

音声の全極型予測残差の極零対フィルターによる分析

伏木田 勝信

KF 研究所

1. まえがき

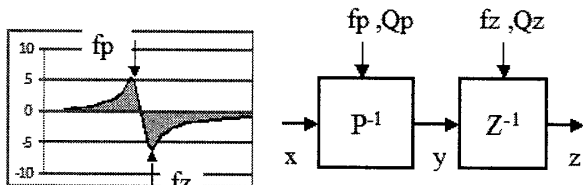
音声の周波数スペクトルの全極型の近似には限界があり、予測残差情報を繰り返すことで音韻性が向上する[1]。また、全極型近似誤差の原因としては、声道中の分岐、鼻腔結合、声帯の開閉などによる零点・極零対の出現などが知られている[2],[3]。これらは、調音系や聴覚系のモデルの高度化・精密化が必要であることを示唆している。ここでは、全極型の予測フィルターの残差成分を極零対フィルターを用いて近似することにより近似精度の向上を検討した結果について述べる。

2. フォルマント (全極型) 分析方式

フォルマント (極) 抽出方式としては、自己相関領域での逆フィルタリング方式と多段推定方式 (MSF(Multi-Step-Focusing)-ADIF(Autocorrelation-Domain-Inverse-Filtering)) による AbS (Analysis-by-Synthesis) 的方式を用いた[5],[6]。この際、フォルマント数は可変とし、ダイナミックプログラミング (DP) を用いて最適選択する方式を採った[4]。なお、フォルマントの尖鋭度 Q (フォルマント周波数 / バンド幅) の値は固定 ($Q=10$) とした。

3. 極零対フィルターと利得スペクトル

図-1 に極零対フィルターの周波数スペクトルと逆フィルタリングのブロック図を示す。



↑強度 → 周波数
 $\{fp, fz\}, \{Qp, Qz\}$ はそれぞれ極, 零の周波数と Q を表す

図-1. 極零対フィルターの周波数スペクトル (左) と逆フィルタリングのブロック図 (右)

極零対スペクトルは、極・零周波数が同じ ($fp=fz$) で Q が異なるとき、単一の凹 ($Qz>Qp$)、あるいは凸 ($Qp>Qz$) 特性となり、同一 Q のとき消滅する。

A pole-zero pair analysis for speech predictive residual wave using all-pole formant model.

Katsunobu Fushikida, KF Laboratory
 fushikida@muf.biglobe.ne.jp

極零対の逆フィルタリングは、時間領域で行った。これは、次の(1), (2)式で表わされる。

$$y_n = \alpha_1 x_{n-1} + \alpha_2 x_{n-2} + x_n \quad (1)$$

$$z_n = y_n - \beta_1 y_{n-1} - \beta_2 y_{n-2} \quad (2)$$

ここで、 $\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}$ は、それぞれ、極、零に対応する線形予測係数を表す

極零対フィルターは周波数領域での局所的な近似方式として適している。極零対の近似利得としては、次の(3)式で表わされる G (dB)を用いた。

$$G = 10(\log_{10} \sum_n x_n^2 - \log_{10} \sum_n z_n^2) \quad (3)$$

横軸に極零対の中心周波数 ($fc=(fp+fz)/2$) をとり、縦軸に利得 (G) を用いて表した特性を、ここでは、極零対の利得スペクトル (pole-zero pair gain spectrum) と呼ぶ。 G が負の場合は $G=0$ とした。極零の周波数間隔は $fd=100*(fz-fp)/fc$ (%)で表わす。 fd が負のとき極零の周波数は逆転する。

4. 極零対合成音の分析実験

インパルス列を音源とし、極零対フィルターを通した合成音を作成し、極零対分析実験を行った。一例として、2対の極零対フィルターによる合成音 (サンプリング周波数: 22.05kHz, $fc=720, 2000$ Hz, $Qp=Qz=10$) の極零対の利得スペクトルを式(1), (2), (3)を用いて2通りの fd について算出した結果を図-2に示す。

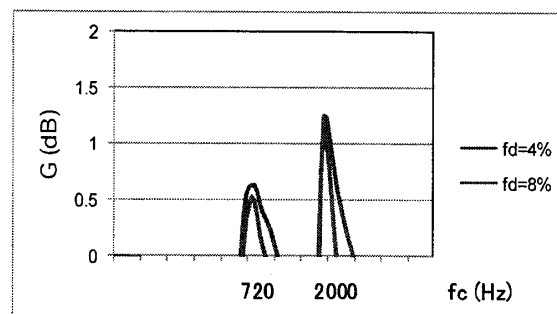


図-2. 極零対合成音の極零対利得スペクトル (fd は中心周波数に対して 4% と 8%)

図-2に示されるように極零対の中心周波数は極零対の周波数間隔 (fd) が変化(4%, 8%)しても、極と零の Q が同じ (対称) の場合、利得スペクトルの極大値として合成の際与えられた値に一致することが確かめられた。

5. 自然音声の分析実験

自然音声サンプルとしては 5 母音と鼻音節連鎖 /mamimumemo/ (男声) を用いた。フォルマント (極) は 2 節の方式により抽出された。図-3 に示す分析システムにより予測残差波形を算出した後、(1), (2), (3) 式を用いて極零対の逆フィルタリングを行い、極零対の利得スペクトルを算出した。

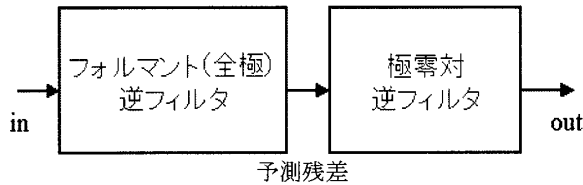


図-3. 全極予測残差の極零対分析システム

極零間の周波数間隔 (fd) と Q を変化させて重ね書きした時の、極零対の利得スペクトル例を図-4 に示す。極零の Q 値としては、非対称 (Qz>Qp) の方が対称の場合より利得が大きい傾向が見られた。(-20< fd <24%) とし、0.1~6kHz の範囲内で比較的大きい利得の極大値を与える数個(5~8)の極零対を選択し極零対組とした。

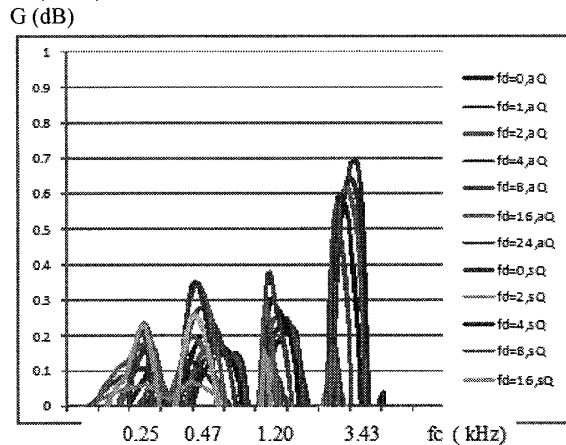


図-4. 全極予測残差の極零対利得スペクトル例 (m-e 過渡部、縦軸は利得G、横軸は極零対の中心周波数 fc を表す。24>fd (%)>0, Qp<Qz, Qp=Qz)

表-1. 鼻子音部と母音部における極零対の近似利得

/mamimumemo/ の鼻子音部と母音部におけるフォルマントと極零対組の近似利得 (dB)

	m(a)	(m)a	m(i)	(m)i	m(u)	(m)u	m(e)	(m)e	m(o)	(m)o	average
Formant	25.1	22.7	26.5	22.6	27.5	25.1	26.3	21.9	26.7	26.1	25.1
Pole-zero	2.36	1.43	1.58	2.61	1.29	1.36	2.14	1.28	1.76	1.04	1.69

図-5. フォルマント周波数、極零対中心周波数の分析例 /mam. ./ の/a/の部分のフォルマント周波数と極零対中心周波数の抽出結果 (Hz)

Formant			664	1205			2975	3395	3854	
Pole-zero pair	274	404		986		1453	2704			4309

表-1 に鼻子音部と母音部における極零対組の近似利得を示す。表-1 に示すように平均して 1.7dB 程度の利得の向上が認められた。各部における組の近似利得は組内の各極零対利得の総和の約 80% であった。このことは、各極零対間の相互の影響が小さく、極零対の利得スペクトルにより極零対の近似特性が (視覚的に) 把握し易いことを示している。低周波数領域 (約 1kHz 以下) で抽出された極零対の利得は、鼻子音部において比較的大きい傾向が見られた。

離散発声された母音/a/, /i/, /u/, /e/, /o/ についても同様の実験を行った。その結果、平均 1.8dB 程度の利得が得られ、極零対の近似効果が認められた。

フォルマントと極零対の分析結果例として、図-5 に /mamimumemo/ 中の/a/の部分において抽出されたフォルマント周波数と極零対中心周波数の分布を示す。極零対中心周波数はフォルマント周波数と相関があるように見受けられる。[7]

6. 合成実験

抽出された極零対中心周波数およびフォルマント周波数を用いて極零対フィルターの有無に対して合成音を作成した。非公式な比較聴取実験により極零対組による音質の若干の改善が認められた。

7. あとがき

音声波形から AbS 的方法を用いて抽出されたフォルマント (全極型) の予測残差波形を極零対フィルターにより分析、近似することにより近似精度を改善する検討を行った。極零対合成音および自然音声の分析実験により極零対の利得スペクトルによる推定法の有効性を示し、自然音声の近似利得が改善されることを確かめた。

参考文献

- [1] 伏木田他, 日本音響学会, 音声研究会資料, S74-23, 1974.
- [2] 本多他, 電子情報通信学会技研報告, SP2003-188, 2004.
- [3] 藤村, 音声科学原論, 2.10-2.11, 岩波書店, 2007.
- [4] 伏木田, 電子情報通信学会総合大会, D-14-25, 2006-3
- [5] 伏木田, 日本音響学会, 音声研究会資料, S81-41, 1981.
- [6] Fushikida, Proc. IEEE ICASSP, E2.8:2260-2263, 1988.
- [7] www5b.biglobe.ne.jp/~hfyf/f/Laboratory/Laboratory.htm