

及川靖広 渡辺秀明 山崎芳男 (早大理工)

1. まえがき 一般化調和解析は1958年にウィナーにより提案された周波数解析手法の一つである¹⁾。最近では、東山、平田らがこの手法をピアノの音や人の歌声などの非定常信号の解析に導入した²⁾。また、我々は高速1bit信号の量子化雑音の解析や高速1bit信号の高能率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、ポコーダは1939年に提案された音声分析合成方式の一つであるが、近年、河原らによりポ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³⁾。

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ポコーダの原理に基づき音声を合成するという、正確な分析、単純な合成を特徴とする音声分析合成法を実現したので報告する。

2. 一般化調和解析 一般化調和解析は観測区間内で原波形から残差が最小となる純音を逐次抽出し、残差成分に対して同様の処理を行うことにより解析を行う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. 声分析合成法 ポコーダは、複数のバンドパスフィルタのゲインを調整することにより声道の特性を表すフィルタを作成し、有声音部分では音声の基本周期で発生するパルス列、無声音部分では白色雑音を音源としてそのフィルタを駆動することにより音声を合成する。我々は、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、ポコーダの原理により合成した。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し、周波数成分の周波数、振幅を求める。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。

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した。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した。また、エネルギーがある閾値以下の場合は無音区間とした。

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、それを駆動音源とした。

フィルタについては、聴覚特性を考慮し作成した。聴覚で行われている音の分析は、バンドパスフィルタを周波数方向にたくさん並べてフィルタバンクを構成して、各周波数に対応するフィルタの出力を得ることにより行われていると考えられている。聴覚フィルタのモデルにGammatoneフィルタがあり、そのインパルス応答は

$$gt(t) = at^{n-1} \exp(-2\pi b t) \cos(2\pi f_c t + \varphi)$$

で表される。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。音声を一般化調和解析により解析した結果とそれにより求めたフィルタの周波数特性を図-1に示す。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. 音声信号の分析と合成 女性アナウンサーの発声による「本日は晴天なり。いかがでしょうか。どうでしょうか。」を音源として、本方式による分析、合成を行った。音源を図-2に示す。

一般化調和解析は1/30秒の解析区間で128成分の周波数を抽出し、1/30秒ずつ移動しながら行った場合と1/300秒ずつ移動しながら行った場合の2通りについて行った。解析結果を図-3に、この結果から求めた基本周波数の時間変化を図-4に、合成音を図-5に示す。1/30秒ずつ移動しながら解析、合成を行った場合は原音とはかなり異なった音となったが、1/300秒ずつ移動しながら解析、合成を行った場合はかなり原音に近い音声を合成することができた。

* Voice analysis and synthesis using generalized harmonic analysis.

By Yasuhiro Oikawa, Hideaki Watanabe and Yoshio Yamasaki (Waseda University)

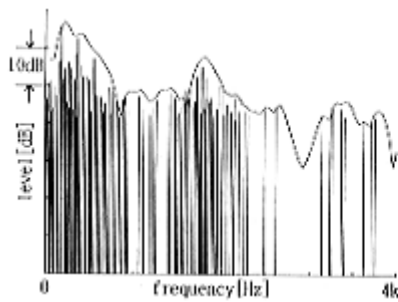


図 - 1 周波数成分とフィルタの特性

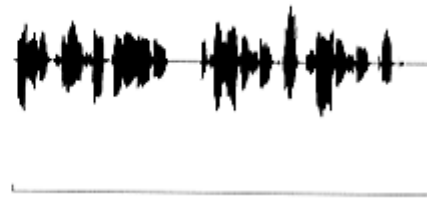


図 - 2 音源信号

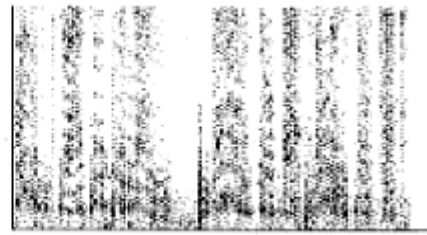
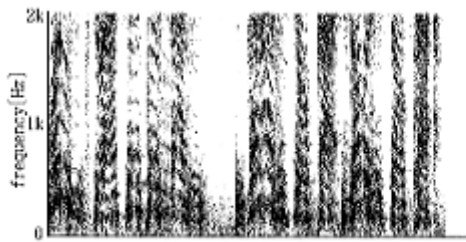


図 - 3 一般化調和解析による解析結果

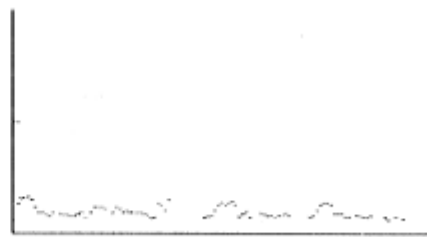
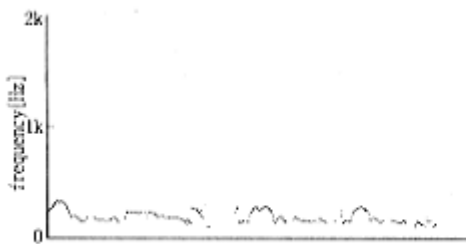


図 - 4 基本周波数の時間変化

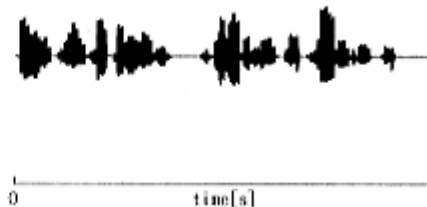
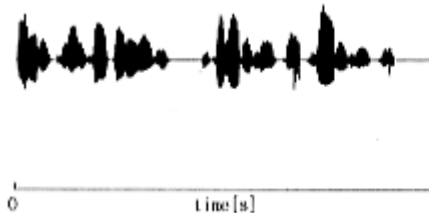


図 - 5 合成信号

5. むすび 今回、我々は分析には一般化調和解析を用い合成はポコーダの原理に基づく音声信号の分析合成を試みた。その結果、解析区間の移動の仕方により音質にかなりの差が生じる結果となったが、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原音に近い音声合成できた。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).