A survey of a method of restoration noisy reverberation for improving speech intelligibility

Shota Morita (0810062)

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

February 9, 2010

Keywords: Speech intelligibility, Listening difficulty, Restoration noisy reverberation .

Speech communication is a fundamental way of information propagation for us. Noisy or/and reverberant environments impair the speech communication in the room. For example, the speech communication is affected by the noise or/and reverberation when we use the hands-free capability of a speech application out of the microphone and we cannot tell the utterance content for the party. We need the method of restoration noise and reverberation to solve the problem for restoring speech intelligibility, but its method has not been proposed.

We set the final goal that is the achievement of smooth voice communications in noisy reverberant environments. Firstly, we surveyed if speech intelligibility can evaluate the smooth speech communication or not by surveying the evaluating method of speech transmission performance. Secondly, we surveyed the previous recovering methods of noise and reverberation. Finally, we surveyed whether the speech intelligibility can be rocovered the method that is noisy reverberant restoration.

We found that we need using the word intelligibility or sentence intelligibility to evaluate the speech communication by surveying the evaluate method of speech transmission performance and we have to feature the familiarity of words when we evaluate the speech intelligibility. “Listening difficulty” was proposed because vocal words of high familiarity had not

difference in speech intelligibility, but it had difference hearings. Therefore it was found that speech communication should be evaluated by both speech intelligibility and listening difficulty. Speech Transmission Index (STI) is objective evaluation of speech intelligibility proposed by Houtgast and Steeneken that is high correlation with speech intelligibility. However, Toida showed that STI is low correlation with speech intelligibility case by case, but Sato shwed that STI is very high correlation with listening difficulty. Therefore, STI is the best method of objective evaluation to evaluate the speech communication.

We surveyed the methods of noise reduction that are Spectral Subtraction (SS) method, Active Noise Canceling, Minimum Mean Square Error-Short Time Spectral Amplitude (MMSE-STSA), Winner filtering method and Max Likelihood method, Real Active SpecTrAl (RASTA). Their methods can reduct the noise well, but it is difficult to dereverberate by their methods because the characteristic of the noise and the characteristic of the reverberation are quite different. The purpose of a lot of noise reduction method is Automatic Speech Recognition (AIR) that it is used recovering feature parameter differ from recovering speech intelligibility. We surveyed methods of dereverberation that are Minimum-phase inverse filtering method, Multiple-input/output inverse theorem (MINT) method, method of acoustic inverse filtering through multi-microphone sub-band processing and Harmonic-based dEReverBeration (HERB). However, the methods other than HERB need to measure the room impulse response (RIR), while the RIR has to be precisely measured before the dereverberation that is non-blind method. HERB can dereverberate on the blind processing by single microphone, but the purpose of the method is ASR that it is used recovering feature parameter, but the purpose of the method is ASR that it is used recovering feature parameter differ from recovering speech intelligibility. In the achievement of the restoration noisy reverberation by using these methods, there is only technique for combining the noise reduction with the dereverberation. Such as methods are proposed that are method of combining the SS method with Linear Prediction (LP) and method of combining the Winner filtering with Linear filter. Either method is sequentially-processing of the noise reduction as subtraction of noise component and dereverberation as inverse filtering of reverberant

component. These methods has limit to recover speech intelligibility because these method do not recover the important parameter for speech.

On the other hand, noise reduction and dereverberation based on Modulation Transfer Function (MTF) have been proposed that the MTF has interrelationship with STI. This method restore the temporal envelope of signal based on MTF that the envelope has important feature of speech intelligibility, but the envelope smeared due to noise and reverberation. STI can predict the Speech intelligibility that STI is calculated by MTF is static approximation of room acoustics (noise and impulse response). STI can be evaluated speech intelligibility in noisy reverberant environments. Therefore the method based on MTF can restore the speech intelligibility and listening difficulty in noisy reverberant environment.

We further surveyed the noise reduction and dereverberation based on MTF. In this time, we put the problem of the method of power envelope inverse filtering based on MTF in reverberant environment and we studied the improvement of restoration accuracy of the restored power envelope. As the result, it was found the proposed method can adequately improve restoration accuracy of the power envelope in the previous method. Improvements are power envelopes, however improvement degree was not bigger as we expected. Speech intelligibility cannot be improved only by recovered envelope of speech, but we knew that speech intelligibility can be improved carrier restoration. However, the method of carrier restoration is carrier regeneration that method needs fundamental frequency estimation and Voice Activity Detection (VAD) in noisy reverberant environment. In the future, we would like to propose these method and finally we are going to achieve to recover the speech intelligibility and listening difficulty based on MTF in noisy reverberant environment.