

Two-Microphone Noise Robust Speech Recognition for Hands-free Voice Control System

Takayuki Arakawa, Takafumi Koshinaka, Osamu Hoshuyama

*Information and Media Processing Laboratory, NEC Corporation
1753, Shimonumabe, Nakahara-ku, Kawasaki-shi, Japan*

t-arakawa@cp.jp.nec.com, koshinak@ap.jp.nec.com, hoshu@bq.jp.nec.com

We present a novel two-microphone noise reduction technique to improve speech recognition accuracy and demonstrate a hands-free voice control system.

Figure 1 shows our noise robust speech recognition system. The two-channel noise canceller reduces directional noise using two microphones. The spectral shaping strongly reduces non-directional noise. This method was originally innovated for nonlinear echo cancellation [1,2], is also effective as a pre-processor for voice activity detection. The multi-feature voice activity detector correctly identifies active voice segments [3]. The model-based Wiener filter suppresses non-directional noise and enhances speaker's voice [4]. This combination eliminates noise disturbance and boosts speech recognition performance.

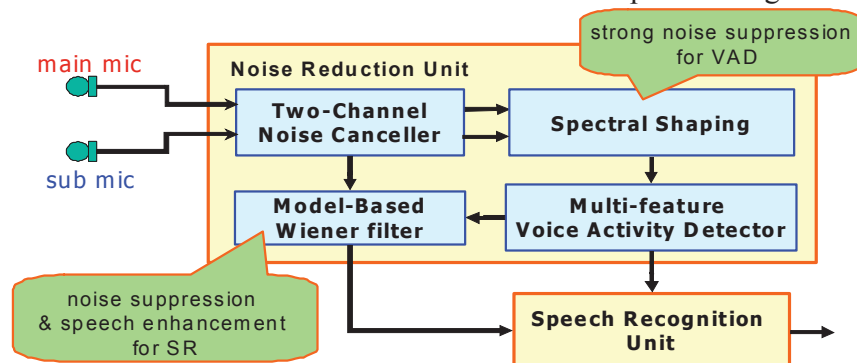


Fig 1. Our Noise Robust Speech Recognition System

Our demonstration system is shown in Fig. 2. Our noise robust speech recognition is installed in a tablet. The tablet has two microphones; one is attached on the front and the other on the back. The tablet recognizes voice commands and controls applications (e.g. TV channel selection). We will show a live demonstration. Our system can recognize 50 commands even in loud TV-sound conditions.

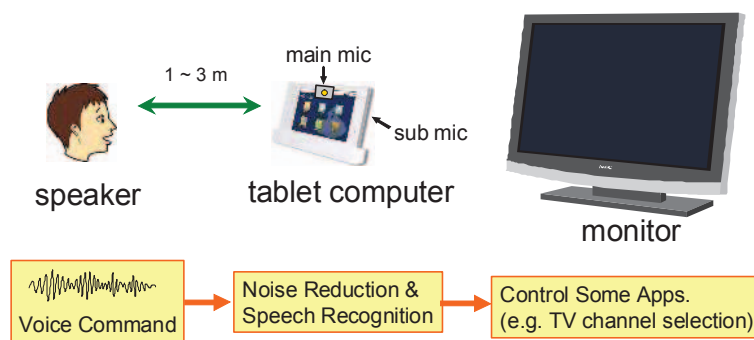


Fig 2. Demonstration: Hands-free Voice Control System

- [1] A. Sugiyama et al., "A versatile echo-and-noise control suite for adverse environments with a user-control capability," ICCE 2009
- [2] O. Hoshuyama and A. Sugiyama, "Nonlinear Echo Cancellation Based on Spectral Shaping," Speech and Audio Processing in Adverse Environments, Signals and Communication Technology, Springer 2008
- [3] T. Arakawa, et al., "Extended Minimum Classification Error Training in Voice Activity Detection," ASRU 2009
- [4] T. Arakawa, et al., "Model-Based Wiener Filter for Noise Robust Speech Recognition," ICASSP 2006