

1.はじめに

聴覚は音声による相互のコミュニケーションのため重要な器官だが、加齢、病気、薬など様々な要因のため聴力が低下する。特に感音性難聴では、周波数に応じて可聴領域が変化するので、線型増幅のみでは音声レベルの変化に対応し聴き取りを改善することが難しい。このような特性を補償するには、音声情報を可聴領域に入れる振幅圧縮処理方式が使用される。先に音声信号に適する方式として単共振分解型補聴方式が考案され、電話音声の圧縮特性の検討、電話帯域での振幅圧縮試料による感音性難聴者、模擬難聴状態の健聴者による聴き取り実験でその効果が確認されている[1][2]。本研究では、通常音声に対する装用可能な単共振分解型補聴器の実現についてソフトとハードの面から検討する。

2.単共振分解型補聴方式

障害により可聴領域が制限されると、限られた範囲の中に健聴者が得ている音響情報を集約しなければならない。単共振分解型補聴方式は音声の性質に着目したもので、音声のホルマント情報を利用し、圧縮に際しての非線形歪が小さいなどの特徴がある。図1に単共振分解型補聴器の基本構成を示す。現在までに、逆フィルタ制御を用いたホルマント周波数抽出部は、サンプリング周波数 12 [kHz]、1 フレーム長 20 [ms]、フレームシフト 10 [ms]で第1～第4ホルマント周波数( $F_1 \sim F_4$ )を約 6[ms]で抽出できることが報告されている[3][4]。A/D変換された音声データは、一方は高域

強調を行った後にホルマント周波数の分析へ、また第5ホルマント以降の高次ホルマントの影響を除くため Fant による高次極補正式を用いた FIR フィルタで高次極除去を行った後  $F_1 \sim F_4$  の各単共振成分抽出部へ送られる[5]。単共振信号の分離抽出は分析したホルマント周波数で各ホルマント成分に対応した逆フィルタの零周波数を順次制御して行い4つの単共振成分を得る。また、音声ホルマント周波数は分析部において各フレーム毎に抽出されるが、音声信号の処理にはフレーム毎の離散値を使用せず、フレーム間を直線内挿している。

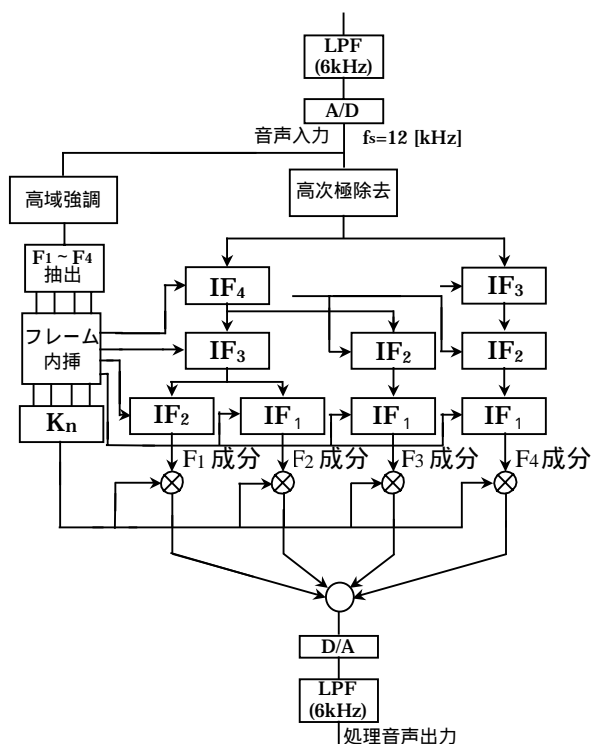


図1 単共振分解型補聴器の基本構成

\*Research on developing the model for the wearable single resonant analysis type hearing-aid. By Hiroyuki Shiotsuka<sup>†</sup> and Takashi Ikeda<sup>‡</sup> (<sup>†</sup>Advanced Engineering School, Kurume National College of Technology, <sup>‡</sup>Department of Electrical and Electronics Engineering, Kurume National College of Technology)

### 3.F3、F4成分を含めたりアルタイム処理

リアルタイムで動作するには10 [ms]のフレームシフトでのホルマント分析の時間を除くと補聴信号処理を4 [ms]以下で終わることが必要になる。

装用装置としてのハードウェア開発の前段階としてPCと組合せたDSPの開発キット(DSK)を用い処理時間について検討する。使用したDSPはTMS320C6713で、処理時間測定の結果、1フレームシフト10[ms]のホルマント分析に約5.8 [ms]、9個の逆フィルタを適用した4つの単共振成分抽出に約0.38 [ms]で処理を終えている。今回使用したDSPは過去に電話音声用として構築されたDSPシステムよりも処理時間を約1/18に短縮している。また、同システムではF1、F2成分の振幅圧縮に9.5 [ms]かかっていたが、今回使用するDSPではF1~F4を約1.1 [ms]で終わると推測でき、ホルマント周波数抽出後の残り時間内に全てのリアルタイム処理が可能と見込める。

### 4. 聴取実験

電話帯域音声から通常音声に拡張するにあたり、F3、F4成分を振幅圧縮出力に含めることが実音声の聴き取りにどのような影響を及ぼすのかを聴取実験で調べる。試料に、母音(実音声)を用い、ホルマントピークレベルを被験者の操作により変更しながら聴こえとホルマントピークの間連について聴取実験を行う。

図2のように各ホルマントの圧縮係数： $K_n$ を被験者が操作するボリュームから取り込み、各ホルマント成分に係数を乗じて圧縮を行い、合成する。実験用の回路構成は図3の通りでボリュームの値を4ch(12bit)のA/DコンバータでA/D変換し、プログラムで図4に示す対数変換後、各単共振成分に乗じる係数を設定する[6]。係数( $K_n$ )の読み込みは処理時間短縮のため、音

声フレーム毎に1つのチャンネルの値をA/Dにて取り込み、その値を図4に示す-40[dB]~20[dB]まで変化するような対数近似式より算出し、各々の単共振成分に乗じる。今回はF2、F3成分のみ被験者がその係数レベル( $K_2$ 、 $K_3$ )を操作し、他は固定とする。

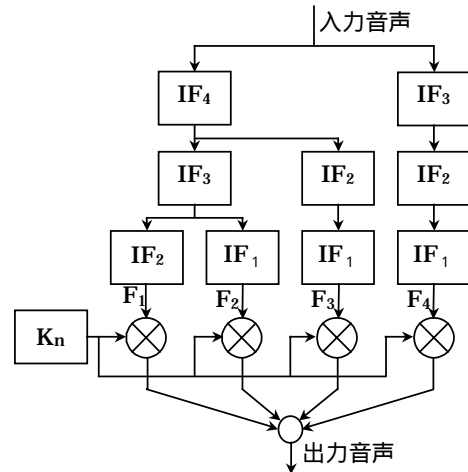


図2 聴取実験の手動係数制御

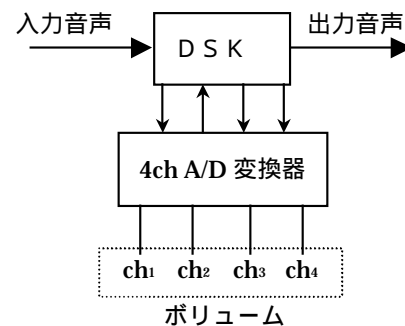


図3 A/D変換器装着時のDSKの構成

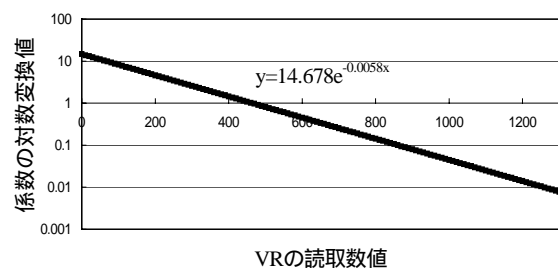


図4 係数の対数近似

図5に実験装置の全体構成を示す。

試料：予め録音した成人男声1名の発声した5母音の試料を編集し1個の単母音試料を無音区間約2 [s]で連結、単母音20回の反復試料とした音声ファイルを作成、

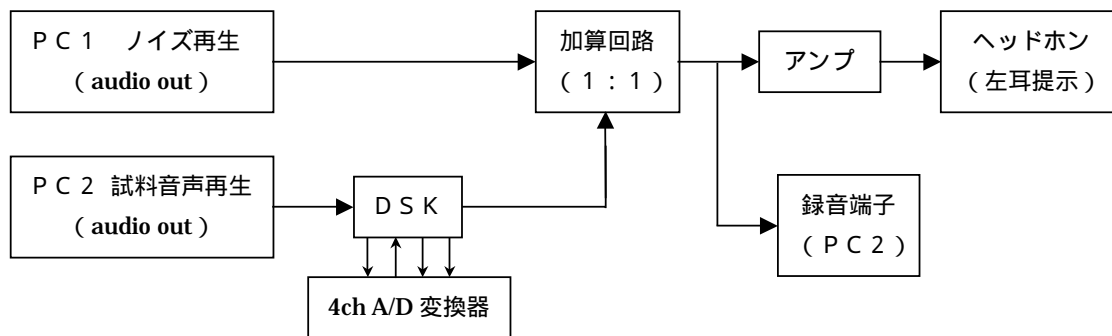


図5 実験装置の構成

PC2で再生する。提示試料はホワイトノイズに聴力レベル(HL)で平坦となる特性のフィルタのインパルス応答を畳み込んで作成したマスクノイズと加算する。

実験方法：PC2で再生した音声試料をDSKに入力し図1と同様に逆フィルタを制御して各ホルマントの単共振成分を分離抽出する。F<sub>1</sub>、F<sub>3</sub>、F<sub>4</sub>の各単共振成分には定数を乗じ、F<sub>2</sub>成分には、加算に際して被験者のボリューム操作による係数を更に乗じる。簡易聴力検査室内に着席した被験者にヘッドホンの片耳から提示する。提示している音声はPC2に録音し、録音ファイルの定常部の平均スペクトル分析によりF<sub>2</sub>のピークレベルを推定した。

実験1：図5の構成で、提示試料にF<sub>3</sub>情報が含まれる時F<sub>2</sub>レベルを上昇、下降させながら以下のレベルを調べる。

- (1)母音が聞こえだした時 :u
- (2)母音がはっきり聞き取れた時 :u
- (3)母音が不明瞭になった時 :d
- (4)再び母音が聞こえだした時 :u

u :F<sub>2</sub>レベル上昇時、d :F<sub>2</sub>レベル下降時

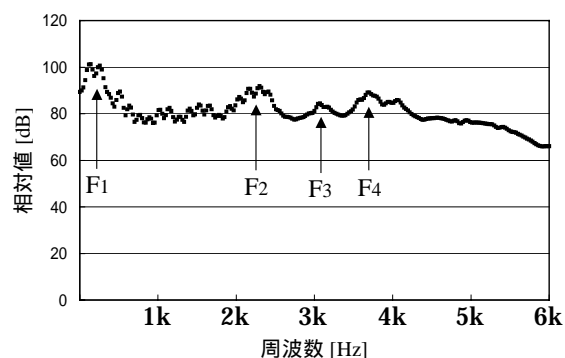
実験2：F<sub>1</sub>、F<sub>4</sub>の加算条件は変えずにF<sub>3</sub>成分をマスクノイズよりも約6 [dB]小さくするように係数レベル(K<sub>3</sub>)を調整、F<sub>3</sub>情報を聴こえないようにして同様の聴取を行う。求めるレベルは次の2つである。

聴こえ始めのレベル：L<sub>s</sub>

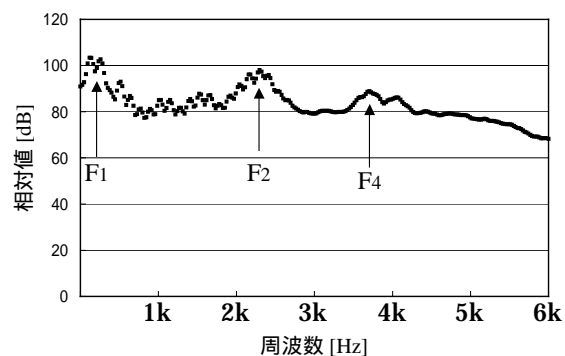
はっきり聞き取れたレベル：L<sub>c</sub>

L<sub>s</sub>は(3)母音が不明瞭になった時のレベルと再び(4)母音が聞こえ出した時のレベル

の間にあると考え、両者の平均をとりL<sub>s</sub>とした。また、L<sub>c</sub>は(2)母音がはっきり聴こえた時のレベルとした。前述の(1)~(4)を1セットとし、各母音で3セットずつ聴取した後、各レベルの平均を取る。図6にマスクノイズと加算した試料/i/の平均スペクトル例を示す。



(a) F<sub>3</sub>ピークが可聴領域に有る場合



(b) F<sub>3</sub>ピークが可聴領域に無い場合

図6 聴取試料の平均スペクトル例 /i/  
実験結果：図7に2人の被験者による実験1、実験2の結果を示す。同図で、F<sub>3</sub>ピークレベルの可聴領域での有無によらず、F<sub>2</sub>のホルマントピークのレベルが2-10[dB]程度マスクノイズより高い場合に母音として聴こえ始めていることが分かる。被験

者 HS では、/i/、/o/、/u/とそれらよりやや高いレベルで聴き取れた/a/、/e/に分かれる傾向が見える。内観では/a/、/e/は Ls と Lc が近いと感じている。また 10-20[dB]程度に Lc が分布している。被験者 OH では/a/、/e/の Lc が Ls と離れて位置しているが、母音の聴こえ初め Ls から母音が判別できていたので手動で F<sub>2</sub> のホルマントピークを上昇させても Lc に至る境界が定めにくかったとの内観を示した。

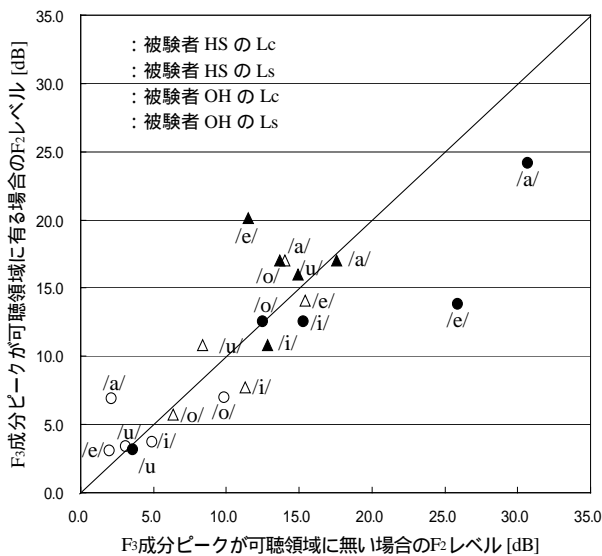


図 7 聴取実験結果

### 5.リアルタイム処理

電話音声用に構築されたと同様に、振幅圧縮係数を算出し、リアルタイムでの動作を確かめる。振幅圧縮係数を、単共振波の実効値、各ホルマントの周波数での HTL、UCL から算出、処理時間を計測した結果を表 1 に示す。単共振成分の分離、圧縮処理の合計時間は 3.46[ms]で、ホルマント周波数の抽出を含め 10[ms]以内に完了できることが分かった。

表 1 演算処理時間の内訳

	処理内容	所要時間 [ms]
1	ホルマント分析	5.80
2	逆フィルタ処理	0.38
3	振幅圧縮処理	3.08
	合計処理時間	9.26

### 6.まとめ

単共振分解型補聴方式を、通常音声に適用するため、マスクノイズで模擬難聴状態にした健聴者 2 名により、F<sub>3</sub> ピークレベルの可聴領域での有無が母音を聴き取る時の可聴領域内の F<sub>2</sub> ピークレベルに与える影響について聴取実験を行った。その結果マスクノイズの 2-10[dB]程度上方に F<sub>2</sub> のピークレベルがある付近から母音として聴き取られる傾向があり、母音をはっきり聴き取れたと感じたレベルはマスクノイズに比べ 10-20[dB]程度上方であった。また可聴領域内での F<sub>3</sub> ピークレベルの有無は顕著に影響しなかった。更に、1 個の DSP で通常音声用の単共振分解型補聴器がリアルタイムで実現できることを示した。

今後、更に、健聴者、感音性難聴者で検討を進め、聴き取りに最も適する圧縮レベルの確認と装用できる大きさの補聴器の実現を図りたい。

謝辞 本研究で用いたリアルタイム動作の DSP ホルマント抽出処理プログラムを提供頂いた熊本大学大学院自然科学研究科上田裕市先生、御討論頂いた熊本大学工学部坂田聡先生に感謝します。

### 参考文献

- [1] 池田隆, 渡邊亮, “単共振分解・合成方式とその補聴器への応用”, 信学会音声研資, SP92 - 1
- [2] 池田隆, 上田裕市, 渡邊亮, “高品質音声を伝達する単共振分解型補聴器”, 日本音響学会 57 326-336 (2001)
- [3] Akira Watanabe, “Formant Estimation Method Using Inverse-Filter Control”, IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING Vol.9 No.4 pp.317-326, May 2001
- [4] Yuichi Ueda “A real-time formant tracker based on the inverse filter control method”, TECHNICAL REPORT Acoust. Sci. & Tech. 28, 4 (2007)
- [5] 電子情報通信学会, “新版 聴覚と音声”, p227, 株式会社 コロナ社(1980年)
- [6] デジタル信号処理シリーズ, “はじめての DSP 活用大全”, p191-p205, C Q 出版社(2006)