

ハイブリッド型人工内耳システムにおける ファントム模擬呈示方式の検討*

佐藤 正幸[†] 坂田 聡^{††} 渡邊 亮[†] 上田 裕市^{†††}

([†]熊本県立技術短期大学校 ^{††}熊本大学工学部 ^{†††}熊本大学大学院自然科学研究科)

1. はじめに

現行の代表的な人工内耳用音声処理方式である SPEAK 方式では、フィルタバンク処理に基づくため、粗い音声ホルマントしか伝達できず、音声了解度や音質の点で改善の余地がある。本研究では、先に開発した高精度のホルマント推定法 (IFC 法)^[1]に基づく FPEAK 方式と SPEAK 方式を併用したハイブリッド方式を提案^[2]し、高品質の音声情報を伝達する音声処理方式の開発を目的としている。

FPEAK 方式では IFC 法による高精度のホルマント周波数が利用可能であるが、人工内耳システムでは推定したホルマント周波数が 20ch へ量子化されるため、精度良く抽出できたとしてもそれを十分にいかしきれていない。本報告では、FPEAK 方式におけるチャンネル量子化の影響に対する改善策として、触感覚において知られる Phantom Sensation を模擬した FPEAK 方式の改良について述べ、FM 音の合成及び母音同定実験によりその効果の確認を行った。

2. Hybrid 型人工内耳用音声処理方式

本方式は、図 1 に示すように、2 系統の音声特徴抽出部と送信パルス信号生成部からなり、無声音区間は SPEAK 方式、有声音区間は FPEAK 方式で音声特徴抽出及び刺激呈示を行う方式で、各々の長所を相補的に利用する方式として提案したものである。また、両方式の切り替えに必要な有聲/無聲判別には、SPEAK 方式で得られる BPF 出力 (20 次元) に対する線形判別法を適用している。

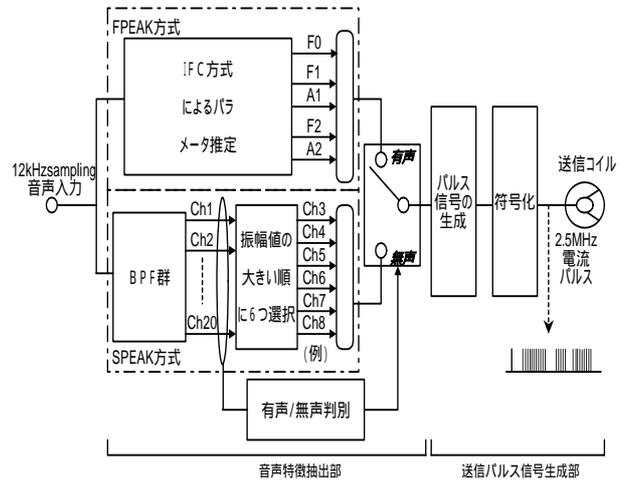


図 1 Hybrid 方式のブロック図

2.1. チャンネル量子化の影響

FPEAK 方式では、図 2 のように 0 ~ 6 kHz 間を等メルスケール間隔で 20ch に分割しているため、近接して抽出されたホルマント周波数は、同じチャンネルに量子化されてしまうことがある。すると、精度良くホルマントを推定できたにも拘らず、実際の知覚としては、どちらも同じチャンネルの周波数となってしまふ。これでは FPEAK 方式の効果を活かしきれず、ホルマント知覚の不連続性といった問題が生じてくる。

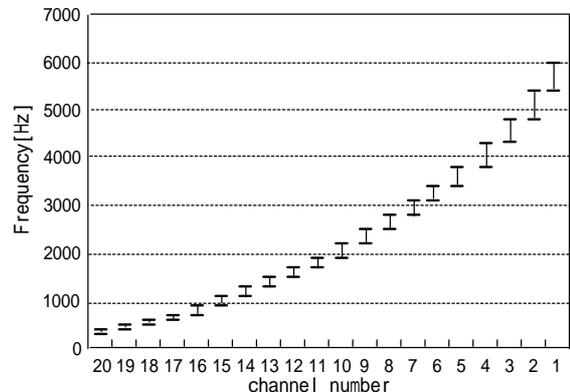


図 2 刺激用チャンネル分割と周波数帯域の関係

* A Study on the Stimulating Method based on Phantom Sensation in a Hybrid Type of Cochlear Implant System .

2.2. Phantom Sensation

Phantom Sensation(以下 Phs)とは、触覚において知られバーチャルリアリティの分野などで広く用いられている現象であり、皮膚上の2点を同時または短い時間間隔で刺激すると、2つの刺激の中間地点への1つの刺激であるかの様に知覚される現象のことである。また、刺激強度の比率を変化させることで、あたかも1つの刺激が2点間を移動するかの様に感じさせることもできる^[3]。

人工内耳システムは、電気刺激により直接聴神経を刺激して音声伝達を行うシステムではあるが、聴神経への電気刺激という点で、Phsを模擬しFPEAK方式の改良を行うことで、チャンネル量子化の影響を解決し、既存のシステムよりも連続的なホルマント遷移が実現でき、これによりさらに高品質な音声を伝達することが可能になると考えられる。

2.3. FPEAK 呈示方式の改良

前節で述べた、Phsを模擬してFPEAK呈示方式の改良を行う際の概念図を図3に示す(Phs-FPEAK方式)。

各チャンネルからの電気刺激により知覚される周波数はそのチャンネルの中心周波数であるため、Phsを模擬して分割する刺激点をチャンネルの中心周波数に設定する。このとき、図中に破線矢印で示したように、FPEAK方式により抽出された値を2チャンネルに分割し、その2点の振幅の比率を制御することでPhsを模擬する。

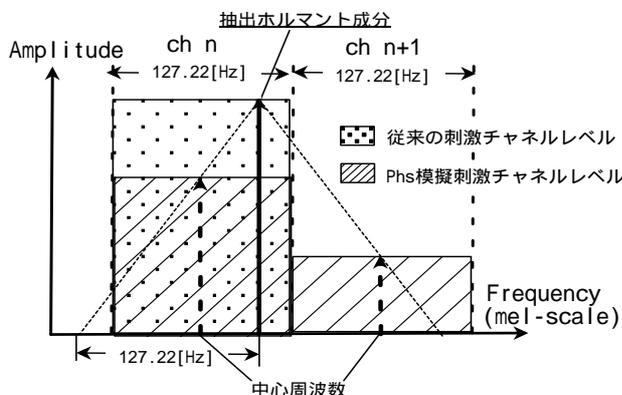


図3 Phsを模擬する呈示方式(Phs-FPEAK方式)の概念

3. FPEAK 改良方式の評価

3.1. FM音を用いた検証

いまFM音を人工内耳システムで処理すると、その中心周波数変化に伴いチャンネルが切り替わっていく。そこで、このFM音を用いて、Phs-FPEAK方式とFPEAK方式の2方式で合成音を作成して比較を行う。

検証方法としては、振幅は一定で周波数だけ変化するパラメータファイルを用いて、合成音の生成を行った。実際には、FPEAK方式では音声の第2ホルマントまでしか抽出表現しないが、ここでは、人工的に周波数を200~3000[Hz]までメルスケール間隔で一定となるように変化させた。また、パラメータ変更は、一定フレーム長(8.36ms)で設定した。図4に両方式で作成したFM合成音の波形とスペクトログラムを示す。

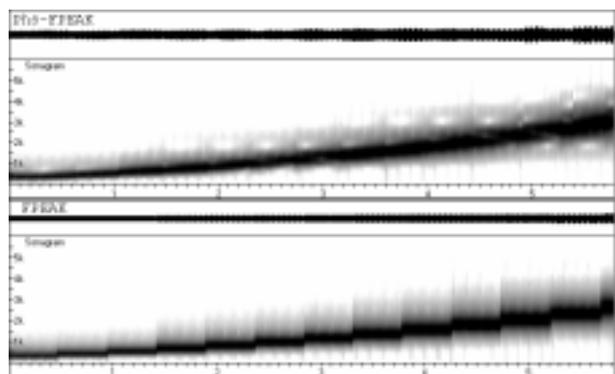


図4 FM音の合成音(上)とスペクトログラム(下)(上段:Phs-FPEAK方式,下段:FPEAK方式)

図より、FPEAK方式では周波数上昇に伴い階段状に合成音の周波数が変化しているのに対して、Phs-FPEAKではそれが改善され、チャンネル間での中心周波数が連続的に変化しているのが確認できる。また、主観評価においても、Phs-FPEAK方式の連続的な変化が聴取できた。

3.2. 母音カテゴリ評価

次に、Phs-FPEAK方式がチャンネル量子化の影響に対してどの程度効果があるのかを確認するために、母音同定実験を行い、母音カテゴリ評価^[4]を行った。以下に実験概要及び結果を示す。

(1) 合成母音の作成

合成母音は、振幅の定常部の前後に直線的に変化する各 50[ms]の立ち上がり、立ち下がりを受け、時間長 500[ms]とした。第 1、第 2ホルマント (F1, F2) は、図 5 に示すように F1- F2 平面上で格子状に配置した合計 60 試料の合成母音刺激として作成した。

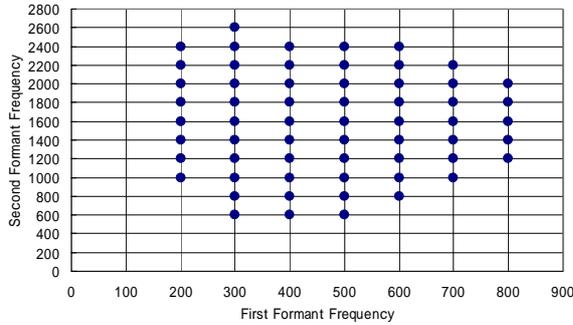


図 5 合成母音試料のホルマント周波数分布

また、ピッチ F0 は 130[Hz]，F3 は F1, F2 を変数とする 2 次の回帰式^[5] から定め、 $F4 = F3 + 1000[Hz]$ ， $F5 = F4 + 1000[Hz]$ に固定した。音源として、ピッチ以外のパラメータを一定に保った Rosenberg 波を用いた。

(2) 実験用試料の作成

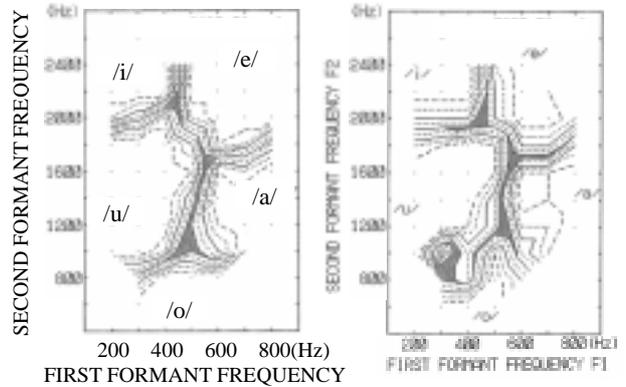
上述の合成音声を原音声 (original) として、FPEAK 方式、Phs-FPEAK 方式、SPEAK 方式により処理した模擬音声 (20ch のパルス信号を 20ch 臨界帯域フィルタバンク出力に通した出力信号の総和) を各 60 試料。計 4 種類 240 試料を作成した。

(3) 実験方法

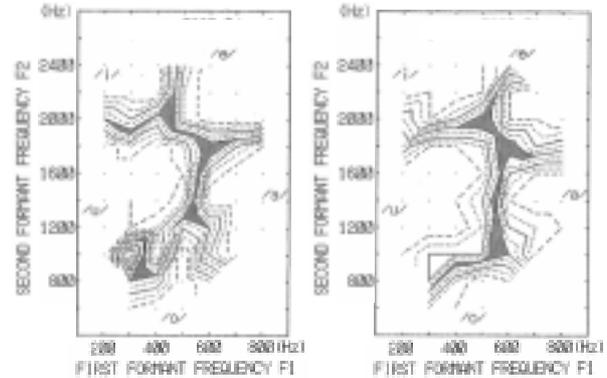
作成した模擬音声を、方式ごとにランダムに並び替え、被験者に対して PC を通してヘッドフォンにて呈示する。被験者は提示された音声に対して、日本語 5 母音のどれに聞こえたかを強制判定しモニタ上の該当するボタンをクリックすることで回答する。呈示回数は各方式 10 回ずつとした。被験者は成人男性 2 名の健聴者 (共に合成音声聴取の未経験者) である。

(4) 実験結果

被験者 MS の応答結果 (母音同定率) より推定した F1- F2 平面上での母音カテゴリ例を図 6 に示す。図の等高線は、母音ごとの同定率 50% ~ 90% を 10% 刻みで示している。



(a) 合成母音(original) (b) FPEAK方式



(c) Phs-FPEAK方式 (d) SPEAK方式

図 6 母音カテゴリ分布

(5) 考察

図 6 (b) から FPEAK 方式では original のカテゴリに比べ、/u/ /i/、/a/ /e/ の遷移が直線となっているのが分かる。これはチャンネル量子化の影響によりホルマント遷移が階段状になっていることを示す結果となっている。即ち、/u/ /i/ での遷移では、F1 が 200[Hz] と 300[Hz] で同じチャンネルに量子化される試料では、F2 が同じであれば同じ音声に聞こえてしまう。(例として、(F1=200[Hz], F2=1600[Hz]) と (F1=300[Hz], F2=1600[Hz]) とでは同じ模擬音声に聞こえる。) このことが原因となり図 6 (b) のような直線的な遷移になったと考えられる。

これに対し，図 6 (c) の Phs-FPEAK 方式では，FPEAK 方式のカテゴリに見られたような遷移の不連続性は見られず，緩やかに遷移している．

このことから，Phs-FPEAK 方式は FPEAK 方式に比べ，ホルマント遷移の連続性という点で大幅な改善がみられる．したがって，触覚における Phantom Sensation を模擬した Phs-FPEAK 方式は FPEAK 方式でのチャンネル量子化の影響を解決するのに有効であるといえる．

次に，SPEAK 方式においては，主観評価では他の 2 方式と比べ，明らかに音質が良くないが，図 6 (d) のようにカテゴリ図においては original との間に大きな相違は見られない．これは，SPEAK 方式ではピッチ周波数を一定にして刺激呈示しているため，抑揚のない音声となり，このことが母音同定率を良くしているものと考えられる．しかし，これまでの実験^[2] (VCV 音節を使用した明瞭度試験) においても明瞭度は他の 2 方式と比べ良くないことから，入力音声信号 (有声音部) の再現性という点で，他の 2 方式の方が優れているといえる．

なお，FPEAK 方式及び Phs-FPEAK 方式の /u/ /o/ での遷移部分にカテゴリの島 (判定の不安定な標本) が見られるが，これは被験者が今回初めての試験であったことから，今後，習熟度を増すことにより改善されるものと思われる．

4. まとめ

FPEAK 方式のチャンネル量子化の影響への解決を図るために FPEAK 方式の改良として触覚における Phantom Sensation を模擬した Phs-FPEAK 方式の提案を行った．そして，その有効性を検証するため，FM 音の合成及び母音カテゴリによる評価を行った．

FM 音を用いた評価では，作成した合成音のスペクトログラムから FPEAK 方式の合成音に見られる階段状の不連続な周波数変化を Phs-FPEAK 方式を用いることで，より連続的で滑らかな周波数変化にできることが確認できた．

また，母音同定による母音カテゴリ評価では，作成した母音カテゴリ図 (図 6) から，母音同定率が緩やかに遷移していることが確認できた．

しかし，Phs-FPEAK 方式における FM 音の合成音や模擬音声からは，わずかではあるが“歪み”が聞き取れた．この原因としては，周波数弁別に起因する 2 チャンネルの分離が問題となっているものと思われる．これは人工内耳システムの仕様上仕方ないことではあるが，今後，この問題が実際の音声聴取において，どのような影響を及ぼすのか，より詳細に検討する必要がある．

参考文献

- [1] A.Watanabe, "Formant Estimation Method Using Inverse-Filter Control", IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING Vol.9 No.4, pp.317-326, 2001
- [2] 佐藤正幸, 西口直宏, 坂田聡, 渡邊亮, 上田裕市, "ハイブリッド型人工内耳用スピーチプロセッサと評価システムの開発", 信学技報, Vol.104, No.7 50, pp.7-12, 2005
- [3] 前野隆司, "ヒトの触覚受容機構 - 力学・アナロジー・錯覚という視点から -", 電気学会センサマイクロマシン部門誌, Vol.122-E, No.10, pp469-473, 2002
- [4] Y.Ueda and A.Watanabe, "Visible/tactile vowel information to be transmitted to the hearing impaired," J.Acoust.Soc.Jpn.(E)8, pp.99-108, 1987
- [5] 渡邊亮, 濱崎良介, 上田裕市, "電話音声の映像化を目的としたパラメータ抽出・推定法", テレビジョン学会誌, 45, pp.233-243, 1991