

論文 / 著書情報
Article / Book Information

論題(和文)	音声スペクトル推定零点に含まれる個人性情報
Title(English)	
著者(和文)	嵯峨山 茂樹, 古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1977年講演論文集, , , pp. 385-386
Citation(English)	, , , pp. 385-386
発行日 / Pub. date	1977, 4

山崎 茂樹 古井 貞熙
(日本電信電話公社 武蔵野電気通信研究所)

1 はじめに

調音器官の個人差には、音声波のスペクトル零点に反映するものがあると考えられてきた。

先に、筆者らは、音声を極零モデル (AR-MA モデル) によりモデル化し、そのパラメータを最尤推定する能率のよいアルゴリズムを開発した。今回は、この音声分析法により推定された零点の持つ個人性情報について調べた結果について報告する。

2 極と零点の最尤推定法

(1) 極零モデル (AR-MA モデル)

音声生成過程を、Fig. 1 に示すような線形デジタルフィルタによりモデル化する。このフィルタは、 p 個の極と q 個の零点を持ち、 $|k_\mu| < 1$, $|h_\nu| < 1$ ($\mu = 1, \dots, p$, $\nu = 1, \dots, q$) のとき最小位相である。入力 $\{e_t\}$ (たは整数) を、平均 0、分散 σ_e^2 の白色過程とすると、出力 $\{x_t\}$ は、いわゆる AR-MA 過程となる。

音声信号の短時間標本列 $\{x_0, \dots, x_{N-1}\}$ を、 $\{x_t\}$ の標本列であるとして、これ

から、パラメータ $\{k_1, \dots, k_p; h_1, \dots, h_q\}$ を推定することを考える。

(2) 初期推定法 (逐次次数増加法)²⁾

Fig. 2 に示すような方法により、パラメータの初期推定を行なう。原理は次のようである。

標本値から、ピリオドグラム (標本スペクトル) を計算する。これを全極形モデルで近似することは易しく (PARCOR 分析)、また、不完全ながら、全零形で近似する方法もいくつかあるが、極零形で近似するには、全極・全零の一方を先に、他方を後に適用する方法には問題がある。そこで、本方法では、モデルの極と零点の次数 p , q を互いに 1 つずつ増やしてゆくことにより、近似的に極と零点の同時推定を行なっている。

この方法は、簡便な極零推定法としてもすぐれている。

(3) 最尤推定法

パラメータ $\{k_1, \dots, k_p; h_1, \dots, h_q\}$ を最尤推定するには、Fig. 3 に示す手順で計算される、 $\{h_1, \dots, h_q\}$ の関数 $J[h]$ を、

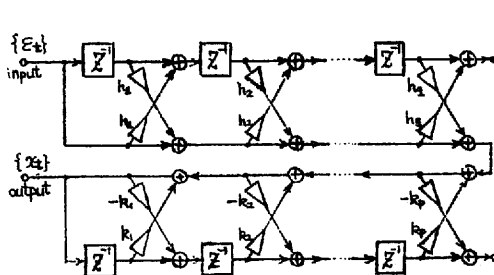


Fig. 1 Linear digital filter model of speech production.

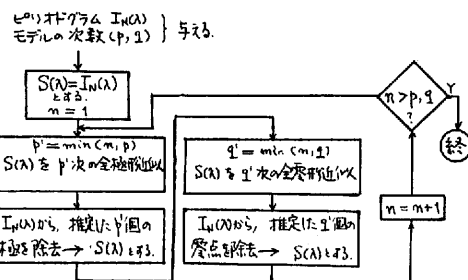


Fig. 2 Initial estimation procedure of k - and h -parameters.

* "On Personal Information in Statistically Estimated Spectral Zeros of Speech," by Shigeki SAGAYAMA and Sadaoki FURUI, Musashino ECL, NTT.

最小位相条件 $|r_v| < 1$ ($v=1, \dots, q$) のもとで最小にするような $\{\beta_v\}$ を求めればよい。

このような非線形最小化問題を解くために、ここでは、上記の初期推定値を出発値とし、Complex法³⁾と呼ばれる直接探索法を利用している。

推定された $\{\beta_v\}$ から零点を計算するには、Fig. 3 中の式④により $\{\beta_v\}$ に変換し、

$$\sum_{v=0}^q \beta_v z^{-v} = 0 \quad (\beta_0 \equiv 1)$$

の根を求めればよい。

③ 最尤推定零点に含まれる個人性

(1) 最尤推定分析の実験条件

音声信号を 4 kHz の low pass filter に通した後、8 kHz でサンプリングし、高域強調のため差分をとり、256 点を 1 フレーム (85 点ずつずらす) として、Hamming 窓をかけ、ピリオドグラムを計算し、これに前述の最尤推定分析を適用する。

(2) 最尤推定分析の多数話者への適用

7 名の話者 (男声) による /na:me/ (名前) 中の /m/ の部分を、 $p=12, q=8$ で分析したものを、Fig. 4 に示す。

(3) Antiformant の分散分析

最尤推定零点のうち、周波数が正の複素零点で、帯域幅の狭いものを anti-formant と仮称しておく。9 名の話者が 3 回ずつ発声した /na:me/ 中の /m/ について $p=12, q=8$ で分析して、定常区間内で平均した anti-formant の周波数と帯域幅について分散分析を行なった。結果 (Fig. 5) から、これらのパラメータが大きい個人性情報を持つことがわかる。

④ 結論

極点モデルによる音声の最尤推定分析法を開発し、多話者に適用して良好な結果を得、また、鼻音の推定零点に、有意な個人差があることがわかった。

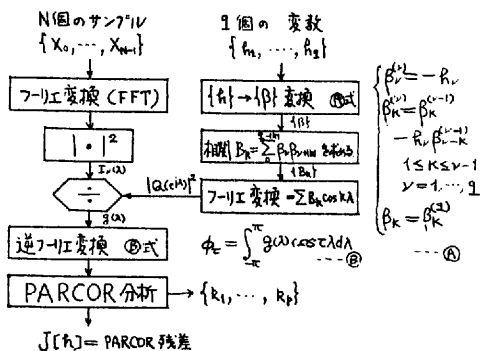


Fig. 3 Procedure for calculation of $J[h]$.

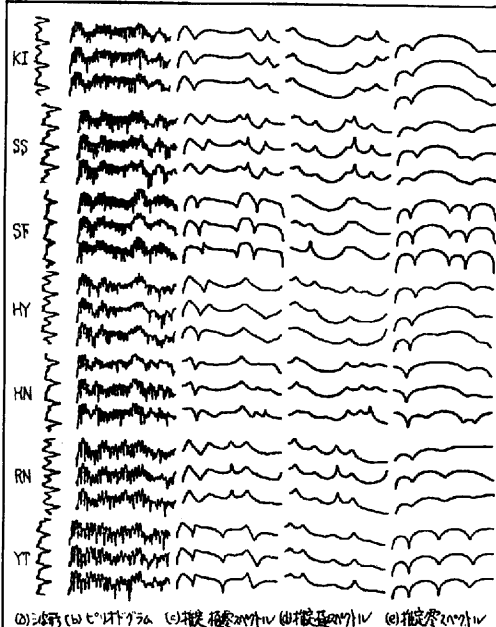


Fig. 4 Pole-zero model maximum likelihood analysis applied to /m/'s of 7 talkers.

Anti-Formant	Freq. B.W.	*	mean value
1		—	800 Hz 330 Hz
2		***	1910 Hz 410 Hz
3		***	3250 Hz 340 Hz

* Significant at 2.5 % false rejection.
 *** Significant at 0.5 % false rejection.

Fig. 5 Results of analysis of variance.

謝辞 日頃御指導頂く斎藤研究室長、熱心に討論された研究室の諸君に感謝します。

参考文献

- (1) 山崎誠山 古井 斎藤, 音学論 3-4-3 (9年10月)
- (2) 山崎誠山 古井, 信学技報 EA96-56 (1977)
- (3) M. J. Box, Computer J., 8 (1965), 42.