Title	残響時間の自動推定による残響波形パワーエンベロープの回復に関する研究
Author(s)	古川，正和
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1553
Rights	
Description	Supervisor:赤木 正人，情報科学研究科，修士

Research on recovery of reverberation waveform power envelope by automatic estimation at reverberation time

Masakazu Furukawa (010103)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 15, 2002

Keywords: power envelope, reverberation time, power envelope inverse filtering, modulation transfer functions.

1 Introduction

The noise and the reverberation exist together in the signal observed in a real environment. The accuracy decreases remarkably when the sound signal processing is done under a real environment. The reverberation cause to decrease the performance of the sound signal processing. A lot of inverse filtering that the original signal is recovered from the reverberation signal have been examined, and the effectiveness is shown. If the impulse response is not measured whenever changing, the recovery accuracy decreases. The impulse response is varied with a change of the environment.

The power envelope inverse filtering based on modulation transfer function (MTF) was proposed by Hirobayashi et. al. The power envelope inverse filtering recover the power envelope of the reverberation signal power envelope. As a result, the influence of the reverberation was able to be reduced without measuring the impulse response. However the constraint condition that the reverberation time is already-known exists.

The aim of this research is

- The reverberation waveform power envelope is recovered by estimating reverberation time from the reverberation waveform.
- The problem of the power envelope inverse filtering is solved.
- The possibility to reduce the influence of the reverberation is examined.

The technique by which the reverberation is effectively suppressed can be constructed.

2 Power envelope inverse filter

The power envelope inverse filtering Hirobayashi propose used to recover the power envelope of the reverberation waveform.a

2.1 Fundamental principle

The original signal and impulse response are assumed to be $x(t)$ and $h(t)$. The reverberation signal $y(t)$ is the convolution of $x(t)$ and $h(t)$. The original signal and impulse response are modeled based on the MTF as follows.

$$x(t) = e_x(t)n_1(t) \tag{1}$$
$$h(t) = e_h(t)n_2(t) \tag{2}$$
$$e_h(t) = ae^{\frac{-6.9t}{T_R}} \tag{3}$$

where $e_x(t)$ and $e_h(t)$ are envelopes of $x(t)$ and $h(t)$, a is constant, and T_R is reverberation time. $n_1(t)$ and $n_2(t)$ are un-correlated white noises.

Assume that the reverberation signal $\mathbf{y}(t)$ is generated from a certain stochastic process. The squared ensemble average of $\mathbf{y}(t)$ is

$$\begin{aligned}
\langle \mathbf{y}(t)^2 \rangle &= \left\langle \left\{ \int_{-\infty}^{\infty} e_x(\tau)^2 e_h(t-\tau)^2 d\tau \right\}^2 \right\rangle \\
&= e_x(t)^2 * e_h(t)^2 \qquad (4) \\
&= e_y(t)^2 \qquad (5)
\end{aligned}$$

The power envelope obtained by (4) is called theoretically value of the reverberation signal power envelope. The original signal transfer function is

$$P_x(z) = \frac{P_y(z)}{P_h(z)} = \frac{1}{a^2}\left\{1 - \exp\left(\frac{-13.8T_s}{T_R}\right) z^{-1}\right\} P_y(z) \tag{6}$$

where $P_x(z), P_y(z), P_h(z)$ are z transformations of $e_x(t)^2, e_y(t)^2, e_h(t)^2$, respectively. The original transfer function is the product of a reverse transfer function of reverberation signal power envelope characteristic $P_y(z)$ and impulse response power envelope characteristic $P_h(z)$.

The restorated signal $\hat{x}(t)$ is

$$\hat{x}(t) = \hat{e}_x(t)\frac{y(t)}{e_y(t)} \tag{7}$$

where $e_y(t)$ is the amplitude envelope of reverberation signal. $\hat{e}_x(t)$ of right side is a recovery envelope requested in inverse z transformation of (6).

Then, the reverberation signal power envelope can be recovered.

2.2 Problem

There is some problems when the reverberation signal power envelope is recovered using the power envelope inverse filtering. (1)The reverberation time used for the recovery processing has to be decided and (2)whether the modulation transfer function can be applied to speech. In this research, these problems are solved and the reverberation waveform power envelope is recovered.

3 Estimation at reverberation time

In the power envelope inverse filtering, the reverberation time has to be already-known. In this research, the reverberation time is presumed automatically and the power envelope is recovered from the reverberation waveform.

The modulation index is considered. In the modulation transfer function, the modulation index of the reverberation signal is decided by the modulation index of original signal. Therefore, the modulation index of the original signal has to be already-known. The silent section dose not have power. Thus it is assumed that the modulation index of original signal is one.

3.1 Estimation method of reverberation time

We propose an estimation method of reverberation time. The reverberation deletes a peak of the power envelope in the right in time and buries a dip of the power envelope. The power envelope inverse filtering emphasizes peaks and dips. The dip is emphasized more than the necessity when the estimated reverberation time is longer than the added reverberation time. The recovery power envelope might become negative. However, the power envelope which is the squared envelope of signal does not negative clearly. It is thought that the reverberation time can be estimated by obtaining the reverberation time in case the power envelope becomes negative because the modulation index of the original signal is assumed to be one.

3.2 Result

The proposed estimation method for the reverberation time is evaluated. The following envelope are used for the original power envelope.

1. The most simple power envelope is
$$e_x(t)^2 = 1 + \sin 2\pi F t \tag{8}$$
where, F is modulation frequency, and $F = 10$ [Hz]. This envelope is be called a sinusoid power envelope.

2. The power envelope containing the frequency from 1 to 20 [Hz] is
$$e_x(t)^2 = \sum_{k=1}^{20} \sin(2\pi k t + \hat{\theta}_k) \tag{9}$$
where, $\hat{\theta}_k$ is random. This envelope is called harmonics power envelope.

3. The power envelope obtained from a random signal passed throug low-pass filtering with the cut off frequency 20 [Hz]. This envelope is called random power envelope.

The carriers are white noises. The added reverberation times are five kinds of $T_R = 0.1, 0.3, 0.5, 1.0,$ and 2.0 [s]. The power envelope is extracted from the signal added the reverberation. The reverberation signal power envelope was obtained by low-pass filtering the absolute value signal after the Hilbert transform. The results of estimation of the reverberation times are shown in Fig.1. (a) in Fig.1 is estimated reverberation times when the theoretical value of the reverberation signal power envelope was used. (b) in Fig.1 is estimated reverberation times when extracted power envelopes were used. The horizonal axis indicates added reverberation times and the vertical axis indicates estimated reverberation times. The correlation values and SNRs of the power envelopes between the recovered the extracted reverberation signal power envelopes $\hat{e}_x(t)^2$ and the original power envelopes are shown in Fig.2. (a) and (b) are the added reverberation time of $T_R = 0.5, 1.0$ [s], respectively. The marks are the estimated reverberation times. In Figs.1 and 2, solid line, dashed line and dashed dot line show the sinusoid power envelope, harmonics power envelope and random power envelope in the original power envelope, respectively.

In Fig.1(a), the reverberation time can be accurately estimated when the reverberation signal power envelope is the theoretical value. In Fig.1(b), the estimation error increases when the reverberation time becomes long. In Fig.2, the degree of power envelope recovery improves when the reverberation signal power envelope is recovered using the estimated reverberation time, the proposed technique is effective.

4 Application to speech

The power envelope of reverberated speech is recovered using the power envelope inverse filtering.

The power envelope of reverberated speech is different in each band. However, the power envelope inverse filtering equates the power envelopes of all bands. In order to recover the power envelope of reverberated speech, the band must be divided. It is thought that the band width can be decided from the frequency domain of the co-modulated power envelope. Thus, the filter bank uses the equal band-pass filter bank of 400 [Hz].

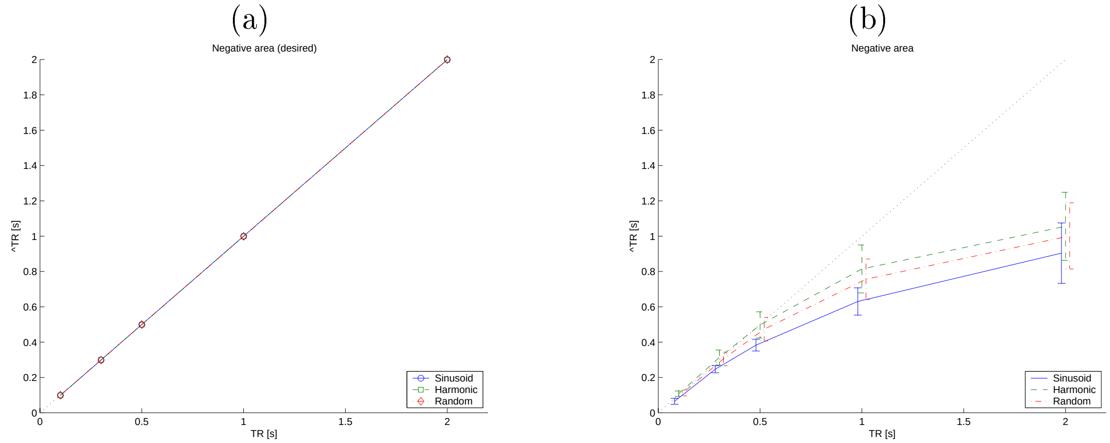

Figure 1: Estimation at the reverberation time by the negative value.

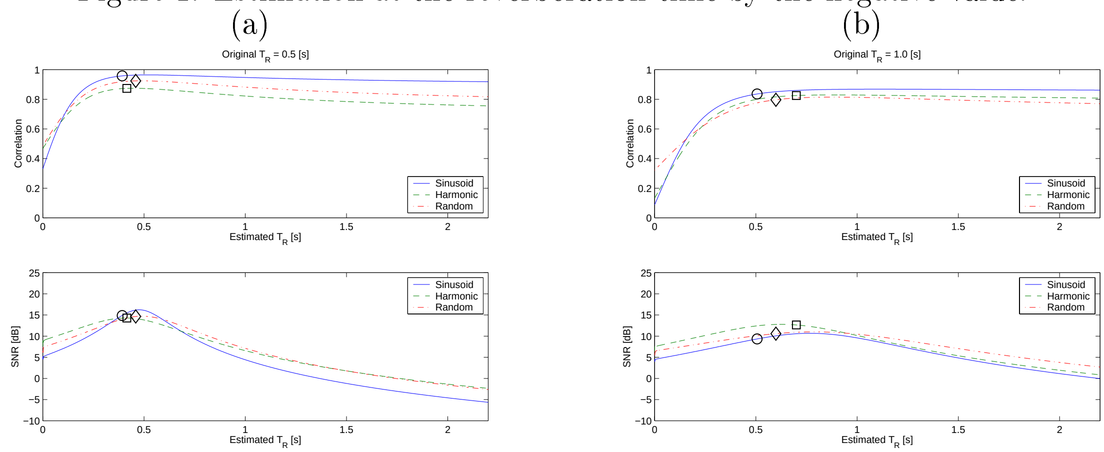

Figure 2: Relation ship between estimated reverberation time and the correlation value,SNR.(a)$T_R = 0.5$ [s],(b)$T_R = 1.0$ [s]. (top)Correlation value,(bottom)SNR

4.1 Result

The original speech wave used a japanese word /aikawarazu/ utterd by a male speaker 'mau' in the ATR speech database.

Improvement index of power envelope distortion(Ip) is defined to evaluate the recovery of the power envelope.

$$\mathrm{Ip} = 10 \log_{10} \frac{\displaystyle\int_0^T \left\{ e_x(t)^2 - e_y(t)^2 \right\}^2 dt}{\displaystyle\int_0^T \left\{ e_x(t)^2 - \hat{e}_x(t)^2 \right\}^2 dt} \qquad [\mathrm{dB}] \tag{10}$$

where, $e_x(t)^2$, $e_y(t)^2$, $\hat{e}_x(t)^2$ is the original power envelope, the reverberation signal power envelope and the recovered power envelope, respectively. T is analyzing time, and about 1.6 [s]. Ip indicates positive value when the

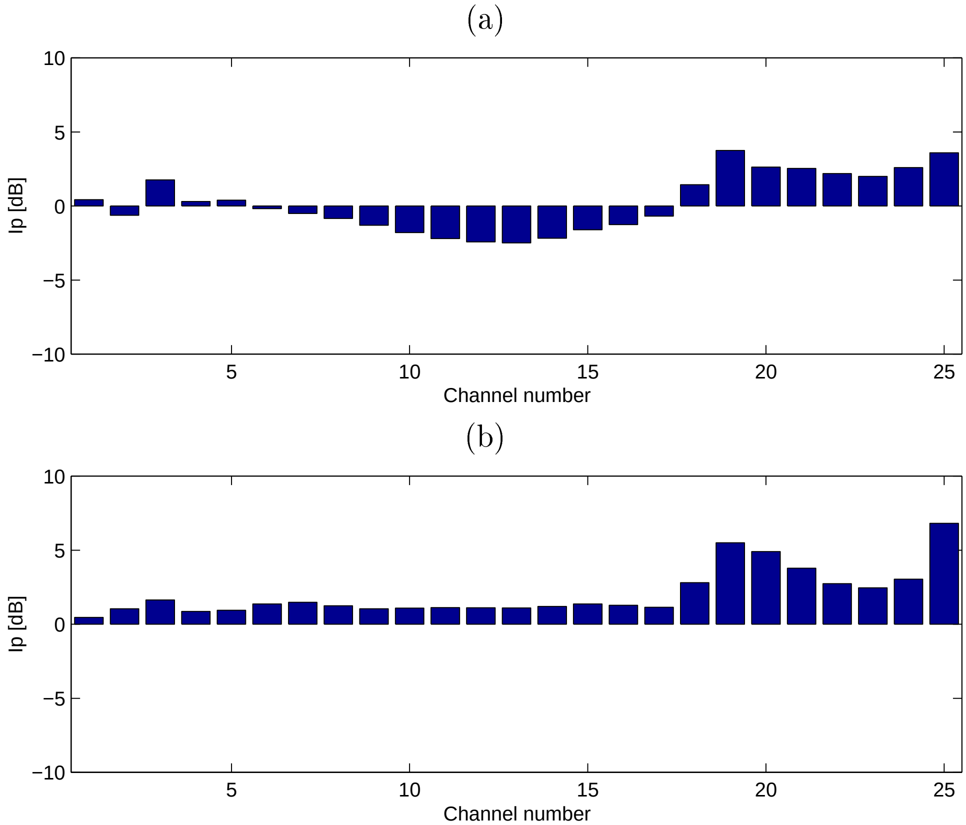

Figure 3: Improvement index of power envelope of each bands.(a)constant reverberation time,(b)estimated reverberation time

power envelope recovers positive value. Ip of each channel of the filter bank is obtained. The results are shown in Figure 3. (a) is a Ip when recovering reverberation time as fixed 0.5 [s] in all bands. (b) is a Ip when estimating reverberation time and recovering in each band.

In Fig.3, the power envelope is not recovered except for high frequency band. In Fig.3 (b), the power envelope is recovered in all bands. For this reason, this technique is effective.

5 Conclusion

This paper showed the power envelope of the reverberation waveform was recovered by the estimation at reverberation time from the reverberation waveform. The problems of the power envelope inverse filtering, the decision of the reverberation time and application to the speech, were examined. Estimation of the reverberation time results showed that the reverberation time was able to be estimated by calculating reverberation time which recovered power envelope does not negative. The power envelope of speech wave was divided by the band width considered to be a co-modulation. The reverberation time of each divided band is estimated. It was shown that the power envelope in each band is recoverable.

References

[1] S.Hirobayashi, H.Nomura, T.Koike, "Speech Waveform Recovery from a Reverberant Speech Signal Using Inverse Filtering of the Power Envelope Transfer Function", IEICE, Vol. J81-A, No10, pp.1323-1330,1998.

[2] S.Hirobayashi, et. al., "Validation of Blind Dereverberation Using Power Envelope Inverse Filtering and Filter Banks", IEICE, Vol. J83-A, No8, pp.1029-1033,2000.

[3] T,Houtgast, H.J.M. Steeneken, and R.Plomp "Predicting speech intelligibility in room acoustics", Acustica, Vol.46, pp.60-72, 1980.