

適応処理時代の ノイズ・キャンセル実験室

第3回 ノイズ除去性能はまあまあ! 確率最大信号を求める…MAP 推定法 川村 新

本連載では、音声のノイズ除去に注目し、さまざまな方式や、それらを実現するプログラムを紹介していきます。

今回は、MAP (Maximum a Posteriori) 推定法に基づくノイズ除去法について説明します。MAP推定法は、事後確率を最大化する方法です。

原理

MAP推定法を実現するブロック図を図1に、効き目(シミュレーション)を図2に示します。信号の流れは、前回(第2回、2016年11月号)のウィナー・フィルタと全く同じです。

MAP推定法は、選択する確率密度関数によってゲインの計算方法が変わります。ほとんどの場合、事後SNRと事前SNRの関数としてゲインを求めることができます。

● FFTで音声を分析する

時刻 n の観測信号 $x(n)$ が、音声 $s(n)$ とそれ以外のノイズ $d(n)$ との和で与えられると仮定します。このとき、

$$x(n) = s(n) + d(n)$$

です。観測信号を N サンプルごとに切り出し、 N 点FFT (Fast Fourier Transform; 高速フーリエ変換) を実行します。切り出した区間をフレームと呼びます。

FFT結果として得られる、 k 番目の複素スペクトルを、

$$X(k) = S(k) + D(k)$$

と書きます。ここで、 k は $0 \sim N-1$ の整数値をとる周波数番号です。

● ハーフ・オーバーラップで分析を進める

FFTによる分析は、 $N/2$ サンプルごとに実行する、ハーフ・オーバーラップを採用します。

ノイズ除去処理後に、IFFT (Inverse Fast Fourier Transform; 逆高速フーリエ変換) で時間領域の信号

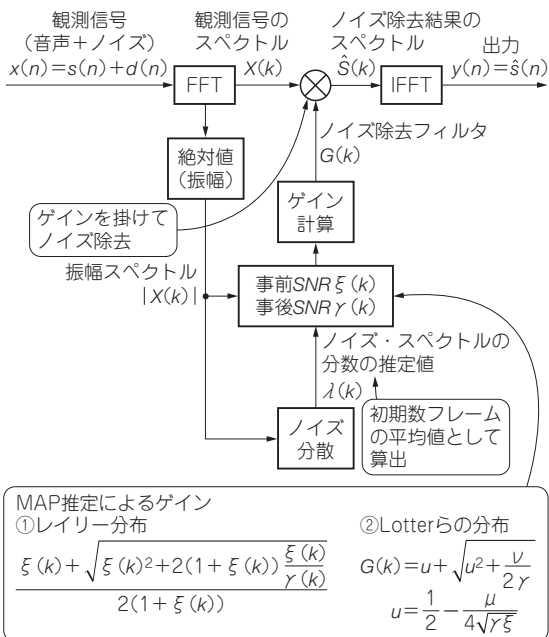


図1 MAP推定法を実現するブロック図

に戻すときに、フレームの半分が重複するので、その部分は加算して最終的な出力を得ます。

● MAP推定法の考え方

MAP推定法は、事後確率を最大化する方法です。事後確率とは、観測信号 $X(k)$ が既に生じた状態(事後)で、音声 $S(k)$ がとる値の確率のことです。その確率密度関数は、事後確率密度関数と呼ばれ、 $p(S|X)$ のように書きます。便宜上、周波数番号 k を省略しています。

S を動かして、 $p(S|X)$ が最大になれば、そのときの S を音声の推定値とすることがMAP推定法です。

問題は、 $p(S|X)$ が未知であることです。

ベイズの定理によれば、 $p(S|X)$ は次のように変形できます。

$$p(S|X) = \frac{p(X|S)p(S)}{p(X)}$$