Title	音環境バリアフリーのためのパワーエンベロープ処理体系
Author(s)	森田，翔太
Citation	
Issue Date	2017-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/14249
Rights	
Description	Supervisor:鵜木　祐史，情報科学研究科，博士

Abstract:
Concept of Ubiquitous speech communication is speech communication for “Anytime”, “Anywhere”, “Anybody”, “Security & Safe”. “Anywhere” in Ubiquitous speech communication is important problem for smooth speech communication in sonic environments. Effects of noise and reverberation in real environments are barrier in sonic environments. I set the big goal in my research that realization of barrier-free for sonic environments. The barrier-free has some approach as speech enhancement, noise reduction and dereverberation. In this paper, the barrier-free is supposed to approach of noise reduction and dereverberation likes cancel out of barrier. Current simultaneous methods for noise reduction and dereverberation is not approach of congruous to sonic environments. Then, these methods cause over and under improvements. Noise and reverberation should be estimated as measurements and index of noise and reverberation. MMSE-STSA as noise reduction needs *a priori* SNR that is estimated as parameter for noise. Almost dereverberation methods, inverse filtering approach by estimated room impulse response (RIR) and preliminary measured RIR is used. In simultaneously methods of noise reduction and dereverberation, effects of noise and reverberation are reduced by noise reduction and dereverberation respectively. The method is combined methods of noise reduction and dereverberation. Therefore, the methods cannot work likes congruous to sonic environments, thus it was caused over and under improvements. Voice activity detection (VAD) is technology for speech coding and preprocessing for automatic speech recognition (ASR). Robust VAD for the barrier-free need fixing the discriminate condition for detecting signal to obtain the required performance of VAD. Noise reduction is used as pre-processing for robust VAD. VAD performance is degraded in not involving supposed sonic environments. Human has stress when they hear the over and under restored speech by the barrier free. Recognition performance of machine is degraded in that case. It is caused by incongruous between sonic environments and human with machine. A goal of my research is congruous processing between sonic environments and human with machine as the barrier-free. Priority element to solve the problem is “congruous to sonic environments.” This paper treat noisy reverberant environments as sound field. The user for the barrier-free is set to machine in this research. The temporal power envelope of signal is treated as target signal in the barrier-free. This paper achieves that power envelope processing systems for barrier-free communication under noisy reverberant environments. As requirements for power envelope processing systems, effects of noise and reverberation are reduced as one sonic environments, and parameters for noise reduction and dereverberation should be used on a feature. Concept of modulation transfer function (MTF) is based to achieve the power envelope processing system for barrier-free communication in noisy reverberant environment. MTF can treat noise and reverberation simultaneously on the MTF. MTF is matched for congruous between sonic environments and human. On MTF concept, modulation index (MI) for power envelope of input signal is 1, MTF is decided by effects of noise and reverberation, and MI for power envelope of output signal is obtained as 1 minus MTF. The system based on MI = 1 for congruous to sonic environments. The power envelope signal processing system is based on the concept. The cohesive speech signal processing system was proposed to achieve power envelope processing system for barrier-free communication in noisy reverberant environments. The cohesive speech signal processing system is consisted of robust VAD, power envelope restoration including power envelope subtraction and MTF inverse filtering, SNR estimation, and reverberation time estimation. Most of these processing is based on MI = 1 of MTF concept. These methods were evaluated in noisy reverberant methods, then the results show these methods can work in noisy reverberant environments. For the application as speech signal processing, the cohesive speech signal processing system was applied to pre-processing of ASR system and STI estimation method. The results show that proposed approach can work as one approach of the barrier-free to sonic environments from the results.

Keywords: power envelope restoration, noisy reverberant environments, modulation transfer function, congruous to sonic environments,