# A Speech Recognition Technique in Noise Environments Using Independent Component Analysis

Seung Ho Choi

*Department of Electronic and IT Media Engineering*

*Seoul National University of Science and Technology, Seoul, Korea*

**Abstract- In this paper, we propose a speech recognition method based on independent component analysis (ICA) that is an algorithm for signal separation. From the speech recognition experiments with noisy environments, we show that the ICA-based approach gives much improved recognition accuracies than conventional methods.**

**Keywords – Independent component analysis (ICA), signal separation, speech recognition**

## I. INTRODUCTION

Speech recognizer shows relatively good performance in a quiet environment, but performance may be deteriorated if it is applied to a real recognition environment. The actual environment in which speech recognition is performed involves factors that degrade recognition performance such as ambient noise, loudness, microphone characteristics, channel distortion, and speaker variation. Most recognizers receive a voice signal using a single microphone (or channel). The research to attenuate the noise from a single micro input signal and obtain a clean speech signal has been carried out steadily, but it has not yet satisfied the recognition performance. Signal separation techniques for separating mixed signals using independent component analysis (ICA) have been studied [1, 2, 3]. This technique is particularly useful when the signal-to-noise ratio (SNR) is low due to the small performance difference depending on the power ratio between the signals. Existing methods such as spectral subtraction [4] often fail to estimate the noise spectrum with low SNR and damage the original speech. This paper deals with noise cancellation using ICA-based signal separation algorithm. In Section II, the noise suppression technique using ICA algorithm is described. Result of speech recognition experiments are given in Section III. Finally, we conclude this paper in Section IV.

## II. NOISE SUPPRESSION BY INDEPENDENT COMPONENT ANALYSIS

The ICA technique is a statistical method that separates the mixtures from two or more sensors into independent components. In order to explain the signal separation process by ICA, first, a model in which signals are mixed is as follows. ICA suppose that each source signal is mutually independent and the mean value is 0, and the signal input to each microphone is a mixture of the original signals.

The signal separation method using ICA technology was proposed by Bell and Sejnowski in 1995 as an information theoretical approach using mutual information [2]. In 1996, Torkkola proposed a time-domain feedback structure Algorithm [3]. This paper is based on a time-domain feedback algorithm which showed the best performance by its own preliminary study on the performance of various ICA-based signal separation algorithms.

In this study, two microphones were used and this is the case in Eq. (1). The stochastic gradient ascent rule was used for the learning rule. Figure 1 shows an example of the noise attenuation using ICA. As shown in the figure, the SNR is greatly improved in the speech interval.

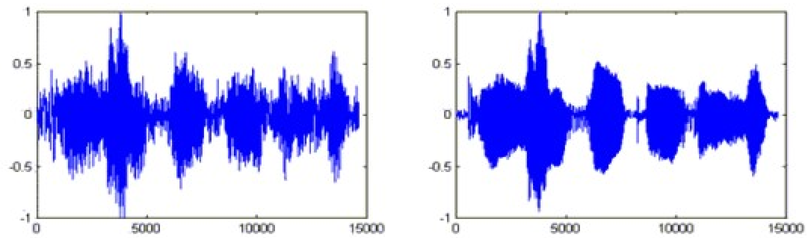


Figure 1. An example of noise suppression. (a) noisy speech signal, (b) ICA result.

## III. Experiment and Result

The In this study, the signal separation experiments were carried out in a laboratory space of 3 m, 5 m, and 3 m in width, height, and height, respectively. Two microphones were used, the distance between the microphones was 60cm, and the straight line distance between the microphone and the signal source was 1m. The algorithm used is a time-domain feedback algorithm and the length of the filter is 128. ICA learning is up to 10 times and the learning rate is 0.0001. First, this experiment deals with speech recognition experiments to examine the signal separation performance by ICA technique.

As a speech recognition system, we used Korean consecutive digits in this experiment. The speaker uses 93 (60 males, 33 females) and 47 (30 males, 17 females) in the training. Each speaker uttered 40 digit strings, and each string of digits is randomly selected from 3 to 7. Sampling and quantization rates are 8 kHz and 16 bits. The feature vector for speech recognition is a total 26-order vector consisting of Mel-Frequency Cepstral Coefficients (MFCC), 12th-order delta MFCC, energy and delta energy. The recognition unit is a word - based unit of numbers and a continuous distribution hidden Markov model (HMM) is used and the output distribution is modeled by 4 mixtures for each HMM state. The noise data used in the noise environment simulation are three kinds of noise, 100 km/h of vehicle traveling noise, white Gaussian noise, and babble noise. Recognition experiments were performed for speech and noise mixed with SNR of 0dB. The mixing matrix for mixing 0dB based on the power of each sound and noise is obtained.

In the table, ‘None’ indicates the case where the mixed signal is used as it is, ‘SE’ indicates the case where the spectral enhancement method is applied to the mixed signal, and ‘ICA’ indicates the case where the ICA algorithm is applied to the mixed signal. The recognition result is 92.67% in the noisy environment. As shown in Table 1, ICA is more effective than SE in noise environments.

Table 1. Result of speech recognition experiments (%)

| **Noise** | **None** | **SE** | **ICA** |
|---|---|---|---|
| **100 km/h Car** | **53.49** | **70.95** | **87.73** |
| **White Gaussian** | **12.29** | **24.66** | **66.05** |
| **Babble** | **21.20** | **34.63** | **84.38** |
| **Average** | **28.99** | **43.41** | **79.39** |

## IV. Conclusion

In this paper, we describe the results of the noise reduction technique using the ICA-based signal separation algorithm to overcome the limitation of the existing method of inputting one channel for speech recognition in a noisy environment. As a result of speech recognition experiment, we confirmed the superiority of ICA technology in noise environments.

## V. ACKNOWLEDGEMENTS

This study was supported by the Research Program funded by the SeoulTech(Seoul National University of Science and Technology).

## REFERENCES

[1] A. Hyvarinen, J. Karhunen, and O. Erkki, Independent Component Analysis, John Wiley & Sons, 2001.

[2] A. J. Bell, T. J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution", Neural Computation, Vol. 7, No. 6, pp. 1129~1159, 1995.

[3] K. Torkkola, "Blind separation of convolved sources based on information maximisation", Proc. IEEE Workshop on Neural Networks for Signal Processing, pp. 423~432, 1996.

[4] Y. Ephraim, D. Malah, "Speech Enhancement using a minimum mean-squre error short-time spectral amplitude estimator", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 32, pp. 1109~1121, 1984.