Speech enhancement technique supporting binaural selective hearing in noisy reverberant environment

Yuuki Sasaki (0910025)

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

Jun 31, 2012

Keywords: Speech Enhancement Technique, Binaural Selective Hearing, Noisy Reverberant Environment, Two–Stage Binaural Speech Enhancement with Wiener Filter, Cepstral Mean Subtraction.

Speech recognition becomes difficult under influence of noise and/or reverberation. Additionally, there are some reports that listening capability of hearing handicapped person declines remarkably in noisy reverberant environments. Therefore, speech enhancement techniques in order to suppress noise and/or reverberation have been introduced into applications like hearing–aids or speech recognition. In speech enhancement techniques proposed until now, some speech enhancement techniques focused on binaural hearing featured of humans.

Frequency domain binaural model (FDBM) based on Lindeman's binaural hearing model was proposed by Usagawa *et al.* This method calculates interaural phase difference and interaural level difference to estimate the direction of the target signal. Then, the received signal is enhanced by FDBM in noisy environment. Two–Stage Binaural Speech Enhancement with Wiener Filter (TS-BASE/WF) was proposed by Li *et al.*, to suppress noise with two–step processing; noise estimation stage and noise suppression one. TS–BASE/WF has excellent noise-reduction performance, because TS–BASE/WF has two–step processing.

When these speech enhancement techniques are used indoors, suppression ability of noise and reverberation simultaneously should be required.

Room impulse responses (RIR) can be divided into early reflection and late reverberation bordering on the time that is dependent on size of the room. Early reflection correlates to the target signal. Late reverberation that is added several reflection sounds have less correlation to the target signal. Moreover, late reverberation diffuses around the room.

FDBM estimate target signal direction by using cross-spectrum. Then, FDBM could not work well under infuluence of early reflection. Noise estimation stage of TS–BASE/WF without using cross-spectrum could work without the influence of early reflection and late reverberation. On the one hand, since noise suppression stage of TS–BASE/WF adopt Wiener filter in which it is assumed there is no correlation between target signal and noise. Hence, enhanced signal could be affected due to early reflection.

Almost all of speech enhancement techniques for supporting binaural selective hearing cannot suppress reverberation. This paper aims at constructing speech enhancement supporting binaural selective hearing in noisy reverberant environment. Performance of TS–BASE/WF in reverberant environment is evaluated. In adition, experiments verifiy whether TS–BASE/WF can suppress early reflection and late reverberation. Results show that TS–BASE/WF can suppress late reverberation. However, early refractions influence enhanced signals by TS–BASE/WF due to using a Wiener filter.

According to the previous experiment results, Cepstral Mean Subtraction (CMS) is used as a frontend for TS–BASE/WF in order to suppress early reflection. Next, experiments are carried out to show whether the modified method is superior to TS-BASE/WF in reverberant and/or noisy environments. Those results indicate that the modified method exceeds TS–BASE/WF in reverberant environments and noisy reverberant. From those results, the speech enhancement technique for supporting binaural selective hearing in noisy reverberant environment was constructed. Aplications like hearing–aids or speech recognition which is introdeced the modified method of TS–BASE/WF will be improved those perfomances.