On Uniqueness of Filter Coefficients and the Learning Method in Stereophonic Acoustic Echo Canceller

by Masahiko Tanaka
(Supervised by Professor Kenji Nakayama and Akihiro Hirano)

ABSTRACT

In multimedia age, multi-channel voice transmission becomes very important to realize presence circumstances. TV conference systems, however, echo reflected in the rooms should be cancelled. In the stereo echo cancellers (SEC), there is more freedom of the parameters used to estimate the echo path characteristics. This causes an uncertainty problem. Namely, even though the residual echo is well reduced, the SECs cannot estimate the echo path characteristics. In order to overcome this problem, some kinds of pre-processing are introduced, which increase the conditions on the SEC adaptation. However, this pre-processing usually degrade voice quality.

In this paper, we derive the condition, under which the unique solution is guaranteed. This condition is imposed on the number of samples in the impulse responses of the room reflection and filter coefficients of the SECs. However, this condition cannot satisfied by the SECs in both directions. Let the rooms A and B be the conference rooms, the SEC for transmission from B to A can be trained, at the same time, however, the SEC in the opposite direction cannot be trained.

Therefore, a new learning algorithm is further proposed in this paper. The above condition is satisfied by dividing the adaptive filter $H(z)$ into the former and latter parts $H_f(z)$ and $H_l(z)$, respectively. $H_f(z)$ and $H_l(z)$ are alternately trained. When $H_f(z)$ is trained, the other part $H_l(z)$ is fixed as the trained in the previous step. $H_f(z)$ cannot converge to the optimum function due to the error in $H_l(z)$, and vice versa. However, if the condition $|\lambda_i| < 1$, λ_i are the eigen values of the matrix related to the room acoustics characteristics, is satisfied, then convergence is guaranteed.

Several examples are simulated. The coefficients of the SECs can move towards the optimum to some extent. The residual echo can be reduced. However, the reduction is not enough compared with the pre-processing methods. So, the learning algorithm itself should be more improved.