

残響環境下における拡声音声の母音部振幅調整による明瞭化の試み

ATTEMPT OF ARTICULATING PUBLIC ADDRESS IN REVERBERANT ENVIRONMENT

BY ADJUSTING AMPLITUDE OF VOWEL PART

堀江祐太
Yuta Horie

竹内太法
Taiho Takeuchi

立蔵洋介
Yosuke Tatekura

静岡大学大学院総合科学技術研究科

Graduate School of Integrated Science and Technology, Shizuoka University

1 まえがき

スピーカーなどで拡声された音声は、残響のある空間において不明瞭になることが知られている。この原因のひとつに、先行する母音の残響成分が後続の子音をマスクする overlap-masking[1] がある。本研究においては、音声の母音部の一部の振幅を調整することで、後続する子音のマスクングを防ぎ、残響中の音声の明瞭化を図れるかを調査する。本稿では、通常発話音声の母音部の一部を無音化することによるマスクング度合いの変化を、客観評価値によって評価した結果について述べる。

2 母音部の無音化

音声中の各モーラにおいて、子音に先行する母音部の振幅調整を行った。今回の調査では、指定した範囲の振幅をゼロとした。これを無音化とする。聴感および波形の変化から子音部と母音部の境界を判断し、境界部分の前部 50 ms を無音化した。

3 マスキングの客観評価

明瞭度を落とす原因となるマスクングの量の定量的評価として、各モーラが残響成分に対してどの程度エネルギーを持っているのかを評価した。評価指標として、SNR と Glimpse Proportion (GP) 基準 [2] を用いた。

SNR は以下の式で与えられる。

$$\text{SNR}[\text{dB}] = 10 \log_{10} \frac{P_{\text{direct}}}{P_{\text{echo}}} \quad (1)$$

ここで、 P_{direct} は直接音の振幅の 2 乗和、 P_{echo} は残響音の振幅の 2 乗和である。

GP 基準は、ある時間フレームにおける音声、背景の雑音よりも大きなエネルギーを持つ周波数成分をどの程度持っているのかを表す指標であり、以下の式で与えられる。

$$\text{GP} = \frac{100}{N_f N_t} \sum_{t=1}^{N_t} \sum_{f=1}^{N_f} \mathcal{L}(y_{t,f}^{(\text{direct})} - y_{t,f}^{(\text{echo})}) \quad (2)$$

ここで、 N_t は時間フレーム数、 N_f は周波数チャンネル数、 $\mathcal{L}(\cdot)$ はロジスティックシグモイド関数であり、 $y_{t,f}^{(\text{direct})}$ 、 $y_{t,f}^{(\text{echo})}$ はそれぞれ分析フレーム t における周波数チャンネル f に対する直接波と残響音の STEP (Spectro Temporal Excitation Pattern) の近似値を表す。

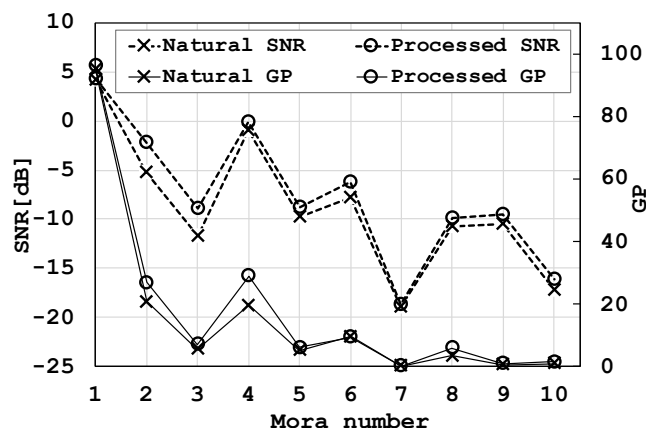


Fig. 1 原音声、および母音の一部無音化処理音声における SNR と GP

4 実験結果

残響音声の作成は、シミュレーションにて作成したインパルス応答と音声の畳み込みによって行った。無音化処理による聴感上的変化を明確にするため、長時間の残響が発生する環境を想定しシミュレーションを行った。シミュレーションの結果、インパルス応答の残響時間は 2.72 s となった。音源信号は、研究話速バリエーション型音声データベース (SRV-DB) [3] の女性話者 PF02 の 5.02 モーラ/s 音声を利用した。発話内容は、「かれは、かいにこだわる」であった。原音声と母音部一部無音化処理音声における、各モーラでの SNR および GP を Fig. 1 に示す。横軸の番号は、発話モーラの時系列順を表す。

今回の調査においては、すべてのモーラにおいて、処理音声の客観評価値が原音声の評価値を上回っていた。これより、本提案手法は overlap-masking の改善において一定の効果をみせる可能性が示唆された。今後の課題としては、母音部無音化の自動化、残響に対する最適な振幅調整法の検討、主観評価実験を通じた本明瞭化手法の妥当性の検討が挙げられる。

参考文献

- [1] A.K.Nabelek, *et al.*, *J. Acoust. Soc. Am.*, vol. 86, no. 4, pp. 1259–1265, 1989.
- [2] C. V. Botinhao, *et al.*, *Proc. ICASSP*, pp. 3997–4000, 2012.
- [3] <http://www.it.ice.uec.ac.jp/SRV-DB/>