

멀티밴드필터에 의한 환경잡음억압 알고리즘

최재승*

Reduction Algorithm of Environmental Noise by Multi-band Filter

Jae-Seung Choi *

요약

본 논문에서는 각 프레임에서의 음성신호 및 비음성신호 구간을 검출하는 음성인식 알고리즘을 제안한다. 그리고 음성신호 및 비음성신호 구간의 검출에 따라서 각 프레임에서 잡음을 제거하는 멀티밴드필터에 의한 환경잡음억압 알고리즘을 제안한다. 이 알고리즘은 음성으로부터 특징 파라미터를 추출하여 필터뱅크의 서브밴드 영역에서 잡음을 제거하는 방법이다. 본 실험에서는 환경잡음억압 알고리즘의 성능을 멀티밴드필터를 사용하여 각 프레임에서 잡음을 제거하는 실험결과를 나타낸다. 잡음에 의하여 오염된 음성에 대하여 스펙트럴 왜곡률을 사용하여 본 알고리즘이 유효하다는 것을 확인한다.

▶ Keywords : 음성인식 알고리즘, 잡음제거알고리즘, 멀티밴드필터, 환경잡음

Abstract

This paper first proposes the speech recognition algorithm by detection of the speech and noise sections at each frame, then proposes the reduction algorithm of environmental noise by multi-band filter which removes the background noises at each frame according to detection of the speech and noise sections. The proposed algorithm reduces the background noises using filter bank sub-band domain after extracting the features from the speech data. In this experiment, experimental results of the proposed noise reduction algorithm by the multi-band filter demonstrate using the speech and noise data, at each frame. Based on measuring the spectral distortion, experiments confirm that the proposed algorithm is effective for the speech by corrupted the noise.

▶ Keywords : Speech recognition algorithm, noise reduction algorithm, multi-band filter, environmental noise.

• 제1저자 : 최재승
• 투고일 : 2012. 2. 6, 심사일 : 2012. 6. 12, 게재확정일 : 2012. 6. 19.
* 신라대학교 전자공학과(Dept. of Electronic Engineering, Silla University)

I. 서론

최근의 실생활의 환경에서 존재하는 배경잡음이 사람이 발생하는 음성의 중요한 특징량 왜곡은 물론 음성 구간의 추출 정밀도를 열화시킴으로써 음성인식률을 저하시켜 인식 시스템의 성능을 나쁘게 하고 있다[1][2]. 이와 같이 다양한 배경잡음이 존재하는 실제 환경에서 음성인식을 구현하기 위해서는 지금까지의 음성인식 알고리즘에 대한 잡음성능의 강건성 향상이 필요하게 되고 있으며, 이에 따라서 근년 이러한 주제의 연구가 지속적으로 증가하고 있는 추세이다[3]. 또한 이러한 주제의 여러 연구 수법이 제안되고 있는데, 주로 스펙트럼 차감법(Spectral subtraction, SS)에 의한 잡음억압법[4], 특징 파라미터 검출방법에 의한 음성인식법[5] 등이 제안되고 있다.

실제 환경에서의 잡음으로는 회화 중에서 상대방이 발생한 음성 이외의 음성, 도로에서 들려오는 자동차잡음, 공항에서의 소음 등 다양한 배경잡음 등이 존재한다. 그러므로 음성인식 장치 및 잡음억제 장치의 성능을 향상시키기 위해서는 이러한 배경잡음 및 사람이 발생한 음성을 충분히 식별하여 실제 환경에서 배경잡음을 억제할 수 있는 시스템의 개발이 필요하다. 이러한 필요성에 의하여 본 논문에서는 전력스펙트럼에 의한 문턱치(Threshold value) 및 첵스트럼(Cepstrum)의 거리 방식을 이용하여 각 프레임에서 정확하게 음성신호(Speech signal) 및 비음성신호(Non-speech signal)를 분류함으로써 불안정한 잡음구간 및 비음성구간 등의 환경에서도 성능이 향상될 수 있는 것을 연구 배경으로 한다. 따라서 본 논문에서 제안하는 멀티밴드필터[6][7] 영역에서 환경잡음을 억압하는 것을 연구 목적으로 하여 이에 대한 유효성을 검증한다.

본 논문에서 제안하는 알고리즘은 시스템에 입력되는 입력 신호가 목적으로 하는 음성신호인지 아니면 음성 이외의 배경잡음이 포함된 비음성 신호인지를 식별하여, 특히 배경잡음을 충분히 억압할 수 있는 특징을 가지고 있다. 음성 및 비음성 신호를 식별하는 방법은 전력스펙트럼에 의한 문턱치 방법, 첵스트럼 거리 측정에 의한 방법 등에 의하여 음성 및 비음성 신호를 식별하는 방법을 제안한다. 그리고 잡음억압 처리방법은 멀티밴드필터(Multi-band filter)에 의한 잡음억압, 각 밴드에서 청각적인 특성을 다르게 적용함으로써 배경잡음 스펙트럼을 억압하는 방법 등을 주로 적용한다. 본 논문에서 제안하는 방법 들은 실제 환경 하에 존재하는 배경잡음의 혼입을 가정하여 음성인식 및 음성신호의 잡음감에 응용하는 것

을 목적으로 하며, 음성인식률의 개선 및 잡음억압의 실현을 검토한다.

본 논문은 먼저 제안하는 음성신호 혹은 비음성신호의 식별을 위한 처리방식의 개요를 기술한다. 그리고 잡음억압 알고리즘을 위한 기초적인 방법의 개요, 시스템 구성 등을 통하여 본 논문에서 제안한 알고리즘이 잡음에 강건하다는 것을 음성인식률 및 객관적인 품질평가 척도에 의하여 잡음억압에 효과적인 것을 실험을 통하여 확인한다.

II. 제안한 방식의 기본 개요

1. 음성신호 및 비음성신호의 인식

기존의 일부 음성인식 시스템에 사용되는 음성 및 비음성 신호를 인식하는 알고리즘은 비음성신호 구간을 함께 음성인식구간으로 인식하는 문제점을 가지고 있었다. 이것은 실제 환경에 존재하는 배경잡음이 음성신호의 스펙트럼의 특징과 거의 유사하기 때문에 서로 오인식을 일으킨다고 본다 [5][8]. 참고문헌 [5]와 [8]에서와 같은 방법은 음성 및 비음성의 분류를 어느 정도 개선시킬 수 있었으나 묵음 등이 포함된 불안정한 환경에서는 분별력이 정확하지 못하는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 첵스트럼 거리 척도 방식을 적용하여 음성 및 비음성구간을 정확히 분류함으로써 음성의 특징 파라미터를 각 구간에서 정확히 추출할 수 있다. 따라서 이러한 특징 파라미터를 사용하여 멀티밴드필터[9]에 적용함으로써 각 서브밴드에서 잡음을 제거하여 첵스트럼 특징을 향상시킬 수 있다. 따라서 본 논문에서는 음성신호 및 비음성신호의 각 특징을 정확하게 분석하여 특징 벡터를 추출한 후에 제안하는 여러 방법을 이용하여 음성 및 비음성신호 구간을 추출하는 처리를 한다.

본 논문에서는 시스템에 입력되는 음성 및 비음성신호를 인식하는 시스템을 구축하는 것을 목적으로 한다. 음성 및 비음성 신호의 식별 방법으로는 여러 방법이 제안되고 있지만, 본 논문에서는 입력신호의 주파수 특성을 이용하여, 전력스펙트럼에 의한 문턱치 방법, 첵스트럼 거리 측정에 의한 방법 등에 의하여 음성 및 비음성신호를 식별하는 방법을 제안한다. 먼저 음성 및 비음성 신호를 식별하기 위하여 각 프레임에서 FFT(Fast Fourier Transform)한 후의 전력 스펙트럼 및 전체 문장의 전력 스펙트럼을 구하여 식 (1)과 같이 음성 및 비음성 신호를 식별한다. 따라서 좌변의 식이 γ 보다 클 경우에는 음성신호로 식별되며, γ 보다 작을 때에는 비음성 신호

로 식별된다.

$$\frac{F_b}{F_a} \sum_{i=0}^{N-1} \left\{ \frac{|Y(\omega)|^2}{|N(\omega)|^2} \right\} > \gamma \dots\dots (1)$$

여기서 $Y(\omega)$ 및 $N(\omega)$ 는 각각 주파수 영역에서의 잡음이 중첩된 음성신호의 스펙트럼 및 잡음 스펙트럼을 나타내며, F_a 는 전 문장의 프레임수, F_b 는 각 문장의 처음의 잡음 구간을 나타내는 5프레임이다. γ 는 문턱값을 나타내며 본 실험에서는 0.5를 채택하였다. 음성 및 비음성 신호를 식별하기 위한 다른 방법은 입력되는 음성데이터를 cepstrum 영역에서의 각 음성데이터의 유클리드(euclid) 거리로 표현되는 식 (2)와 같은 cepstrum 거리식으로 측정한다.

$$dis_{j,*} = \sqrt{\sum_{i=1}^{Cep} [y_{j,i} - C_{*,i}]^2} \dots\dots (2)$$

$$C_{*,i} = \frac{1}{Cep} \sum_{j=1}^{Cep} y_{j,i} \dots\dots\dots (3)$$

여기에서 $y_{j,i}$ 는 잡음이 중첩된 음성신호의 cepstrum 계수, i 는 각 프레임에서의 cepstrum($i = 1, 2, \dots, Cep$)을 나타내며, 본 실험에서의 Cep 수는 14개로 결정하였다. 그리고 $C_{*,i}$ 는 cepstrum의 중심을 나타낸다. 따라서 식 (4)와 같이 cepstrum 거리가 일정한 문턱치 이상일 때에 음성신호로 판정하며, 문턱치 이하일 때에는 비음성신호로 판정한다. 본 실험에서는 문턱값 η 가 0.3 이상이면 음성신호로 판정한다.

$$dis_{j,*} > \eta \dots\dots\dots (4)$$

이상과 같이, 최종적으로 그림 1과 같이 식 (1)의 전력스펙트럼 방식 및 식 (2)의 cepstrum 거리 방식을 이용하여, 각 식에서 사용하는 문턱값 이상일 때에 한하여 음성신호로 판별하고, 문턱값 이하일 때에는 비음성신호로 판별하도록 하는 알고리즘으로 구성하였다.

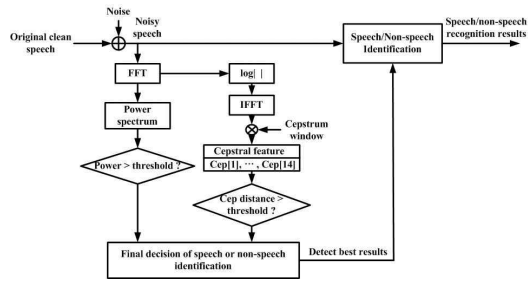


그림 1. 음성 및 비음성신호 인식 시스템
Fig. 1. Speech/non-speech identification system(SNIS).

표 1은 음성신호에 백색잡음이 중첩된 경우에 대하여 cepstrum 계수(Cep)를 10차, 12차, 14차로 하였을 경우의 cepstrum의 거리 $dis_{j,*}$ 를 나타낸다. 표 1로부터 $Cep=14$ 인 경우에 cepstrum 거리가 평균값으로부터 가장 떨어져 있어서 각 음성데이터 간의 중첩이 적은 것을 알 수 있다.

표 1. cepstrum 데이터간의 거리
Table 1. Distance of cepstrum data.

speech	Cepstrum distance		
	Cep=10	Cep=12	Cep=14
data1	0.396	0.398	0.400
data2	0.496	0.497	0.497
data3	0.412	0.413	0.415

2. 청각 모델

청각계에서 음성의 주파수분석에 대해서는 주파수성분이 각각의 청각 신경세포의 활성패턴으로서 인간의 두뇌에 전달된다는 연구들이 나오고 있다. 따라서 기저막이 멀티채널 필터로 구성된다는 이유로, 본 논문에서는 기저막과 청각 말초 신경계가 잡음처리 기강으로 동작하고 있다고 생각하여, 그림 2와 같은 청각모델(Auditory model)을 제안한다. 이 청각 모델은 측억제결합을 가지는 잡음 억압기강으로 알려져 있다. 측억제 모델은 각 프레임의 스펙트럼에 대하여 주파수축 방향에 흥분성 및 억제성의 특징을 조합하여 가중치 계수를 컨볼루션(Convolution)하여 스펙트럼을 강조하는 처리를 한다. 이 측억제 모델의 처리에 의한 잡음억제 및 인식성능을 향상시키는 연구의 보고가 나오고 있다[10]. 그림 2에서와 같이 1개의 흥분영역과 2개의 억제영역으로 분리하여 음성신호를 강조하고, 비음성 신호를 억압하는 기강이며, 본 논문에서는 각 밴드에서의 잡음 강도에 따라서 흥분성분 및 억제성분의 계수들을 각각 다르게 구성한다. 여기에서 P는 청각모델함수의 강도를 나타내며, B는 청각모델함수의 폭을 나타낸다. 특

히, B의 값이 적을수록 각각의 필터의 주파수성분 능력이 높게 되며, 반대로 B의 값이 넓으면 필터의 주파수성분 능력이 낮게 되는 특성을 가진다.

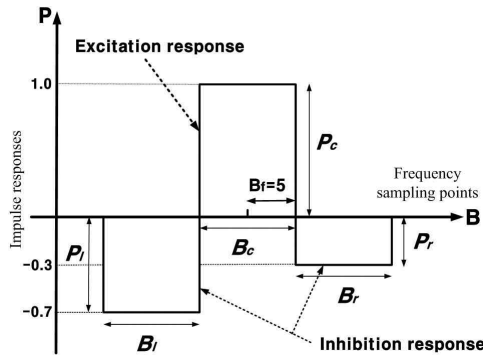


그림 2. 측억제모델의 결합함수
Fig. 2. The conjunction function of the lateral inhibition.

3. 멀티밴드필터 설계

최근에 Male Scale 서브밴드 등 여러 필터뱅크에 의한 방법 들[11][12]이 제안되고 있다. 이러한 필터뱅크는 일정한 주파수 대역폭에 있어서 주파수 성분을 다르게 할당하여도 사람들은 이러한 차이점을 확실하게 구분하지 못한다. 따라서 이러한 특성을 응용하여 공학분야에서 적용하게 되는 방법이 멀티밴드필터라고 할 수 있다. 특히 일반인들과 다르게 난청 인들은 개인별로 청력손실이 많으며 또한 청각능력이 각 주파수 대역별로 다르게 나타나므로 기존의 청각필터로는 이러한 다양한 난청인의 청력손실을 보상해주지 못하는 문제점을 가지고 있다[10][11]. 따라서 본 논문에서는 음성의 입력 크기에 따라 중심주파수를 다르게 적용하여 주파수 밴드별로 따라서 서로 다른 청각필터 형태를 가진 4개의 멀티밴드 필터를 구성한다. 본 논문은 4 kHz의 주파수 대역 중에서 저역에 해당하는 0~2 kHz를 2개의 대역으로, 고역에 해당하는 2~4 kHz를 2개의 대역으로 세분화하여 2.2절에서 기술한 그림 2의 측억제모델의 주파수 능력을 극대화하고자 하는 목적의 연구이다.

본 논문에서는 이러한 특성을 이용하여 멀티필터뱅크에 의하여 잡음환경에 적응시켜 잡음에 강인한 환경잡음 억압 알고리즘을 제안한다. 본 논문에서는 4 kHz의 주파수 대역 (8kHz의 샘플링 주파수)에 대하여 표 2와 같이 전체 주파수 대역을 4개의 다중 대역으로 나눈다.

표 2. 다중 서브밴드의 주파수 범위
Table 2. Frequency range of multi sub-band.

다중 서브밴드	주파수 대역 범위(Hz)
Band 1	0 Hz ~ 1 kHz
Band 2	1 kHz ~ 2 kHz
Band 3	2 kHz ~ 3 kHz
Band 4	3 kHz ~ 4 kHz

III. 제안한 환경잡음 억압 알고리즘

본 논문에서 제안한 기본적인 내용에 기초한 환경잡음 억압 알고리즘의 블록도를 그림 3에 나타낸다.

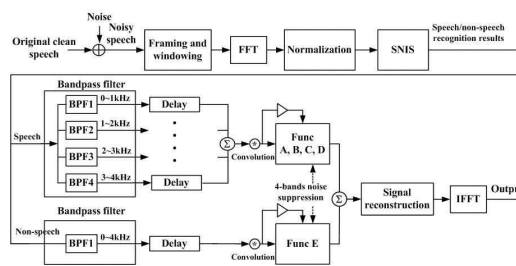


그림 3. 제안한 환경잡음억압 알고리즘
Fig. 3. Proposed noise suppression algorithm.

시스템 입력으로 들어오는 음성 및 잡음신호는 8 kHz의 샘플링 주파수를 각 프레임에서 256 샘플로 프레임링(Framing)한다. 잡음이 중첩된 음성신호는 해밍창(Hamming window)을 통과한 후에 FFT한다. 제안한 시스템에 입력된 음성신호는 FFT에 의해 4 kHz까지의 진폭 스펙트럼을 사용하게 되고 위상 성분은 IFFT(Inverse Fast Fourier Transform)할 때에 사용된다. 이 후에 정규화 과정을 통하여 평균적인 신호로 변환된 후에 그림 1에서 제안한 음성 및 비음성신호 인식 알고리즘에 의하여 음성 및 비음성신호로 분리된다. FFT에 의해 추정된 스펙트럼은 음성 및 비음성신호 구간에 대하여 각각 표 2의 다중 서브밴드 주파수 범위로 나누어진다. 이 후에 음성신호 구간에서는 그림 2의 측억제 모델의 결합함수의 흥분영역을 강조한 각기 다른 결합함수를 적용하고, 비음성 구간에서는 측억제 모델의 억제영역을 확장한 결합함수를 이용하여 가중치가 적용된 잡음억제 방법을 사용한다. 따라서 음성신호 구간에서는 그림 2의 측억제 모델 결합함수의 흥분영역 계수 Pc를 강조한 각각 다른 함수들(Func A, B, C, D)이 사용되고, 비음성신호 구간에서는 그림 2의 측억제모델 결합함수의 억제영역 계수 Pi 및 Pr를 강조함 함수(Func E)가 사용됨으로써 표 2의 다중 서브밴드의 주파수 범위에 따라서 각각 다르게 적용된다. 마지막으로

음성 및 비음성신호를 결합하여 진폭 및 위상성분을 사용한 IFFT에 의하여 잡음이 억제된 출력신호를 구하게 된다.

IV. 실험 및 결과

1. 실험 환경

본 논문에서는 음성 및 비음성신호의 구간 인식 및 잡음억제 알고리즘 구현을 위하여, 제안한 알고리즘을 이용하여 실험을 한다. 실험환경은 Intel(R) Core(TM)2 Quad 2.4GHz CPU와 3.25GB RAM이 장착된 IBM 호환 컴퓨터에서 Windows XP OS를 사용하여 Microsoft VC++ 6.0으로 구현하였다.

본 실험에서 사용한 음성신호는 8 kHz의 샘플링 주파수를 가진 환경에서 녹음된 음성 데이터를 사용하였으며, 일본인 남성화자에 의한 단어를 사용하였다. 본 논문에서 사용한 음성 데이터로서는 "fukui(data1)", "aioi(data2)", "hachioji(data3)"를 사용하였다. 또한 본 실험에서 사용한 잡음데이터는 컴퓨터에 의해서 작성된 가우스 백색잡음(white noise)의 배경잡음을 사용하여 평가하였다. 본 실험에서 사용한 음성데이터는 약 11 프레임에서 52 프레임 정도를 나누워지며, 1 프레임은 256 샘플로 구성된다. 본 논문에서 제안한 잡음제거의 평가방법으로는 음성 명료도와 관계가 깊은 스펙트럴 왜곡률(Spectral distortion, SD)를 사용한다. 이 SD는 각 주파수 대역에서 입력신호의 대수 스펙트럴(dB)와 출력신호의 대수 스펙트럴(dB)의 차를 측정구간의 프레임 수에 대하여 구한 것으로 식 (5)와 같이 정의한다.

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{W} \int_0^w S_x(w) - S_y(w)^2 dw} \dots\dots (5)$$

단, $S_x(w)$, $S_y(w)$ 는 입출력 신호의 대수 스펙트럼(dB)이며, N 은 측정구간의 프레임 수, W 는 신호의 대역폭이다. 또한, SNR_{input} 으로서는 식 (6)과 같이 음성 $S(n)$ 과 잡음 $N(n)$ 의 전체에 해당하는 전력의 비율로서 정의되는 전역 SNR_{input} 을 사용하였다. 여기에서, N 은 음성데이터의 샘플수이다.

$$SNR_{input} = 10 \cdot \log_{10} \left(\frac{\sum_{n=1}^N S(n)^2}{\sum_{n=1}^N N(n)^2} \right) \dots\dots\dots (6)$$

2. 실험결과

제안한 시스템은 4.1절에서 기술한 일본인 음성데이터 및 백색잡음을 사용하여 인식을 및 잡음제거의 실험을 실시하였다.

표 3은 음성신호 및 백색잡음에 대하여 식 (1)에 의한 캡스트럼 거리 측정 방법 및 식 (2)에 의한 전력 스펙트럼의 문턱치 방법을 사용하여 음성과 비음성신호를 식별하는 식별률의 실험 결과를 나타낸다. 표 4는 음성신호 및 백색잡음을 대하여 그림 3에서 제안한 환경잡음억압 알고리즘의 블록 내에 있는 4개의 멀티밴드 필터(BPF1, BPF2, BPF3, BPF4)에 대한 SD에 의한 실험결과를 나타낸다. 표의 결과는 음성 식별된 신호에 대하여 각 밴드별 출력 결과에 측정제함수 A, B, C, D를 각각 컨벌루션하였으며, 제안한 알고리즘이 입력 SD와 비교하였을 때 data1에 대해서는 평균 2.96dB, data2에 대해서는 평균 2.87dB, data3에 대해서는 평균 6.28dB이 각각 개선되었다. 따라서 본 논문에서 제안한 멀티밴드 필터에 의한 알고리즘이 표 5에서 나타난 바와 같이 전체적으로 SD값이 향상됨을 알 수 있다. 표 5는 여러 잡음레벨(SNRin=20dB, 15dB, 10dB)에 대하여 학습에 사용한 음성신호 및 백색잡음을 사용하여, 10개의 잡음이 중첩된 음성데이터에 대하여 평가한 경우의 SD에 의한 잡음제거 실험이다. 본 논문에서 제안한 알고리즘(표의 Prop.)은 참고문헌 [4]의 기존의 스펙트럼 차감법(SS algorithm)과 본 논문에서 제안한 방법을 출력 SD를 사용하여 비교 분석하였다. 표 5의 결과로부터 제안한 알고리즘이 입력 SD와 비교하였을 때에는 최대 6.37dB 정도(Input SD=21.58dB 일 때), SS algorithm과 비교하여 최대 7.39 dB 정도(Input SD=12.61dB 일 때) 향상되는 것을 확인할 수 있었다.

표 3. 음성 및 비음성 신호의 식별률(%)
Table 3. Discrimination rates for speech/non-speech data.

Data	음성 및 비음성신호의 식별률(%)	
	캡스트럼 거리 측정 방법	전력스펙트럼에 의한 문턱치 방법
data1	93.5%	90.2%
data2	94.7%	91.4%
data3	95.1%	92.6%

표 4. 백색잡음인 경우의 멀티밴드 필터별 SD 실험 결과
Table 4. The experimental results for each multiband filter in the case of white noise.

Data	Input SD[dB]	Output SD [dB]			
		BPF1	BPF2	BPF3	BPF4
data1	12.61	9.55	9.59	9.67	9.78
data2	17.79	14.73	14.82	14.95	15.17
data3	21.58	15.11	15.19	15.35	15.53

표 5. 백색잡음인 경우의 SD 실험 결과
Table 5. The experimental results in the case of white noise.

Data	Input SD[dB]	Output SD [dB]		
		SS	Prop.	Impr.
data1	12.61	17.00	9.61	7.39
data2	17.79	16.25	14.87	1.38
data3	21.58	15.99	15.21	0.78

지금까지의 실험결과로부터, 음성 및 비음성신호 구간 인식 시스템에 의한 잡음제거를 각 프레임에서 실시함으로써 잡음이 많이 중첩된 음성에 대해서도 SD의 값을 경감하는 것을 알 수 있었다. 이상의 결과로부터 입력 SNR이 12dB까지의 잡음이 중첩된 음성신호에 대해서도 본 방식에 의한 잡음제거의 효과가 있는 것을 알 수 있었다.

IV. 결 론

본 논문에서는 음성신호 구간 및 비음성신호 구간에 대하여 각 프레임에서의 구간 검출에 대한 검출 시스템을 제안하였다. 또한 본 논문에서는 음성 및 비음성 구간의 인식시스템에 의하여 음성 및 잡음 구간의 검출에 따라서 각 프레임에서 잡음을 제거하는 환경잡음 억압 알고리즘을 제안하여, 다른 기존의 방법들과 비교하였다. SNR_{input}이 12dB, 17dB, 21dB로 입력된 음성 및 잡음이 중첩된 음성에 대해서 본 알고리즘이 유효하다는 것을 실험에서 확인하였다.

이상으로 잡음이 중첩된 음성신호 중에서 잡음을 제거하기 위해서 인식시스템 및 환경잡음 억압알고리즘이 백색잡음에 대해서 유효하다는 것을 알 수 있었다. 향후의 연구과제로는 여러 가지 비정상적인 잡음을 사용하여 인식시스템의 향상 및 잡음제거를 검토할 필요가 있다고 본다. 본 연구를 더욱더 발

전시킴으로써 청각실험에 도움이 된다고 기대한다.

참고문헌

- [1] J. P. Haton, "Automatic recognition of noisy speech," Springer Verlag Heidelberg, Speech Recognition and Coding-New Advances and Trends, Vol. 147, Part I, pp. 3-13, 1995.
- [2] C. Y. Lee and Y. G. Lee, "Study of the noise processing to technique speech recognition system", Journal of the Korea Society of Computer & Information, Vol. 7, No. 2, pp. 73-78, June, 2002.
- [3] K. H. Lee, "Speaker adaptation using linear transformation network in speech recognition", Journal of the Korea Society of Computer & Information, Vol. 5, No. 2, pp. 90-97, June, 2000.
- [4] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Processing. Vol. 27, No. 2, pp. 113-120, 1979.
- [5] Y. Wu and Y. Li, "Robust speech/non-speech detection in adverse conditions using the fuzzy polarity correlation method," IEEE International Conference on Systems, Man, and Cybernetics, Vol. 4, pp. 2935-2939, 2000.
- [6] J. Chen, K. K. Paliwal and S. Nakamura, "Sub-band based additive noise removal for robust speech recognition," in Proc. Eurospeech, pp. 571-574, 2001.
- [7] L. Lin, W. H. Holmes and E. Ambikairajah, "Subband noise estimation for speech enhancement using a perceptual Wiener filter," IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, pp. I-80 - I-80-3, 2003.
- [8] A. Ishida and H. Gobata, "Speech/non-speech discrimination under real life environments," The Journal of the Acoustical Society of Japan, Vol. 47, No. 12, pp. 911-917, 1991 (in Japanese).
- [9] P. Chaisakul, N. Tangsangiumvisai and P.

- Luangpitakchumpon, "Number-of-band Selection for Acoustic Feedback Cancellation in Multi-band Compression Hearing Aids," IEEE Region 10 Conference, pp. 1-4, 2006.
- [10] J. H. L. Hansen and S. Nandkumar, "Robust estimation of speech in noisy backgrounds based on aspects of the auditory process", The Journal of the Acoustical Society of America, Vol. 97, No. 6, pp. 3833-3849, 1995.
- [11] T. Etame, R. Le Bouquin Jeannes, C. Quinquis, L. Gros and G. Faucon, "Towards a new reference impairment system in the subjective evaluation of speech codecs", IEEE Transactions on Audio, Speech, and Language Processing, Vol. 19, No. 5, pp. 1301 - 1315, 2011.
- [12] L. Yi-Lin and W. Gang, "Speech emotion recognition based on HMM and SVM", Proceedings of 2005 International Conference on Machine Learning and Cybernetics, Vol. 8, pp. 4898 - 4901, 2005.

저자 소개



최재승

1989년 조선대학교 전자공학과 공학사
 1995년 일본 오사카시립대학 전자정보공학부 공학석사
 1999년 일본 오사카시립대학 전자정보공학부 공학박사
 2000년~2001년 일본 마쯔시다 전기산업주식회사 (현, 파나소닉 주식회사) AVC사 연구원
 2002년~2007 경북대 디지털기술연구소 책임연구원
 2007년~현재 신라대학교 전자공학과 교수

※ 관심분야 : 음성신호처리, 신경회로망, 적응필터와 잡음제거처리 등

Email : jschoi@silla.ac.kr

