

信号の相対的な振幅変化に着目した残響抑圧

正員 広林 茂樹* 非会員 林 智也*
正員 山淵 龍夫*

Dereverberation using variance of amplitude between past and current signals

Shigeki Hirobayashi*, Member, Tomonari Hayashi*, Non-member, Tatsuo Yamabuchi*, Member

We propose a method to recover the original sound waveform from the observed sound waveform without using the inverse filtering of the room transfer function. Reverberation is caused by the echo element of the original sound signal. In this study, short time frame signals subdivided reverberant speech by time windowing, and we attempted to reduce the echo influence by subtracting previous frame signals from the current frame signal. Weighing coefficient and adjusted time for the subtractions are determined by correlation between past and current signals. In the case that reverberation time is about 1 sec, the result of simulation shows good recovery of original speech.

キーワード：デコンボリューション，相関係数，信号処理，音源回復，残響抑圧

Keywords: deconvolution, correlation coefficient, signal processing, source waveform recovery, dereverberation

1. はじめに

音声認識や拡声系会議通話において，室内反射音は音源波形に歪みを与え音源の特徴を埋もれさせる原因となっている。この音場における逆問題，すなわち残響抑圧では，これまで様々な手法が開発されているが，その殆どは室内インパルス応答の測定を伴う逆フィルタ理論である^{(1)~(6)}。そこで，著者等は室内インパルス応答の測定を必要としない残響抑圧法について検討してきた⁽⁷⁾⁽⁸⁾。正確な逆フィルタリングに及ばないものの，単話者の条件下では音源音声の信号包絡線に着目したエンベロープ回復処理が高い改善効果を有することを確認した。特に音声信号中の語尾を人工的に短くすることで，明瞭性が向上することがわかった。そこで著者等は，時間的に変化する相対的な振幅変化から，残響抑圧を行うことを考えた。観測された残響音声をそれぞれ短時間窓で切り出し，時間的に過去に遡った窓内信号の大小関係から小さい方を逐次残響成分とみなし，残響成分を減算する手法である。本報告では，多くの反響を伴う環境下でも，室内インパルス応答の測定を行わない残響抑圧法実現の可能性を実験的に調べる。

2. 相関関係とインパルス応答

原音声を $x(t)$ ，室内インパルス応答を $h(t)$ とすると残響音声 $y(t)$ は次式の畳み込み演算で表されることが知られている。

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = x(t) * h(t) \cdots (1)$$

ただし， $*$ は畳み込み演算子である。室内インパルス応答 $h(t)$ は，1つの直接音と複数の反射音で構成されるため，まず直接音と1つの反射音の和で表される単純なインパルス応答モデルで回復アルゴリズムを検討する。まず，残響信号 $y(t)$ は，音源信号 $x(t)$ とこの遅延信号 $ax(t-\tau)$ の和で表されることから，反射波の影響を取り除き音源信号 $x(t)$ を復元できる可能性がある。

$$\hat{x}(t) = y(t) - ax(t-\tau) \cdots \cdots \cdots (2)$$

$\hat{x}(t)$ は回復信号， τ は直接音と反射音の間における時間遅れである。また，複数の反射波に関しても式(2)を単純に拡張できると考えると，次式により残響歪の回復の可能性はある。 N は反射波の個数である。

$$\hat{x}(t) = y(t) - \sum_{k=1}^N a_k \times x(t-\tau_k) \cdots \cdots \cdots (3)$$

本報告では式(3)に示される回復アルゴリズムを近似式により導出し，その回復効果を定量的に検討する。

3. 残響抑圧法

一般的に音源発生直後の観測信号は平衡状態の観測信号に比べ，残響の影響が相対的に少ない。そこで観測信号 $y(t)$ を微小区間で区切ると，音源発生直後の区間信号は残響の影響が少なくほぼ直接音のみで構成されたと考え，音声信号 $x(t)$ として近似できると仮定した。すなわち，音源発生直後の観測信号 $y(t)$ を音源信号 $x(t)$ として，残響の影響が取り除かれた回復信号 $\hat{x}(t)$ を順次式(3)の第2項 $x(t-\tau_k)$

* 富山大学工学部 〒930-8555 富山県富山市五福 3190 番地
Faculty of Engineering, Toyama University,
Toyama-shi, Toyama, 930-8555 Japan

に代入し、次式のように逐次的に回復信号 $\hat{x}(t)$ を求める。

$$\hat{x}_1(t) = y_1(t)$$

$$\hat{x}_k(t) = y_k(t) - \sum_{m=1}^{k-1} a_m \hat{x}_m(t - \tau_m) \quad \dots\dots (4)$$

なお、フレーム処理のため、微小区間に分別された音源信号 $x(t)$ 、観測信号 $y(t)$ 、回復信号 $\hat{x}(t)$ をそれぞれ $x_k(t)$ 、 $y_k(t)$ 、 $\hat{x}_k(t)$ とした。また、 $y_k(t)$ は観測信号のフレーム番号 k における微小区間(k 番目のフレーム)であり、 \hat{x}_k は回復信号の k 番目の微小区間である。本報告では a_m 、 τ_m を \hat{x}_m と式(4)の第2項との各相互相関関数で求められた最大値とその時間遅れ τ とする。また、反射音は直接音に比べその大きさが小さくなるため、時間前方の比較対象フレー

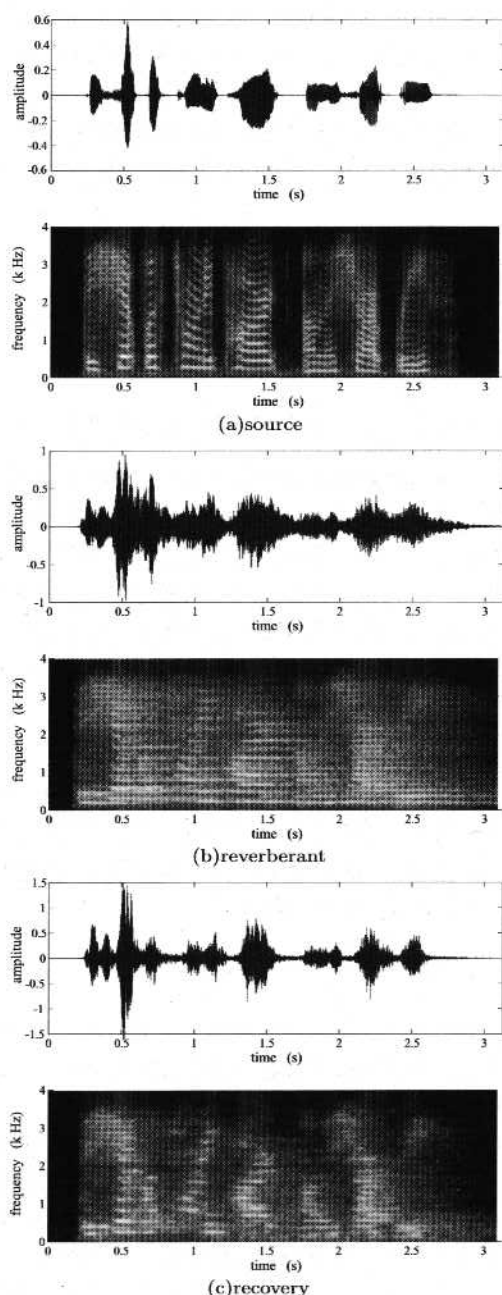


図1 波形とスペクトログラム

Fig.1. Waveforms and spectrograms

ムのエネルギーに対し、比較フレームのエネルギーが大きい場合のみ処理を行うこととする。

4. 残響抑圧実験

実験においては音声信号処理でよく用いられているように、信号のサンプリング周波数を8kHzとし、暫定的に16個の帯域に等分割するフィルタバンクを用いて、帯域毎に独立に回復処理を行った。また、フレーム長は64msで、シフト長は32msとする。残響時間約1秒の残響音声の回復例を図1に示す。図1(a),(b),(c)は、それぞれ音源音声、残響音声、回復音声の時間波形(上)とスペクトログラム(下)である。式(4)に示す単純な処理でも音源波形の概略が復元できることがわかる。また、一般的に残響音声における残響感と音声の包絡線変化には高い相関があり、残響回復を定量的に検証するため信号包絡線に対する次式の二乗誤差を用いた⁽⁷⁾。

$$I_p = 10 \log_{10} \frac{\int_0^T \{e_x(t) - e_y(t)\}^2 dt}{\int_0^T \{e_x(t) - \hat{e}_x(t)\}^2 dt} \quad \dots\dots (5)$$

式(5)によると、図1における回復効果は約4.3dBであることが確認できた。エンベロープ回復法⁽⁸⁾に比べ、残響時間等の情報も必要ないため、実用機器の環境調整等の作業の軽減もはかれる。

5. まとめ

本報告では室内インパルス応答の測定を必要としない逆問題について検討した。残響信号における時間的な振幅変化から残響成分を特定し、残響の影響を取り除き音源信号を推定する解法を提案した。実音声を用いた回復実験では、フィルタバンクを用いて広域信号を狭域化して、本提案処理を試みた。今後は、明瞭度評価やフィルタバンクにおける狭域化の影響、他手法との比較実験などを調べる予定である。本研究の一部は、SCAT研究助成金による。ここにお礼申し上げる。

(平成15年3月31日受付、平成15年5月2日再受付)

文 献

- (1) S.T. Neely and J.B. Allen: "Invertibility of a room impulse response", *J. Acoust. Soc. Am.*, Vol.66, No.1, pp.165-169 (1979)
- (2) M. Miyoshi and Y. Kaneda: "Inverse filtering of room acoustics", *IEEE Trans. ASSP*, Vol.36, pp.145-152 (1988)
- (3) H. Wang and F. Itakura: "Realization of acoustic inverse filtering through multi-microphone sub-band processing", *IEICE Trans. Fundamentals*, Vol.E75-A, No.11, pp.1417-1483 (1992)
- (4) K. Furuya and Y. Kaneda: "Two-channel blind deconvolution of nonminimum phase FIR systems", *IEICE Trans. Fundamentals*, Vol.E80-A, No.5, pp.804-808 (1997)
- (5) H. Nakajima, M. Miyoshi, and M. Tohyama: "Sound field control by indefinite MINT filter", *IEICE Trans. Fundam.* E80-A, pp.821-824 (1997)
- (6) 神沼允伸・伊勢史郎・鹿野清宏:「周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計」, 音響学会誌 Vol.57, pp.175-183 (2001)
- (7) 広林茂樹・野村博昭・小池恒彦・東山三樹夫:「パワーエンベロープ伝達関数の逆フィルタ処理による残響音声の回復」, 電子情報通信学会論文誌 A, **J81-A**, 10, pp.1323-1330 (1998.10)
- (8) 広林茂樹・山淵龍夫:「帯域分割を用いたパワーエンベロープ逆フィルタ処理の残響抑圧効果」, 電子情報通信学会論文誌 A, **J83-A**, 8, pp.1029-1033 (2000.8)