

母音の線形予測分析

Fujita, Masato / Sakakura, Mamoru / 板倉, 守 / 藤田, 正
人

(出版者 / Publisher)

法政大学工学部

(雑誌名 / Journal or Publication Title)

法政大学工学部研究集報 / 法政大学工学部研究集報

(巻 / Volume)

19

(開始ページ / Start Page)

49

(終了ページ / End Page)

54

(発行年 / Year)

1983-03

(URL)

<https://doi.org/10.15002/00004094>

母音の線形予測分析

藤 田 正 人*・坂 倉 守**

Linear Prediction of Vowels

Masato FUJITA and Mamoru SAKAKURA

Abstract

Five vowels of Japanese are analysed by Linear Prediction Method. It is shown that the orders of Prediction filter are important to find collect Formant frequencies.

§ 1. 緒 言

母音は、声帯の振動による有声音であり、声道は比較的安定し、鼻腔との音響的結合がないという3つの特徴があり、音声現象的に安定である。

これらのことから、線形予測分析モデルを用いるための分析次数、分析区間の検討を行い、日本語母音の識別について述べる。

§ 2. 原 理

音声合成モデルとして、図1⁽¹⁾のような線形音声合成モデルを示す。これを離散化し、 z 変換

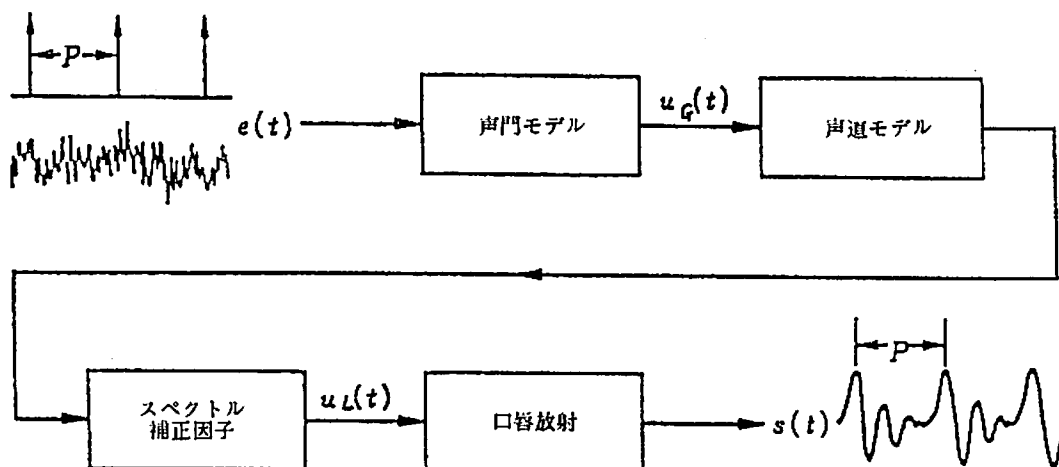


図1 線形音声合成モデル

* 電気工学科計測制御専攻大学院

** 電気工学科計測制御専攻

の概念を用いて表すと次式のようにになる。

$$S(z) = E(z)G(z)V(z)L(z) \quad (1)$$

$E(z)$ は声門モデルへの励振関数であり、周期 P をもつインパルス列である。 $G(z)$ は声門モデルである。 $V(z)$ は数個の共振点（ホルマント）をもつ声道モデルである。 $L(z)$ は口唇放射モデルである。

式(1)を更に簡単化された離散化モデルで表すと次式のような全極形合成モデルになる。

$$S(z) = E(z) \frac{1}{A(z)} \quad (2)$$

ただし

$$A(z) = \sum_{i=0}^M a_i z^{-i} \quad (a_0=1) \quad (3)$$

$$\approx \frac{1}{G(z)V(z)L(z)}$$

である。フィルタ $A(z)$ は全零形の逆フィルタである。

式(2)の両辺に $A(z)$ をかけると次式のような分析モデルになる。

$$E(z) = S(z)A(z) \quad (4)$$

この分析モデルと合成モデルとの関係を理想的な有声音の場合について表したものが図2⁽¹⁾である。

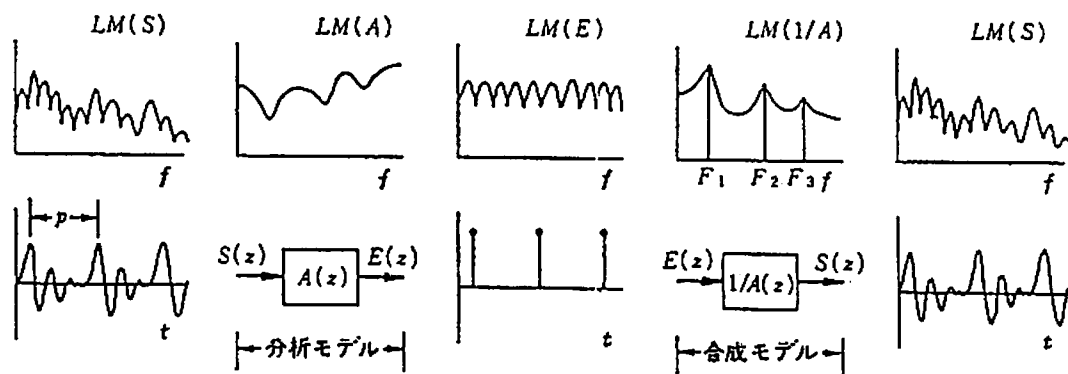


図2 有声音の理想化波形と分析・合成モデル

記号 $LM[\cdot]$ は、対応する離散信号から得られた対数振幅スペクトルを表している。例えば、

$$LM(S) = 10 \log_{10} |S(e^{j\theta})|^2$$

ここで

$$S(e^{j\theta}) = S(z)|_{z=e^{j\theta}}$$

であり、 θ は正規化周波数、 $\theta = 2\pi f/f_s$ 、 f_s はサンプリング周波数、 f は連続的な周波数変数である。

音声合成モデルを決定するパラメータは、式(3)の $A(z)$ の係数 a_i 、($i=1, 2, \dots, M$) と、 $E(z)$

のパラメータのピッチ周期 P である。

§3. 方 法

線形予測を適用して $A(z)$ の係数を直接決定する方法について述べる。

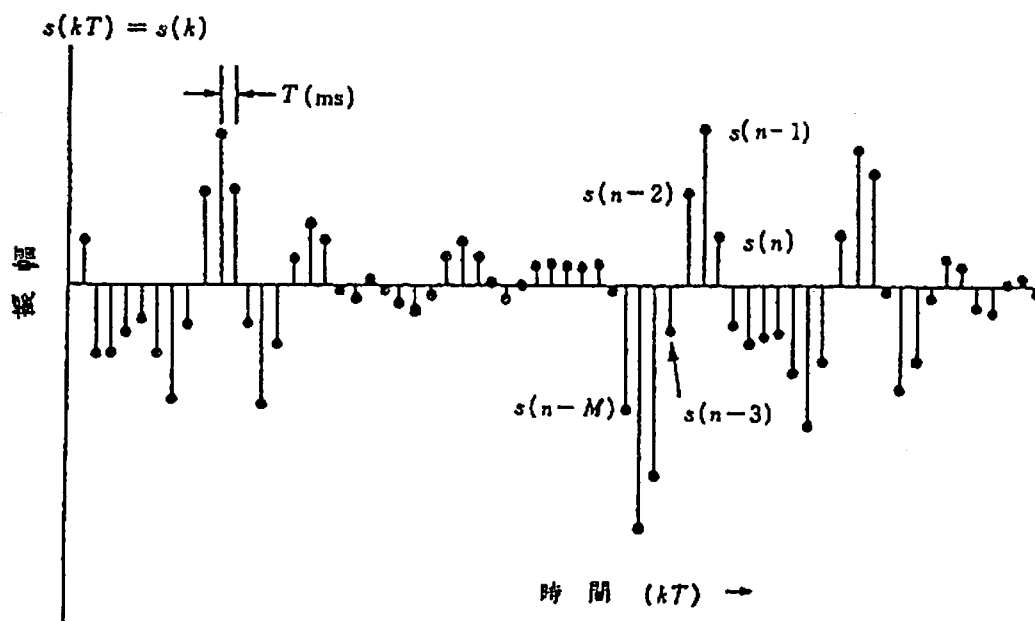


図3 サンプルしたデータ表示, 母音波形の一部

図3 (1)に示すような離散化された母音波形を考える。離散化されたデータは、分析モデル式(4)から次式のように書ける。

$$\begin{aligned} e(n) &= \sum_{i=0}^M a_i s(n-i) \\ &= s(n) + \sum_{i=1}^M a_i s(n-i) \end{aligned} \quad (5)$$

ここで、 $A(z)$ は式(3)から定義されている。サンプル値 $s(n)$ は、線形予測分析では先行する M 個のサンプル値の線形結合で表せる。ここで、 $\hat{s}(n)$ を予測値として定義すると、式(5)は次式のようになる。

$$e(n) = s(n) - \hat{s}(n) \quad (6)$$

ただし

$$\hat{s}(n) = -\sum_{i=1}^M a_i s(n-i) \quad (7)$$

である。式(6)で、 $e(n)$ の項は、実際のサンプル値 $s(n)$ と予測値 $\hat{s}(n)$ との予測誤差である。更に、式(7)は a_i ($i=1, 2, \dots, M$) の項は予測係数である。ここで、式(7)を z 変換すると次式のようになる。

$$\hat{S}(z) = F(z)S(z) \quad (8)$$

ここで

$$F(z) = -\sum_{i=1}^M a_i z^{-i} \quad (9)$$

である。 $F(z)$ は線形予測フィルタである。線形予測フィルタ $F(z)$ と逆フィルタ $A(z)$ との関係は図4に示す。

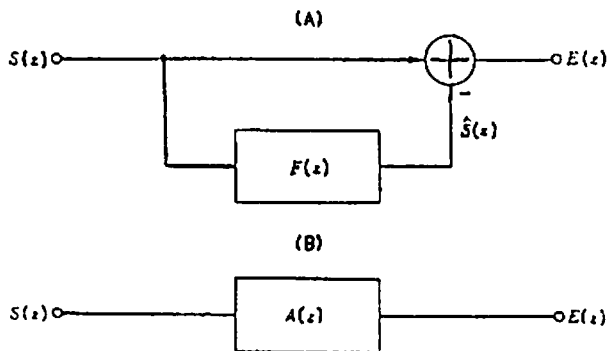


図4 (A) 予測器 $F(z)$ を明示したモデルの表示
(B) 逆フィルタ $A(z)$ による等価的表示,
 $A(z) = 1 - F(z)$

図4で、音声サンプルの線形予測は音声合成の線形モデルと等価になる。線形予測は、フィルタ $A(z)$ の係数 a_i , ($i=1, 2, \dots, M$) を、式(6)の予測誤差に最小2乗規準を適用することによって、音声波形から直接求めることができる。

§4. 方法の検討

音声、特に母音の特徴をスペクトル的に見た場合、その特徴はホルマントである。線形予測を用いて信頼性の高いホルマントの抽出するための分析次数と分析区間の選定について述べる。

4.1 分析次数の選定

L を声道の長さとし、 C を音速とすると、フィルタ $A(z)$ の記憶時間は、音波が声門から口唇までの伝搬するのに要する時間 (L/C) の2倍に等しくなる。典型的値として、 $C=34$ cm/ms, $L=17$ cm とすると、必要な記憶時間は1 ms である。サンプリング周波数 20 kHz のときは、フィルタの次数 M は少なくとも20必要である。

図5に示されるように、フィルタ $A(z)$ の記憶時間にかかる分析次数 $M=14$ の場合は、ホル

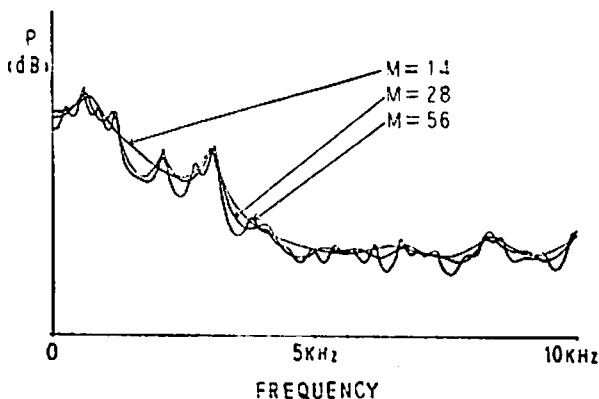


図5 $LM[1/A]$ 対数振幅スペクトル, サンプリング周波数 20 kHz, 分析区間 10 ms

マントの欠落が見られる。また、記憶時間を大きく超える分析次数 $M=56$ 場合は、ホルマント以外の共振点を多く含んでいる。

このように、分析次数はフィルタ $A(z)$ の記憶時間とサンプリング周波数によって決定される次数に、求めるホルマント数を加えた数に選定するとよい。

4.2 分析区間の選定

音声合成モデルより明らかに分析区間は、

基本周期を1周期以上含むことが必要であることが分かる。

男性の場合、基本周期は10～15 ms 程度である。図6に示されるように、基本周期1周期分を含まない $T=5$ ms では、ホルマントがはっきりと強調されず。基本周期を数周期含む $T=25$ ms の場合は、ホルマントがはっきりと強調されている。

このように、分析区間は基本周期を確実に含む必要がある。

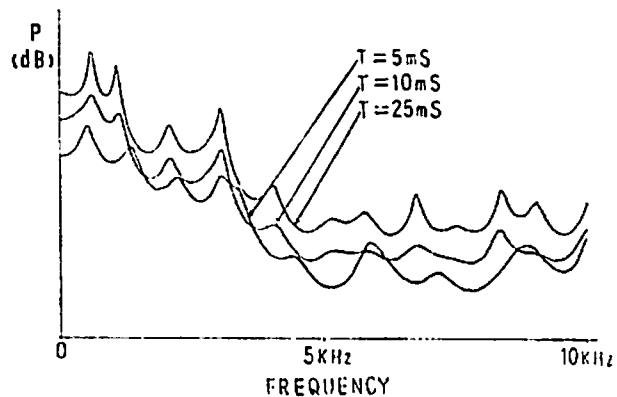


図6 LM[1/A]対数振幅スペクトル，サンプリング周波数 20 kHz，分析次数 28

§5. 結 果

適当な分析次数，分析区間を選定し，日本語母音を線形予測分析を行った結果を図7に示す。

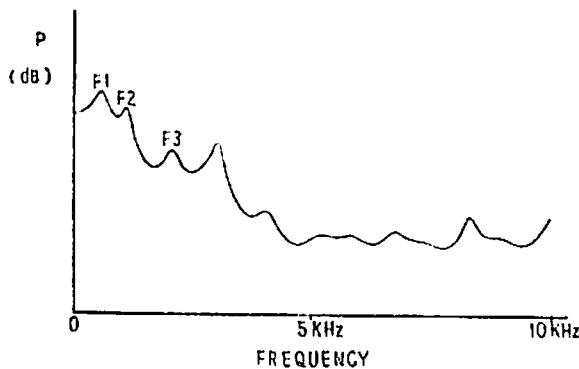


図7 LM[1/A]対数振幅スペクトルサンプリング周波数 20 kHz，分析次数 28，分析区間 10 ms

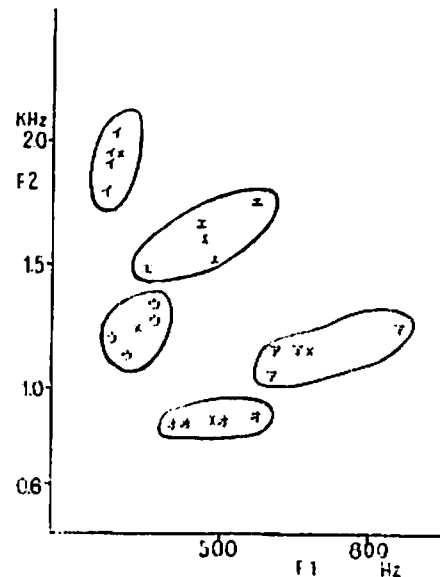


図8 日本語五母音のF1-F2マップ

図7は幾つかのホルマントをもっている。これを周波数の低い方から第1ホルマント(F1)，第2ホルマント(F2)，……と呼ぶ。エネルギーが5 kHz以下に集中し，特にF1とF2が高いことが分かる。このことから，F1とF2の関数として日本語母音の分布をみると図8のようになる。このようにF1とF2は，日本語母音を識別するよいパラメータになっていることが分る。

§6. あとがき

今後，被試験者数を増し，性別，年齢，出身地などの分類を加えることにより，日本語母音がより正確に分析評価できると考える。

参 考 文 献

- 1) J.D. Markel and A.H. Gray, Jr.: Linear Prediction of Speech (1976), Springer-Verlag.
- 2) 藤村 靖(編), 大泉充郎(監): 音声科学 東京大学出版会。