

黙声認識における 冗長ウェーブレット係数の集合選択に依存した周波数特徴の違い

Difference of Frequency Characteristics
Depending on Selecting a Set of Redundant Wavelet Coefficients in Inaudible Speech Recognition

永井 秀利[†]
Hidetoshi Nagai

1.はじめに

我々は、声を出さずに自然に発声(いわゆる口パク)した内容を口唇周辺から頸部の表面筋電に基づいて認識することを黙声認識と呼び、その実現を目指している。

周波数特徴に基づいて信号の時間変化を捉えようとする場合には窓フーリエ解析を用いることが多いが、自然な発声時に継続的に生じる非周期的で急激な筋活動変化を追跡するには、より時間解像度が高いウェーブレット解析の方が適する。しかし、離散ウェーブレット解析手法として一般的な多重解像度解析では、得られるウェーブレット係数が対応する時間位置と幅とは周波数帯域によって変化する。そのため、ある時点における特徴を捉えるためのウェーブレット係数集合の選択方法には複数の候補が存在する。

本稿では、すべてのサンプル時刻を起点とするすべてのウェーブレット係数を算出するような冗長離散ウェーブレット解析を行った場合において、ウェーブレット係数集合の選択と得られる筋活動特徴の違いについて述べる。

2.表面筋電信号の獲得

自然な速度での日本語発声では、子音+母音を1音の単位として、1秒間に5～10音程度を発すると想定できる。この場合、1音当り100～200ms程度となるが、子音に伴う筋活動変化を捉えるには各音先端の数msから十数ms程度の区間での変化を追跡する必要がある。

一般的には、表面筋電の周波数特徴を捉えるには1000～1500Hz程度までの成分で十分とされている。それゆえ、サンプリング周波数を2000～3000Hz程度に設定することが多い。しかし、この周波数で得られるサンプル数では、特に子音発声に伴う変化を追跡するには十分とは言えない。サンプル数が少なければ、値の変化が活動変化なのかノイズ的なものなのかを判断することも難しい。

そこで本研究では、一般的な表面筋電計測に比べて遥かに高い周波数である10k～20kHzでサンプリングしたものを計測データとして使用する。

3.冗長ウェーブレット解析

ウェーブレット解析では、ウェーブレットの形状が解析対象信号の成分波形に近いほど、より感度良く特徴を捉えることができる。表面筋電は筋繊維の励起によって生じた電位変化の重畳であるので、サポートが狭く単発の励起波形にできるだけ近い形状のウェーブレットを用いるのが望ましい。よって本研究では、解析用ウェーブレットとしてDaubechies 4-tap (N=2)を用いた多重解像度解析を行う。

/u//me/と連続発声した際の口角下制筋の表面筋電を20kHzでサンプリングしたデータに対し、多重解像度解析と、ウェーブレット縮退によるノイズ低減処理[1, 2]とを施した結果の時間-周波数平面を図1に示す。図の上部に描かれた波形は、解析の対象となった表面筋電信号の波形である。活動が活発化するにつれて、湧き上がるように高周波成分が増大している様子を観測できる。

多重解像度解析にはシフト不変性の欠如という問題があることは良く知られている。そこで本研究では、すべての

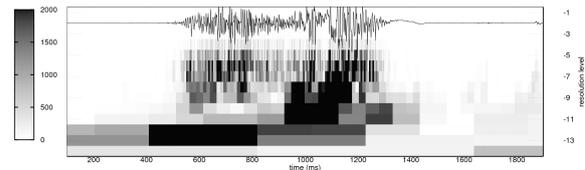


図1: 連続発声/u//me/の下唇下制筋の表面筋電の時間-周波数平面

サンプル時刻においてその時刻を起点とする多重解像度解析を実行することで、算出可能なすべてのウェーブレット係数を求める冗長ウェーブレット解析を行う。必要な計算量[3]は解析を行う周波数帯域レベルの数 L に依存する。計算の途中過程を保持するために領域計算量は $O(L \cdot 2^L)$ となるが、サンプル入力毎の時間計算量は $O(L)$ である。

4.ウェーブレット係数集合

多重解像度解析で得られるウェーブレット係数は、対応する時間-周波数領域の時間幅が周波数帯域によって異なる。そのため、ある時刻の周波数特徴を調べる際にかなるウェーブレット係数集合を用いるかが問題になる。解析の計算過程を考えた場合、共通の先端位置を持つウェーブレット係数集合とするのが基本であり、次いで共通の末端位置を持つ係数集合とするのが候補となろう。冗長ウェーブレット解析を行った場合、すべてのサンプル時刻でこれらの係数集合を得ることができる。

しかしながら、計算上の直接的な関係という制約を取り払い、対応する時間-周波数領域の値として単純に捉えるなら、各周波数帯域におけるウェーブレット係数の選択をより柔軟にすることが可能となる。

レベル-1の係数の時間幅が2サンプル分であり、レベル- n の時間幅はレベル-($n-1$)の2倍幅であることを鑑みると、規則性を持った係数選出位置のタイプは図2に示す7種となる。それぞれ、図の $T=0$ のサンプル位置での特徴算出に用いることを想定している。

位置タイプに加え、オフセットタイプとして図3に示すように、レベル毎に-2～2サンプル時刻だけずらしつつ選出することも可能である。ただし、位置タイプとオフセットタイプのすべての組が有効というわけではなく、図4に示すように、不適切となる組もある。表1に位置タイプとオフセットタイプとの組み合わせの可否を示す。レベルごとにオフセット幅を変える等、表1以外の変則的な選択とすることはもちろん可能だが、その意味や意義の定義が難しくなるため、可とされた25種をウェーブレット係数集合の全パターンと定義しても差し支えないと考える。

表1: 位置タイプとオフセットタイプの組み合わせ可否

	a:-2	b:-1	m:0	x:+1	z:+2
A:全先行	×	×	○	○	○
B:半先行	×	○	○	○	×
H:先端揃え	○	○	○	○	○
M:中央揃え	×	○	○	○	×
T:末端揃え	○	○	○	○	○
X:半遅延	×	○	○	○	×
Z:全遅延	○	○	○	×	×

[†]九州工業大学, Kyushu Institute of Technology

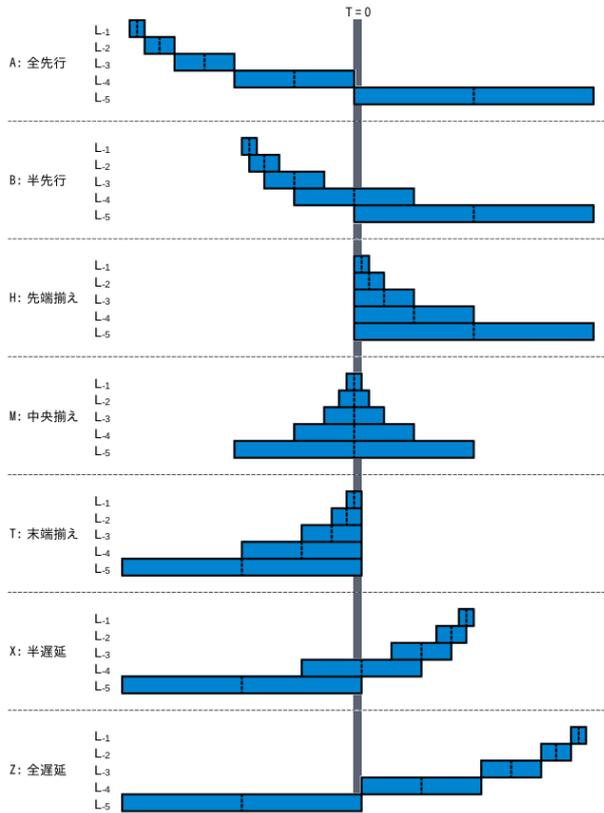


図2: ウェーブレット係数集合の選出位置タイプ

図5に、図1の筋電信号(図5の最上段)を L_{-11} まで解析した際の係数集合のウェーブレット重心[4]を示す。上から順に、オフセット0の全先行(Am)、先端揃え(Hm)、全遅延(Zm)の結果である。係数集合間の時間遅れのような関係は、計算に要する領域量の増大を代償に図2の $T=0$ を L_{-1} の係数の位置に取る(オフセットありも同様に扱う)ようにすれば解消されるが、その差異を除いても、全先行は活動低下に、全遅延は活動上昇に、より高い感度を持つ傾向が見て取れる。また全遅延の場合は、1音の発声区間での値の推移が他と比べて平坦である傾向が強いように見える。

5. おわりに

非周期的で短時間に生じる筋活動変化を追跡するために、冗長離散ウェーブレット解析によって得たウェーブレット係数に対し、各時刻の特徴抽出の基盤となる係数集合のパターンを規定した。ウェーブレット係数集合によって呼応しやすい活動特徴には違いが存在する。そのため、複数パターンを同時に利用することにより、特徴抽出能力を向上できる可能性がある。今後は、ウェーブレット係数集合ベースでの信号強度特徴なども含めた係数集合の特性の更なる分析に加え、発声変化位置の検出のような具体的課題において係数集合の特徴差を活用する手法の開発などを目指す。

謝辞 本研究の一部は科研費(課題番号:16H03211)の助成を受けた。謹んで感謝の意を表す。

参考文献

[1] D. L. Donoho, "De-Noising by Soft-Thresholding," IEEE Trans. Information Theory, Vol. 41 No. 3, pp. 613-627 (1995)
 [2] 永井, 中村, 野村: "無発声ないし微発声音声認識のための表面筋電波形からのノイズ低減手法", 情報処理学会九州支部「火の国シンポジウム2003」, pp.1-8 (2003)

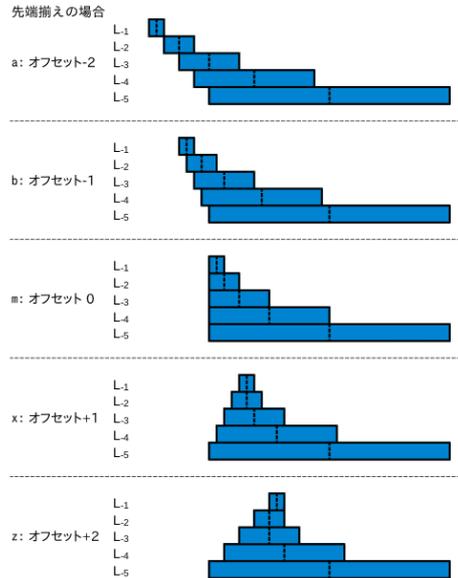


図3: ウェーブレット係数集合のオフセットタイプ

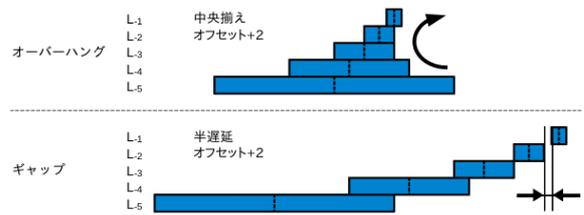


図4: 不適切なウェーブレット係数集合選出パターン

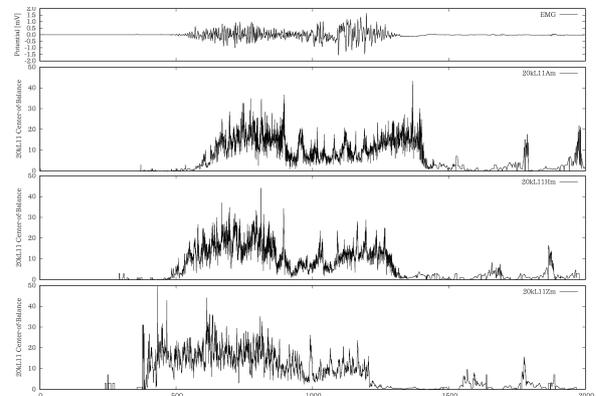


図5: 係数集合によるウェーブレット重心の違い

[3] H. Nagai, "Computational Complexity in Continuous Estimation of Muscular Activity based on Redundant Wavelet Coefficients of Surface EMG", Proc. of Life Engineering Symposium 2015, pp. 329 - 334 (2015)
 [4] 永井, 宮武, 中村: "黙声認識のための口唇周辺表面筋電波形におけるウェーブレット係数の重心推移法に基づく特徴", FIT2011講演論文集 vol.3, pp.41-46 (2011)
 [5] S. G. Mallat, "Zero-crossings of a wavelet transform," IEEE Transactions on Information Theory, 37-4, pp. 1019-1033 (1991)
 [6] 戸田, 川畑, 章: "最新ウェーブレット実践講座 入門と応用", ソフトバンククリエイティブ (2005)
 [7] 福永哲夫編: "筋の科学事典 —構造・機能・運動—", 朝倉書店 (2002)