッチ, 半ピッチと思われる誤りが多くみられるが, 提案法ではこうした誤りが少なくなっている。

更に表3の結果から, 提案法では従来法よりも gross pitch error が少なくなっていることがわかる。特にSN比が-5dB以下の低SN比音声に対しては, 雑音に強い自己相関関数法よりも3.4~7.2%改善されている。また, 特に女声に対する改善が大きくなっているのも特徴である。

一方, 表4, 5の fine pitch error の抽出誤差を見ると, 提案法による誤差の平均値は従来法に比べて劣る結果となった。標準偏差について見ると, まず男声では従来法にやや劣っている。また女声については, 高SN比では従来法と同程度, 低SN比では優れている結果となった。

以上の結果から, 提案法が雑音の有無に関わらず安定したピッチ抽出を実現できる方式であることがわかった。特に, 提案法は低SN比音声のピッチ抽出に威力を発揮し, 基本周波数の高い女声に対して有効である。

3.5 抽出精度向上のための検討

実験の結果から, 提案法では fine pitch error の抽出精度が従来法より低下することがわかった。提案法においては, 音声信号をFFTによってスペクトルの

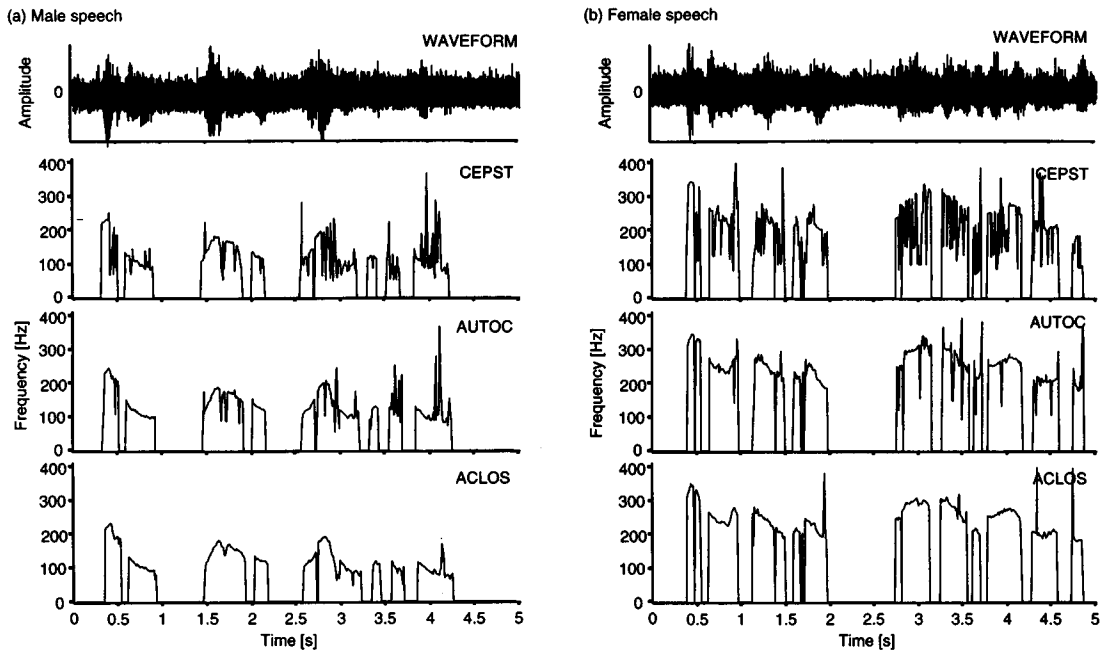


図5 0 dB の音声に対する各ピッチ抽出法による基本周波数抽出結果 (a) 男声, (b) 女声
Fig.5 Examples of pitch contour extracted by several methods (SNR=0 dB). (a) Male speech. (b) Female speech.

表3 ピッチ抽出誤り率 (gross pitch error) (%)
Table 3 Gross pitch error at various SNR (signal-to-noise ratio) (%).

Speaker	Method	SNR							
		∞	20 dB	15 dB	10 dB	5 dB	0 dB	-5 dB	-10 dB
Male	CEPST	0.39	1.28	2.95	7.28	17.51	35.23	57.49	73.07
	AUTOC	1.46	1.63	2.01	2.57	6.59	16.05	38.61	65.20
	ACLOS	0.26	0.47	0.77	1.97	6.04	14.38	35.19	59.42
Female	CEPST	2.65	2.86	5.47	12.19	23.38	37.66	54.62	65.94
	AUTOC	1.26	1.13	1.30	1.74	4.21	10.15	28.03	46.68
	ACLOS	0.04	0.13	0.26	1.00	3.08	9.72	21.30	39.44

表4 Fine pitch error の平均値 (ms)
Table 4 Mean of fine pitch error (ms).

Speaker	Method	SNR							
		∞	20 dB	15 dB	10 dB	5 dB	0 dB	-5 dB	-10 dB
Male	CEPST	-0.001	0.014	0.019	0.024	0.030	0.027	0.017	0.004
	AUTOC	-0.036	-0.035	-0.037	-0.041	-0.046	-0.041	-0.028	-0.016
	ACLOS	0.063	0.068	0.084	0.121	0.162	0.192	0.151	0.088
Female	CEPST	-0.009	-0.016	-0.017	-0.015	-0.016	-0.008	-0.007	-0.009
	AUTOC	-0.021	-0.023	-0.027	-0.039	-0.065	0.100	-0.096	-0.061
	ACLOS	0.043	0.061	0.082	0.110	0.144	0.161	0.157	0.108

データに変換した後に自己相関関数を計算する。このとき、スペクトル信号の標本点の分解能は 9.7 Hz となる。提案法による抽出精度が良くない理由として、まずこの数値が粗かったことが考えられる。

そこで、FFT のポイント数を増やし、スペクトルの分解能を向上させたときの抽出精度を調べるためにシミュレーション実験を行った。FFT のポイント数を 2^n とし、付加する 0 の数を増やして提案法による

表 5 Fine pitch error の標準偏差 (ms)
Table 5 Standard deviation of fine pitch error (ms).

Speaker	Method	SNR							
		∞	20 dB	15 dB	10 dB	5 dB	0 dB	-5 dB	-10 dB
Male	CEPST	0.112	0.147	0.171	0.203	0.253	0.299	0.374	0.451
	AUTO	0.151	0.150	0.153	0.166	0.189	0.235	0.307	0.400
	ACLOS	0.127	0.146	0.181	0.220	0.269	0.341	0.424	0.505
Female	CEPST	0.065	0.117	0.144	0.180	0.223	0.291	0.365	0.443
	AUTO	0.095	0.102	0.112	0.134	0.176	0.220	0.290	0.395
	ACLOS	0.074	0.094	0.116	0.139	0.172	0.210	0.265	0.359

表 6 FFT のポイント数 2^n と fine pitch error の平均値 (ms) の関係
Table 6 Relation between 2^n and mean of fine pitch error (ms).

n	Male	MS	MV	Female	FS	FV
9	0.0634	-0.0226	0.0223	0.0429	-0.0012	0.0012
10	0.0658	-0.0142	0.0179	0.0434	-0.0003	0.0012
11	0.0671	-0.0125	0.0162	0.0441	-0.0004	0.0012
12	0.0680	-0.0117	0.0160	0.0446	-0.0004	0.0012
13	0.0685	-0.0116	0.0160	0.0449	-0.0004	0.0012

表 7 FFT のポイント数 2^n と fine pitch error の標準偏差 (ms) の関係
Table 7 Relation between 2^n and standard deviation of fine pitch error (ms).

n	Male	MS	MV	Female	FS	FV
9	0.1265	0.0243	0.0748	0.0736	0.0063	0.0047
10	0.1305	0.0140	0.0685	0.0757	0.0030	0.0046
11	0.1335	0.0113	0.0669	0.0770	0.0014	0.0045
12	0.1350	0.0102	0.0668	0.0776	0.0011	0.0045
13	0.1358	0.0100	0.0668	0.0779	0.0011	0.0045

ピッチ抽出を行う。このとき、式 (1) における N の値とフィルタ次数は、FFT のポイント数に応じて同じ周波数帯域の相関関数を計算するように変化させた。

使用した音声試料は、3.1 で述べた実音声と合成音/a/である。合成音は、基本周波数を 25 Hz ごとに一定とする定常の男声 MS (基本周波数の範囲は 75~175 Hz) および女声 FS (200~375 Hz) と、基本周波数が 300 Hz/秒で変化する男声 MV (75~175 Hz) および女声 FV (175~375 Hz) である。なお、ピッチ抽出を行うに際しては、gross pitch error は生じなかった。

FFT のポイント数を 2^n にしたときの抽出誤差の平均値と標準偏差を表 6, 7 に示す。この結果を見ると、合成音については n を 10, 11 とすることによって抽出精度が向上しているが、実音声では精度の向上が見られずやや劣化している。また、合成音についてみると、時間変化している音声の方が改善度が小さくなっていることに気づく。

このように、実音声については FFT のポイント数を増やすことが必ずしも抽出精度向上に結びつかないことがわかる。実音声は合成音とは異なり、明確な調

波構造をもたず、ゆらぎが生じたりする。このような調波構造の乱れが、自己相関関数で計算したときに微妙な誤差として影響するものと考えられる。従って、実音声に対しての抽出精度の向上については、今後の課題として残される。

4. む す び

本論文では、対数スペクトルの自己相関関数を利用した音声の新しいピッチ抽出法を提案し、その性能について評価した。本方式では、ケプストラム法と自己相関法を組み合わせることによって、雑音に robust なピッチ抽出を実現することができた。

シミュレーション実験の結果、提案法では自己相関関数法やケプストラム法よりも gross pitch error を改善することができた。特に、-5 dB 以下の音声に対しては、雑音に強い自己相関関数法よりも gross pitch error を 3.4~7.2%改善しており、低 SN 比において威力を発揮することがわかった。一方、提案法の抽出誤差の平均値は自己相関法やケプストラム法に劣るといふ結果となった。これは、実音声における調波構造の乱れが影響しているものと考えられるが、抽出精度向

上については今後の課題とした。

また、本方式の耐雑音性を向上させるための手段としては、クリッピングレベルの設定法が挙げられよう。今後、他の耐雑音性に優れたピッチ抽出法との比較とあわせて検討したい。

今回は、白色雑音が重畳した場合についてのみ検討を行ってきた。しかしながら、本手法では対数スペクトル上での周期性を検出するため、正弦波のような周期雑音や調波構造をもたない雑音に対しても robust であると考えられる。また本方式の原理から、電話音声などの低周波数の高調波成分が失われた狭帯域信号や SSB において離調された音声に対しても有効であると考えられる。こうした本方式の可能性について今後検討する予定である。

謝辞 本研究の予備的検討を行った本研究室の斎藤修氏（現（株）日本ブレインウェアトラスト）および抽出精度向上に関する分析に協力された本研究室大学院生金子信一郎氏に感謝致します。また、日ごろから御討論して頂く本学の八嶋弘幸助教授に感謝致します。なお、本研究の一部はテレコム先端技術研究支援センター（SCAT）の援助により行われたものである。

文 献

- [1] W.J. Hess, Pitch determination of speech signals, Springer-verlag, Berlin, 1983.
- [2] W.J. Hess, "Pitch and voicing determination," in Advances in speech signal processing, ed. S. Furui and M.M. Sondhi, pp.3-48, Marcel Dekker, Inc, New York, 1992.
- [3] 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," 音響誌, vol.51, no.7, pp.509-518, July 1995.
- [4] 瀬藤美有紀, 鈴木誠史, 島村徹也, "音声スペクトルの調波構造を利用した基本周波数の評価," 音講論, 2-Q-7, pp.277-278, March 1994.
- [5] A. Moreno and J.A.R. Fonollosa, "Pitch determination of noisy speech using higher order statistics," IEEE Int. Conf. Acoust., Speech, Signal Process., San Francisco, U.S.A., vol.I, pp.133-136, March 1992.
- [6] M. Lahat, R.J. Niederjohn, and D.A. Krubsack, "A spectral autocorrelation method for measurement of the fundamental frequency of noise-corrupted speech," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-35, no.6, pp.741-750, June 1987.
- [7] 李 金忠, 鈴木誠史, 島村徹也, "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, Jan. 1995.
- [8] P. Chen and S. Ando, "Pitch from zeros of bank-filtered signals," IEEE Int. Conf. Acoust., Speech, Signal Process., Minneapolis, U.S.A., pp.530-533, April 1993.
- [9] 船田哲男, 鈴木達也, "帯域フィルタ対バンクによる音声ピッチ抽出," 信学論 (A), vol.J72-A, no.3, pp.466-474, March 1989.
- [10] 阿部敏彦, 小林隆夫, 今井 聖, "瞬時周波数に基づく雑音環境下でのピッチ推定," 信学技報, SP95-79, Nov. 1995.
- [11] M.S. Andrew, J. Pincone, and R.D. Degroat, "Robust pitch determination via SVD based cepstral methods," IEEE Int. Conf. Acoust., Speech, Signal Process., Albuquerque, U.S.A., no.S4b.10, pp.253-256, April 1990.
- [12] L. Hodgson, M.E. Jernigan, and B.L. Wills, "Nonlinear multiplicative cepstral analysis for pitch extraction in speech," IEEE Int. Conf. Acoust., Speech, Signal Process., no.S4b.11, pp.257-260, April 1990.
- [13] 加藤誠二, 三輪譲二, "移動平均と帯域制限を用いたケプストラム型基本周波数抽出とその応用," 信学技報, SP94-95, Feb. 1995.
- [14] A.M. Noll, "Cepstrum pitch determination," J. Acoust. Soc. Am., vol.41, no.2, pp.293-309, Feb. 1967.
- [15] B. Gold and L. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain," J. Acoust. Soc. Am., vol.46, no.2, pp.442-448, Aug. 1969.
- [16] N.J. Miller, "Pitch detection by data reduction," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-23, no.1, pp.72-79, Feb. 1975.
- [17] L.R. Rabiner, "On the use of autocorrelation analysis for pitch detection," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-25, no.1, pp.24-33, Feb. 1977.
- [18] M.J. Ross, H.L. Shaffer, A. Cohen, R. Freudberg, and H.J. Manley, "Average magnitude difference function pitch extractor," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-22, no.5, pp.353-362, Oct. 1974.
- [19] M.R. Schroeder, "Period histogram and product spectrum: new methods for fundamental-frequency measurement," J. Acoust. Soc. Am., vol.43, no.4, pp.829-834, April 1968.
- [20] D.J. Hermes, "Measurement of pitch by subharmonic summation," J. Acoust. Soc. Am., vol.83, no.1, pp.257-264, Jan. 1988.
- [21] A.M. Noll, "Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and a maximum likelihood estimate," in Symposium on Computer Processing in Communication, ed. Microwave Institute, vol.19, pp.779-797, University of Brooklyn Press, New York, 1970.
- [22] 吉谷清澄, 鈴木誠史, 田中良二, "自己相関関数を利用した音声処理方式 (SPAC) の SN 比改善特性," 信学論 (A), vol.J61-A, no.3, pp.217-223, March 1978.
- [23] C.A. McGonegal, L.R. Rabiner, and A.E. Rosenberg, "A semiautomatic pitch detector (SAPD)," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-23, no.6, pp.570-574, Dec.1975.
- [24] L.R. Rabiner, M.J. Cheng, A.E. Rosenberg, and C.A. McGonegal, "A comparative performance study of several pitch detection algorithms," IEEE Trans. Acoust.,

Speech & Signal Process., vol.ASSP-24, no.5, pp.399-417, Oct. 1976.

[25] 文献[2]のp.20.

(平成8年2月23日受付, 7月22日再受付)



國枝 伸行 (学生員)

平4埼玉大・工・情報卒。平6同大大学院博士前期課程了。現在、同大大学院博士後期課程に在学中。音声情報処理に関する研究、特に音声強調や雑音に埋もれた信号の処理に関する研究を進めている。日本音響学会、IEEE各会員。



島村 徹也 (正員)

昭61慶大・理工・電気卒。平3同大大学院博士課程了。工博。同年埼玉大・工・助手。スペクトル解析および適応信号処理に関する研究に従事。計測自動制御学会、システム制御情報学会、日本音響学会、IEEE各会員。



鈴木 誠史 (正員)

昭30東工大・工・電気卒。同年郵政省電波研究所(現、通信総合研究所)入所。昭42同通信機器部音声研究室長、昭57同情報処理部長、昭60同総合通信部長、昭63通信総合研究所長。平元埼玉大・工・教授、現在に至る。昭41~42、MIT客員研究員。

この間、音声合成、音声認識、通信品質の改善、音声伝送方式などの音声情報処理、通信システムの研究に従事。工博。アメリカ音響学会、日本音響学会、テレビジョン学会、情報処理学会、IEEE各会員。