Study on Estimation of Speech Transmission Index and Room Acoustics without Measurement of Acoustics in the Room

Kyohei Sasaki (1110028)

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

February 6, 2013

Keywords: Modulation transfer function, room impulse response, speech transmission index, reverberation time, deutlichkeit.

The speech transmission index (STI) is an objective measure that is used to assess the quality of speech transmission in room acoustics. The STI has been proposed by Houtgast and Steeneken and this calculation has been standardized in IEC60268-16. STI is calculated from modulation transfer function (MTF). In addition, MTF is represented as the normalized power spectrum of the room impulse response (RIR). However, people must be excluded for the measurement of RIRs as the viewpoint of hearing protection because the input signal with high sound pressure level is used to measure RIR in a room. Therefore, it is impossible to measure STI and room acoustics in a room where people cannot be excluded. This study aims to blindly estimate STI and room acoustics without measuring RIR to solve the above-mentioned problem. Unoki and others have previously proposed a simplified method of blindly estimating STI in room acoustics based on the MTF concept. This method effectively estimates STI from the observed reverberant AM signal by estimating inverse MTF in which the RIR is modeled as the Schroeder's artificial RIR. However, the simplified method has three problems. The first problem is estimated precision. The simplified method could not correctly estimate STI from the observed AM

signal if the measured RIR could not be approximated as the Schroeder's RIR model. The second problem is the use to an observed signal. The simplified method does not know that it has the same precision of estimates when it use a personal sound. The third problem is an estimation in the actual environment. STI and room acoustics can be estimated in the environment where people are in the room, however it has not yet been verified.

This study aimed to solve three problems. This paper proposed an improved method of blindly estimating STI and room acoustics by assuming the modified Schroeder's RIR (generalized RIR) to solve the first problem. In the proposed method, MTF is derived by estimating two parameters (reverberation time and the order of n) of the generalized RIR model from the reverberant AM signal. STI is calculated from the estimated MTFs by using IEC 60268-16. This paper verified an estimate of STI and room acoustics using observed speech signal and improved it as needed to solve the second problem. As for the observed speech signal, speech is used, and investigated the applicable law of the STI estimate from the characteristic of the speech. STI and room acoustics is measured to solve the third problem. In addition, the proposed method was estimated STI and room acoustics in the environment where people are in the room. Estimation of STI and room acoustics without measurement of acoustics in the room was proposed by abobe methods.

The proposed method was evaluated by estimating STI and room acoustics from reverberant AM signal, reverberant speech signal and measured reverberant speech signal. The evaluation of estimating STI and room acoustics from reverberant AM signal revealed that the estimated STI and room acoustics by the proposed method has higher precision of estimates than those by the previous method. The evaluation of estimating STI and room acoustics from reverberant speech signal revealed that the method of estimating STI and room acoustics using speech signal has precision as high as the method of estimating STI and room acoustics using AM signal. The evaluation of estimating STI and room acoustics from measured reverberant speech signal revealed that the estimated STI and room acoustics by the proposed method has the same precision of estimates as measured STI and room acoustics. In addition, the proposed method can correctly

estimate STI where people are in the room. Thus, the proposed method could solve all problems of the previous method. Therefore, these results revealed that the proposed method can be used to blindly estimate STI from the observed speech signal in real environments in which people are in the room.