



Title	Single Channel Noise Suppression Based on Speech and Noise Spectral Models
Author(s)	Thanhikam, Weerawut
Citation	大阪大学, 2012, 博士論文
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/27505">https://hdl.handle.net/11094/27505</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

氏 名	タニカーン THANHIKAM	ウィーラウット WEERAWUT
博士の専攻分野の名称	博 士 (工学)	
学 位 記 番 号	第 2 5 6 8 7 号	
学 位 授 与 年 月 日	平成 24 年 9 月 25 日	
学 位 授 与 の 要 件	学位規則第4条第1項該当	
	基礎工学研究科システム創成専攻	
学 位 論 文 名	Single Channel Noise Suppression Based on Speech and Noise Spectral Models (音声とノイズのスペクトルモデルに基づく単一チャネルノイズ除去)	
論 文 審 査 委 員	(主査) 教 授 飯國 洋二	
	(副査) 教 授 乾口 雅弘 教 授 大塚 敏之	

## 論 文 内 容 の 要 旨

The purpose of this research is to achieve single channel noise suppression based on speech and noise spectral models. This thesis consists of two main parts. The first part describes stationary noise suppression and the second part describes impulsive noise suppression.

First, a stationary noise suppression algorithm using Maximum a Posteriori (MAP) estimation with a speech spectral amplitude probability density function (speech PDF) is investigated. An estimated speech spectrum is given as a MAP solution which is obtained from the speech PDF. A useful speech PDF has been established and is entirely characterized by two shape parameters. Speech can be efficiently extracted when these parameters are properly applied. However, the speech property should be considered as a time-variant function. In this case, the fixed speech PDF cannot track the property change. In this research, under the assumption that the speech PDF changes according to signal to noise ratio (SNR), the author proposes adaptive shape parameters which are derived from real-speech PDFs in various narrow SNR intervals. The proposed adaptive shape parameters can track the change of the speech property, and give an appropriate MAP solution which is identical to the estimated speech spectrum. The effectiveness of the proposed method was examined and compared to conventional algorithms. The simulation results showed that the proposed method improved segmental SNR around 6 and 9 dB when the input speech signal was corrupted by white and tunnel noise signals at input SNR of 0 dB, respectively.

Second, an impulsive noise suppression method is investigated. This method utilizes a zero phase (ZP) signal which is defined as the IDFT of a spectral amplitude. In the impulsive noise suppression research, we assume that a speech signal has periodicity in a short observation, i.e., its spectral amplitude has values at equally spaced frequencies. In this case, the corresponding ZP signal becomes also periodic. This assumption is especially appropriate for a voiced speech which is mainly arisen

in speech signals. On the other hand, we assume that a noise spectral amplitude is approximately flat. In this case, its ZP signal takes nonzero values only around the origin. Actually, many impulsive noise signals have such property. Under these assumptions, the ZP signal of a speech signal embedded in impulsive noise in an analysis frame becomes a periodic signal except around the origin. Hence, replacing the ZP signal around the origin with the ZP signal in the second or latter period, we get an estimated speech ZP signal. Taking DFT of it gives the estimated speech spectral amplitude. The IDFT of the estimated speech spectral amplitude with the observed spectral phase provides the estimated speech signal in time domain. The major advantage of this method is that it can suppress impulsive noise without a prior estimation of the noise spectral amplitude, while the a prior estimation of the noise is indispensable in most stationary noise suppression methods. Moreover, it is shown that the proposed impulsive noise suppressor can also be available to suppress stationary wide-band noise. Simulation results showed that the proposed noise suppressor improved the SNR more than 5dB for stationary tunnel noise and 13dB for impulsive clap noise in a low SNR environment.

## 論文審査の結果の要旨

単一チャネルノイズ除去は、ひとつのマイクロフォンで取得した1系列の観測信号から、ノイズを除去し、音声信号だけを取り出す技術であり、携帯電話、TV会議システム、音声認識装置などの応用分野において重要な役割を果たす。ノイズ除去は、音声とノイズのスペクトルモデルが適切でない場合には十分なノイズ除去性能が得られないという問題がある。本論文では、観測信号に含まれるノイズが、定常的に存在する場合、突発的に生じる場合の二つの場合について、それぞれに有効なスペクトルモデルを導出し、そのモデルに基づいたノイズ除去法を示している。

定常的に存在するノイズに対しては、ガウス分布によるスペクトルモデルがよく適合することが知られており、ノイズよりも、音声のスペクトルモデルの設定が重要となる。まず、実際の音声信号のヒストグラムをパラメトリック関数を用いて表現することで、現実的かつ有用なスペクトルモデルを導出している。その際、スペクトルモデルをSNR (signal to Noise Ratio) に応じて複数導出している。そして、この結果に基づき、SNRに応じてスペクトルモデルを変更するノイズ除去法を導出し、従来法よりもノイズ除去性能が改善されることを明らかにしている。

突発的に生じるノイズに対しては、ほぼ平坦なスペクトルモデルを利用している。このモデルは、ゼロ位相信号と呼ばれる領域では、原点付近にのみ値をもつ信号に変換される。一方、音声に対しては、周期的なゼロ位相信号を与えるスペクトルモデルを仮定している。両者の性質から、突発性ノイズ除去アルゴリズムを導出し、その有効性を実際のノイズを用いたシミュレーションによって確認している。

本論文は、実際の音声およびノイズを効果的に表現するスペクトルモデルを導いたという点に新規性が認められる。また、その有効性をアルゴリズム解析およびノイズ除去シミュレーションにより確認している。よって、博士（工学）の学位論文として価値のあるものと認める。