Redução de Ruído em Sinais Digitais de Voz

Fernando Raganhan Barbosa

Orientador: Marcelo Gomes de Queiroz
Instituto de Matemática e Estatística
Universidade de São Paulo

15 de novembro de 2009

Sobre o Trabalho

- Utilizar versões modificadas do filtro de Wiener para reduzir ruído em sinais de voz.
- Importância: telecomunicação e VoIP.

Ruído

- Definição: qualquer sinal indesejado $\mathbf{r}$ que contamina o sinal $\mathbf{x}$.
- Sinais digitais, na prática, estão contaminados por ruído.
- Ruído tem diversas origens como eletromagnético e ambiente.

Algumas considerações sobre o ruído

- O ruído é aditivo: $\mathbf{y} = \mathbf{x} + \mathbf{r}$.
- O ruído, $\mathbf{r}$, e o sinal limpo, $\mathbf{x}$ não são correlacionados.
- O ruído é praticamente estacionário.

Sobre o Filtro

- Formulado em 1940 por Norbert Wiener.
- É um filtro linear de mínimos quadrados.
- Grande número de aplicações em processamento de sinais.

Importância do Filtro de Wiener

- É a base de diversas técnicas de processamento de sinais, como predição linear.
- Tem um relacionamento muito forte com a maioria das técnicas de redução de ruído, como os filtros de Kalman.
- Uma das técnicas mais tradicionais de redução de ruído.

Filtro FIR

- Equação de Filtro:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k)$$

- ou

$$\hat{\mathbf{x}} = \mathbf{w}^t \mathbf{y}$$

Erro

O erro do filtro é dado por:

$$e(m) = x(m) - \mathbf{w}^t \mathbf{y}$$

Em um segmento de N amostras:

$$\mathbf{e} = \mathbf{x} - Y\mathbf{w}$$

Considerações sobre N e P

- $N = P$: w pode não existir se o ruído não for linearmente separável do sinal, computacionalmente caro.
- $N < P$: mesmos motivos de $N = P$ só que ainda mais caro computacionalmente.
- $N > P$: não tem solução com $\mathbf{e} = 0$, mas sempre é possível encontrar um $\mathbf{w}$ que minimiza o erro quadrático.

Como Calcular w

$$
\begin{aligned}
E^2[e(m)] &= E^2[x(m) - \mathbf{w}^t \mathbf{y}] \\
&= E^2[x(m)] - 2\mathbf{w}^t E[\mathbf{y}x(m)] + \mathbf{w}^t E[\mathbf{y}^t \mathbf{y}]\mathbf{w} \\
&= r_{xx}(0) - 2\mathbf{w}^t \mathbf{r_{yx}} + \mathbf{w}^t R_{yy} \mathbf{w} \\
\frac{\partial}{\partial w} E^2[e(m)] &= -2\mathbf{r_{yx}} + 2\mathbf{w}^t R_{yy} \\
\mathbf{w} &= R_{yy}^{-1} \mathbf{r_{yx}}
\end{aligned}
$$

Calculando $\mathbf{r}_{\mathbf{yx}}$

$$
\begin{aligned}
r_{yx}(m) &= E[y(m)x(m)] \\
&= E[y(m)(y(m) - r(m))] \\
&= E[y(m)y(m)] - E[(x(m) + r(m))r(m)] \\
&= r_{yy}(m) - E[x(m)r(m)] - E[r(m)r(m)] \\
&= r_{yy}(m) - r_{xr}(m) - r_{rr}(m) \\
&= r_{yy}(m) - r_{rr}(m)
\end{aligned}
$$

Equação de Wiener-Hopf

Para ruído aditivo não correlacionado com o sinal desejado temos **w** ótimo dado por:

$$\mathbf{w} = R_{yy}^{-1}(\mathbf{r_{yy}} - \mathbf{r_{rr}})$$

Limitações do Filtro de Wiener

- Assume que o sinal limpo é estacionário, o que não é o caso dos sinais de voz.
- Introduz uma distorção no sinal original, quanto mais ruído é removido maior é a distorção é introduzida.

Ideia

- Como o filtro ótimo reduz ao máximo o ruído ele também causa uma distorção máxima.
- Encontrar um filtro sub-ótimo que ofereça o melhor trade-off entre redução de ruído e distorção.

Equação do Filtro Sub-Ótimo

Seja $\mathbf{w_o}$ o filtro ótimo e $\mathbf{u_1} = [1, 0, 0, 0, ...]$ o filtro canônico que não muda entrada e α um número real.

$$\mathbf{w_s} = \mathbf{u_1}(1 - \alpha) + \alpha \mathbf{w_o}$$

Alguns Índices

- Índice de distorção de voz:

$$i_{dv}(\mathbf{w}) = \frac{E^2[x(m) - \mathbf{w}^t\mathbf{x}(m)]}{E^2[r(m)]}$$

- Fator de redução de ruído:

$$f_{rr}(\mathbf{w}) = \frac{E^2[x(m) - \mathbf{w}^t\mathbf{y}(m)]}{E^2[r(m)]}$$

α ótimo

- função de custo:

$$c(\alpha) = \frac{1 - f_{rr}(\mathbf{w_s})}{1 - f_{rr}(\mathbf{w_o})} - \frac{i_{dv}(\mathbf{w_s})}{i_{dv}(\mathbf{w_o})} = 2(\alpha - \alpha^2)$$

- $\alpha = \frac{1}{2}$ atinge o melhor trade-off segundo $c(\alpha)$.

Filtro de Wiener Para Sinais Não Estacionários

- Adaptações da formulação clássica que atualizam $\mathbf{w}$ de alguma maneira refletindo as mudanças no sinal limpo.

Filtro de Wiener Para Sinais Não Estacionários

- Adaptações da formulação clássica que atualizam **w** de alguma maneira refletindo as mudanças no sinal limpo.
- BLS: Block Least Square
- RLS: Recursive Least Square
- LMS: Least Mean Squared Error

Block Least Square

- Em uma janela pequena o suficiente um sinal de fala pode ser considerado estacionário.
- A ideia então é recalcular **w** a cada janela como se a janela fosse todo o sinal.

$$\mathbf{w} = (Y^t Y)^{-1} Y^t \mathbf{x}$$

$$Y = \begin{bmatrix} y(0) & y(-1) & \cdots & y(1-P) \\ y(1) & y(0) & \cdots & y(2-P) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-P) \end{bmatrix}$$

Recursive Least Square

- Utiliza uma atualização recursiva do valor de w, baseada na minimização do erro médio quadrático, que introduz um fator de esquecimento no cálculo da matriz de correlação e dos vetores de correlação cruzada.

$$
\begin{aligned}
R_{yy}(m) &= \lambda R_{yy}(m-1) + \mathbf{y}(m)^t \mathbf{y}(m) \\
\mathbf{r_{yy}}(m) &= \lambda \mathbf{r_{yy}}(m-1) + \mathbf{y}(m) y(m) \\
\mathbf{w}(m) &= w(m-1) + Atualizacao(m)
\end{aligned}
$$

Least Mean Squared Error

- Utiliza uma atualização simplificada de w que consiste em um único passo do método de máxima descida, usando a derivada de erro instantâneo.

$$\mathbf{w}(m) = \mathbf{w}(m-1) - \lambda[\frac{\partial e^2(m-1)}{\partial \mathbf{w}(m-1)}]$$

$$e(m) = x(m) - \mathbf{w}^t(m)\mathbf{y}(m)$$

Referências

Y. Huang e S. Doclo J. Chen, J. Benesty.
New insights into the noise reduction wiener filter.
2006.

F. Richard Moore.
Elements of Computer Music.
P T R Prentice Hall, 1990.

Saeed V. Vaseghi.
Advanced Signal Processing and Noise Reduction.
John Wiley and Sons Ltd., 4th edition, 2008.