

スペクトルサプレッション法における雑音スペクトル推定の改善

On Noise Spectrum Estimation in Noise Cancelers Based on Spectrum Suppression Approach

鈴木 大和[†]

中山 謙二[†]

平野 晃宏[†]

Hirokazu Suzuki

Kenji Nakayama

Akihiro Hirano

[†] 金沢大学大学院 自然科学研究科 電子情報工学専攻

Division of Electronics and Computer Science

Graduate School of Natural Science and Technology, Kanazawa Univ.

1 まえがき

本論文ではスペクトルサプレッション法を用いた単一マイクノイズキャンセラにおける雑音スペクトル推定の改善について検討する．音声が存在する区間でも雑音推定が可能であり，3GPP(The Third Generation Partnership Project) が定める携帯電話の雑音除去性能を満たしている重み付き雑音推定 [1] において無音区間のフレームを検出することにより雑音スペクトル推定を補正することによって改善を行なった．

2 スペクトルサプレッション法

音声と雑音ともにスペクトル成分において統計的独立で平均 0 であるとする．時間領域でのクリアな音声を $x(n)$ ，雑音を $d(n)$ ，雑音混入音声を $y(n)$ とおくと，

$$y(n) = x(n) + d(n) \quad (1)$$

この信号の周波数領域での表示を次のように表す．

$$Y(k, m) = X(k, m) + D(k, m) \quad (2)$$

k はフレーム内周波数番号， m はフレーム番号を表す．雑音の分散を $\lambda_d(k, m)$ とすると，事前 SNR(クリーン音声対雑音比)，事後 SNR(雑音混入音声対雑音比) はそれぞれ次の式で表せる．

$$\xi(k, m) = \frac{|X(k, m)|^2}{\lambda_d(k, m)} \quad (3)$$

$$\gamma(k, m) = \frac{|Y(k, m)|^2}{\lambda_d(k, m)} \quad (4)$$

実際に利用可能なものは，雑音混入音声のみで，事前 SNR，事後 SNR は推定しなくてはならない．事前 SNR $\xi(k, m)$ は，次式で推定できる [1]．

$$\hat{\xi}(k, m) = \alpha\gamma(k, m-1)G(k, m-1)^2 + (1-\alpha)P[\gamma(k, m)-1] \quad (5)$$

$P[x]$ は半波整流関数であり， α は忘却係数を表す．

本論文では音声スペクトルをガウス分布と仮定してスペクトルゲインを求める MMSE 法 [1] と，より実際の音声スペクトルの分布に近いスーパーガウス分布と仮定して求める Joint MAP 法 [2] を用いた．また，ミュージカルノイズの抑制にその周波数成分が高域にあることを利用して高域抑圧 [3]，雑音を過剰に抑圧するためによる音質の劣化を防ぐために原音(観測信号 = 音声 + 雑音)をある割合で付加した [3]．

3 雑音スペクトル推定の改善

重み付き雑音推定 [1] において推定された雑音スペクトルと実際の雑音スペクトルとは誤差が生じる．特に無音区間で誤差が大きい場合，雑音が大きく残ってしまう．そこで，無音区間を検出することにより雑音スペクトル推定及び事後 SNR $\gamma(k, m)$ の補正を行なう．

無音区間の検出に，各々のフレームに対する全周波数における事後 SNR $\gamma(k, m)$ の平均値 $\bar{\gamma}_m$ を用いる [4]．定常で独立な雑音に対して無音区間であるフレームにおける平均値 $\bar{\gamma}_m$ はおよそ 1 となる．そこで，各フレームにおいて $\bar{\gamma}_m$ を計算することによって適切な閾値を設定することにより無音区間を検出することができる [4]．無音区間のフレームにおいては，重み付き雑音推定における重み係数 $W(k, m) = 1$ ， $\gamma(k, m) = 1$ として $\lambda_d(k, m)$ を計算する．

4 シミュレーション

入力信号としてサンプリング周波数 10kHz の男性音声を用いた．出力では，信号を 12.8ms の区間に分割し，各区間の SNR の平均値であるセグメンタル SNR で評価した．正確な雑音スペクトルを用いた場合を理想値とした．背景雑音が白色雑音の場合のシミュレーション結果を表 1 に示す．

表 1 SNR 改善量 (出力 SNR - 入力 SNR) - 白色雑音 -

入力 SNR[dB]	-6 dB	0 dB	6 dB
MMSE(理想値)	11.68 dB	8.55 dB	5.76 dB
MMSE(従来法)	9.47 dB	6.94 dB	4.08 dB
MMSE(提案法)	10.76 dB	7.81 dB	4.87 dB
JointMAP(理想値)	11.45 dB	8.43 dB	5.85 dB
JointMAP(従来法)	9.93 dB	7.15 dB	4.32 dB
JointMAP(提案法)	10.83 dB	7.79 dB	4.88 dB

5 まとめ

従来の重み付き雑音推定 [1] において，無音区間の検出を行なうことにより雑音スペクトル推定の改善を行ない，白色雑音の場合 SNR が大きく向上した．

参考文献

- [1] M.Katou, A.Sugiyama and M.Serizawa, "Noise suppression with high speech quality based on weighted noise estimation and MMSE STSA," IEICE Trans. Fundamental, vol.E85-A, no.7, pp.1710-1718, Jul.2002.
- [2] T.Lotter and P.Vary, "Noise Reduction By Joint Maximum A Posteriori Spectral Amplitude And Phase Estimation With Super-Gaussian Speech Modeling," in Proc.of EUSIPCO-04, pp.1447-60, Sep. 2004.
- [3] 鈴木大和, 中山謙二, 平野晃宏 "スペクトルサプレッション法によるノイズキャンセラの音質改善," 信学技法, SIP2005-11, May. 2005.
- [4] R.Martin, D.Malah, V.Cox and J.Accardi "A Noise Reduction Preprocessor for Mobile Voice Communication," EURASIP Journal on Applied Signal Processing, pp.1046-1058, Aug.2004.