Title	振幅変調特性に着目した雑音残響に頑健な基本周波数推定法
Author(s)	三輪，賢一郎
Citation	
Issue Date	2018-12
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/15758
Rights	
Description	Supervisor:鵜木　祐史，情報科学研究科，博士

Abstract

Fundamental frequency (F0) has a very important role in various speech/music information processing scenes including the speech communications in which a machine intervenes, and depending on whether or not the fundamental frequency can be accurately grasped, the processing accuracy of these various sound/music information processing is influenced.

Estimation of the fundamental frequency, F0, of a target sound signal has been widely studied in the literature on speech signal processing, and many methods have been proposed over the last half century. The conventional methods can be divided into processing in the time domain or frequency domain, or both. Most of these methods use periodic features in the time domain (e.g., autocorrelation) or harmonic features in the frequency domain (e.g., comb filtering). The others focused the information of sound source by the vocal tract filter effect based on the source-filter model. Some methods that precisely estimate the F0 of target noiseless speech have been established. However, they cannot precisely estimate F0 under noisy conditions, reverberant conditions, or both conditions. Other methods have been proposed that robustly estimate the F0 of observed speech under noisy conditions. On the other hand, a few methods to estimate the F0 of a target signal robustly and accurately have been studied in reverberant environments. However, it has not yet been clarified whether all these methods can precisely estimate F0 in very noisy reverberant environments. In other words, there is no method that can accurately estimate F 0 in an environment with the both of noise and reverberation.

Therefore, in this research, we got a hint from the behavior of the pitch perception for AM signal, and focused on the modulation component of the sound signal. We proposed a novel F0 estimation method robust to noise and reverberation by using an unprecedented approach to simulate human pitch perception by amplitude modulation demodulation technology. Focusing on the modulation component of the sound signal has several merits. Considering the amplitude modulation signal from the viewpoint of the modulation transfer function, the influence of disturbance such as noise and reverberation can all be put in a simple diagram of a decrease in modulation index. Therefore, it is possible to grasp the influence of the disturbance by the observing of the modulation index without going through complicated signal processing such as examining the autocorrelation of the waveform or examining the frequency spectrum information. Although the amplitude of the modulation component decreases due to disturbance, the modulation component period is retained, so the modulation component itself of the sound signal is a signal component robust to disturbance. In addition, it is also possible to add the waveform restoration function with modulation index as a parameter, if necessary.

Simulation results have shown that the proposed method can reliably cope with time-variant signals and has robustness for noise and reverberation. Furthermore, we consider the F0 estimation of musical instrument sound as an example of an application that matches the characteristics of the proposed method, and shown that the proposed method can be applied sufficiently to a signal such as instrument sound.

Throughout this research, we have proposed F0 estimation method based on pitch perception for AM signal concerning F0 estimation in noisy reverberant environment, we show that F0 estimation method robust to noisy reverberation can be constructed by using amplitude modulation demodulation technique.

There should certainly be a field where the proposed method can contribute. For example, by implementing the proposed method as a part of other F0 estimation methods, we can expect roles that complement other F0 estimation methods robustly. Thus, the essence of the proposed method will contribute greatly to speech/music information processing technology in the future.

Keywords: F0 estimation, pitch perception, amplitude modulation/demodulation, noisy reverberant environments