Title	残響音声からの音声特徴量抽出法と 音源波形再合成に関する研究
Author(s)	酒田，恵吾
Citation	
Issue Date	2003-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1672
Rights	
Description	Supervisor:赤木 正人，情報科学研究科，修士

A study on extracting speech features from reverberant speech and re-producing speech signal

Keigo Sakata (110049)

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

February 14, 2003

Keywords: Modulation transfer function, Power envelope inverse filtering, Reverberation time, filterbank.

1 Introduction

Recovery of the original speech from reverberant speech is an important issue such as speech emphasis for remote meeting system and preprocessing for speech recognition system. A lot of dereverberation methods that use inverse filter of the impulse response of room acoustics have been proposed to solve this important issue. However, it is difficult to utilize these methods because it has to measure the impulse response of the room acoustics which characteristics temporally vary with various factors such as temperature.

On the other hand, Furukawa et al. proposed the method for recovering the power envelope from the reverberant speech band on the MTF concept[1], without measuring impulse response[3]. However, there are two problems

This paper proposes a speech dereverberation method without measuring the implulse response of room acoustics. In this model, the envelopes and carriers are decomposed from speech signal using a filterbank and then they are processed by each block induvidualy. About processing of the envelope, we also use the power envelope inverse filtering method based on the MTF concept. This is improved to resolve the problems of Furukawa et al.'s method. About processing of the carrier, we propose the method for reproducing carrier using the estimated fundemental frequency (F0) from the reverberant speech. In this paper, it is assumed that the estimated F0 is knownd. The dereverberated speech signal is reconstructed using filterbank from the dereverberated envelopes and the reproduced carrier.

2 The problems of the dereverberation method for power envelope

The dereverberated method for power envelope by proposed Furukawa et al. have two problem. We examine these problems.

1. Appling MTF concept is not inquired in narrow bandwidth, when the sub-band bandwidth on the filterbank is determined
2. The recovery effect of low Frequency region are small

2.1 Examination about the determination of adequate sub-band bandwidth

When determinating of adequate sub-band bandwidth, It is necessary to take and two focus point into consideration. one is the co-modulation for power envelope , other is appling the MTF concept. The method by Furukawa 'et al. examine about co-modulation for power envelope, but is don't examine appling the MTF concept. Therefore We examine these points. First, Examination about co-modulation for power envelope results that more narrower sub-band bandwidth, It can be regarded as co-modulated power envelope envelope. Second, Examination about appling the MTF concept result that more narrower sub-band bandwidth, It can be not regarded as appling the MTF concept. It turns out that the examination result of two points to two focus point has the relation of a trade-off, and to adequate sub-band BandWidht is from 300 to 400 Hz.

2.2 Examination about the recovery effect of low Frequency region

2.2.1 The cause that the recovery effect of low Frequency region are small

The cause that the recovery effect of low Frequency region are small is explained. In low Frequency region, the situations that the silent section was too long are often seen. Such a situation, the reverberant power envelope's modulation index is 1.the estimation method of Reverberation Time(T_R) proposed by Furukawa et. al is inapplicable to such a situation. threfore the method is assumed that original power envelope's modulation index is 1. Therefore, We propose the new Estimation method of T_R which can be applied even when the such situation.

2.2.2 Propose the new Estimation method of T_R applied the situations that the silent section are too long

We propose the new method focused on the amount of move change of the power envelope. The power envelope inverse filtering [2] has work that move processed power envelope in the direction of reverse time. While corresponds with T_R ,Reverberation Time is Estimated by Searching the point that the amount of move change of the power envelope (D) falls rapidly as a boundary conditions that inverse filter is filtering adequatly. This new method can apply the situation that the silent section are too long, and improved recovery effect in low frequency region.

3 the method of carrier reproduciong

We reproduce carrrier devided the voiced sound interval and unvoiced sound interval. About the voiced sound interval, carrier implemented by harmonics using the PIFM source model using the estimated F0. About the unvoiced sound interval carrier implemented by random noise.

4 Simulation to Evaluate proposed model

Simulation is carried to Evaluate proposed model. Reverberant speech signal that covolute of Original speech (mau /sinbun/)and Impluse respoonse room acoustic (T_R=0.5) , is processed by proposed model. BandWidths of filterbank is 400 Hz. As evaluation measures, we use log Spectral Disstrition(LSD) to evaluate as a speech. To measure three speech signal more exactly, three speech signal are resconstructed, This three signal's envelope are diffrent 'Original','Reverberant',and 'Dereververant'. carrier is equally to reproduced by the carrier reproduciong method. Masuring to LSD each conbination one is 'Original','Reverberant', other is 'Original','Dereveberant',we evalute Improved LSD. As a result, it was shown Improving 1 dB in speech exist.

5 conclusion

This paper proposed a speech dereverberation method without measuring the implulse response of room acoustics. This method is composed of the pwower inverse filtering method and carrier reproducing method. In the processing of the envelope, Power envelope inverse filtering method was improved to solve the problems of Furukawa et al.'s method. Improvements were to determine the adequate bandwidths of the filterbank considering relationship between co-modulation within channels and the MTF concept with channels, and to determine adequatly reverberant time even if long silence exist in. The carrier reproduction model was implemented using the PIFM source model. Simulation was carried out to evaluate the proposed method. As a result, it was shown that the proposed method was precisely dereverberated reverberant speech.

References

[1] Houtgast,T.,Steenken,H.J.M., "A review of the MTF concept in room acoustic and its use for extimating speech intelligibility in audioria," J.Acoust.Soc.Am Vol.77, No.3, March 1985. (In Japanese)

[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. (In Japanese)

[3] Furukawa, Unoki, Akagi "A method for recovering the power envelope from the reverberant speech based on MTF", IEICE, SP2002-15, pp. 49-54, 2002.