

1. まえがき 一般化調和解析は1958年にウイナーにより提案された周波数解析手法の一つであり、東山、平田らにより音響信号の解析に導入された。我々は高速1bit信号の量子化雑音の解析や高速1bit信号の高エネルギー符号化に一般化調和解析を導入しその有用性を示した。

一方、ボコーダは1939年に提案された音声分析合成方式の一つである。近年、河原らによりボコーダの原理そのもので高音質で自然な音声分析合成方式 (STRAIGHT) が提案されるなど、その技術がもう一度見直されつつある<sup>3)</sup>。我々も窓の影響を受けることなく周波数解析が可能な一般化調和解析を用いて音声を分析しボコーダの原理に基づき合成するという、正確な分析、単純な合成を特徴とする音声分析合成法を報告した<sup>4)</sup>。

今回、一般化調和解析による信号分析とボコーダの原理に基づいた合成法を併用することにより、ピッチ変換、時間軸圧伸、声質変換、雑音中の信号の抽出等のシグナルエンハンスメントを試みたので報告する。

2. 一般化調和解析 一般化調和解析は観測区間で原波形から残差が最小となる純音を引き、その残差成分に対して同様の処理を行うことにより周波数成分を逐次抽出していく周波数解析手法である。窓の影響を受けることなく非周期的信号の解析を行うことが可能である等の特徴がある。

3. 分析合成法 一般化調和解析により音声を分析し、その結果から聴覚特性を考慮しフィルタと駆動音源を作成し、ボコーダの原理により合成する。具体的処理を以下に記す。

まず、音声信号をある長さの解析区間で一般化調和解析し主要周波数成分を抽出する。この操作を解析区間を移動しながら行うことにより、音声信号の周波数成分の時間変化を求める。

駆動音源については、解析区間において最も大きなエネルギーを持つ成分でかつ周波数が60~800Hzのものを基本周波数とする。基本周波数が見つからなくかつその解析区間でのエネルギーがある閾値以上の場合は無声音区間とする。また、エネルギーがある閾値以下の場合は無音区間とす

る。有声、無声の判別、基本周波数をもとに、有声音区間では基本周波数で発生するパルス列を、無声音区間ではホワイトノイズを生成し、それを駆動音源とする。

フィルタについては、聴覚特性を考慮し作成した。聴覚フィルタのモデルに Gammatone フィルタがあり、各周波数成分を中心周波数とした Gammatone フィルタのゲインをその周波数成分の大きさにあわせ、加え合わせることで駆動フィルタを作成する。

このフィルタを駆動音源で駆動することにより音声信号が合成される。

4. 声質の変換 ピッチ変換、時間軸圧伸、声質変換について検討を加えた。

ピッチ変換では、基本周波数を変化させて生成した駆動音源で、周波数を変化させて生成したフィルタを駆動し、時間を変えずにピッチのみを変化させた音声を得る。

時間軸圧伸では、解析区間の移動幅より大きな移動幅で生成した音源でフィルタを駆動することにより時間軸を伸張できる。逆に、解析区間の移動幅より小さな移動幅で生成した音源でフィルタを駆動することにより時間軸を圧縮できる。

声質変換については、駆動フィルタのゲインの増減、ピッチの変換を行ない変換する。

声質の変換についての例を図-1に示す。一般化調和解析は1/30秒の解析区間で64成分の周波数を抽出し、1/300づつ移動しながら行った。(a)に原信号、(b)に解析結果、(c)に合成信号、(d)にピッチ変換(1.5倍)した信号、(e)時間軸圧縮(0.5倍)した信号を示す。いずれの場合も、かなり自然な音声の変換ができた。声質変換については女性の声を男性の声に、またはその逆のような異性への声への変換については容易であったが、同性間の微妙な声質の変換については現在のところ必ずしも十分な結果を得るにいたっていない。

5. 雑音中の信号の抽出 音声成分が多い区間では前後の駆動フィルタ間に強い相関があり、雑音成分が多い区間では相関の値が小さい。

基本周波数が抽出できず、かつその解析区間や

\*Signal enhancement using generalized harmonic analysis and vocoder method.  
By Yasuhiro Oikawa, Hisao Hattori and Yoshio Yamasaki (Waseda University).

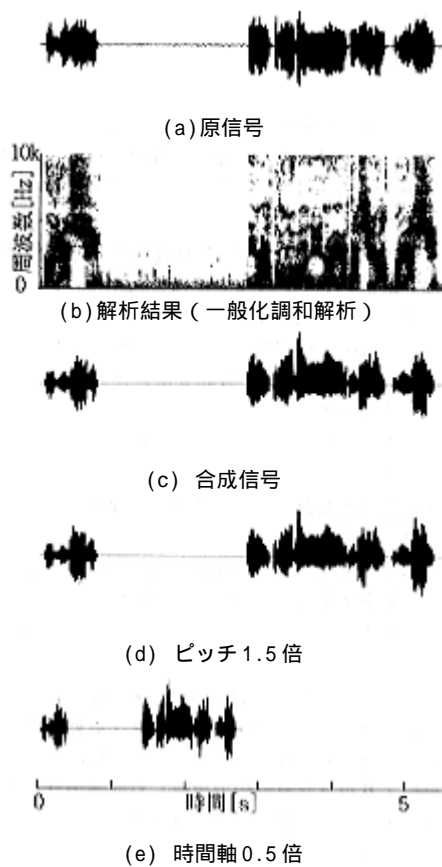


図 - 1 音質の変換

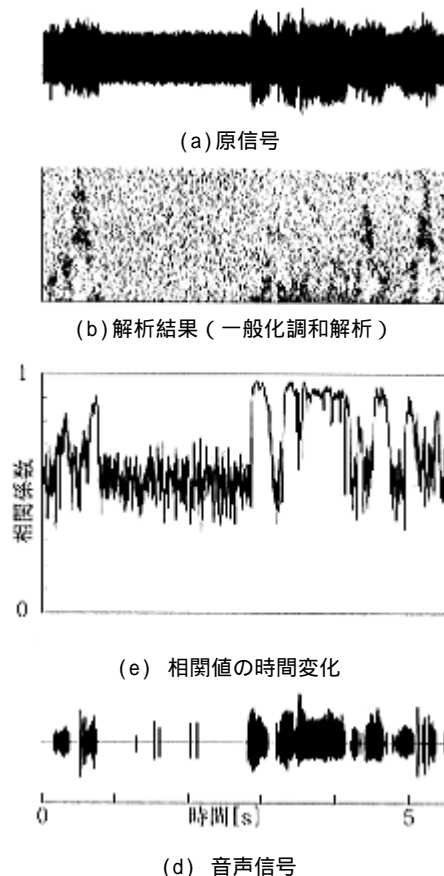


図 - 2 雑音中の音声の抽出

の駆動フィルタのインパルス応答と前区間での駆動フィルタのインパルス応答の相関係数がある閾値以下の場合には雑音区間とし、雑音区間以外の区間で音声を合成する。音声がある区間に含まれる雑音については、窓の影響を受けずに大きな成分から解析を行っていくという一般化調和解析の特徴から雑音成分を抑圧したフィルタが生成され、雑音抑圧された音声が合成される。

雑音中の音声信号の抽出についての例を図 - 2 に示す。(a)に音声信号にホワイトノイズをSN比0dBで加えた信号、(b)に解析結果、(c)に相関係数の変化の様子、(d)に抽出した音声信号を示す。およそ雑音中から音声の抽出を行うことができたが、雑音区間と無声音の判別ができず合成されない無声音が生じてしまった。

6.むすび 今回、一般化調和解析による信号分析とポコーダの原理に基づいた合成法を併用する

ことにより、ピッチ変換、時間軸圧伸、声質変換、雑音中の信号の抽出等のシグナルエンハンスメントを試みた。その結果、解析区間の移動を音声の基本周期と同じ程度として解析を行った場合かなり自然な音声の変換、および信号の検出ができた。今後は、一般的な音響信号への適用についての検討を行う。

文献

- 1) N.Wiener, "The Fourier Integral and Certain of Its Application," Dover Publication Inc (1958).
- 2) T.Terada, H.Nakajima, M.Tohyama and Y.Hirata, "Nonstationary Waveform Analysis and Synthesis Using Generalized Harmonic Analysis," IEEE-SP, 429-432 (1994).
- 3) 河原英紀, 増田郁代, "時間周波数領域での補間を用いた音声の変換について," 信学技報, EA96-28, 9-16 (1996).
- 4) 及川靖広, 渡辺秀明, 山崎芳男, "一般化調和解析を用いた音声分析合成法," 音講論集, 445-446 (1999.3).