A SINGLE-CHANNEL NOISE REDUCTION FILTERING/SMOOTHING TECHNIQUE IN THE TIME DOMAIN

Ningning Pan1, Jacob Benesty2, and Jingdong Chen1
1 Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, Xi’an, Shaanxi 710072, China
{nnpan@mail.nwpu.edu.cn, jingdongchen@ieee.org}
2 INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Suite 6900, Montreal, QC H3A 1K6, Canada
benesty@emt.inrs.ca

INTRODUCTION

Noise reduction: recover a speech of interest from noisy observation.
The proposed method:
• It first applies a time smoothing window to the noisy signal.
• Noise reduction filters are applied to the smoothed noisy signal.

PERFORMANCE MEASURES

Input SNR
\[ \text{SNR} = \frac{\sigma_d^2}{\sigma_n^2} \]

Output SNR
\[ o\text{SNR} (h) = \frac{\sigma_d^2}{\sigma_n^2} - \frac{h^H R_{xx} h}{h^H R_{yy} h} \]

\( h \) should be found in such a way that \( o\text{SNR} (h) > \text{SNR} \).

Signal distortion index
\[ v_l(h) = \frac{E\{[x(k) - h^H X(k)]^2\}}{\sigma_d^2} \]

MSE CRITERION

Error signal:
\[ \epsilon(k) = x(k) - h^H Y(k) = x(k) - x_h(k) + x_o(k) \]
\[ \epsilon_o(k) = h^H X(k) - x(k) \]
\[ \epsilon_h(k) = h^H Y(k) - x_h(k) \]

MSE criterion:
\[ J(h) = E\{[\epsilon(k)]^2\} = J_h(h) + J_o(h) \]
\[ J_h(h) = E\{[\epsilon_h(k)]^2\} \]
\[ J_o(h) = E\{[\epsilon_o(k)]^2\} \]

LINEAR FILTERING/SMOOTHING FOR NOISE REDUCTION

MICROPHONE SIGNAL IN THE TIME DOMAIN:
\[ y(k) = x(k) + n(k) \]

CONCLUSIONS

• A linear estimation approach to single-channel noise reduction in the time domain with the ability to smooth and filter the observation signal at the same time is presented.
• Three different filters, maximum SNR, Wiener, and tradeoff filters, are derived.
• Results show advantages of smoothing.

SIMULATIONS

Minimum distortion filter
Minimize \( J_h(h) \) subject to \( J_o(h) = 0 \)
\[ h_{TD} = (R_{xx} + \beta R_{yy})^{-1} R_{yx} \]

Wiener filter
Minimize \( J(h) \)
\[ h_{WD} = R_{yx}^{-1} R_{yy} \]

Implementation:
- clean speech
- noise smoothing window
- 20 sentences from TIMIT database
- recorded in a sedan car
- Sampling rate: 8 kHz
- Hanning window
- \( \text{SNR} = 10 \text{ dB} \)

OPTIMAL FILTERS

Maximum SNR filter
\[ h_{\text{max}} = \frac{E\{X(k)\}}{E\{X(k)\}^H R_{xx} X(k)} \]

Substituting \( h_{\text{max}} \) into the distortion-based MSE, we get
\[ J_o(h_{\text{max}}) = \sigma_o^2 - 2 \sigma_o^2 \frac{R_{yx} R_{yy}}{R_{xx}} + \sigma_o^2 \frac{R_{yx} R_{yy} h_{\text{max}}}{R_{xx} h_{\text{max}}} \]

The optimal value of \( c \):
\[ c_{\text{opt}} = \frac{R_{yx} R_{yy} h_{\text{max}}}{R_{xx} h_{\text{max}}} \]

The performance of the Wiener, tradeoff, and maximum SNR filters (without the time smoothing technique, i.e., \( N = 1 \)) as a function of the filter length, \( L \). The PSQ score of the noisy signal is 2.375.