

DSPを用いた蝸牛外刺激型人工内耳

平田恵啓* 伊福部達* 坂尻正次* 泉隆* 松島純一**

*北海道大学電子科学研究所

**北海道大学医学部

我々は単一電極蝸牛外刺激型人工内耳の開発を行っている。今回、DSPを利用して音声を分析し刺激パターンに変換する信号処理装置を開発した。本人工内耳システムは刺激電極が一つだけである。このため、音声符号化方式として音声からピッチ周波数と第2ホルマント周波数を抽出し、時系列信号に変換する方式を採用した。この装置は、従来のアナログフィルタ群による構成とは異なり、A/D・D/A・DSP (TMS320C30: テキサス・インスツルメンツ製)を主な構成とするデジタル回路でできている。装置の評価を日本語5母音の合成音声で行ったところ、約14ミリ秒で音声の分析及び刺激パターンの作成を行うことができた。

A speech signal processor using a DSP for an extra-cochlear prosthetic device for the profoundly hearing impaired

Yoshihiro HIRATA†, Tohru IFUKUBE†, Takashi IZUMI†,

Masatsugu SAKAJIRI† and Jun'ich MATSUSHIMA‡

†Research Institute for Electronic Science, Hokkaido University.

‡School of Medicine, Hokkaido University.

We have developed a speech signal processor for an extra-cochlear prosthetic device. As our device has only one electrode, we have designed a speech processor which can convert the pitch frequency (F0) and the second formant frequency (F2) into a time sequential pattern. The processor consists of an A/D, a D/A and a DSP (TMS320C30: Texas Instruments) in place of a conventional analog filter bank. In order to check the efficacy of our processor, the five Japanese synthetic vowels were used for input signals. As the result, it could detect F0 and F2 within 14 msec conversion time.

1. はじめに

ヒトが周囲の状況を把握したり、他の人とのコミュニケーションをはかるうえで聴覚は重要な役割をはたしている。聴力が衰えると、通常は補聴器等が用いられる。しかし難聴の原因が内耳の機能障害による場合は、これらの機器の効果は期待できない。人工内耳は患者にまだ正常に機能する聴神経が残されている場合に、この神経を直接電気刺激して患者の聴力を再建しようとする試みである。これまでにアメリカ・ヨーロッパ等でいくつかのシステムが開発され[1]、現在わが国においてもオーストラリアのコクレア社の製品が臨床で評価されている[2,3]。現在、筆者らは患者への安全性を考慮し、内耳の外部より電気刺激を行う蝸牛外刺激型人工内耳システムの開発を進めている。本稿では高速数値演算の可能なDSPを用いた信号処理装置の設計と評価について述べる。

2. 蝸牛外刺激型人工内耳システム

本人工内耳のシステム概念図を図1に示す。本システムは内耳の外部より効率的に聴神経を刺激できるように内耳の正円窓膜に設置する電極部、言語概念の獲得に大切な小児・幼児期にも適用ができるように小型化した信号伝送部、そして今回報告する音声信号

処理部からなる。本システムは電極部及び信号伝送部の制約により、現在は信号伝送路は1チャンネルのみである。そのため聴神経への刺激パターンは1次元であり、伝送可能な情報量を増加させるためには、従来の単なるアナログフィルタを用いただけの信号処理装置を見直し、言葉の認識に重要な情報を選択して伝送する必要がある。近年、高速の数値演算を実行できるDSPが開発され、従来は計算量の多さのため実時間では処理できなかった音声の分析アルゴリズムを利用することが可能になった。また、ソフトウェアによる処理系は患者の聴力の状況によってその処理の内容を柔軟に変化させることが可能である。そこで筆者らはDSPを用いた信号処理装置を提案してきた[4]。

2. 1 単一チャンネル人工内耳のための音声符号化方式

現在の設計では多くの情報を伝送できない本人工内耳システムにおいて、音声の中のどの情報を選択しどの様に符号化するかは非常に重要な問題である。これまでに、この音声の符号化として、筆者らは音声中のピッチ周波数(F0)と母音の認識に重要であるといわれる第二ホルムント周波数(F2)を組み合わせたものを提案している[4]。実際には図

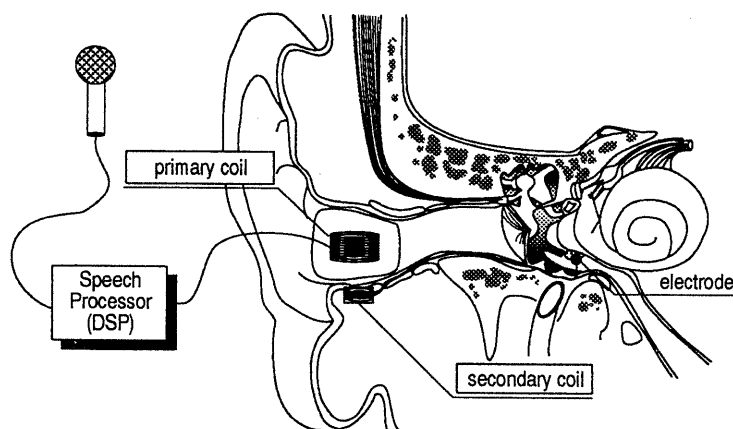


図1 蝸牛外刺激型人工内耳システム図

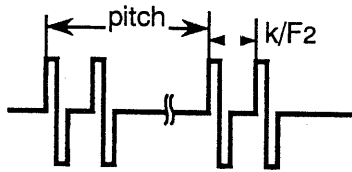


図2 音声の符号化方式

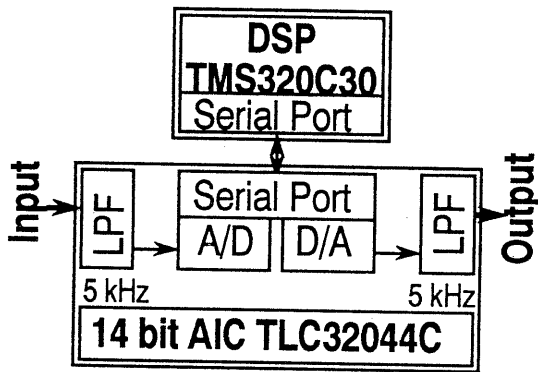


図3 信号処理装置の構成

2に示すように、 F_0 に同期してピッチパルスが発生し、このピッチパルスから k/f_2 (k :定数)の時間遅れでもう1つのパルスが発生して、連続的に聴神経を刺激する。これまでにこの符号化方式の有効性については報告している[4]。

3. DSPを用いた音声信号処理装置

音声データを実時間で分析し、2. 1に述べた音声符号化方式を実現するために、テキサス・インスツルメンツ社製の32ビット浮動小数点DSPのTMS320C30(以降C30と呼ぶ)を用いた信号処理装置を使用した。従来より筆者らは日本電気製のDSPを用いてきたが、豊富なそのDSPファミリーと高級言語(C言語)を用いることのできる開発環境等を考慮して、使用するDSPチップをC30に変更した。

図3にこのDSPボードの概略を示す。マイクロホンからの音声データは、アナログ・

インターフェイス回路(AIC:TLC32044C、テキサス・インスツルメンツ社製)内の14ビットA/Dコンバータで10kHzサンプリングされシリアルポートを介してDSP内部に転送される。また処理結果のデータは10kHz毎にAIC内のD/Aコンバータに転送、出力される。つまり0.1ミリ秒毎に割り込み信号がAICのタイマで発生して、DSPとAICとのデータの入出力が実行される。我々が以前用いていたDSPは割り込み信号が発生するとレジスタの値の破壊が行われ、複雑なプログラム処理が要求されるのに比べ、ソフトウェアの設計が非常に容易になった。またAICはサンプリング周波数に対応して変化するアンチ・エイリアジングフィルタを入出力の双方に内蔵しているため、ハードウェアの構成がより簡単になっている。

3. 1 DSP内部のデータの流れ

図4にDSP内部での大まかな信号の流れを示す。全ての信号処理は実時間で連続的に行われる必要がある。しかしデジタル信号処理では分析にある程度のフレーム長のバッファとそのフレームを分析するための時間が必要である。そこでメモリ内部に作業領域とは別に入力用と出力用の2つのリングバッファを設け、DSP外部のAICとのデータのやりとりはこのバッファを介して行うように設計した。このため、ある程度の時間遅れ(フレーム長+分析時間・・・およそ40から50ミリ秒)で連続処理ができるようになった。出力用リングバッファの更新に要する時間は約0.1ミリ秒(1サンプリング)である。そこで出力のデータをなめらかにつなぐ工夫として、音声信号の分析・処理が終わった時点で、次のパルス出力ポイントを調べ、その部分のデータのみを残して、残りの全てのバッファを書き換えるようにした。また、前のフレームと今回のフレームとの間で検出したピッチや第2ホルマント周波数が変化した場合には、更新した直後の一つめのピッチ及びそれに続くパルスのタイミングをこの二つの



図4 DSP内部のデータの流れ

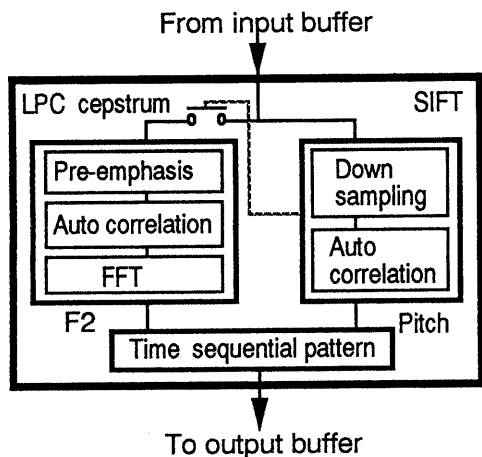


図5 DSP内部の処理手順

フレームの平均値になるように設定した。

3. 2 音声分析アルゴリズム

前回の報告[5]と同様、ピッチの抽出には変形自己相関法のSIFTアルゴリズム[6]を、第2ホルマント周波数の検出にはLPCケプストラムを用いた。このため同じ自己相関の計算アルゴリズムを用いてプログラム領域を小さくすることができた。ただし、前回はメモリの制約のため、新たな回転因子の書き込みが難しく256ポイントの複素実数2のFFTをしていたが、今回はこの制約がなくなったため、256ポイントの実数基数2のFFTをプログラムすることで処理速度を増すことが可能になった。図5にはDSP内部での信号処理の流れを示す。まず入力バッファから作業領域に取り込まれた信号はピッチの抽出が行われ、ここでピッチがあると判

表1 各ルーチンの処理時間

サブルーチン	処理時間(msec)
バッファのデータ取り込み	0.16
ピッチの抽出	6.86
第2ホルマント周波数の推定	4.61
バッファへのデータ書き込み	0.09
A/D,D/A へのデータの入出力	1.98
合計	13.7

断されると、続いて第2ホルマントが推定される。次の段階では、これらの情報をもとに2. 1に示した音声符号化を行い、この結果を直接、出力用のリングバッファに書き込む。

3. 3 合成音声による信号処理装置の評価

本装置に移植したプログラムの評価を行うために日本語5母音の合成音声を作成し、分析に要する時間及び計算結果の精度を求めた。

使用した合成母音はklatt型音声合成器[7]を用い、入力に用いたホルマント周波数とバンド幅は平原ら[8]の報告に示されているデータを用いた。またホルマント周波数は第1から第3ホルマント周波数までとした。

結果

表1に各処理ルーチン毎の計算に要した時間を、表2には入力した合成母音及びDSPの分析結果のピッチと第2ホルマント周波数を示す。処理時間は、それぞれデータの作業

表2 合成母音の入力に対するDSPの出力

	合成母音		分析結果	
	F0	F2	F0	F2
/i/	150	2281	152	2305
/e/	150	2031	152	2031
/a/	150	1187	152	1211
/o/	150	781	152	781
/u/	150	1219	152	1250

領域への取り込みが0.16ミリ秒、ピッチの抽出に6.86ミリ秒、第2ホルマントの抽出に4.61ミリ秒、データの出力に0.09ミリ秒、A/D及びD/Aの入出力に2.0ミリ秒の合計13.7ミリ秒であった。この結果は前回報告した合計の11.5ミリ秒より大きな値である。しかしこの値は、処理のフレーム長30ミリ秒とあわせて43.7ミリ秒の遅れ時間でありADFの現象が発生しはじめる50ミリ秒以内には全ての処理が納まっている。5母音すべてに同じ150Hzのピッチを用いたが、すべて同じ152.3Hzの出力が得られた。入力した合成母音の第2ホルマントにたいして、分析した結果はほぼ正確であり、得られた第2ホルマント周波数の差はいちばん大きい/u/で31Hzであった。本符号化方式においては第2ホルマントの逆数が用いられるが、この際の比例乗数を3とすると最も差の大きい/a/で2サンプリング分であった。よってこの程度の誤差であればさほど問題はないと考える。

4. 考察

今回の分析に要した時間が前回報告したものよりも長くなった。具体的に見ると、ピッチの抽出時間が2.8ミリ秒から11.5ミリ秒と大幅に増加している。一方、第2ホルマントの抽出では8.5ミリ秒から4.6ミリ秒と短縮された。

この主な原因としてプログラミング方法の違いが考えられる。前回用いたDSP (μ PD77230) においては全てのプログラミングをアセンブリ言語で記述したのに対し

て、今回はその大部分をC言語によって作成した。通常、高級言語よりもアセンブリ言語のほうが効率的なプログラミングが可能である。実際、第2ホルマントの抽出ルーチンに用いたFFTアルゴリズムはC30のユーザー・マニュアルを参考にアセンブリ言語で作成し、非常に高速化された。ピッチ抽出のルーチンについてもアセンブリ言語で記述すると処理速度を上げることができると予想できるが、筆者らの知識不足のため今回は断念した。この部分の改善は今後の課題とした。

DSPを μ PD77230 (日本電気製) からC30に変えた理由は両者のアーキテクチャの違いに由来する。前者がメモリ空間内に全く独立した2つのデータ用RAM空間を持ち、個々のRAM空間に対応した機械語の命令体系を持つものに対して、C30は16Mワード (実装されているのは6Kワード+1Mビット) のリニアなメモリ空間を有し、高級言語でのプログラミングを可能としている。このため μ PD77230の場合、プログラムの変更時にはかなりの労力を必要とするが、C30ではまず高級言語によるプログラミングで動作をさせてから最適化をはかることができる。実用段階でのポータブル化の為の消費電力を考えると、ASIC化すべきであるが、本人工内耳システムのための音声符号化及びアルゴリズムの研究・開発を行っている現状ではC30を用いるほうが有利である。またテキサス・インスツルメンツはDSPファミリの多角化にも力をいれており、今後同様の命令体型の高速・低消費電力のデバイスが供給され、本研究がそのまま実用化される可能性もある。

5. まとめ

今回、蝸牛外刺激型人工内耳の信号処理装置としてDSPを用いた装置を試作した。このシステムに母音情報を含んだ音声の符号化方式として筆者らが提案している方式を移植した。前回試作した処理装置に比べ、処理時

間ではほぼ同等の性能であるが、プログラムの容易さで改善されたシステムとなった。本信号装置の評価の第一段階として、日本語母音の合成音声を用い、十分な精度でピッチ及びホルマントの抽出を行うことを確認した。

今後は入力に実際の音声を用い、出力をクリック音として健聴者に聞かせ、より効率的な音声符号化を模索すると共に、どの程度の音声認識の精度が得られるのかを調べる。また、ホルマント周波数は個人差があるがこの影響が符号化された音声の認識に度のように影響を与えるかについても検討したい。一方、実際に難聴の人に協力してもらい、臨床での評価も行いたいと考えている。

謝辞

本システムのハードウェアに関してご協力いただいた、日立・中央研究所の禰寝、池田の両氏に感謝いたします。

本研究は平成4年度科学研究費補助金（特別研究員奨励費）の一部によって行われた。

参考文献

- [1] C.W.Parkins: The bionic ear: principles and current status of cochlear prostheses, *Neurosurgery*, 16-6; 853-865 (1985).
- [2] 城間、河野、船坂、熊川：人工内耳の語音聴取に及ぼす因子の検討。 *Audiology Japan*, 33; 545-546 (1990).
- [3] 藤原、柗山、森：人工内耳に関する臨床検査学的一考察。 *Audiology Japan*, 34; 393-394 (1991).
- [4] T.Ifukube, Y.Hirata and J.Matsushima: A new model of auditory prosthesis using a digital signal processor. *J.Microcomputer Applications*, 13; 219-227 (1990).
- [5] 平田、伊福部、松島、柗内：蝸牛外電気刺激のためのDSPを用いた音声信号処理。 *信学技報*, SP90-95; 17-22(1991).
- [6] J.D.Markel: The SIFT algorithm for fundamental frequency estimation, *IEEE*

Trans., Audio Elect., AU-20; 367-377 (1972).

[7] D.H.Klatt: Software for a cascade/parallel formant synthesizer, *J.acoust.Soc.Am.*, 67(3); 971-995 (1980).

[8] 平原、加藤：定常母音の知覚における基本周波数の役割。 *信学技報* SP88-126; 33-40 (1988).