

|              |                                                                                 |
|--------------|---------------------------------------------------------------------------------|
| Title        | 変調伝達関数に基づいた雑音抑圧に関する研究                                                           |
| Author(s)    | 山崎, 悠                                                                           |
| Citation     |                                                                                 |
| Issue Date   | 2009-03                                                                         |
| Type         | Thesis or Dissertation                                                          |
| Text version | author                                                                          |
| URL          | <a href="http://hdl.handle.net/10119/8099">http://hdl.handle.net/10119/8099</a> |
| Rights       |                                                                                 |
| Description  | Supervisor: 鶴木 祐史, 情報科学研究科, 修士                                                  |

# A study on the noise suppression method based on the MTF concept

Yutaka Yamasaki (0710073)

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

February 5, 2009

**Keywords:** Modulation Transfer Function, Temporal power envelope, Noisy environments, Noise suppression.

In real environments, significant features of speech are smeared due to noise and reverberation so that sound quality and intelligibility of the observed speech are drastically reduced. Improvement of noisy reverberant speech (noise suppression and dereverberation) is, therefore, needed in various speech signal processing, such as hearing aid systems and preprocessing for automatic speech recognition systems.

There are several well-known suppression methods which can be used to remove the effects of noise or reverberation in either noisy or reverberant environments. There are, for example, spectral subtraction method proposed by Boll, Kalman filtering method proposed by Paliwal and Basu, minimum-phase inverse filtering method proposed by Neely and Allen, and multiple input/output inverse theorem (MINT) method proposed by Miyoshi and Kaneda. Although these methods can work well in either noisy or reverberant environments, it cannot work in both noisy and reverberant environments, simultaneously. Recently, Kinoshita *et al.* studied a strategy of the speech enhancement method in noisy reverberant environments, by considering two sequential processes: noise reduction using spectral subtraction for noisy reverberant speech and then dereverberation using linear prediction for noise-reduced reverberant speech. However, it seems to be more complex in Kinoshita's modeling. We thought that the

best solution should be able to deal with both additive noise and reverberant effect simultaneously.

On the other hand, Houtgast and Steeneken proposed a prediction method that can assess the effects of the enclosure on speech intelligibility in both noisy and reverberant environments by using the modulation transfer function (MTF). The MTF concept makes the suppression method of both noise and reverberation simultaneously realization.

Unoki *et al.* proposed the temporal power envelope inverse filtering method based on the MTF concept. Their method assumed an environment that is the reverberation. These methods improved about 30% of the reduction in speech intelligibility caused by the reverberation. If we could propose a noise suppression method based on MTF concept, we can propose an MTF-based speech enhancement method which can suppress noise and reverberation simultaneously, so that this method can improve the loss of speech intelligibility suffered from additive noise and reverberation. The goal of our work is to propose a speech enhancement method for noise reduction and dereverberation.

We propose a noise suppression method by restoring the smeared MTF. The modulation index and the averaged power in the temporal output power envelope ( $e_y^2(t)$ ) are affected by noise. We restore the averaged power levels and modulation index using MTF to suppress the noise effects.

Input information is needed for the calculation of MTF. The average of temporal noise power envelope ( $\overline{e_n^2}$ ) is estimated from non-speech sections. The average of temporal input power envelope is estimated by subtracting  $\overline{e_n^2}$  from the average of temporal output power envelope. MTF can be calculated by using this SNR.

We carried out three simulations on noise reduction to evaluate whether the proposed method can adequately restore the desired speech from noisy speech signals. Three Japanese sentences uttered by ten speakers (five males and five females) were used in the evaluations. Three types of noise, white, pink, and babble noise signals, were used in the simulations. Signal to noise ratios (SNRs) were fixed at 20, 10, 5, 0, and -5 dB. In these simulations, correlation (Corr), SNR, and log spectrum distortion (LSD) and weighted LSD of speech intelligibility were used as evaluation measures to show the improvement the restoration accuracy achieved through our

method. As the results of simulations, the maximum improvement in LSD was about 31 dB, the maximum improvement in weighted LSD was about 8 dB, the improvement in Corr was constant and the improvements in SNR increased as SNRs decreased. These results show that the proposed method can improve the temporal power envelope and the waveform of the input signal from the noisy signal. We found that the proposed method can be used to adequately restore the temporal power envelopes and to suppress the noise effects of noisy signal.