

ET4147: Signal Processing for Communications

Homework 3

Erik Hagenaars
4272404

Teun de Smalen
4321014

I. INTRODUCTION

II. SYSTEM

A. Signal model

$$Y_t[n] = S_t[n] + N_t[n] \quad (\text{time domain}) \quad (1)$$

$$Y_k[l] = S_k[l] + N_k[l] \quad (\text{frequency domain}) \quad (2)$$

$$R_{S_t N_t}(n, m) = 0 \quad (\text{uncorrelated}) \quad (3)$$

$$R_{Y_t Y_t}(n, m) = R_{S_t S_t}(n, m) - R_{N_t N_t}(n, m) \quad (4)$$

$$R_{Y_t Y_t}(n, m) = R_{Y_t Y_t}(m - n) \quad (\text{wide-sense stationary}) \quad (5)$$

$$P_{YY,k} = \lim_{L \rightarrow \infty} \sum_{m=-L/2}^{L/2} R_{Y_t Y_t}(m) e^{-j2\pi \frac{km}{K}} \quad (6)$$

$$= P_{SS,k} + P_{NN,k} \quad (7)$$

$$\hat{P}_{YY,k}^P(l) = \frac{1}{L} |Y_k(l)|^2 \quad (8)$$

$$\hat{P}_{YY,k}^B(l) = \frac{1}{M} \sum_{m=l-M+1}^l \hat{P}_{YY,k}^P(m) \quad (9)$$

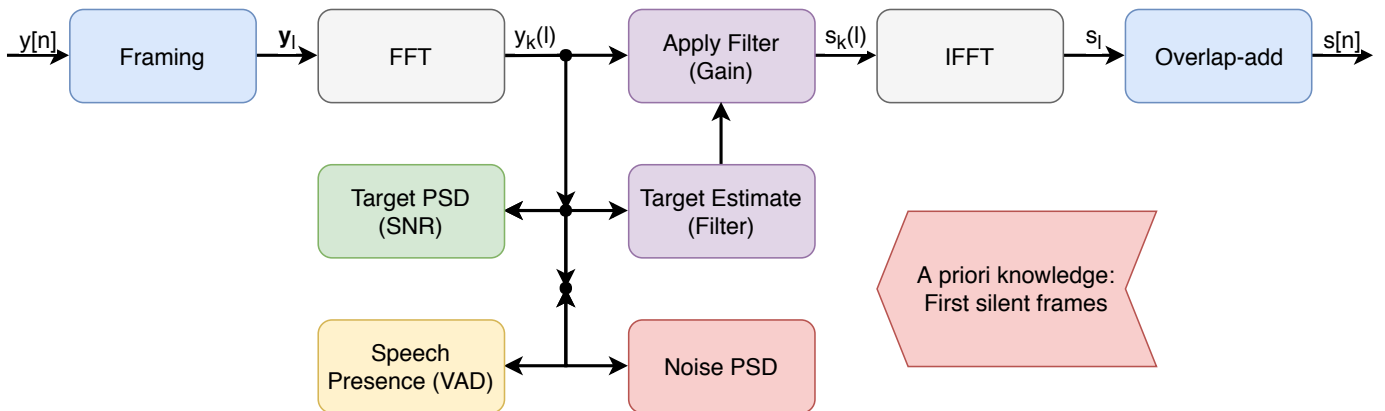


Fig. 1: Overview of the system.

B. System overview

III. FRAMING & OVERLAP ADD

$$y_l[n] = y[n + lR]w[n], \quad n = 0, \dots, N - 1 \quad (10)$$

$$y[n] = \sum_{l=1}^k y_l[n]/w[n] \quad (11)$$

IV. NOISE ESTIMATION

$$H_0 : Y_K(l) = N_k(l) \quad (\text{speech absence}) \quad (12)$$

$$H_1 : Y_K(l) = S_k(l) + N_k(l) \quad (\text{speech presence}) \quad (13)$$

$$f(x) = \begin{cases} \alpha \hat{\sigma}_{N,k}^2(l-1) + (1-\alpha) |y_k(l)|^2 & \text{when } H_0(l) \\ \hat{\sigma}_{N,k}^2(l-1) & \text{when } H_1(l) \end{cases} \quad (14)$$

$$\mathbf{Q} = \{P_{YY,k}(l-M+1) \quad \dots \quad P_{YY,k}(l)\} \quad (15)$$

$$\hat{\sigma}_{N,k}^2(l) = Q_{min} \quad (16)$$

$$\widehat{\sigma}_N^2(l) = \alpha \widehat{\sigma}_N^2(l-1) + (1-\alpha) E \left[|N_k(l)|^2 |y_k(l)| \right] \quad (17)$$

$$E \left[|N_k(l)|^2 |y_k(l)| \right] = P(H_{0,k}(l)|y_k(l)) E \left[|N_k(l)|^2 |y_k(l)|, H_{0,k} \right] + P(H_{1,k}(l)|y_k(l)) E \left[|N_k(l)|^2 |y_k(l)|, H_{1,k} \right] \quad (18)$$

$$P(H_{0,k}(l)|y_k(l)) = 1 - P(H_{1,k}(l)|y_k(l)) \quad (19)$$

$$E \left[|N_k(l)|^2 |y_k(l)|, H_{0,k} \right] = |y_k(l)|^2 \quad (20)$$

$$E \left[|N_k(l)|^2 |y_k(l)|, H_{1,k} \right] = \widehat{\sigma}_N^2(l-1) \quad (21)$$

$$P(H_{1,k}(l)|y_k(l)) = \frac{P(H_{1,k}(l)) p_{Y|H_1}}{P(H_{1,k}(l)) p_{Y|H_1} + P(H_{0,k}(l)) p_{Y|H_0}} \quad (22)$$

$$p_{Y|H_0} = \frac{1}{\widehat{\sigma}_N^2 \pi} \exp \left(-\frac{|y^2|}{\widehat{\sigma}_N^2} \right) \quad (23)$$

$$p_{Y|H_0} = \frac{1}{\widehat{\sigma}_N^2 (1 + \xi_{H_1}) \pi} \exp \left(-\frac{|y^2|}{\widehat{\sigma}_N^2 (1 + \xi_{H_1})} \right) \quad (24)$$

V. SNR ESTIMATION

$$\xi = \frac{\sigma_{S,k}(l)^2}{\sigma_{S,k}(l)^2} = \frac{P_{SS,k}}{P_{NN,k}} = \frac{E \left\{ |S_k(l)|^2 \right\}}{E \left\{ |N_k(l)|^2 \right\}} \quad (25)$$

$$\xi_k(l) = \frac{E \left\{ |Y_k(l)|^2 \right\}}{E \left\{ |N_k(l)|^2 \right\}} - 1 \quad (26)$$

$$= \frac{\hat{P}_{YY,k}^B(l)}{\frac{1}{L} E \left\{ |N_k(l)|^2 \right\}} \quad (27)$$

$$\xi_k(l) = \alpha \frac{E \left\{ |S_k(l)|^2 \right\}}{E \left\{ |N_k(l)|^2 \right\}} + (1 - \alpha) \left(\frac{E \left\{ |Y_k(l)|^2 \right\}}{E \left\{ |N_k(l)|^2 \right\}} - 1 \right) \quad (28)$$

$$|S_k(l)|^2 = \left| \hat{S}_k(l-1) \right|^2 \quad (29)$$

$$\frac{E \left\{ |Y_k(l)|^2 \right\}}{E \left\{ |N_k(l)|^2 \right\}} - 1 = \max \left[\left(\frac{|Y_k(l)|^2}{E \left\{ |N_k(l)|^2 \right\}} - 1, 0 \right) \right] \quad (30)$$

VI. VOICE ACTIVITY DETECTION

$$T(l) = \frac{1}{L} \sum_{k=1}^{k=L} \log(\Lambda_k(l)) \underset{H_0}{\underset{H_1}{\geq}} \lambda \quad (31)$$

$$\Lambda_k(l) = P_{YY,k} \quad (32)$$

VII. TARGET ESTIMATION

$$P_{SS,k}(l) = P_{YY,k}(l) - P_{NN,k}(l) \quad (33)$$

$$\left| \widehat{S_k(l)} \right|^2 = |Y_k(l)|^2 - |N_k(l)|^2 \quad (34)$$

$$\left| \widehat{S_k(l)} \right|^2 = \left(\max \left\{ 1 - \frac{E \left[|N_k(l)|^2 \right]}{|y_k(l)|^2}, \epsilon \right\} \right)^{\frac{1}{2}} |y_k(l)| \quad (35)$$

$$S_k(l)^2 = \left| \widehat{S_k(l)} \right|^2 e^{j\angle y_k(l)} \quad (36)$$

$$\hat{S}_k = H_k \dot{Y}_k \quad (37)$$

$$H_k = \frac{P_{SY,k}}{P_{PY,k}} \quad (38)$$

$$= \frac{SNR_k}{SNR_k + 1} \quad (39)$$

VIII. CONCLUSION

REFERENCES

- [1] R. C. Hendriks, T. Gerkmann, and J. Jensen, *DFT-Domain Based Single-Microphone Noise Reduction for Speech Enhancement: A Survey of the State of the Art*. Morgan & Claypool, 2013. [Online]. Available: <https://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6813348>

APPENDIX A

MATLAB CODE