

残響環境下における音声明瞭度改善のための前処理 -DSPを用いた簡易版定常部抑圧処理の実装-*

© 高橋慶, 安啓一, 程島奈緒, 荒井隆行 (上智大・理工), 網谷智博, 栗栖清浩 (TOA)

1 はじめに

講義や講演が行われる際, 残響によって音声の明瞭性が下がる場合がある. その原因の1つとして, 先行する音素に付加された残響が後続する音素に影響を与える *overlap-masking* が挙げられる^[1-2].

荒井ら^[3-4]は, 残響による *overlap-masking* を軽減するための前処理として, 定常部抑圧処理を提案しており, Hodoshima *et al.*^[5]は, 特定の残響環境下において定常部抑圧処理が有効であることを示している. また, Takahashi *et al.*^[6]は, 定常部抑圧処理を DSP に容易に実装するために安武ら^[7]の前後フレームのレベル差に着目した子音判定法を応用し, 定常部判定の際に行われる複雑な処理の簡素化を行い, 簡易版定常部抑圧処理を提案した.

本報告では, Texas Instruments (TI) 社の TMS320 C6713DSK に簡易版定常部抑圧処理を実装し, 処理時間の計測及び, 定常部を判定する閾値の再設定を行った.

2 簡易版定常部抑圧処理

定常部抑圧処理は, 音声知覚に必要な情報が遷移部と比べて少ない定常部を予め抑圧しておくことで *overlap-masking* の影響を軽減し, 音声の明瞭性を改善することができる^[3-4].

本報告にて DSP に実装する簡易版定常部抑圧処理は, 先行研究^[6]で用いたものと同じである. 始めに A/D 変換器により 48kHz で標準化した音声信号をフレームに分割した. その際の, フレーム長は 30ms とし, フレーム毎のオーバーラップはなしとした.

次に前後フレーム (前フレーム: $W1$, 後フレーム: $W2$) のレベル差 ($W1/W2$) を比較した. その差 (これ以降, フレーム間レベル差) が閾値である -0.72dB と 6dB の間であれば定常部とみなし, 抑圧処理が施された音声信号を出力する. そして, それ以外の部分は遷移部とみなし, 処理を施さずに音声信号をそのまま出力した (Fig. 1). これらの閾値は, 先行研究^[6]で従来法^[3-5]と同様の定常区間を検出するように設定した値である. また, 抑圧率は, 従来法^[3-5]と同

様に 40% とし, それに相当する 7.95dB だけ抑圧した. 処理を施した音声と施さない音声を Fig.2 に示した.

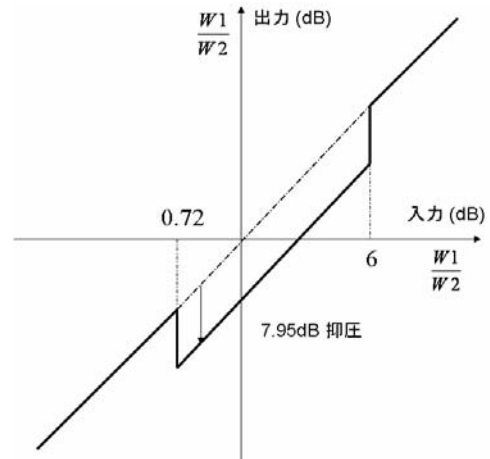


Fig. 1 簡易版定常部抑圧処理の定常部判定

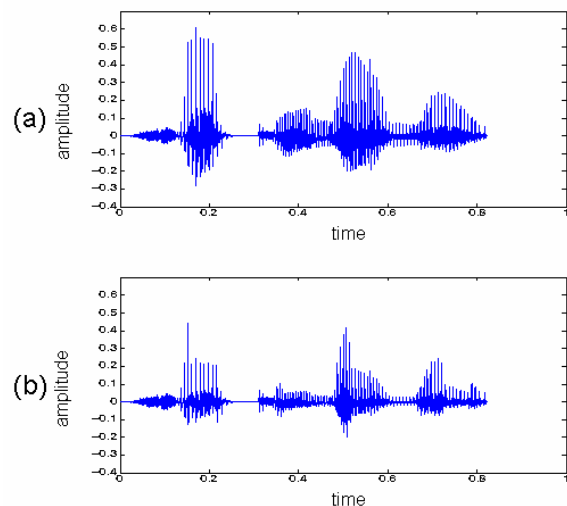


Fig.2 (a) 簡易版定常部抑圧処理を施さない音声 (b) MATLAB 上で簡易版定常部抑圧処理を施した音声

3 TMS320 C6713DSK への実装

ここからが本報告の検討である.

3.1 ハードウェアの仕様

標準化周波数は, A/D 変換器, D/A 変換器ともに 48 kHz に設定した. また, 高速にデータを伝達する McBSP からの割り込みにより, CPU を使わずにデータのやり取りができる

* A Preprocessing Technique for Improving Speech Intelligibility in Reverberant Environments: Implementation of Simple Algorithm for Steady-state Suppression with DSP, by Kei Takahashi, Keiichi Yasu, Nao Hodoshima, Takayuki Arai (Sophia Univ.), Tomohiro Amitani and Kiyohiro Kurisu (TOA)

EDMA を動かしデータの入出力を行った。その際、ダブルバッファリングを用い、音声を入力すると同時に音声処理をする設計にした。これらより、理論上の処理時間の上限は 30 ms となった。

3.2 実装の正確性

簡易版定常部抑圧処理を MATLAB 上で動作させた際の結果と DSP の開発環境である Code Composer Studio ver.2.21 の Simulator 機能(これ以降, CCS Simulator)を用いて, 処理を動作させた時の結果を比較し, アルゴリズム通りに実装されていることを確認した。確認方法としては, 単音節 (Hodoshima *et al.*^[5] で用いられているものと同じ 14 個の子音+母音/a)を用いて, フレーム毎に計算されるフレーム間レベル差が MATLAB 上と CCS Simulator が一致することを調べた。その結果, dB 値の誤差が 0.001% 未満であることからアルゴリズム通り実装できていることが確認できた。

3.3 処理時間の計測

CCS Simulator を用いて, 処理に要するサイクル数を測定し, 処理時間を計算した。その結果, サイクル数は 147,806 であり, 処理時間は C6713 DSK の動作周波数 225MHz より, 約 0.7ms であることがわかった。この値は, 処理時間の上限である 30ms を大きく下回っていることから, 安定して動作していることが確認された。

3.4 閾値の再設定

DSP 実機では, 伝送路雑音が付加されるため CCS Simulator とは異なった dB 値が出力されてしまう。そこで, 定常部判定を行う閾値の再設定を行った。設定方法としては, 音声に対し, CCS Simulator での結果と同等の箇所を抑圧するように閾値を -6dB から 6dB の範囲で 0.1 dB ずつ変化させ, 遷移部だけを残し抑圧箇所の変化を目視で観測した。実際は, 定常部の判定区間を MATLAB と DSP 実機で一致させたかったが, 伝送路雑音の影響で一致させることができなかった。そこで, 目視で確認して最も一致度が高いとみられる値である -0.7 dB と 1.8 dB を上限と下限の閾値として設定した。Fig.3 は, -0.7 dB と 1.8dB を閾値に設定した際の単語「しゃくなげ」に MATLAB 及び DSP 実機で簡易版定常部抑圧処理を施したものである。Fig.3(a), (b)をみてもわかるように MATLAB と DSP 実機で抑圧箇所に差はみられてしまうが, 今後は DSP 実機と MATLAB での明瞭度にも差が現れるかどうかを確認したいと考えている。

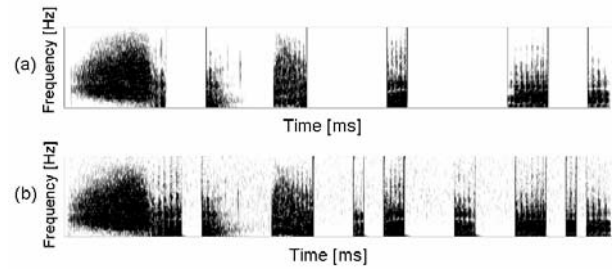


Fig. 3(a) MATLAB で簡易版定常部抑圧処理を施した時の音声信号 (b) DSP 実機で簡易版定常部抑圧処理を施した時の音声信号

4 まとめ

簡易版定常部抑圧処理を TI 社の TMS320 C6713DSK に実装し, 処理時間から安定して動作することを確認した。しかし, 伝送路雑音により, DSP 実機と MATLAB の定常部検出箇所が多少変化してしまうことがわかった。今後は, 本研究で用いた DSP 実機と MATLAB 上での簡易版定常部抑圧処理で明瞭度に差が現れるかを聴取実験で調べると共に伝送路雑音の軽減について考え, MATLAB の定常部判定に近づきたい。

謝辞

本研究は, 科学研究費補助金(A-2, 16203041) および, 文部科学省私立大学学術研究化推進事業上智大学オープン・リサーチ・センター「人間情報科学研究プロジェクト」の支援を受けて行われた。

参考文献

- [1] Nabelek and Pickett, *J. Speech Hear. Res.*, 17, 724-739, 1974.
- [2] Bolt and MacDonald, *J. Acoust. Soc. Am.*, 21(6), 577-580, 1949.
- [3] 荒井他, 音講論 (春), 449-450, 2001.
- [4] Arai *et al.*, *Acoust. Sci. Tech.*, 23, 229-232, 2002.
- [5] Hodoshima *et al.*, *J. Acoust. Soc. Am.*, 119(6), 4055-4064, 2006.
- [6] Takahashi *et al.*, *Proc. ICA*, 2007.
- [7] 安武, 中島, 信学技法, HIP2005-94, 79-84, 2005.