A Learning Method of Echo Canceller in Double Talk

by Shinichiro Ushimaru
(Supervised by Professor Kenji Nakayama and Akihiro Hirano)

ABSTRACT

In multimedia age, group communication systems, such as TV conference systems, are very important. In some cases, such as TV conference systems and satellite communication systems, however, echo is generated due to reflection in the conference rooms and impedance discordance. If delay time is longer than 30 mSec, it can be heard as echo, which should be cancelled to achieve high quality voice transmission. For this purpose, echo cancellers (ECs) are used.

One problem in adjusting the ECs is double talk. Let A and B be the conference rooms, the voice of the speakers in the room A is reflected in the room B, and returned to the room A. This is the echo. The EC is used to cancel this echo. When the speakers in the room B do not talk, then only the echo is transmitted to the room A. This situation is called 'single talk'. In this situation, The EC can be adjusted using only the echo. However, at the same time, if the speakers in the room B talk, then the signal transmitted from the room B to A includes the echo and the near-end voice, which is used to adjusts the EC. The near-end voice should not be reduced by the EC. However, the EC cannot distinguish the echo and the near-end voice, and they are both minimized resulting in bad estimation of the real echo path characteristics. This problem can be overcome by using variable step size. However, how to control the step size still remain as an open question.

In this paper, we introduce a method to estimate the cross-correlation between the echo and near-end voice. The replica of the echo is used instead to estimate the cross-correlation. The step size is controlled as inversely proportional to the cross-correlation. The learning process is controlled by the far-end voice power. Furthermore, two kinds of the cross-correlations estimated in short and long intervals are combined to follow nonstationary inputs. by this method, the step-size can be quickly adjusted for quick change in the cross-correlation. Furthermore, it is stable for slow change intervals. Still, it is rather difficult to exactly follow many kinds of changes, for instance, from the single-talk to teh double-talk, and vice versa, and echo path changes. For this reason, a double filter method is further developed. Adaptive filters AF_1 and AF_2 are used. AF_2 is adjusted using the variable step-size. The filter coefficients in AF_1 are transferred to AF_1, when some conditions are satisfied. Namely, the far-end voice power exceeds some level, and the cross-correlation stays under some threshold. When these conditions continue within some period, the AF_2 coefficients are transferred to AF_1.

The proposed methods are evaluated though computer simulation using several examples. The last method demonstrates good results, including stable adaptation in the double talk periods, fast adaptation for quick echo path change, and stable and moderate convergence speed for echo path change in the double talk.