

デジタルシグナルプロセッサによる音声の基本周波数の実時間抽出

メタデータ	言語: jpn 出版者: 明治大学工学部 公開日: 2012-05-24 キーワード (Ja): キーワード (En): 作成者: 鎌田, 弘之, 松本, 直樹, 石田, 義久, 小川, 康男 メールアドレス: 所属:
URL	http://hdl.handle.net/10291/12406

Ⅱ-320 デジタルシグナルプロセッサによる 音声の基本周波数の実時間抽出

鎌 田 弘 之*
松 本 直 樹**
石 田 義 久**
小 川 康 男**

Pitch Frequency Detector on Real Time by Digital Signal Processor

Hiroyuki KAMATA*
Naoki MATSUMOTO**
Yoshihisa ISHIDA**
Yasuo OGAWA**

Abstract

This paper deals with pitch frequency detection of human voice by using digital signal processor (DSP).

The purpose of this study is as follows;

- (1) realization of detection system in a very small scale circuit by using digital signal processor.
- (2) detection of pitch frequency with highly precision and high speed (real time).
- (3) development of a digital filter with less round-off error.

1. ま え が き

半導体技術の進歩には、目を見はるものがある。中央演算処理装置 (CPU) のような極めて応用範囲の広いものだけでなく、ある特定の機能を重視して設計された回路をも、盛んに LSI 化されるようになってきた。

その成果として、汎用のマイクロプロセッサでは困難であった分野、例えば音声帯域信号のリアルタイム処理にも適用可能な高速シグナルプロセッサが開発されている。

ところで、筆者らは以前から、聴覚障害者のための発話訓練機の研究開発を行い、ろう教育の場に提供してきた¹⁾²⁾。

* 明治大学大学院工学研究科博士後期課程

** 明治大学工学部

この訓練機における問題点として、1) 分析結果の理解しやすい画面表示、2) 訓練機機能の多様化、3) 経済性の向上、などがある。

1)は、教育効果が最も期待できる幼児、児童の訓練に重要な事柄であり、音声のもつ特徴をより簡単に表現することが子供達の理解を深めるために必要である。

又、2)は、障害の程度や年齢に応じた最適な教育を遂行するために、要求される事項である。

さて、3)の経済性の問題は、ろう学校における発話訓練機の設置台数を高めるための重要事項であり、1人当りの訓練時間の延べ数を増加させ、教育効果の向上を図るために特に重要である。

そこで、コスト的に大きなウエイトを占めるハードウェア部を簡素化し、経済性のすぐれた発話訓練機を実現させる一つの試みとして、高速シグナルプロセッサを用

いた訓練機を開発することにした。

本研究では、その第1段階として、音声の基本周波数（ピッチ周波数）を検出するピッチ抽出回路を、シグナルプロセッサによって構成した。その結果、従来のアナログの手法に勝る良好な特性が得られた。

更に、基本周波数成分を抽出するデジタルフィルタの構成法について検討したところ、ダイナミックレンジの広いフィルタの状態方程式を導出することができた。

このフィルタは、次の特徴を有する。

- (1) 内部状態変数が直接、所望の出力となるので、演算過程が簡素である。
- (2) フィルタの遮断周波数を制御しやすい。
- (3) 有限ビット長演算における丸め誤差の影響を受けにくい。

以下、シグナルプロセッサを用いたピッチ抽出回路の原理、構成、及び実験結果を述べる。

2. シグナルプロセッサを用いたシステム構成

本研究では、図1のような内部構成を有するデジタルシグナルプロセッサ（NEC, μ PD 7720; PR \ddot{O} Mタイプ、以後 DSP と略す）を用いた。これは、音声帯域信号の高速処理を目的として開発された、1チップマイクロプロセッサである⁹⁾。

内部には、次のものが含まれている。

- (1) インストラクション R \ddot{O} M

(512 [words] \times 23 [bits])

- (2) データ R \ddot{O} M

(512 [words] \times 13 [bits])

- (3) データ RAM

(128 [words] \times 16 [bits])

- (4) 高速乗算器

(16 [bits] \times 16 [bits])

インストラクションコードの解読、及びそれに従った命令の実行は、内部で並列処理されており、極めて能率的な演算処理を可能にしている。

次に、この DSP の周辺システムを、図2に示す。

本システムは、LSI 内部のインストラクション R \ddot{O} M等に格納するプログラムの開発支援システム（ブロック(a)）と、DSP チップから構成される実験システム（ブロック(b)）とから成っている。

3. DSP によるピッチ周波数の抽出

図3は、従来の発話訓練機に内蔵してきたピッチ抽出回路の構成図である¹⁾²⁾。

これは、初段の低域通過フィルタ（以後、LPF と略す）の遮断周波数を、入力音声の基本周波数に一致させる方向に制御することにより、音声のピッチ周波数を抽出しようとするものである。

本論文において示す、DSP を用いたピッチ抽出に関しても、ほぼ同様の原理を採用している。但し、実用化

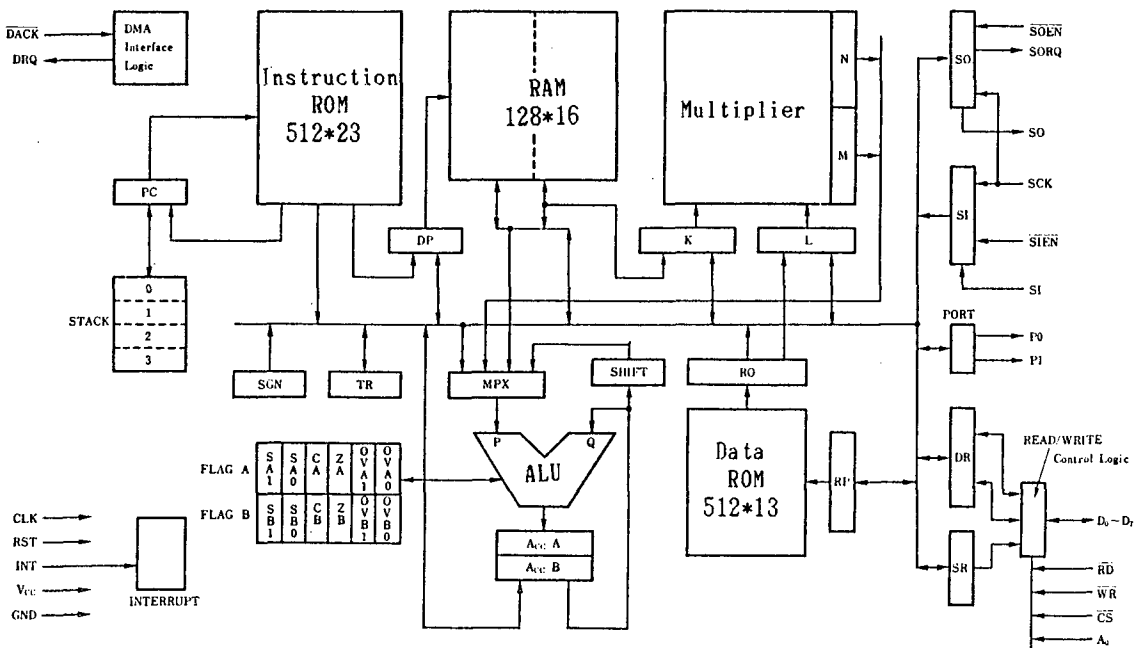
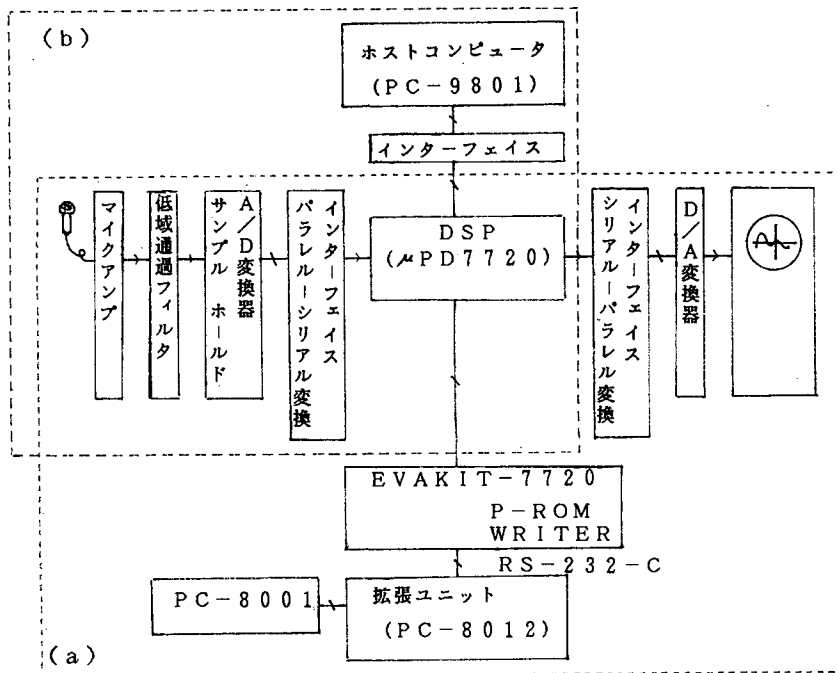


図1 DSP の内部構成 (μ PD 7720 ユーザーズマニュアルより)



(a) DSP 内プログラムの開発支援システム (b) 実験システム

図 2 DSP の周辺システム

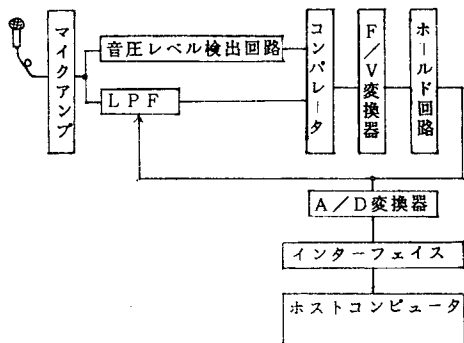


図 3 ピッチ抽出回路の構成図

に際して重要な次の事項に留意して、ソフトウェアの開発を行った。

- 1) ピッチ周波数の実時間抽出の実現
- 2) 有限ビット長の数値計算により生じる丸め誤差累積の緩和
- 3) 他機能を付加するための、DSP 内部メモリの有効利用

これらの条件を満足すべく、検討を加えたピッチ抽出アルゴリズムを図 4 に示す。

図中の LPF には、遮断周波数が可変の 2 次ディジタル

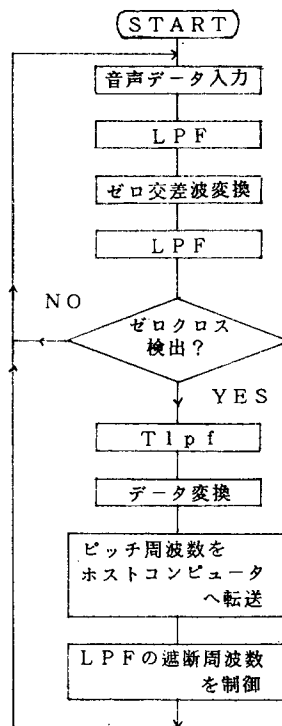
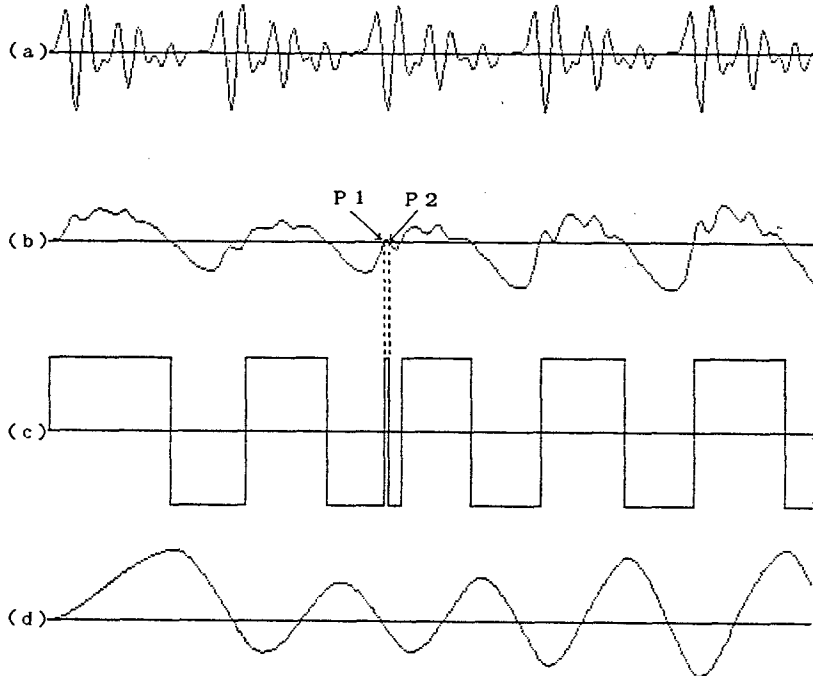


図 4 ピッチ周波数抽出アルゴリズム



(a) 自然音声 (b) LPF の出力 (初段)
(c) ゼロ交差波 (d) LPF の出力 (終段)

図 5 アルゴリズム中の信号波形

表 1 ゼロ交差波の効果 (FFT によるパワースペクトル, ピッチ周波数 $f=117.4$ [Hz] の振幅を基準)

	ピッチ分析開始時からの経過時間					
	0 [msec]		51.8 [msec]		102.4 [msec]	
ゼロ交差波への 変換の有無	無 変 換	変 換	無 変 換	変 換	無 変 換	変 換
周 波 数						
0 (直流)	- 7.7	-44.9	-22.8	-26.7	-24.8	-28.7
19.6 [Hz]	-13.5	-11.2	-74.9	-70.2	-74.2	-80.2
78.3 [Hz]	37.3	-53.3	-67.7	-73.3	-59.2	-44.4
117.4 [Hz]	0.0	0.0	0.0	0.0	0.0	0.0
176.1 [Hz]	39.5	-41.1	-65.2	-99.1	-57.6	-57.6

ルフィルタを2段直列接続したものをを用いている。(f₀の初期値は, 75 [Hz] としている。)

まず, 初段の LPF で, 入力信号の高調波成分を除去する。(図 5-(b)) しかし, 図 5-(b) の点 P1, P2 のように, 微妙にゼロラインを交差する部分を完全に除くことは困難である。

そこで, 初段の LPF の出力信号を, +1, -1 の値をもつ矩形波であるゼロ交差波に変換し, その後, 次段の LPF に入力する。(図 5-(c)) この過程の効果を, LPF (終段) の出力信号の周波数成分で比較したもの

を, 表 1 に示す。

同表の“ゼロ交差波への変換の有無”の項で, “無変換”は, 初段の LPF の出力をそのまま終段の LPF へ入力した場合, “変換”は, 一度ゼロ交差波に変換した後, 終段の LPF へ入力した場合である。

この表から, 特に過渡成分を強く含む区間における, ピッチ周波数の尖鋭度の強調には, 効果的であるという結果が得られている。

次に, 終段の LPF の出力波形において, ゼロクロス点を検出し, その周期を計測する。

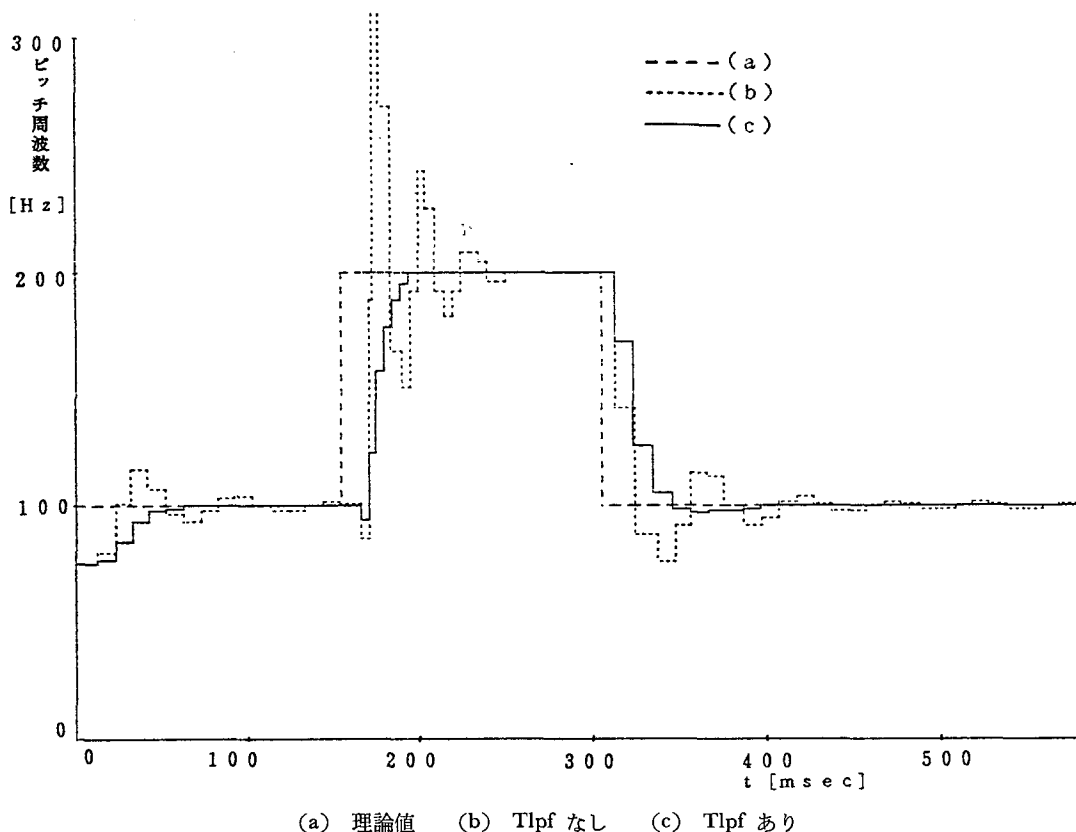


図 6 Tlpf の効果

このようにして得られた周期からピッチ周波数を求め、ホストコンピュータに転送すると同時に、前段の LPF にフィードバックする。

さて、アルゴリズム中の Tlpf は、次の目的で用いている。

本アルゴリズムは、終段の LPF の出力信号からピッチ周波数を検出し、その周波数に従って LPF の遮断周波数を変化させる、一種のフィードバック制御系を構成している。そのため、信号の入力開始時、及びピッチ周波数に急激な変化が起きた時に過渡現象を生じる。

しかし、LPF そのものによる過渡現象を抑えることは事実上不可能なので、本アルゴリズムでは、得られたピッチ周波数の平滑化を行い、この値によって LPF の遮断周波数を制御している。(この平滑化に用いたフィルタを Tlpf と呼んでいる。)

図 6 の (b)、(c) は、この Tlpf をアルゴリズム中に挿入しない場合 (b) と、挿入した場合 (c) において、コンピュータシミュレーションを行った結果である。

同図を見てもわかるように、Tlpf を有する方は、オーバーシュートも小さく、収束も速くなっている。

4. デジタルフィルタ⁴⁾⁵⁾⁶⁾

本研究では、ピッチ周波数を検出するフィルタとして、新たに開発したデジタルフィルタを用いている。

これは、実時間分析が要求される本アルゴリズムに対して、有効なフィルタである。以下に、その詳細を述べる。

まず、基本となるアナログ帯域通過フィルタの伝達関数 $H(s)$ を (4.1) 式に示す。

$$H(s) = \frac{sw_0}{s^2 + \frac{w_0}{Q}s + w_0^2} \quad (4.1)$$

Q : 尖鋭度

w_0 : 固有角周波数

これを、ステップ応答不変法に従い、 Z 変換する。

(4.1) 式の伝達関数におけるステップ応答 $\hat{Y}(t)$ は、

$$\hat{Y}(t) = L^{-1} \left\{ \frac{w_0 s}{\left(s + \frac{w_0}{2Q}\right)^2 + w_0^2 \left(1 - \frac{1}{4Q^2}\right)} \cdot \frac{1}{s} \right\}$$

$$= \frac{1}{\sqrt{1 - \frac{1}{4Q^2}}} \exp\left(-\frac{w_0 t}{2Q}\right) \sin\left(\sqrt{1 - \frac{1}{4Q^2}} w_0 t\right) \quad (4.2)$$

となる。これを、間隔 T [sec] でサンプリングし、(従って、 $t=nT(n=0, 1, 2, \dots)$)

$$r = \exp\left(-\frac{w_0 T}{2Q}\right) \quad (4.3)$$

$$\theta = \sqrt{1 - \frac{1}{4Q^2}} w_0 T \quad (4.4)$$

とおくと、(4.2) 式は

$$\hat{Y}(nT) = \frac{w_0 T}{\theta} r^n \sin n\theta \quad (4.5)$$

となる。

こうして得られた(4.5)式をもとに、ここで新たに、ステップ応答が

$$Y(nT) = r^n \sin n\theta \quad (4.6)$$

であるシステムを考える。

このようなシステムのインパルス応答を $h(n)$ とおくと、

$$\left. \begin{aligned} h(0) &= 0 \\ h(n) &= Y(nT) - Y((n-1)T) \\ &= r^n \sin n\theta - r^{n-1} \sin(n-1)\theta \quad n \geq 1 \end{aligned} \right\} (4.7)$$

となる。そこで、インパルス入力に対し、状態 $x_1(n)$, $x_2(n)$ が、

$$\left. \begin{aligned} x_1(0) &= x_2(0) = 0 \\ x_1(n) &= r^n \cos n\theta - r^{n-1} \cos(n-1)\theta \\ x_2(n) &= h(n) \\ &= r^n \sin n\theta - r^{n-1} \sin(n-1)\theta \end{aligned} \right\} n \geq 1 \quad (4.8)$$

となるシステム

$$\left. \begin{aligned} \mathbf{x}(n+1) &= \mathbf{A}\mathbf{x}(n) + \mathbf{B}\mathbf{u}(n) \\ \mathbf{y}(n) &= \mathbf{x}(n) \end{aligned} \right\} (4.9)$$

を考える。但し、

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} \quad (4.10)$$

$$\mathbf{A} = \begin{bmatrix} r \cos \theta & -r \sin \theta \\ r \sin \theta & r \cos \theta \end{bmatrix} \quad (4.11)$$

$$\mathbf{u}(n) = \delta(n) \quad (4.12)$$

であり、制御行列 \mathbf{B} を未知のベクトルとする。

ここで、システム行列 \mathbf{A} を(4.11)式に選んだの

は、丸め誤差を抑えることを目的としたものである。

更に、(4.9) 式の解は、

$$\left. \begin{aligned} \mathbf{x}(0) &= \mathbf{0} \\ \mathbf{x}(n) &= \mathbf{A}^{n-1} \mathbf{B} \quad n \geq 1 \end{aligned} \right\} (4.13)$$

である⁷⁾ ことから、(4.8), (4.11) 式より

$$\left. \begin{aligned} \mathbf{B} &= \mathbf{A}^{1-n} \mathbf{x}(n) \\ &= \begin{bmatrix} r \cos \theta - 1 \\ r \sin \theta \end{bmatrix} \end{aligned} \right\} (4.14)$$

となる。以上により、状態方程式

$$\left. \begin{aligned} \mathbf{x}(n+1) &= \begin{bmatrix} r \cos \theta & -r \sin \theta \\ r \sin \theta & r \cos \theta \end{bmatrix} \mathbf{x}(n) \\ &+ \begin{bmatrix} r \cos \theta - 1 \\ r \sin \theta \end{bmatrix} \mathbf{u}(n) \\ \mathbf{y}(n) &= \mathbf{x}(n) \end{aligned} \right\} (4.15)$$

となる。

一方、(4.15) 式より、伝達関数 $\mathbf{H}(z)$ を求めると、

$$\left. \begin{aligned} \mathbf{H}(z) &\triangleq \begin{bmatrix} H_1(z) \\ H_2(z) \end{bmatrix} \\ &= \frac{1}{z^2 - 2r \cos \theta z + r^2} \\ &\quad \begin{bmatrix} (z-r)(r \cos \theta - 1) - r^2 \sin^2 \theta \\ (z-1)r \sin \theta \end{bmatrix} \end{aligned} \right\} (4.16)$$

となる。

この $H_1(z)$, $H_2(z)$ の周波数特性を求めると、図7、図8となり、 $H_1(z)$ がLPF、 $H_2(z)$ がBPFとなることがわかる。

図9に、(4.15) 式より得られるブロック図を示す。同図のように、このフィルタは、一般に丸め誤差に強いと評価される結合型で構成できる。

又、Tlpf の伝達関数を次式に示す。

$$H_{\text{Tlpf}}(z) = \frac{m}{m+1-z^{-1}} \quad (4.17)$$

式中の変数 m は、ピッチ周波数への収束の速さについて検討を加えた結果

$$0.75 \leq m \leq 1.00 \quad (4.18)$$

の範囲の値が適当である事がわかった。図10に、(4.17) 式の周波数特性を示す。

5. テスト波形によるアルゴリズムの評価

本章では、ホストコンピュータで合成したテスト波形2種と自然音声に対して行った、ピッチ分析プログラム

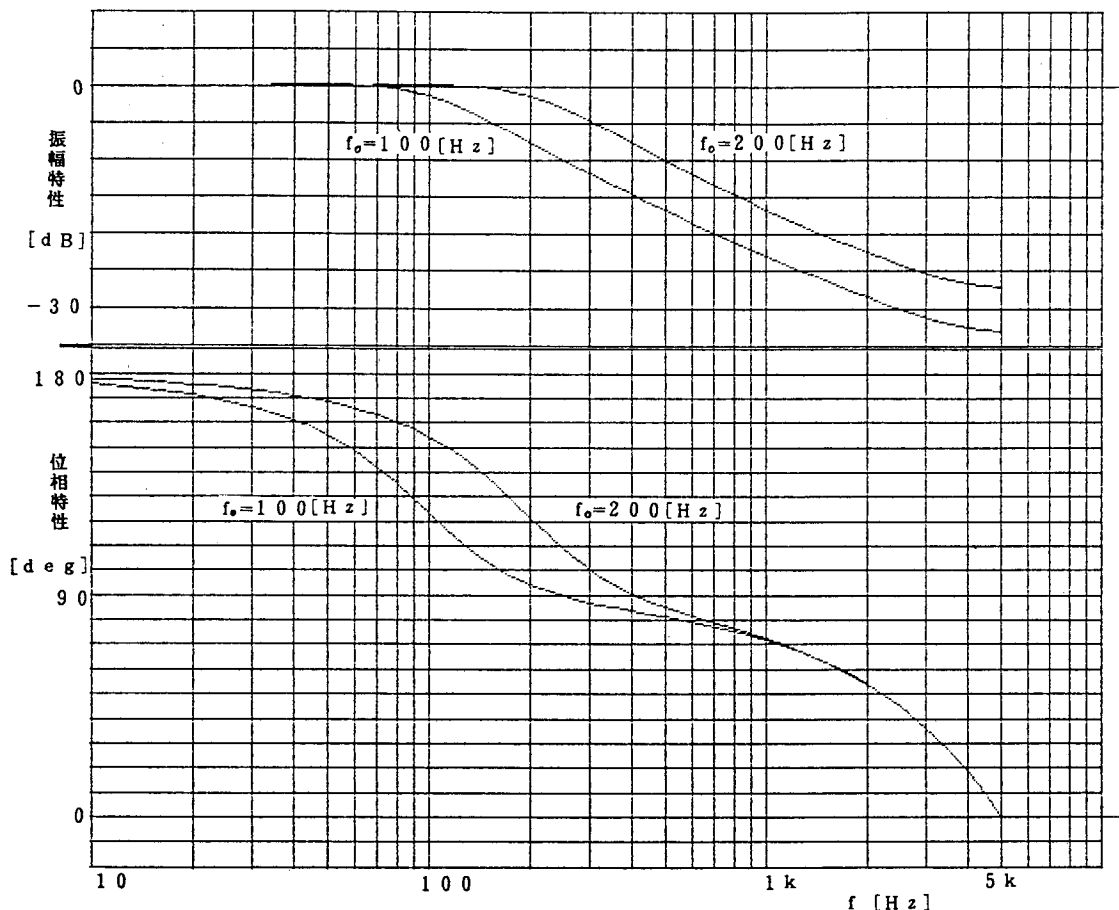


図 7 LPF の周波数特性

のシミュレーションの結果を示す。但し、実験に際しては、以下の設定を施した。

- ・サンプリング間隔 $T=100$ [μ sec]
- ・LPF の尖鋭度 $Q=0.71$
- ・Tlpf の係数 $m=0.75$
- ・テスト波形 $=f_p+5f_p+9f_p+$ 乱数
(但し、 f_p は基本周波数)

図11の(1)~(2)が、シミュレーションの結果である。

同図を見てもわかるように、基本周波数 f_p が急激に変化する場合は、収束までに5~7ピッチの時間を要する。それに対し、 f_p が滑らかに変化する場合、及び自然音声の場合は、ほぼ確実に追従している事がわかる。

従って、本研究により作成されたピッチ抽出プログラムの有効性が確認された。

6. むすび

本論文において、シグナルプロセッサを用いたピッチ抽出について述べた。この研究を要約すると、次のようになる。

- ・高速シグナルプロセッサを用いることにより、数値計算によるピッチ周波数の実時間分析が可能となった。
- ・新たに導出したデジタルフィルタの状態方程式により、高速且つ高精度の、低減フィルタ及び帯域フィルタを実現し得た。
- ・ゼロ交差波及びTlpfを用いることにより、ピッチ周波数の検出精度及び検出速度の向上を図った。

現在、DSPの他の応用、例えば、FFTに関する検討中である。それらの内容については、機会を改めて述べたい。

最後に、試作装置、部品の寄贈を賜った焼結金属工業㈱、アジアエレクトロニクス㈱、日本電気㈱超LSI開発事業部の方々に、謝意を表します。更に、御指導下

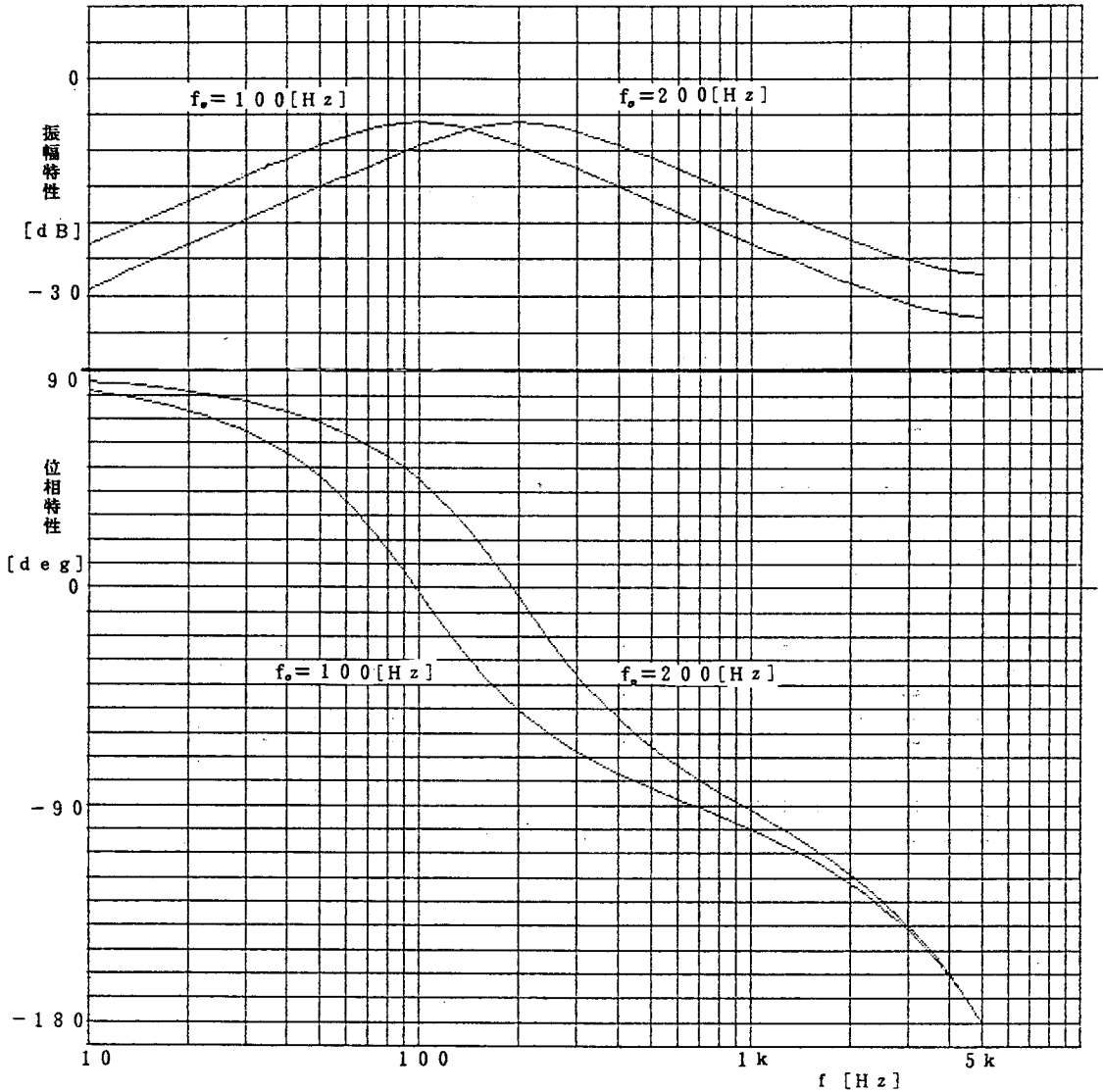


図 8 BPF の周波数特性

さった本学電気工学科教授本多高先生に、感謝の意を表します。又、実験に際し、御協力頂いた本学計測制御研究室修士課程 2 年森信明君に感謝致します。

参 考 文 献

- 1) 鎌田, 石田, 小川他: “マイクロコンピュータの特殊教育機器への応用”, 電気学会資料 IP-82-22, 1982
- 2) 鎌田, 石田, 小川: “発話訓練機のカラー化と多重化”, 電子通信学会資料 ET 82-9, 1983
- 3) 丸田: “音響信号処理用 LSI”, 日本音響学会誌

Vol. 39, No. 11, 1983

- 4) 安居院, 中嶋: “コンピュータ音声処理”, 産報出版, 1982, p. 101~163
- 5) Oppenheim, Schafer, “Digital Signal Processing” Prentice-Hall, Englewood Cliffs, N. J., 1975
- 6) 森, 松本, 石田, 小川: “シグナルプロセッサを用いたピッチ周波数検出器”, 日本シミュレーション学会・コンファレンス資料, 1984-6, p. 31~34
- 7) 計測自動制御学会編: “自動制御ハンドブック(基礎編)”, コロナ社, 1983, p. 365~407

デジタルシグナルプロセッサによる音声の基本周波数の実時間抽出

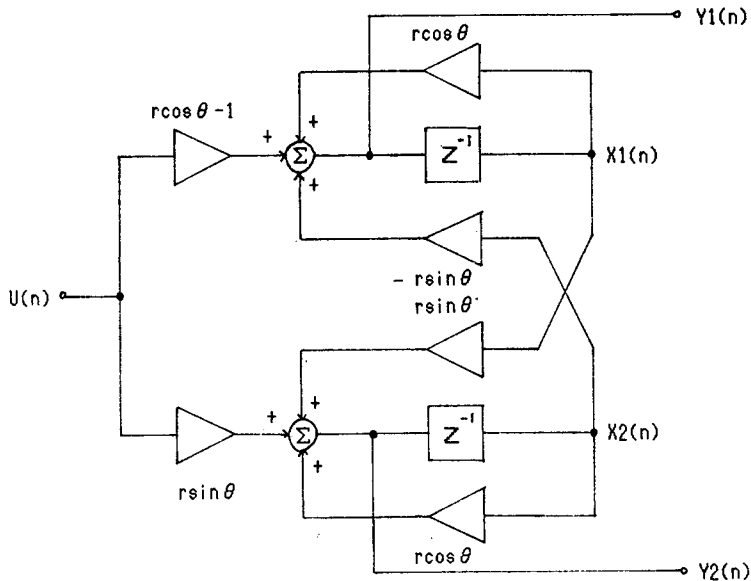


図9 デジタルフィルタのブロック図

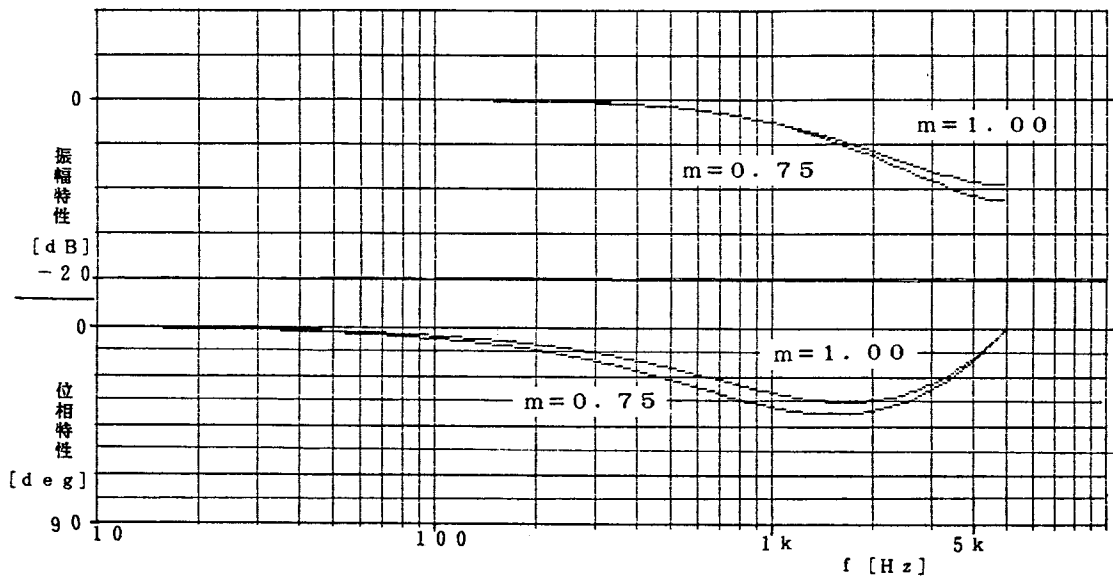
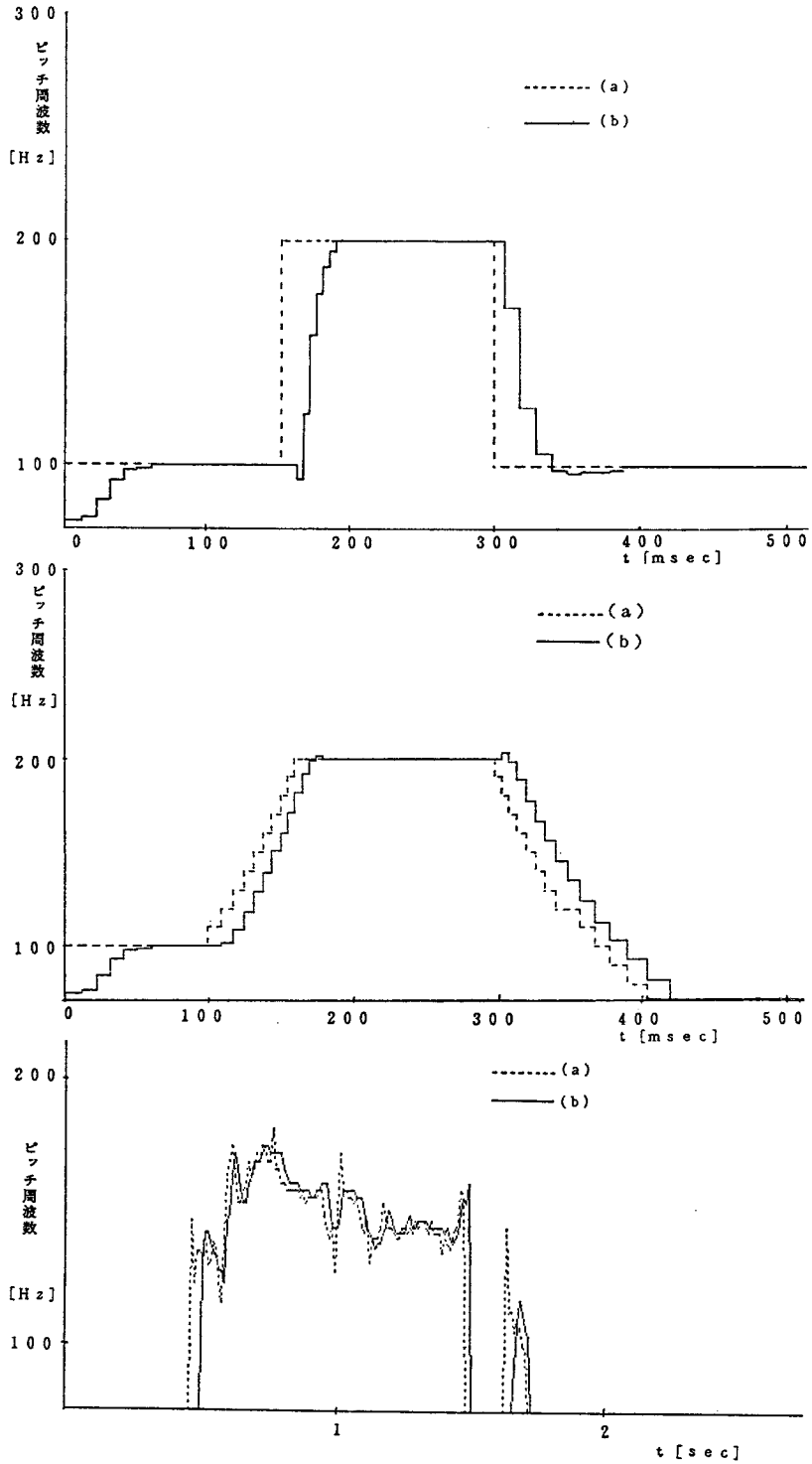


図10 Tlpf の周波数特性



- (1) ピッチが急激に変化する場合 (a) 理論値
- (2) ピッチが滑らかに変化する場合 (b) 検出値
- (3) 自然音声 (“おはようございます”)

図11 アルゴリズムの評価