

線形予測法による 音声信号処理

本章では、サウンドデータの中でも音声に焦点を絞り、音声信号処理に頻繁に利用される線形予測法について説明します。線形予測法は音声の生成モデルと密接に関係しており、線形予測法を適用することで、フォルマントに代表される音声の特徴パラメータを抽出することができます。また、他方、こうした特徴パラメータから音声を合成することもできます。

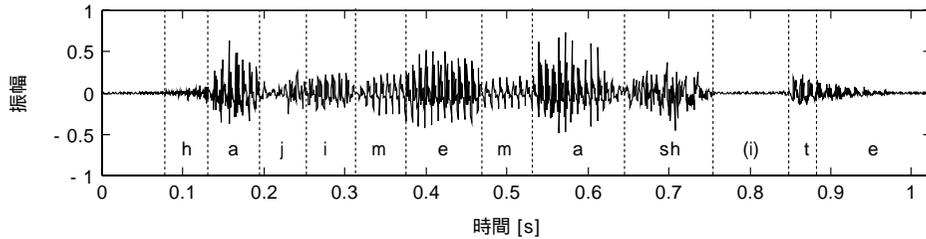
4.1 音声の生成モデル

人間のコミュニケーション手段である音声は、私たちにとって最も身近な音の代表です。図4.1に示すように、音声は時々刻々と信号の性質が変化するサウンドデータですが、概念上、「音素」と呼ばれる基本要素に分解することができます。図4.2に示すように、たとえば日本語にはわずか30個程度の音素しかありませんが、これらを組み合わせることで、様々な言語情報を相手に伝達することができます。

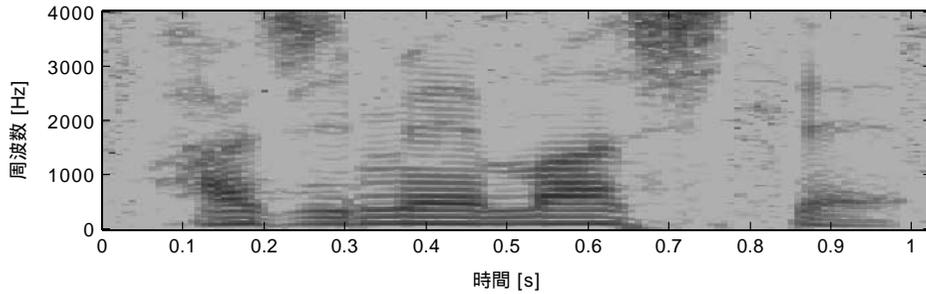
音素は、信号の性質の違いから「有声音」と「無声音」の2種類に大きく分類することができます。図4.3(a)に示すように、有声音は周期的な信号となっています。有声音は周期的な「声帯振動」を音源として生成される音で、舌の位置などで決定される声道の音響特性に応じて、たとえば[a]などの「母音」になります。

一方、図4.3(b)に示すように、無声音は非周期的な信号となっています。無声音は声帯振動をとまなわず、声道の隙間を呼気が通過するときに発生する「氣息雑音」を音源として生成される音で、声道の音響特性に応じて、たとえば[sh]などの「子音」になります。なお、氣息雑音が発生する場所は子音の種類によって異なり、[sh]では歯と舌を狭めることで発生する氣息雑音を音源としています。

このように、有声音と無声音はそれぞれ音源の種類が異なりますが、声道の音響特性によって音色が特徴づけられるという点ではまったく同じです。こうした音声の生成メカニズムは、図4.4のように、声帯振動や氣息雑音といった音源を入力信号、声道をフィルタとするデジタル・フィルタとしてモデル化することができます。



(a) 音声信号



(b) スペクトログラム

図4.1 音声信号【hajimemashite】

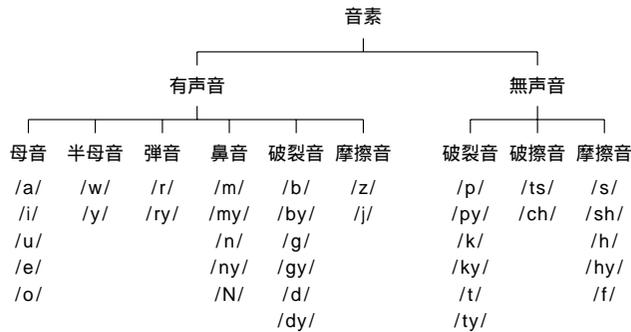


図4.2 日本語の音素

「線形予測法 (Linear Prediction)」は、こうしたデジタル・フィルタを具体的に実現する一つの手法です。線形予測法は、次のような IIR フィルタによって音声の生成モデルを定義しています。

$$s(n) = -\sum_{m=1}^M a_M(m)s(n-m) + f_M(n) \quad \dots \dots \dots (4.1)$$

ここで、 $s(n)$ は音声信号、 $a_M(m)$ はフィルタ係数、 $f_M(n)$ は入力信号、 M は線形予測の次数です。なお、フィルタ係数は「線形予測係数 (LPC : Linear Prediction Coefficient)」と呼ばれます。

ところで、式(4.1)を変形すると次のような FIR フィルタとなります。

$$f_M(n) = \sum_{m=0}^M a_M(m)s(n-m) \quad \dots \dots \dots (4.2)$$

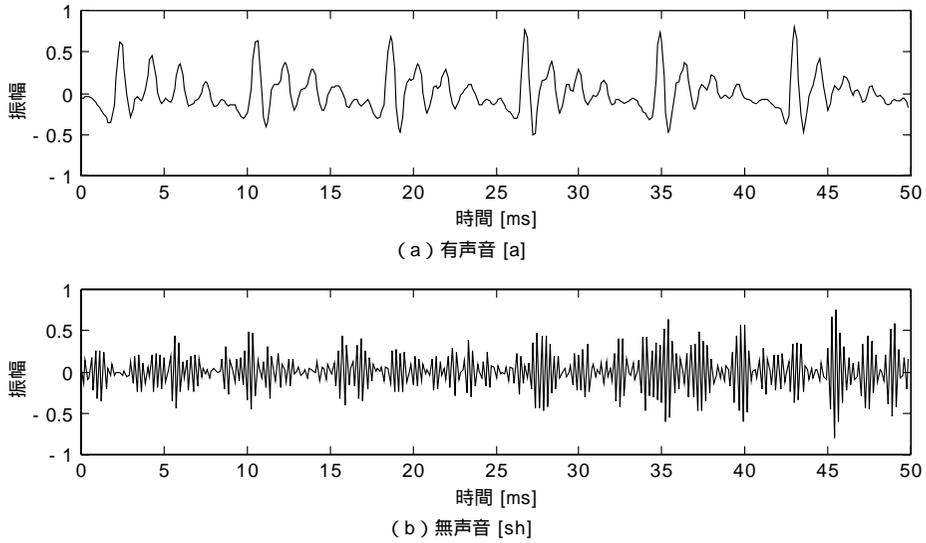


図4.3 有声音と無声音

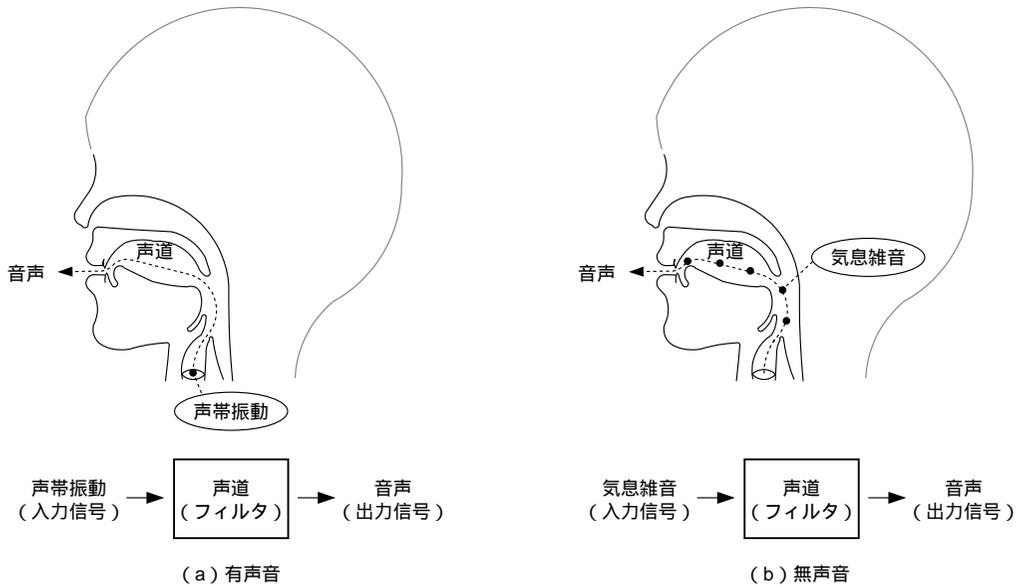


図4.4 音声の生成モデル

じつは、式(4.1)は線形予測法の「合成フィルタ」、式(4.2)は線形予測法の「分析フィルタ」と呼ばれ、両者は表裏一体の関係となっています。本章では次節以降、分析フィルタの導出を出発点として、線形予測法について詳しく説明していくことにします。

4.2 前向き線形予測

線形予測法は、過去あるいは未来の信号からそれよりも未来あるいは過去の信号を予測するデジタル信号処理の手法です。図4.5(a)に示すように、時刻 n の信号をそれよりも過去の M サンプルから予測することを「前向き線形予測」と呼びます。また、図4.5(b)に示すように、 $n-M$ 時刻の信号をそれよりも未来の M サンプルから予測することを「後向き線形予測」と呼びます。

ここでは、まず前向き線形予測について考えてみることにしましょう。 M 次の前向き線形予測の定義は次のとおりです。

$$\hat{s}(n) = \sum_{k=1}^M h_M(k) s(n-k) \dots\dots\dots (4.3)$$

ここで、 $s(n)$ は音声信号、 $h_M(k)$ は線形結合の重み係数、 $\hat{s}(n)$ は $s(n)$ の予測値となっています。式(4.3)は予測値を出力するFIRフィルタになっており、ブロック図として表すと図4.6になります。

さて、前向き線形予測における予測誤差は次のようになります。

$$f_M(n) = s(n) - \hat{s}(n) \dots\dots\dots (4.4)$$

この予測誤差が小さくなればなるほど、予測が適切に行われていることになります。そのため、線形予測法では予測誤差の自乗平均が最小となるように重み係数を決定しています。

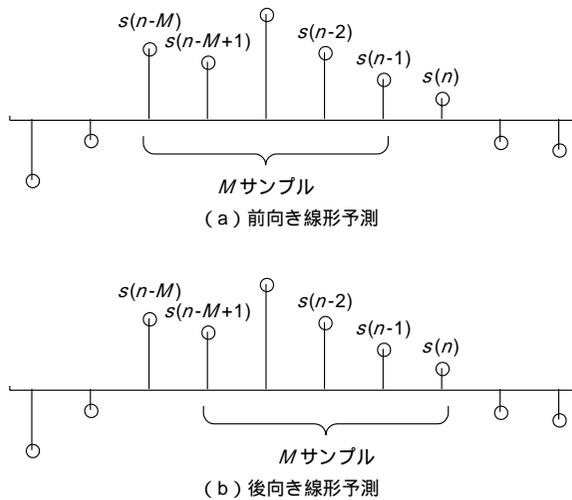


図4.5 線形予測法

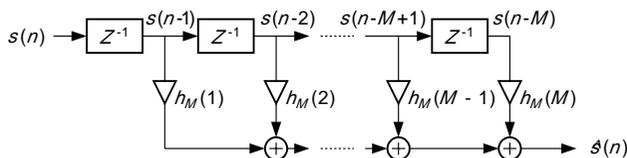


図4.6 前向き線形予測