Title	Noise Reduction Based on Microphone Array and Post-filtering for Robust Hands-free Speech Recognition in Adverse Environments
Author(s)	,
Citation	
Issue Date	2006-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/973
Rights	
Description	Supervisor: , ,

Noise Reduction Based on Microphone Array and Post-filtering for Robust Hands-free Speech Recognition in Adverse Environments

Junfeng Li

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

March, 2006

Abstract

This research proposes a noise reduction system using microphone array and post-filtering with the goal of improving the recognition accuracy and robustness of hands-free speech recognition systems in adverse environments.

Acoustic noise signals dramatically degrade the performance of many speech applications, such as speech recognition system, in practical environments. Though the problem of dealing with acoustic noises has been researched for several decades and is still a challenging topic due to the complex and time-varying characteristics of signals and acoustic environments. Noises with different characteristics from various kinds of sources make it difficult to construct an effective noise reduction system. Moreover, only the system with small physical size is preferable because of the limited space, e.g., in car environments.

To deal with localized noise, we propose a hybrid noise estimation technique by combining the multi-channel estimation approach and a single-channel estimation approach. The estimation accuracy of this hybrid technique is further improved by integrating a *robust and accurate speech absence probability* (RA-SAP) estimator. The estimated spectrum of localized noise is then compensated and suppressed from that of noisy observation on each microphone. Moreover, we further develop a generalized subtractive beamformer by relaxing the assumption of a perfectly coherent noise field to the one of an arbitrary noise field. The theoretical analysis is also presented to show the linkage between these two beamformers and to show the theoretical performance of the generalized algorithm in the theoretically defined noise fields.

To further deal with the residual non-localized noise post-filtering is normally used at beamformer output. We propose a hybrid post-filter for microphone arrays with an assumption of a diffuse noise field. In the proposed hybrid post-filter, a modified Zelinski post-filter is applied to the high frequencies; a single-channel Wiener filter is applied to the low frequencies. The proposed hybrid post-filter was proven to be able to deal with both high-correlated and low-correlated noise components in a diffuse noise field.

The performance of the proposed noise reduction system is finally investigated using speech recognition results. The speech recognition results show that the proposed noise reduction algorithm outperforms the other traditional ones in improving the speech recognition performance in the tested adverse environments.

Compared with other traditional noise reduction algorithms, this proposed algorithm demonstrates some advantages: (1) in theory, it provides the optimal solution to the problem of multi-channel noise reduction for broad-band inputs in *minimum mean square error* sense; (2) it is able to deal with various kinds of noise signals, including localized and non-localized noise, stationary and non-stationary (e.g., sudden) noise; (3) it avoids the problems of slow convergence rate and low stability in practical environments; (4) it can be implemented in real-time mode; (5) it is successful in improving the performance of hands-free speech recognition systems in adverse environments.

In addition to hands-free speech recognition systems, the noise reduction system proposed in this thesis is also useful and preferable to many other applications. For example, for hearing aid, it is able to provide more clean and intelligible speech, enhancing the performance of hearing aid to hearing impaired with a small-size microphone array at a low computational complexity in adverse conditions.

Key Words: **Hands-free speech recognition, noise reduction, microphone array, beamforming, post-filtering, multi-channel Wiener filter.**