

# 補聴器のための低遅延簡易残響低減手法

## Low-Latency Dereverberation Method for Hearing Aids

1W130440-3 野崎 琴代 指導教員 及川 靖広 教授

NOZAKI Kotoyo

Prof. OIKAWA Yasuhiro

概要：難聴者にとって、残響時間の長い環境下における言葉の聞き取りは、特に困難であることが知られている。補聴器のような小さなデバイスでも実装可能な計算コストの少ない残響低減手法が春原らによって提案されている。一方、周波数分析に伴う処理遅延を改善が必要であった。そこで本研究では、先に述べた残響低減手法において、高域の遅延が少なく聴覚特性に応じた帯域分割とサンプル毎の処理が可能な分析合成系を導入する。その評価として、客観評価には SRMR を用い、残響低減に対する本手法の有効性を確認した。また、パラメータの影響が少ないことを明らかにした。主観評価では、音声の「残響感」と「自然さ」について 3 つの残響時間と 2 つの帯域について、残響が低減されていないと不自然に聞こえるという結果となった。

キーワード：残響抑圧、聴覚フィルタ、SRMR、シェッフエの一対比較法

Keywords: dereverberation, auditory filter, SRMR, Scheffe's paired comparison

### 1. ま え が き

残響は難聴者にとって音声の聞き取りを妨げる要因となり [1]、補聴器のような小さなデバイスでの残響低減を行うには計算コストと入出力の低遅延化が必要となる。本研究では、その問題に対して再帰的な処理により演算コスト低減が可能な簡易残響低減手法 [2] と、高域の遅延が少なく聴覚特性に応じた処理が可能な分析合成系をそれぞれ適用し、それらを組み合わせたものを提案手法とする。その評価として、残響低減度合いを調整するパラメータの影響を明らかにするため、客観評価及び主観評価を行った。

### 2. 指数平滑平均を用いた簡易残響低減手法

残響成分を簡易的に推定するアタック・リリース付き指数平滑平均を用いた簡易的な残響低減アルゴリズム [2] を簡単に説明する。時刻フレーム  $l$ 、帯域  $\omega$  におけるフィルタバンク分析出力信号  $X(l, \omega)$  より、以下の式によって入力信号のエネルギー

$$|\hat{X}(l, \omega)|^2 = \beta |X(l, \omega)|^2 + (1 - \beta) |\hat{X}(l-1, \omega)|^2, \quad (1)$$

を求める。ただし、 $\beta \in (0, 1]$  は、指数平滑平均における重みを表すパラメータである。 $|\hat{X}(l, \omega)|^2$  に含まれる残響成分のエネルギー  $|\hat{Z}(l, \omega)|^2$  を、

$$|\hat{Z}(l, \omega)|^2 = \gamma(l) |\hat{X}(l, \omega)|^2 + \{1 - \gamma(l)\} |\hat{Z}(l-1, \omega)|^2, \quad (2)$$

のように、再帰的な処理で求めることで演算量を低減する推定方法が、春原らによって提案されている [2]。このうち、時定数を制御するパラメータ  $\gamma$  は

$$\gamma(l) = \begin{cases} \gamma_{\text{at}}, & \{|\hat{X}(l, \omega)|^2 > |\hat{Z}(l-1, \omega)|^2\} \\ \gamma_{\text{re}}, & (\text{otherwise}) \end{cases}, \quad (3)$$

のように与えられる。式 (2) の推定残響エネルギー、観測信号のエネルギーより、フィルタバンク分析出力信号  $X(l, \omega)$  に乗算するゲイン  $G(l, \omega)$  は、

$$G(l, \omega) = \begin{cases} \sqrt{\frac{|\hat{X}(l, \omega)|^2 - |\hat{Z}(l, \omega)|^2}{|\hat{X}(l, \omega)|^2}}, \\ \{\text{if } |\hat{X}(l, \omega)|^2 > |\hat{Z}(l-1, \omega)|^2\} \\ \eta, & (\text{otherwise}) \end{cases} \quad (4)$$

と求められる。ただし、ゲインの下限值として  $\eta$  を設ける。求めたゲインには指数平滑平均によるスムージング処理を行う。

### 3. 客観評価

#### 3.1 Speech to Reverberation Modulation energy Ratio

SRMR は、Falk らによって提案された変調エネルギーを利用した残響音声の客観評価指標である [3]。入力信号となる残響音声  $x[n]$  の  $j$  番目 ( $j \in \mathbb{N}, 1 \leq j \leq 23$ ) のガンマトーンフィルタ出力信号  $x_j[n]$  をヒルベルト変換して ( $\mathcal{H}\{\cdot\}$ )、時間包絡  $e_j[n]$  を得る。変調周波数  $f$ 、時間フレーム  $m$  における変調エネルギー  $E_j(m, f)$  は、時間包絡のフーリエ変換  $\mathcal{F}$  の二乗によって、

$$E_j(m, f) = |\mathcal{F}(e_j[m, n])|^2, \quad (5)$$

のように得られる。 $E_j(m, f)$  の時間方向での平均を  $\overline{\mathcal{E}_{j,k}}$  とすると、各変調周波数における変調エネルギー  $\mathcal{E}_k$  は聴覚フィルタ方向に平均をとることで得られる。SRMR は、高変調周波数のエネルギーに対する低変調周波数のエネルギーの比をとり、

$$\text{SRMR} = \frac{\sum_{k=1}^4 \overline{\mathcal{E}_k}}{\sum_{k=5}^8 \overline{\mathcal{E}_k}}, \quad (6)$$

と求めることができる。残響時間が長くなるにつれて、高い変調周波数でのエネルギー  $\overline{\mathcal{E}_k}$  が大きくなり SRMR の値が小さくなることが知られている [3]。そのため、残響が低減されていると SRMR が大きくなる。

#### 3.2 評価結果

図 1 に、残響音声及びゲイン下限値による処理後音声の SRMR の変化を示す。ゲイン下限値  $\eta$  には、-36 dB、-27 dB、-18 dB、-9 dB のレベルに対応する値を用いた。曲線が

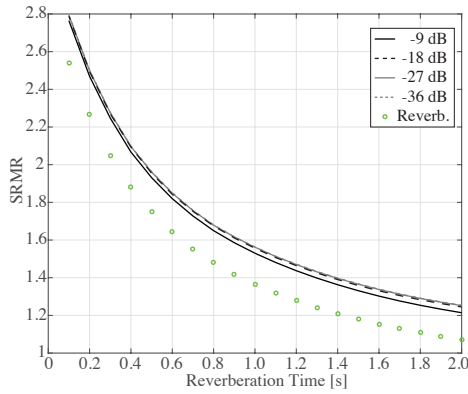


図-1 残響付加信号及びゲイン下限値による処理後音声のSRMR

それぞれの下限值での処理後音声、丸印が残響音声のSRMRを表している。残響時間1.5秒の処理後音声及び残響時間1.0秒の残響音声のSRMRに注目すると、ほぼ同じ値になっている。つまり、0.5秒分の残響を低減したのと同等の効果がみられた。一方、ゲイン下限値による差はあまり見られなかった。

## 4. 主観評価

### 4.1 実験方法

本実験では、音源を0-800 Hz, 800-2700 Hzの二つの周波数帯域に分けて、それぞれに残響付加の処理を行った。残響時間は、1.0秒, 1.5秒, 及び2.0秒の3種類とし、それぞれを提案システムで処理を施した。

結果の分析にはシェッフェの対比較法 [4] を用い、それぞれの指標についての平均嗜好度を算出した。順序を考慮して刺激音の対を提示し、提示順序はランダムとした。1つ目の音声と比較した2つ目の音声について、それぞれの指標で5段階で評価してもらった。実験参加者は健聴者22名(男性17名, 女性5名)とし、そのうち「残響感」の実験に11名(男性9名, 女性2名), 「自然さ」の実験に11名(男性7名, 女性4名)であった。刺激音はヘッドホン (AKG: K271 MkII) で提示し、提示音圧レベルは LAeq で 80 dB となるようにダイヘッド (Brüel&Kjaer: Type45BB) を用いて校正した。

### 4.2 実験結果

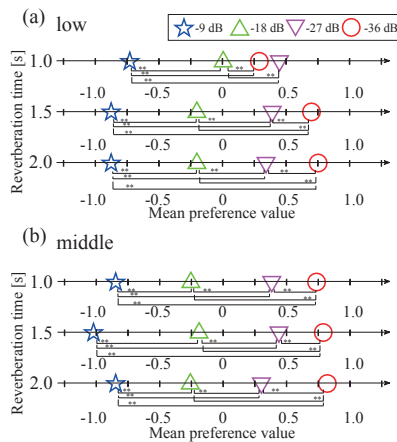


図-2 音声の残響感における平均嗜好度 ((a)0-800 Hz, (b)800-2700Hz)

図-2 と図-3 は、残響時間 1.0 秒, 1.5 秒及び 2.0 秒の音声に

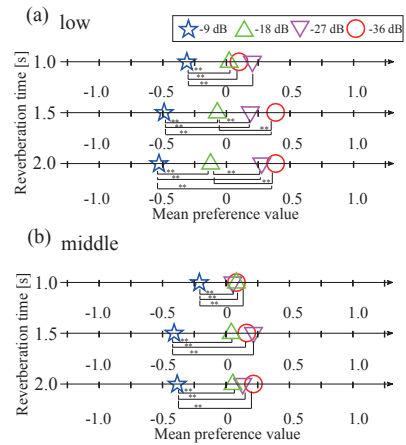


図-3 音声の自然さにおける平均嗜好度 ((a)0-800 Hz, (b)800-2700Hz)

おける各ゲイン下限値での「残響感」と「自然さ」の平均嗜好度を示している。平均嗜好度が大きいことは「残響が減っている」、「自然である」と参加者が感じたことを意味している。

図-2 より、残響時間 1.0 秒のときの低域の結果を除いて、どの残響時間においてもゲイン下限値が下がるにつれ、概ね統計的に有意に平均嗜好度が大きくなっている ( $p < 0.01$ )。したがって、ゲイン下限値を下げると残響がより減っていると感じられたことを示している。

図-3(a)(b)のどちらにおいても、ゲイン下限値-9 dB のときの平均嗜好度が統計的に有意に下がっていることから、不自然だと判断されている。また、統計的に有意ではなかったが、ゲイン下限値-36 dB や -27 dB のとき平均嗜好度はおおむね上がっているため、下限値を下げた方がより自然に聞こえることが示唆された。図-3(a)における残響時間 1.0 秒の結果について、ゲイン下限値-36 dB より -27 dB のときの音声の方が自然であるという傾向が見られることから、残響時間が短い場合に残響を過度に低減され音声の自然さが損なわれてしまう恐れがあることが示唆された。

## 5. むすび

本研究では、残響付加音声の聴き取り改善にむけて、処理遅延を考慮した補聴器のための残響低減手法を提案し、客観評価と主観評価によって残響低減度合いのパラメータによる影響について評価した。客観評価からは残響低減の有効性及びパラメータの影響が小さいこと、主観評価からは残響が低減されていないと音声不自然に聞こえるということがわかった。今後は、実環境を想定した実験に加え提案手法の総合的な評価をする必要がある。

### 参考文献

- [1] A. K. Nabelek and J. M. Pickett, "Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners," *J. Speech, Lang., Hear. Res.*, vol.17, no.4, pp.724-739, 1974.
- [2] 春原政浩, 中市真理子, 近藤幸弘, "補聴器のための簡易残響低減手法の検討," *音講論集*, pp.853-854, Mar. 2015.
- [3] T. H. Falk, C. Zheng and W.-Y. Chan, "A Non-Intrusive Quality and Intelligibility Measure of Reverberant and Dereverberated Speech," *IEEE Trans. Audio Speech Lang. Process.*, vol.18, no.7, pp.1766-1774, Sep. 2010.
- [4] 佐藤信, *統計的官能検査法*, 日科技連, pp.253-262, Mar. 1985.