

ケプストラムによる音声のピッチ周波数推定 — 評価信号の提案と分析可能範囲の考察 —

Cepstral Pitch Extraction of Continuous Speech — Performance evaluation by proposed signal —

半谷 猛[†]
Takesi HAN'YA

井 研治[‡]
Kenji INOMOTO

Abstract *A measuring system for fundamental pitch frequency of human speech signal was considered. Pitch frequencies can be estimated by using cepstrum technique, and a frequency modulated sawteeth signal is proposed to evaluate the accuracy of the cepstrum pitch estimator. The results of calculated pitch frequency of continuous speech are shown. Running pitch frequency is also used by averaging every instantaneous pitch frequencies with exponentially decreasing weights for realtime computer processing of human speech signal.*

1. はじめに

情報通信技術が高度にデジタル化された現状においても、音声信号による情報通信は依然として大きな役割を担っている。

特に航空機と多くの人命をあずかる空港管制官は、離発着頻度の上昇と共に疲労が急増し、管制能力の低下を招来すると言われている。

本論文は、音声の基本周波数の連続測定からこのような危険を速やかに判断し、要員の投入や交替などの判断のための基礎データを得るため、ケプストラムによるピッチ周波数の測定法とその精度について論じたものである。

2. ケプストラム

この方法は、観測された波形から周期を測定する方法として、地震波の解析に用いられたのが最初であるが、その後、音声波からのピッチ

の抽出、さらに、スペクトルの包絡線の抽出に広く用いられるようになった手法である。

ケプストラム^{1) 2) 3)}は、信号のパワースペクトルの対数のフーリエ逆変換（又は、フーリエ変換）として定義されている。

声帯の音源信号を $g(t)$ とし、唇の放射特性を含めた調音器官のインパルスレスポンスを $h(t)$ で表すと、観測される音声信号 $x(t)$ は、次式で表される。

$$x(t) = \int_0^t g(\tau)h(t-\tau)d\tau$$

$x(t)$ 、 $g(t)$ および $h(t)$ のフーリエ変換をそれぞれ $X(f)$ 、 $G(f)$ および $H(f)$ とすると、上式は次のようになる。

$$X(f) = G(f) \cdot H(f)$$

声帯の開閉によって発声する音源信号 $g(t)$ は、一般に調波構造を有する周期関数であり、 $|X(f)|$ は、その周期の逆数の周波数毎の線スペクトルとなる。

¹愛知工業大学 工学研究科 電気電子工学専攻 (豊田市)

[‡]愛知工業大学 情報通信工学科 (豊田市)

一方、 $\log |X(f)|$ は、

$$\log |X(f)| = \log |G(f)| + \log |H(f)|$$

となる。

次に、周波数 f を変数としてこれをフーリエ逆変換（又はフーリエ変換）する。これがケプストラムであり、フーリエ逆変換を \mathcal{F}^{-1} の記号で表すと、

$$\mathcal{F}^{-1} [\log |X(f)|] = \mathcal{F}^{-1} [\log |G(f)|] + \mathcal{F}^{-1} [\log |H(f)|] \quad (1)$$

となる。ケプストラムでは、通常、絶対値を用いるため、フーリエ変換の虚部の符号には依存しない。従って、2 回目のフーリエ逆変換は、フーリエ変換でも置き換えが可能である。

これを、次式⁴⁾で示す離散的フーリエ逆変換 (IDFT) で求める場合は、波形の標本化の場合と同様に折り返しが生ずるので、変換の基底の大きさ N を十分に大きくとる必要がある。

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} \log |X(k)| e^{j \frac{2\pi}{N} kn}$$

$$(0 \leq n \leq N-1)$$

“ケプストラム (cepstrum)” は、スペクトルを逆変換するという意味を含めて、スペクトル (spectrum) をもじって作った造語であり、その横軸はフリクエンシ (frequency) をもじってケフレンシ (quefrequency) と呼ばれる。ケフレンシのディメンジョンは、周波数領域からの逆変換であるから時間の次元になる。式 (1) の右辺第 1 項は音源に依存し、スペクトル上での微細構造に依存する比較的細かい周期のパターンである。第 2 項は伝達系を反映しており、周波数による変化のゆるやかなパターンであって、スペクトル包絡をなしている。従って、フーリエ逆変換（又はフーリエ変換）後の両者にはパワーの分布に大きな違いがあり、第 1 項は高ケフレンシ部に、また第 2 項は、音声の場合 0 から 2 ~ 4ms 程度の低ケフレンシ部に集中する。この性質を利用すれば、高ケフレンシ部のピークによって音源 $g(t)$ のピッチが求められる。高ケフレンシ成分を分離する操作を、リフタ (lifter) と呼ぶ。

ケプストラム法によるピッチ周期の抽出手順を図 1 に示す。

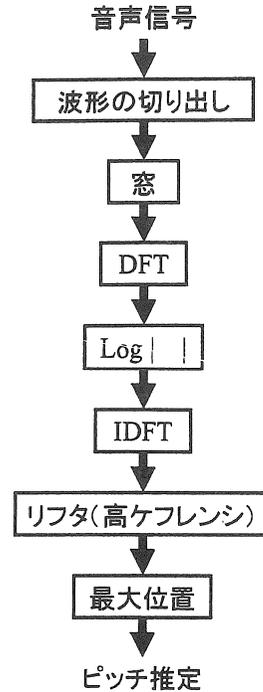


図 1: ケプストラム法のブロック図

3. 評価信号

線形システムの特性測定には、正弦波が用いられる。

ケプストラムによるピッチ周期推定を実際に論議するためには、周期あるいは周期の変動が即知の信号を用いて、その振る舞いを調べるのが基本である。

ピッチ推定の精度を測定するための評価信号には、どのような信号が適しているか簡単に論じてみる。

3・1 正弦波

まず、線形システムに多く用いられる正弦波信号について考えてみる。周期が $T (= 1/f_0)$ である正弦波は、

$$y(t) = \sin \frac{2\pi}{T} t \quad (2)$$

で表される。本章では信号の振幅は特に問題にしないので、振幅は式 (2) のように、1 に選んで

話を進める。この正弦波に、図1における最初のフーリエ変換を実行すると、そのパワースペクトルは図2のように一本のスペクトルが立つだけで、それ以外の高調波成分が現れない。つまり、パワースペクトルに周期性が現れないため、求めたケプストラムからピッチを推定することができない。

このような観点から、正弦波はケプストラムの性能を評価するには向かない。性能を適切に評価するには、高調波を多く含む信号が適していると言える。

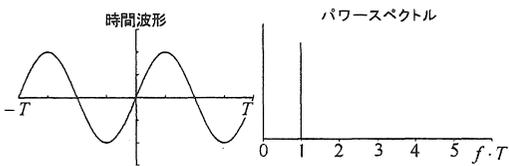


図 2: 正弦波とそのパワースペクトル密度

3・2 三角波

高調波を含む信号として次に、三角波を取り上げる。図3に示すような三角波は、そのスペクトルが次式で示される。

$$y(t) = \frac{8}{\pi^2} \left\{ \sin \frac{2\pi}{T}t - \frac{1}{3^2} \sin 3 \frac{2\pi}{T}t + \frac{1}{5^2} \sin 5 \frac{2\pi}{T}t - \frac{1}{7^2} \sin 7 \frac{2\pi}{T}t + \dots \right\} \quad (3)$$

これから求めたパワースペクトルは、図3に示すように周期性を示すものの、その間隔は $2f_0$ である。ケプストラムではこの値が検出され、ピッチ周波数は2倍、いわゆる倍ピッチとなり精度の測定には向かない。

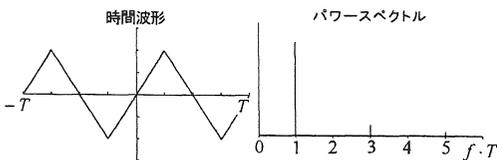


図 3: 三角波とそのパワースペクトル

3・3 方形波

また図4に示す方形波であっても、そのフーリエ変換が、

$$y(t) = \frac{4}{\pi} \left\{ \sin \frac{2\pi}{T}t + \frac{1}{3} \sin 3 \frac{2\pi}{T}t + \frac{1}{5} \sin 5 \frac{2\pi}{T}t + \frac{1}{7} \sin 7 \frac{2\pi}{T}t + \dots \right\} \quad (4)$$

となることから、三角波と同様に推定値は、倍ピッチである。

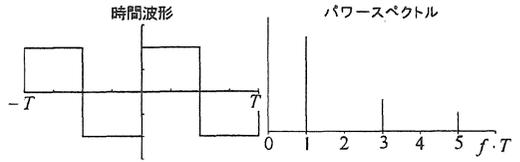


図 4: 方形波とそのパワースペクトル密度

3・4 鋸歯状波

図5に示す鋸歯状波は、これをフーリエ展開すると、

$$y(t) = \frac{2}{\pi} \left\{ \sin \frac{2\pi}{T}t - \frac{1}{2} \sin 2 \frac{2\pi}{T}t + \frac{1}{3} \sin 3 \frac{2\pi}{T}t - \frac{1}{4} \sin 4 \frac{2\pi}{T}t + \dots \right\} \quad (5)$$

となり、そのパワースペクトルは、周期が f_0 である。ケプストラムが示すピークは $1/f_0$ であり、鋸歯状波では正確なピッチが検出されることになる。

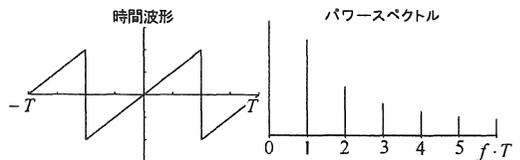


図 5: 鋸歯状波とそのパワースペクトル密度

以上のことをまとめると表1のようになる。この表からわかるように、取り上げたいいくつかの信号の中で、ケプストラムの性能評価に用いる

ことができるものとしては、鋸歯状波が好ましいと言える。

信号	パワースペクトルの周期性	検出周期
正弦波	なし	—
三角波	あり	倍ピッチ
方形波	あり	倍ピッチ
鋸歯状波	あり	ピッチ

表 1: パワースペクトルの周期性とケプストラムで検出される周期

3・5 鋸歯状波を用いた推定可能範囲の測定

このようにして求められた鋸歯状波を用いて、ケプストラムによりピッチ周波数を推定し、その誤差を求めた。

分析条件 (鋸歯状波) :

鋸歯状波の周波数変化範囲

: 2 [周期/分析窓] ~ 1.1k [Hz]

アンチエイリアシングフィルタ

: 8k [Hz]

サンプリング周波数

: 22k [Hz]

サンプル点数 : 128, 256, 512, 1024, 2048 [点]

窓関数 : Hanning

リフト : 1k [Hz]

サンプル点数に対する推定誤差 1%以下となる周波数範囲を図 6 に示す。

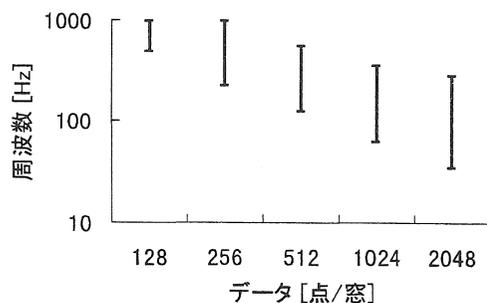


図 6: 鋸歯状波信号に対するケプストラムの推定誤差 1%以下の範囲

まず、推定誤差の下限について考えてみる。ケプストラムは DFT の結果を更に IDFT して求めているため、それぞれの下限は約 3 [周期/分析窓] となっている。最初の DFT の結果が示す周期性が最も激しいときは、周波数軸におけるサンプル値が、交互に増減する場合であり、これは鋸歯状波が分析窓内に 2 周期存在するときである。従って、ケプストラムによるピッチが正確に推定できるためには 2 [周期/分析窓] が下限になる。

一方、誤差の上限については、リフトの設定値である 1kHz 付近となっているはずであるが、分析点数が増すにつれて上限が下降している。

従って鋸歯状波のみを用いても周波数の上限におけるケプストラムの推定精度を正しく測定できない。この要因は次章で考察し、鋸歯状波を用いて、この問題を解決することを考えてみる。

4. FM 鋸歯状波

4・1 スペクトルの duty factor

ケプストラムによる推定は、図 7 に示すように①に見られる周期性を②のケプストラム、つまり、2 回目のフーリエ逆変換で検出することであり、このためには①に調波関係を必要としている。

そのため、①の波形とそのスペクトル②について考えてみる。

分析した①の幅が広いほど、つまり、duty factor が大きいほど、②は減衰する系列となり、その結果、1 番目のピークが顕著になり検出しやすくなる。反対に、①がインパルス列状なら、②も高さが等しいインパルス列となり、1 番目のピーク検出を誤ることがある。これが、ピッチ周波数をその整数分の 1 の周波数に誤る理由である。

従って、ケプストラムの性能を正しく測定するには①のスペクトルにおいて duty factor が大きな信号が好ましい。つまり、正しくピッチ測定が行える評価信号には、そのケプストラムが減衰パルス系列になることが要求される。

ところで、周波数が一定の鋸歯状波では図 5 に見られるように、そのパワースペクトルの幅は

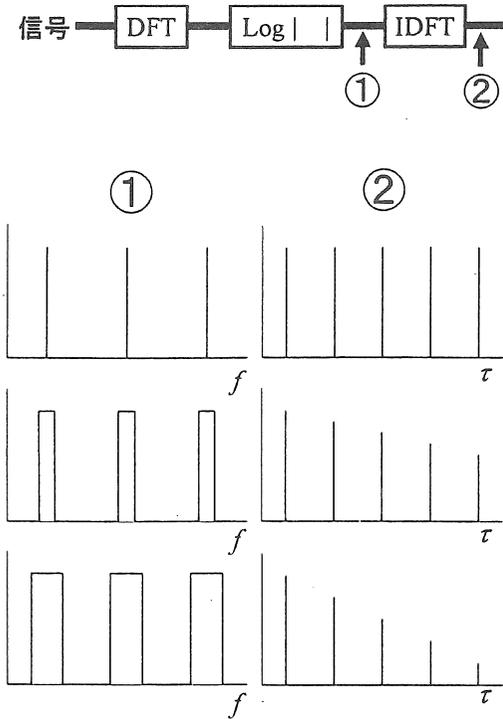


図 7: 対数パワースペクトル①とそのケプストラム②の関係

インパルス状であるため、何らかの方法でこれを広げる必要がある。次節では、これに周波数変調を付加することにより、パワースペクトルの幅を広げることを考えてみる。

4・2 FM 信号 ⑥

普通に発声される音声のピッチ周波数は、多少揺らいでいる。つまり、音声は周波数変調 (FM: Frequency Modulation) を受けている。

簡単のため被変調波、変調波を正弦波として考える。FM によって得られる波である FM 正弦波の瞬时周波数 $f(t)$ は、

$$f(t) = f_0 + f_d A_m \sin \omega_m t \quad (6)$$

となる。ここで、 ω_m は変調波の角周波数、 A_m は変調波の振幅、 f_d は係数、そして f_0 は被変調波のピッチ周波数である。

式 (6) より、瞬时位相角 $\phi(t)$ は、

$$\begin{aligned} \phi(t) &= 2\pi \int f(t) dt \\ &= 2\pi f_0 t - \frac{2\pi f_d A_m}{\omega_m} \cos \omega_m t \end{aligned} \quad (7)$$

となるので、FM 正弦波 $y(t)$ は、

$$\begin{aligned} y(t) &= A_c \sin \phi(t) \\ &= A_c \sin (2\pi f_0 t - \beta \cos \omega_m t) \end{aligned} \quad (8)$$

として得られる。

ただし、

$$\beta = \frac{2\pi f_d A_m}{\omega_m}$$

である。

式 (6) より、瞬时周波数の偏移は $f_d A_m$ であるから、ピッチ周波数 f_0 に対する周波数変化率 γ は、次式になる。

$$\gamma = \frac{f_d}{f_0} A_m \quad (9)$$

ところで周波数偏移は $f_d A_m$ であるから、スペクトル占有帯域幅 B は、

$$\begin{aligned} B &= 2f_d A_m \\ &= 2\gamma f_0 \end{aligned} \quad (10)$$

となり、スペクトルを広げることができる。

鋸歯状波のケプストラムを減衰系列にするため、この性質を利用して鋸歯状波のスペクトルを広げること考えてみる。

ここで簡単のため、調波構造を有するパワースペクトルの対数 $P(f)$ が、式 (11) で表されるとする。

$$P(f) = h(f) * x(f) \quad (11)$$

ここで、演算子 $*$ は畳み込み演算であり、 $h(f)$ 、 $x(f)$ は、それぞれ式 (12) と式 (13) に示すような矩形波とインパルス列である。また、図 8 に $P(f)$ の概形を示す。

$$h(f) = \begin{cases} 1 & (|f| < f_1) \\ \frac{1}{2} & (|f| = f_1) \\ 0 & (|f| > f_1) \end{cases} \quad (12)$$

$$x(f) = \sum_{n=-\infty}^{\infty} \delta(f - n f_0) \quad (13)$$

ただし、

$$\delta(f) = \begin{cases} 0 & (f \neq 0) \\ 1 & (f = 0) \end{cases}$$

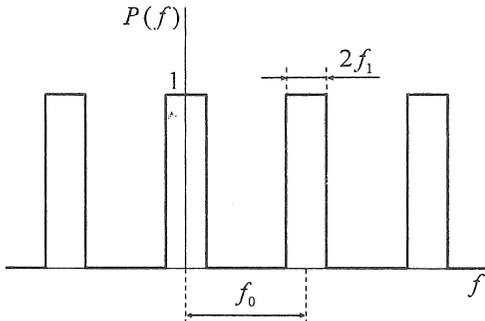


図 8: 対数パワースペクトル

また、 $P(f)$ の duty factor k は、

$$k = \frac{2f_1}{f_0} \quad (14)$$

であり、ケプストラムは、

$$\begin{aligned} p(\tau) &= \mathcal{F}^{-1} P(f) \\ &= \mathcal{F}^{-1} h(f) \cdot \mathcal{F}^{-1} x(f) \\ &= \frac{2f_1}{f_0} \cdot \frac{\sin(2\pi f_1 \tau)}{2\pi f_1 \tau} \sum_{n=-\infty}^{\infty} \delta(\tau - \frac{n}{f_0}) \\ &= k \frac{\sin(2\pi f_1 \tau)}{2\pi f_1 \tau} \sum_{n=-\infty}^{\infty} \delta(\tau - \frac{n}{f_0}) \quad (15) \end{aligned}$$

となる⁵⁾。

つまり、ケプストラムは、インパルス列に図 9 に示す $\sin(2\pi f_1 \tau)/2\pi f_1 \tau$ が包絡線となって乗じられた形になっている。

対数パワースペクトルにおけるインパルス間隔 f_0 と変調波の占有帯域幅 $2f_1$ の比、つまり、duty factor k が大きくなれば、ケプストラムは減衰系列となり、図 9 に示すようにその減衰率が増す。つまり、最初のインパルスが最大値を示すため、インパルス列の最初の最大位置 $1/f_0$ を探せば、ただちにピッチ周波数 f_0 を求めることが可能になる。

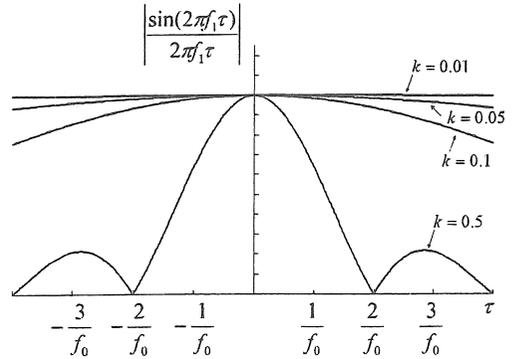


図 9: k によるケプストラムの包絡線

4・3 実験方法と結果⁷⁾

FM 鋸歯状波を用いてケプストラムによるピッチ周波数の推定誤差を求めた。

まず、鋸歯状波の FM には、周波数変化率 γ の値を 0.1 とした。これは、普通に発声される単独母音のピッチ周波数成分の占有帯域幅 B が、男性の場合ピッチ周波数 $f_0 = 90 \sim 130\text{Hz}$ に対しては $20 \sim 50\text{Hz}$ 、女性の場合 $f_0 = 250 \sim 330\text{Hz}$ に対しては $60 \sim 90\text{Hz}$ 、と報告されていることによる⁸⁾。

FM 鋸歯状波は、正弦波によって周波数変化率が 10%、つまり $\gamma = 0.1$ の周波数変調を付加した。また、変調波は、分析区間幅で少なくとも 1 周期の周波数変動を受けるようにするため、正弦波の周波数は $1/\text{分析区間幅} [1/\text{s}]$ とした。その他の分析条件は、3.5 節と同じである。

サンプル点数に対する推定誤差 1% 以下の範囲を図 10 に示す。

ここで、誤差は FM を伴う鋸歯状波の平均周波数に対するケプストラム法で測定された周波数で定めた。それぞれの下限は、FM なしの鋸歯状波と同様に約 3[周期/分析窓] となっている。上限については、リフタの設定通り 1kHz 付近まで伸ばすことができ、FM によりスペクトル幅を広げたときの有効性を示唆している。

このように、FM によりスペクトル幅を広げた結果によって、ケプストラム法で推定される範囲が図 10 より、広がったことがわかる。この結果を基にして、次章の測定でのサンプル点数を

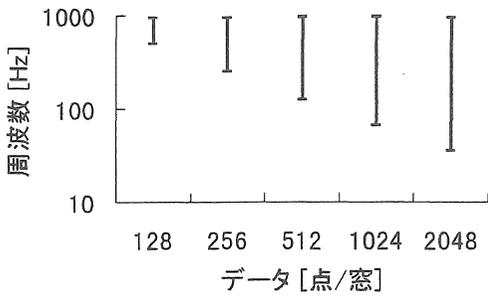


図 10: FM 鋸歯状波信号に対するケプストラムの推定誤差 1%以下の範囲

決定した。

5. ケプストラムによる連続音声のリアルタイムピッチ周波数推定

本法で用いたリアルタイムなピッチ周波数推定のブロック図を図 11 に示す。

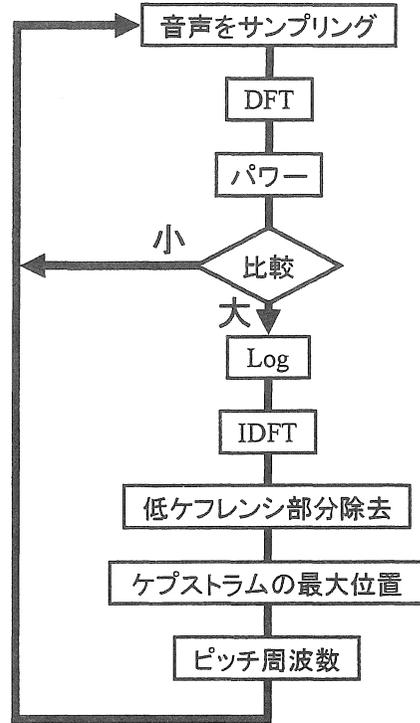


図 11: 測定系のブロック図

5・1 リフタ

リフタの値を定めるために、先ず約 1 分間にわたる種々の音声サンプルを用いて、リフタの値を変化させて、推定されたピッチ周波数の平均値を求めた。図 12 にその一例を示す。

その結果、話者が男女にかかわらず、1/リフタが 400 から 600[1/s] あたりで安定していることから、リフタの値として 1/500[s] を採用した。

5・2 有声音と無声音

無音声部分はもちろん、無声音も有声音とは異なりピッチ周波数を持たない。従って、これら無声音、無音声部分のピッチ周波数を推定しても間違ったものであり、何も意味をなさない。

音声信号がすべて有声音で構成されていれば問題は無いが、一般的には無声音、無音声部分を含むので、ピッチ測定のためにこれらを判別する必要がある。

有声音のエネルギーの大部分は 3kHz 以下の低い周波数域に存在し、一方、無声音のエネル

ギーは高い周波数域に分布するので¹⁰⁾、低い周波数域でのパワーから有声音と他のものとを判別できる。つまり、1 回目のフーリエ変換で求められた $\pm 3\text{kHz}$ 以内のパワースペクトルの総和を求め、それをしきい値と比較することにより判別を行う。よって本法では、これにより有声音として判別された場合、そのまま計算を進めてピッチ周波数を推定し、無声音又は無音声部分として判別された場合は、ピッチ周波数を推定せず、すぐに次のサンプリングを行うことにした。

そのしきい値は、いくつかの音声信号から値を定めた。これは、分析窓を通過して分析できる最大パワーの 1/25 に選ばれている。実際の入力振幅において、このしきい値を、FM 鋸歯状波を用いた実験により換算しておく。その結果、推定可能な最低信号電圧は、最大許容電圧の約 8% であり、これ以下のときは無声音又は無音声部分に、反対に約 8% 以上のときは有声音として判別された。ちなみに過去のデータによると、無音声部分の実効値の平均は、有声音である母音

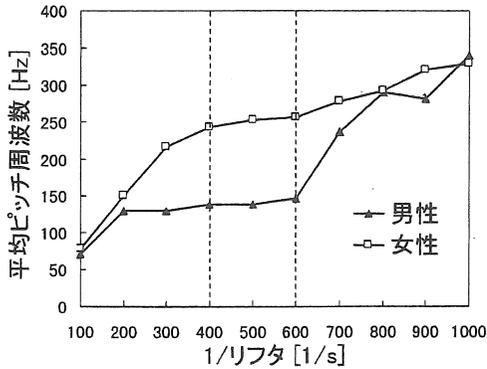


図 12: リフトと平均ピッチ周波数の一例

の 0.041%である⁹⁾。

5・3 分析条件

分析条件 (連続音声):

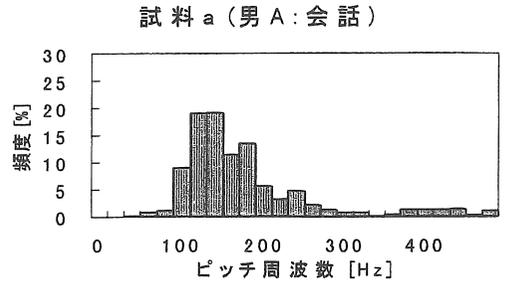
- アンチエイリアシングフィルタ : 8k [Hz]
- サンプリング周波数 : 16k [Hz]
- 量子化ビット数 : 16 [bit]
- 分析点数 : 1024 [点]
- 窓関数 : Hanning
- 有声音判別のしきい値 : 8 [%]
- リフト : 1/500 [s]

5・4 分析結果

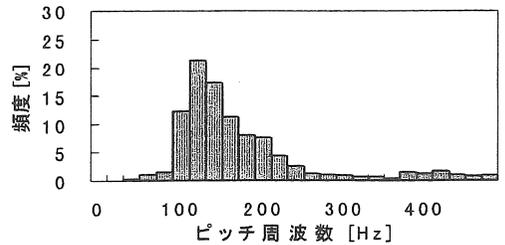
各試料についての平均ピッチ周波数 μ および、ピッチ周波数の標準偏差 σ を表 2 に示す。

また、図 13、14 に各試料におけるピッチ周波数の頻度分布を示す。

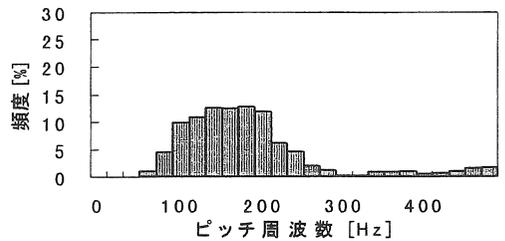
どの試料においても、それぞれの平均ピッチ周波数付近での頻度が高く、平均ピッチ周波数から離れるに連れて小さくなっている。



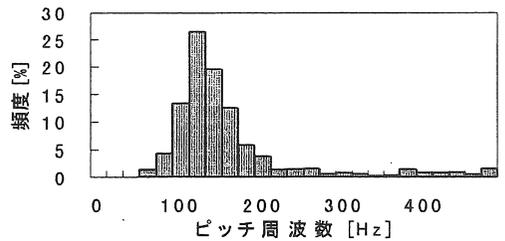
試料 a (男 A: 会話)



試料 c (男 B: 天気予報)



試料 d (男 C: 天気予報)



試料 e (男 D: ニュース)

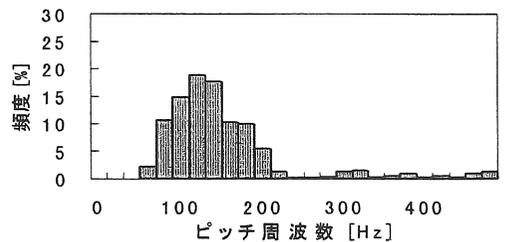


図 13: 話者:男性のピッチ周波数頻度

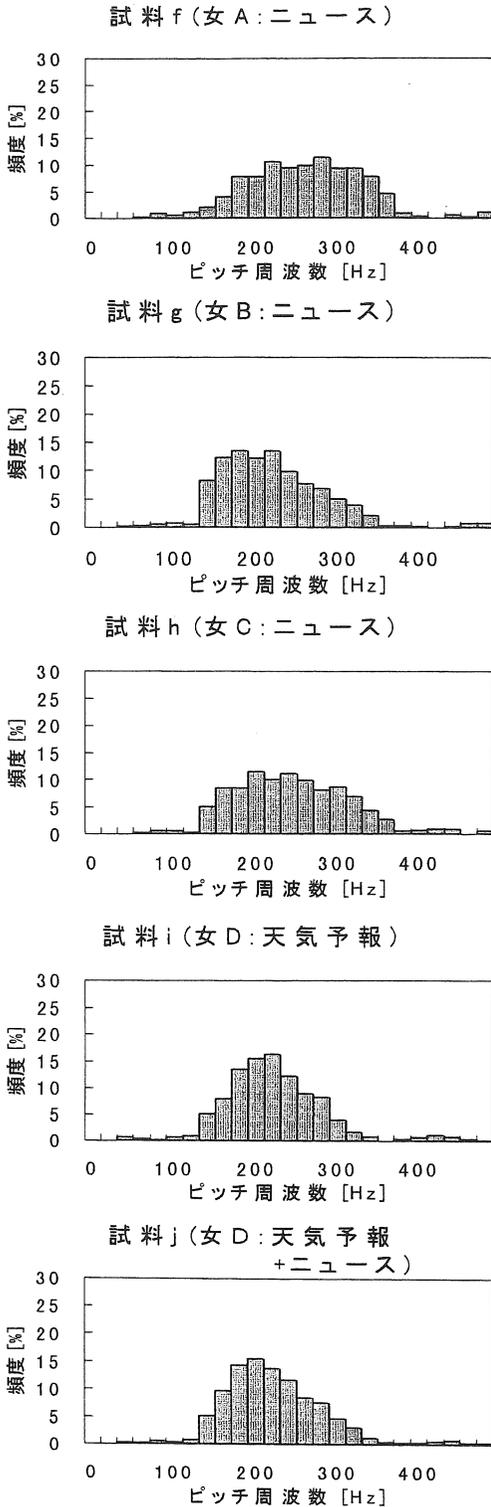


図 14: 話者:女性のピッチ周波数頻度

試料	分析時間 [s]	話者	内容	μ [Hz]	σ [Hz]
a	62	男 A	会話	187.0	82.82
b	793	男 A	会話	184.9	87.68
c	62	男 B	天気予報	195.5	89.42
d	62	男 C	天気予報	169.3	83.02
e	62	男 D	ニュース	163.8	80.80
f	62	女 A	ニュース	269.6	70.41
g	62	女 B	ニュース	230.5	66.62
h	61	女 C	ニュース	254.2	69.41
i	61	女 D	天気予報	233.3	61.12
j	608	女 D	天気予報+ニュース	233.7	64.49

表 2: 平均ピッチ周波数 μ および標準偏差 σ

5.5 ランニングピッチ周波数

ケプストラム法によって測定された瞬時のピッチ周波数を x_i とするとき、ランニングピッチ周波数 μ_i を、指数重み a ($0 < a < 1$) を用いて以下のように定義した。

$$\begin{aligned} \mu_i &= (x_i + ax_{i-1} + a^2x_{i-2} + \dots)(1-a) \\ &= \{x_i + a(x_{i-1} + ax_{i-2} + \dots)\}(1-a) \\ &= (1-a)x_i + a\mu_{i-1} \end{aligned} \quad (16)$$

これは、アナログ回路における CR 積分回路に相当するもので、ここでは等価的な時定数を 100[ms] に定め、指数重み a を決定した。指数重み a と時定数との関係は式で与えられ、図 15 のグラフのようになる。時定数を 100[ms] するには、指数重み a を約 0.53 に定めればよい。

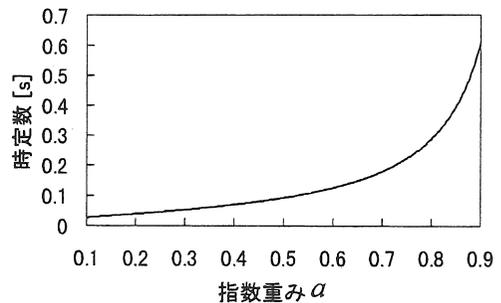


図 15: 指数重み a と時定数との関係

時定数の確認を行うため、インデンシャル応答、つまり、FM 鋸歯状波の周波数を 160kHz から 320kHz にステップ状に変化させ、計算される

ランニングピッチ周波数をプロットし、これにより時定数が定めた通りであることを確認した。

これを図 16 に示す。時定数は約 105[ms] となっており、ほぼ設計通りに動作していることが確認できた。

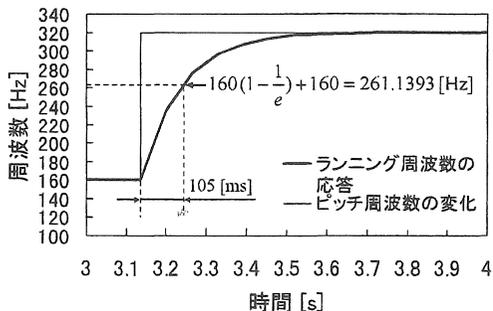


図 16: インデンシヤル応答と求められた時定数

実際の音声をケプストラムにより、リアルタイムで連続測定した結果を次に示す。用いた音声試料は、女性によるニュースの朗読、および男性による全国高校野球における選手宣誓であり、これらのランニングピッチ周波数を図 17、図 18 に示す。

ニュース朗読の音声では、ピッチ周波数の変化を明瞭に抽出できています。一方、単調で力強い口調の全国高校野球における選手宣誓では、ピッチ周波数も変化に乏しく単調であることがわかる。

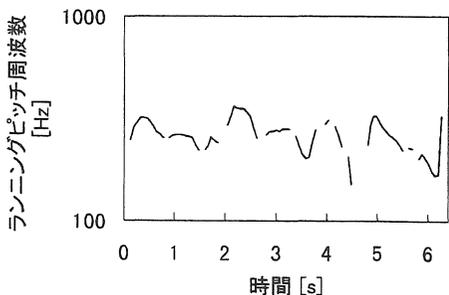


図 17: ニュース朗読のランニングピッチ周波数

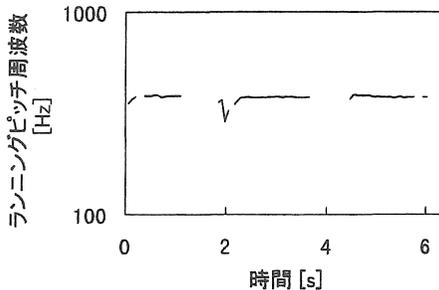


図 18: 全国高校野球における選手宣誓のランニングピッチ周波数

6. まとめ

ここでは、ケプストラム法によるピッチ周波数推定可能範囲について検討した。まず、ピッチ周波数の推定可能範囲を測定するための評価信号について考察を行った。正弦波、三角波、方形波などはケプストラム法のアルゴリズム上、本法では正確なピッチが計算できないことを示した。そして、スペクトルの幅についての理論的考察から、鋸歯状波に周波数変調したものを評価信号に用いることを提案し、これによってピッチ周波数の推定可能範囲を議論した。

次に、ケプストラム法を用いたリアルタイムでのピッチ周波数推定法とその測定条件などについて述べた。

実際に、話者が男性、女性、それぞれの連続音声のピッチ周波数を推定し、そのピッチ周波数の頻度分布を示した。

更に、ピッチ周波数の変動を見るためにランニングピッチ周波数を提案した。そして、特徴の異なった2つの音声についてランニングピッチ周波数の変動を示した。

以上のように、本研究では音声信号からのピッチ周波数の推定をすることを主眼に置き、種々の観点から議論した。今後の課題としては、先ず、多くのサンプルについてピッチ周波数の平均値と標準偏差のデータを蓄積しなければならない。しかし、そのためには標準偏差を逐次計算するアルゴリズムとその性能をきちんと評価しておくことが必要である。

参考文献

- 1) 日野幹雄. スペクトル解析. 朝倉書店. pp. 280, 1977.
- 2) 安居院猛, 中嶋正之. コンピュータ音声処理. 産報出版. pp. 159, 1980.
- 3) 城戸健一. デジタル信号処理入門. 丸善. pp. 154, 1975.
- 4) E.O.Brigham. The Fast Fourier Transform. Prentice-Hall. pp. 98, 1974.
- 5) *ibid.* pp. 62.
- 6) 半谷猛, 深田邦之, 井研治. ケプストラムによるピッチ周波数推定精度の検討. 電気関係学会東海支部連合大会. No. 487, 1997.
- 7) 半谷猛, 井研治. ケプストラムによる連続音声のピッチ周波数変動特性. 電気関係学会東海支部連合大会. No. 486, 1997.
- 8) 電子通信学会. 聴覚と音声. 電子通信学会. pp. 286, 1966.
- 9) *ibid.* pp. 248.
- 10) 太田光雄. 情報音響工学. 朝倉書店. pp. 132-133, 1992.

(受理 平成10年3月20日)