

# 차량에서의 음성인식을 향상을 위한 전처리 기법

## Preprocessing Technique for Improvement of Speech Recognition in a Car

김현태\*, 박장식\*\*

동의대학교 멀티미디어공학과\*, 동의과학대학 디지털정보전자과\*\*

Hyun-Tae Kim(htaekim@deu.ac.kr)\*, Jang-Sik Park(jsipark@dit.ac.kr)\*\*

### 요약

본 논문에서는 차량에서의 자동 음성인식 시스템과 같이 신호대잡음비가 낮은 잡음 환경에서의 음성 인식에 적합한 변형된 스펙트럼 차감법을 제안한다. 기존의 스펙트럼 차감법은 스펙트럼에서 낮은 신호대 잡음비(SNR)를 갖는 부분은 감쇄되고, 신호대잡음비가 높은 부분은 강조되는 신호대잡음비에 의존한다. 그러나 이러한 구성은 높은 신호대잡음비를 갖는 환경에서는 적절하나 차량 환경과 같이 낮은 신호대잡음비를 나타내는 환경에서는 매우 부적절하다. 제안하는 방법은 낮은 신호대잡음비를 갖는 잡음 환경을 위해 음성우세영역을 강조하여 불필요하게 음성영역이 과차감되지 않도록 방지한다. 차량용 음성명령어 어휘를 대상으로 한 실험 결과에서 제안하는 방법이 기존의 방법에 비해 우수한 것을 확인하였다.

■ 중심어 : | 스펙트럼 차감법 | 음성우세영역 | 고조파 강조 | 가중치함수 |

### Abstract

This paper addresses a modified spectral subtraction schemes which is suitable to speech recognition under low signal-to-noise ratio (SNR) noisy environment such as the automatic speech recognition (ASR) system in car. The conventional spectral subtraction schemes rely on the SNR such that attenuation is imposed on that part of the spectrum that appears to have low SNR, and accentuation is made on that part of high SNR. However, such postulation is adequate for high SNR environment, it is grossly inadequate for low SNR scenarios such as that of car environment. Proposed methods focused specifically to low SNR noisy environment by using weighting function for enhancing speech dominant region in speech spectrum. Experimental results by using voice commands for car show the superior performance of the proposed method over conventional methods.

■ keyword : | Spectral Subtraction | Speech Dominant Region | Harmonics Enhancement | Weighting Function |

## I. 서론

차량 내에서의 음성인식은 운전자의 안전과 편의성

증진을 위해 그 필요성이 대두되고 있다. 그러나 차량  
에서와 같이 주변잡음이 심각한 환경은 음성인식의 성  
능을 크게 저하시키는 요인이 되고 음성인식에서 혼련

\* " 본 논문은 2006년도 동의대학교 교내연구비에 의해 연구되었습니다(과제번호 : 2006AA185)"

접수번호 : #080827-001

접수일자 : 2008년 08월 27일

심사완료일 : 2008년 09월 30일

교신저자 : 김현태, e-mail : htaekim@deu.ac.kr

시 조건과 테스트시 조건이 서로 불일치하는 문제점을 낳는다. 따라서 이러한 환경에서의 음성인식을 위해서는 음성인식 과정 중에 잡음처리에 관한 처리 기술이 반드시 포함되어야 한다. 음성인식에서의 잡음처리 기술은 단계별로 음성향상(또는 음질개선으로도 사용됨), 특징보상, 모델적응 등으로 분류될 수 있으며, 이 중 음성향상은 잡음이 부가된 입력신호로부터 잡음의 영향을 줄이고 원래의 깨끗한 음성 성분을 추정하는 기술을 말한다[1]. 특징보상 기법은 음성의 품질을 직접적으로 향상시키는 것이 아니라 음성인식을 위하여 추출된 특징들을 개선시키는 것이다. 모델적용 기법은 인식 모델을 변환하여 마치 적용된 모델이 현재의 잡음이 섞인 음성으로부터 학습된 것처럼 하는 것이다. 일반적으로 음성인식에 사용되는 모델들은 잡음이 없는 깨끗한 음성 데이터베이스를 사용하여 학습하므로 음성 품질을 직접적으로 향상시키는 음성향상 기법이 음성인식은 물론 음성부호화 및 잡음이 있는 상황에서의 원활한 음성통신을 위하여 오래 전부터 널리 연구되어 왔다. 음성향상 기법에는 스펙트럼 향상, 칼만필터 기법 등이 있으나, 스펙트럼 향상 기법이 다른 기법에 비하여 쉽게 구현될 수 있고 다양한 잡음에 잘 대처한다는 장점을 지니기 때문에 가장 널리 사용되고 있다. 스펙트럼 향상 기법은 음성이 존재하지 않는 구간에서 잡음 스펙트럼을 추정한 후, 이를 잡음이 부가된 입력신호의 스펙트럼으로부터 차감하는 방식이다. 이 방식에는 스펙트럼 차감법, 비선형 스펙트럼 차감법, 그리고 Wiener 필터와 같은 단구간 스펙트럼 크기를 이용하여 잡음을 억제하는 스펙트럼 차감 형태의 알고리즘이 있다. 그러나 이러한 방법들은 모두 신호대잡음비라는 평범한 기준을 적용하기 때문에 상대적으로 잡음이 많은, 즉 낮은 신호대잡음비 환경에서 음성 영역이 과차감되는 문제점을 가진다. 그 외 고전적인 스펙트럼 차감법의 문제는 차감 과정에서 불필요하고 귀에 거슬리는 뮤지컬 노이즈(musical noise)가 발생하는 것이며, 이를 줄이기 위한 연구도 진행되고 있다[2][3]. 본 논문은 고전적인 스펙트럼 차감법의 문제점을 스펙트럼 상에서 음성 영역 별로 차등하여 차감하는 새로운 형태의 스펙트럼 차감 알고리즘이며, 잡음에 의해 음성이 많이 왜곡된

경우에도 음성의 고조파 스펙트럼의 형태를 어느 정도 유지할 수 있다[4][5].

스펙트럼 차감 형태의 알고리즘에 대해 2장에서 살펴보고, 제안하는 방법은 3장에서 기술한다. 실험 결과를 통해 제안하는 방법이 고전적인 방법에 비해 효과적임을 4장에서 보이고 결론은 5장에 나타내었다.

## II. 스펙트럼 차감 형태의 알고리즘

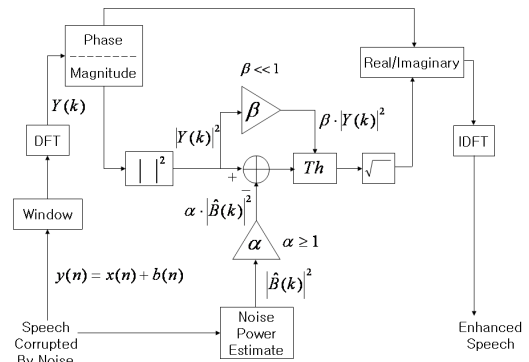


그림 1. 스펙트럼 잡음 차감법의 블록도

[그림 1]은 전형적인 스펙트럼 차감 알고리즘의 블록도이다. 음성 신호  $x(n)$ 는 배경 잡음  $b(n)$ 에 의해 오염될 때 오염된 음성은 다음과 같이 표현된다.

$$y(n) = x(n) + b(n) \quad (1)$$

음성과 잡음이 서로 상관이 없다고 가정하면 주파수 영역에서는 다음과 같이 표현된다.

$$|Y(k)|^2 = |X(k)|^2 + |B(k)|^2 \quad (2)$$

여기서  $k$  는 주파수 bin의 인덱스이다.

### 1. 스펙트럼 차감

전력 스펙트럼 차감의 경우, 잡음 저감된 음성 신호의 단구간 전력 스펙트럼  $|\hat{X}(k)|^2$ 은 오염된 음성으로

부터 잡음의 추정 전력  $|\hat{B}(k)|^2$ 을 차감하여 얻어진다. 모든 절차는 프레임 단위로 처리되어진다[6]. 전력 스펙트럼 대신 크기 스펙트럼으로도 사용가능하다. 일반적인 전력 스펙트럼 차감 알고리즘은 식 (3)과 같다.

$$|\hat{X}(k)|^2 = \begin{cases} |Y(k)|^2 \left(1 - \alpha \frac{|\hat{B}(k)|^2}{|Y(k)|^2}\right), \\ \text{if } |Y(k)|^2 - \alpha |\hat{B}(k)|^2 > \beta |Y(k)|^2. \\ \beta |Y(k)|^2, \\ \text{otherwise} \end{cases} \quad (3)$$

식 (3)에서 과차감 요소  $\alpha$  와 기저 요소  $\beta$  를 가지는 이 알고리즘은 잡음 저감과 잔류 잡음 사이에 trade-off 를 가진다. 식 (3)의 첫 번째 식에서 괄호 성분은  $G(k) = (1 - \alpha |\hat{B}(k)|^2 / |Y(k)|^2)$  로 두면, 이 부분은 일종의 전력 감쇄 성분으로 전력이 음수를 가질 수 없으므로 보통  $0 \leq G(k) \leq 1$  의 범위를 가지며, 과차감 요소  $\alpha$ 와 식 (4)로 주어진 후위(posteriori) SNR 값에 따라 감쇄 정도가 결정된다[7].

$$SNR_p = |Y(k)|^2 / |\hat{B}(k)|^2 \quad (4)$$

## 2. 비선형 스펙트럼 차감

비선형 스펙트럼 차감법은 기존의 스펙트럼 차감법이 프레임 단위로 처리되기 때문에 각각의 주파수 성분마다 상이한 신호대잡음비를 반영할 수 없는 문제점을 개선한 것으로 식 (5)와 같이 주어진다[8].

$$|\hat{X}(k)|^2 = \begin{cases} |Y(k)|^2 \left(1 - \frac{\Phi(k)}{|Y(k)|^2}\right), \\ \text{if } |Y(k)|^2 - \Phi(k) > \beta |Y(k)|^2. \\ \beta |Y(k)|^2, \\ \text{otherwise} \end{cases} \quad (5)$$

식 (5)에서 함수  $\Phi(k)$ 는 각 주파수 성분마다 신호대 잡음비를 조사하여 주파수 단위로 낮은 SNR 영역에는

상대적으로 큰 차감을 적용하고 높은 SNR 영역에는 보다 적은 차감의 개념을 구현할 수 있는 함수이면 임의로 선택가능하다.

일반적으로 음성 구성 성분, 특히 스펙트럼 상의 고조파의 크기 스펙트럼은 잡음 또는 고조파 사이의 부엽(side lobe)의 그것보다 크므로 이러한 알고리즘은 상당히 큰 SNR (약 15~20dB)을 갖는 잡음 환경의 음성인 경우에 적합하다. 그러나 낮은 SNR(약 0~10dB)을 갖는 환경에서는 잡음의 크기가 음성의 스펙트럼 크기에 근접하여 불필요하게 차감된 음성 영역이나 보다 적게 차감된 잡음 영역이 발생한다. 이러한 문제점을 스펙트럼에 가중치를 부여하여 개선하는 방법도 제안되었다 [9]. 본 논문에서는 음성 우세 영역을 찾아 영역별 차감 정도를 차등하는 스펙트럼 차감법을 제안한다.

## III. 제안하는 알고리즘

### 1. 음성 우세 영역 분류

음성에서 유성음 구간은 성대의 진동에 따라 스펙트럼 상에서 주기적인 피크 값을 가진다. 이러한 피크 점들은 음성인식에 있어서 매우 중요한 부분을 차지한다. [그림 2]는 이러한 피크 점들을 볼 수 있는 예이며, 잡음이 없는 환경(실선)과 10dB의 잡음이 부가된 환경(점선)에서 여성이 /oh/ 에 대해 발음한 경우의 스펙트럼이다[9]. 이러한 피크 점들을 살펴보면, 이 피크 점들의 크기가 이 점들 사이의 인접한 작은 피크들 또는 부엽들보다 훨씬 크다는 것을 알 수 있다. 또한 잡음에 의해 훼손된 정도가 피크점들 주변이 부엽 주변의 점들보다 훨씬 적다는 사실이다. 이러한 관측으로부터 스펙트럼 차감 방법 적용 시, 잡음 스펙트럼을 음성 우세 영역에는 보다 적게 차감하고 잡음 우세 영역에는 보다 많이 차감하여 음성 스펙트럼이 불필요하게 차감되어 음성 인식의 성능이 저감되는 것을 방지하는 간단한 방법을 제안한다.

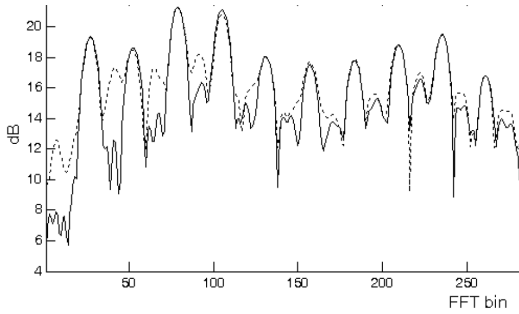


그림 2. 음성 구간의 스펙트럼 예(여성 음성 /oh/)

## 2. 음성 우세 영역 지정

기본 주파수 성분은 대략 자기상관으로 구할 수 있으며, 이것으로부터 피크 점들을 찾을 수 있다. 스펙트럼 상의 고조파는 기본 주파수의 배수가 되는 주파수에 위치한다고 알려져 있다[10]. 먼저, 자기상관 함수를 이용하여 국소 최대값들의 인덱스를 구한다. 이들 값들 중에서 최대값을 나타내는 인덱스의 역수를 취하여 기본 주파수  $f_0$  를 구할 수 있다. 그러나 충분한 주파수 해상도를 가지지 못하면 보다 정확한  $f_0$  값을 얻지 못하므로 충분한 길이의 주파수 해상도를 확보한다. 다음으로 유성음 구간내의 연속 프레임간 자기 상관의 평균 값을 구하여 보다 정확한 피치 정보로부터 실제 기본 주파수에 가깝도록  $f_0$ 를 구한다[9]. 자기상관 함수는 식 (6)과 같이 표현된다[11].

$$\Phi(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+\tau) \quad (6)$$

이렇게 구한 기본 주파수  $f_0$ 에 일치하는 주파수 bin의 인덱스가  $k_0$ 일 때,  $h_1 = k_0$ , 첫 번째 고조파는  $h_2 = 2k_0, \dots$ , 마지막 고조파는  $h_L = Lk_0$ 로 두며 기본 주파수를 포함하여 모두  $L$  개의 고조파에 대해 [그림 3]과 같이 피크 점으로부터 좌우 각각 약 6dB 정도 떨어지는 주파수까지를 음성 우세 영역으로 분류한다.

남성과 여성의 기본 주파수 범위는 다소 차이가 있으므로 기본 주파수 값의 범위에 따라 피크 점을 기준으

로 한 좌우 6dB 구간 대역폭을 달리 적용한다.

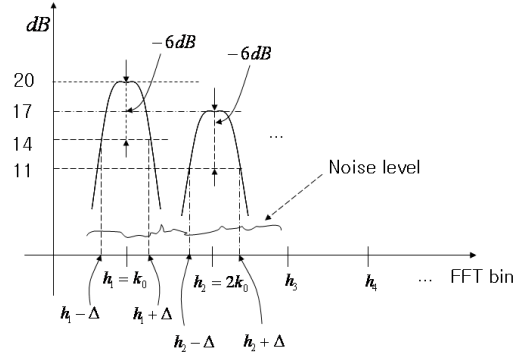


그림 3. 음성 우세 영역 분류 방법

## 3. 음성우세영역을 고려한 가중치 함수

기존 스펙트럼 차감법의 불필요한 음성 영역 차감을 방지하기 위해 앞 절에서 구한 음성우세영역을 고려한 가중치 함수를 구한다. 이때 적용하는 가중치 함수는 음성우세영역에서는 낮은 값을 가지며 잡음우세영역에서는 기존의 스펙트럼 차감법에서 적용하는 정도의 잡음이 차감될 수 있도록 구성한다. 제안하는 가중치 함수의 예는 식 (7)과 같다.

$$\Phi(k) = \begin{cases} 1 - 0.25 \left( 1 - \cos \left( \frac{2\pi k}{2\Delta} \right) \right), & \text{if } h_i - \Delta \leq k \leq h_i + \Delta \\ & \text{for } i = 1, \dots, L \\ 1 & \text{otherwise} \end{cases} \quad (7)$$

식 (7)의 가중치 함수를 살펴보면, 음성 스펙트럼의 각 피크 값에서 0.5의 값을 가지며 대역폭의 끝부분으로 가면 1의 값을 가지는 것을 알 수 있다. 제안하는 가중치 함수를 스펙트럼 영역에 적용한 예는 [그림 4]와 같으며, 가중치 함수는 입력 음성의 기본 주파수에 따라 스펙트럼 상에서 함수 폭과 위치가 변한다. 가중치 함수를 특정 입력 음성에 적용한 예는 [그림 5]와 같다. [그림 5]에서처럼 음성의 기본 주파수에 따라 피크 점인 고조파들의 주변 영역인 음성 우세 영역에서 잡음 스펙트럼을 적절히 억제할 수 있을 것으로 기대된다.

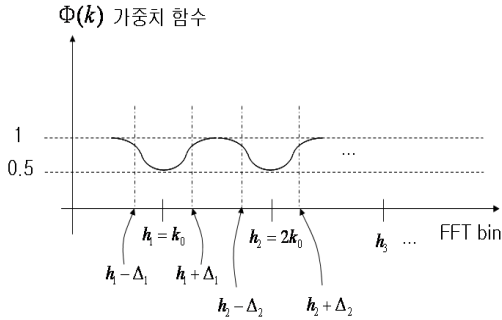


그림 4. 가중치 함수를 잡음 스펙트럼에 적용한 예

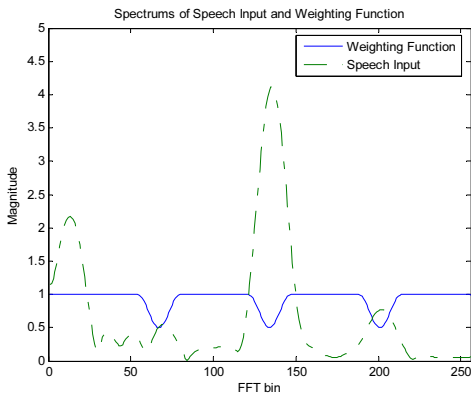


그림 5. 특정 입력 음성(여성, /Ye/)의 스펙트럼에 대응하는 가중치 함수

이상에서 언급한 제안하는 알고리즘을 포함한 변형된 스펙트럼 차감법에 대한 전체 블록도는 [그림 6]과 같다.

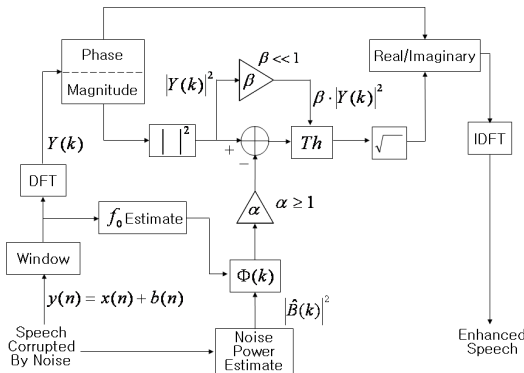


그림 6. 제안하는 스펙트럼 차감법의 블록도

## IV. 실험 결과

### 1. 실험 조건

제안하는 알고리즘은 성격 상 음성인식의 전처리 과정에 해당하나, 화자독립, 화자 종속 등 음성인식 본 과정에서 사용하는 다양한 인식 알고리즘에 따라 인식이 달라질 수 있어 알고리즘에 대한 평가는 전처리 과정 자체에 대한 보편타당한 평가 기준을 따른다. 이에 따라 알고리즘의 성능은 잡음이 부가되어 오염된 음성 신호를 제안하는 방법 및 기존 방법으로 잡음 처리한 후의 음성신호가 오염되기 전 원래의 깨끗한 음성 신호와 얼마나 유사한 지에 대해 각각 평가하고자 한다. 이를 위해 원 음성 신호에서 추출한 음성 특징 벡터에 대해 잡음 처리한 후의 음성신호에서 추출한 음성 특징 벡터간의 유클리디안 거리를 측정하여 알고리즘의 성능을 평가한다. 유클리디안 거리는 식 (8)과 같다.

$$d_{ED,i} = \sum_{j=1}^P (x_i(j) - y_i(j))^2 \quad (8)$$

식 (8)에서 j 는 특징 벡터의 차수이며, i 는 특정 음성 구간을 의미하고 x 와 y 는 원 입력 음성 특징 벡터와 잡음 제거 후 음성의 특징 벡터를 각각 의미한다. 식 (8)의 유클리디안 거리가 적을수록 원 음성의 특징 벡터와 유사해져 음성인식률의 향상을 기대할 수 있다. 각 음성에서 추출한 LPC(linear prediction coefficients) 계수 및 캡스트랄 계수에 대해 평가한다[12]. 각 계수의 차수는 10 차로 둔다. 실험은 차량용 네비게이션 제조사로부터 제공 받은 음성 D/B 에서 발췌한 차량용 음성 명령어를 대상으로 하였으며, 목록은 [표 1]과 같다. 무향실에서 남녀 각각 10명이 발음한 음성이며, 원 음성 데이터의 샘플링 주파수는 24kHz 이며, 다운 샘플을 통해 8kHz로 줄여 사용하였다. 입력 음성에 대한 프레임 길이는 25ms 이며 12.5ms 씩 중첩한다. FFT 포인트 수는 기본 주파수 및 그 고조파에 대한 추정 오차를 줄이고 스펙트럼 해상도를 높이기 위해 1024로 두었다. 기존의 스펙트럼 차감법과 비교하기 위해 알고리즘의 계수  $\alpha = 5, \beta = 0.05$  는 모두 동일하게 두었으며 제안하

는 알고리즘의 주엽으로 분류하는 단축 대역폭  $\Delta$  는 기본 주파수가 남성과 여성에 따라 차이가 있어 100Hz 이하의 기본 주파수에서는 3개의 주파수 빈, 200Hz 이하는 6개, 400Hz 이하는 12개로 주파수에 비례하여 차등을 두었다. 음성 검출기는 자기상관 방법을 이용하였다[9].

실험에 사용한 잡음은 두 가지 종류이며 먼저, 대역 제한(3kHz 이내) 된 백색 잡음을 [표 2]에 언급한 SNR 비율에 따라 크기를 달리하여 원 음성에 더한 후 제안하는 방법 및 기존 방법에 적용하였으며 그 결과를 역시 [표 2]에 언급하였다. 이 잡음을 적용한 이유는 차량 실내와 같은 환경에서 유입되는 잡음은 대부분 3kHz 이내에 신호의 에너지가 분포하기 때문이며, 그 다음으로 [그림 7]에 나타낸 바와 같이 실제 차량에서 측정된 잡음을 사용하여 동일한 방법으로 적용하였으며 그 결과를 [표 3]에 언급하였다.

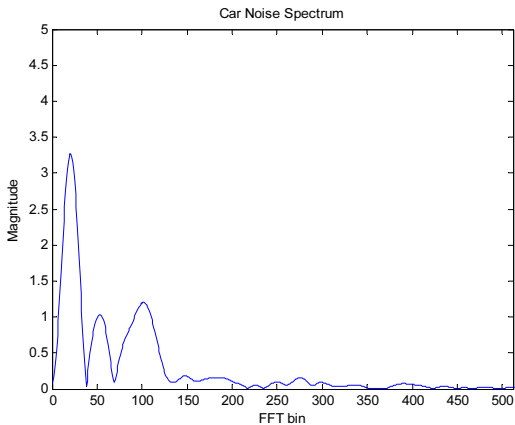


그림 7. 주행 중 차량 실내에서 측정된 잡음 스펙트럼

[그림 7]은 실제 차량에서 측정된 잡음의 스펙트럼이며, 텍트론사의 레코딩 장비인 RT Pro Dynamic Signal Analysis 를 이용하여 중형 승용차 실내 운전자 위 선 바이저 중앙에서 에어컨을 켜 상태로 시속 약 60 km 속도로 시내 주행 시 측정하였다.

표 1. 어휘 목록

차량용 음성 명령어(29 단어)				
디브이디	재생	창문 내려	이전 채널	창문 올려
중지	다음 채널	테이프	라디오	소리 크게
빨리 감기	에스비에스	에어콘 온	텔레비전	소리 작게
되감기	케이비에스원	에어콘 오프	실행	종료
케이비에스투	에어콘	약하게	자동 선국	엠비씨
에어콘 세계	밴드	채널 검색	와이프 멈춤	와이프 동작

2. 실험 결과 검토

[표 2] 및 [표 3]의 결과로부터 대역제한된 임의의 잡음 및 특정 차량 잡음에 대해 제안하는 방법이 기존의 스펙트럼 차감법보다 잡음 제거 성능이 우수한 것을 확인할 수 있다. 이는 제안하는 알고리즘이 음성 우세 영역에서 잡음에 의해 과차감되는 것을 어느 정도 줄일 수 있다는 것을 의미한다. 이러한 사실은 [그림 8]에서 제안하는 전처리 과정 전·후 스펙트럼의 비교를 통해서도 확인이 가능하다. 또한 SNR 이 높을수록 개선 폭이 커져, 보다 잡음이 많은 환경에서 더 효과적인 것도 알 수 있다. 전반적으로 유클리디안 거리가 다소 큰 것은 객관적인 비교를 위해 과차감 요소  $\alpha$  를 매우 큰 값인 5 로 두었기 때문이다. 특히 [표 3]에서 캡스트랄 계수의 유클리디안 거리가 대역 제한된 백색 가우스 잡음을 사용한 것 보다 큰 것은 [그림 7]에서와 같이 실제 차량 잡음의 스펙트럼이 저주파수 영역에서 매우 큰 값을 가져 원 신호 영역을 많이 훼손하였기 때문인 것으로 판단된다.

표 2. 유클리디안 평균거리(대역 제한된 백색 가우스 잡음 사용)

방법 \ SNR	기존 스펙트럼 차감법		제안하는 방법	
	LPC 계수	캡스트랄 계수	LPC 계수	캡스트랄 계수
clean speech	0.000	0.000	0.000	0.000
SNR=20 dB	0.4425	7.113	0.3186	5.867
SNR=15 dB	0.4493	7.449	0.4432	7.233
SNR=10 dB	1.2198	7.715	1.0051	7.553
SNR=5 dB	2.2614	19.105	1.9647	18.235
SNR=0 dB	3.1828	33.597	3.1016	29.312

표 3. 유클리디안 평균거리(실제 차량 잡음 사용)

SNR	기존 스펙트럼 차감법		제안하는 방법	
	LPC 계수	캡스트랄 계수	LPC 계수	캡스트랄 계수
clean speech	0.000	0.000	0.000	0.000
SNR=20 dB	0.362	17.895	0.281	15.624
SNR=15 dB	0.495	22.243	0.394	19.730
SNR=10 dB	0.661	27.547	0.537	25.137
SNR=5 dB	0.892	36.805	0.708	31.953
SNR=0 dB	1.247	54.077	0.932	41.957

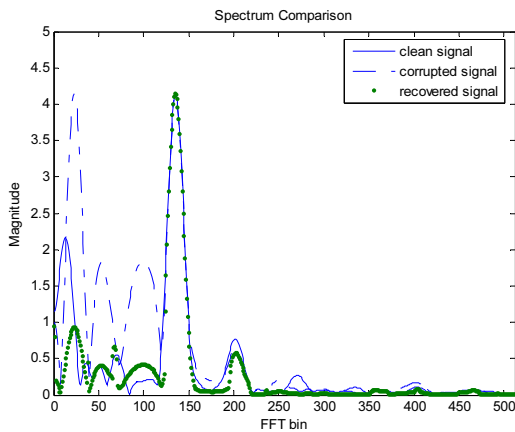


그림 8. 입력 신호(여성, /Ye/)에 대한 전처리 전후 스펙트럼 비교

[그림 8]은 특정 음성(여성, /Ye/)에 대해 원래의 깨끗한 음성과 이 음성에 실제 차량 잡음[그림 7]을 SNR 10dB 비율로 부가하여 오염된 음성, 그리고 제안하는 방법으로 잡음 저감시킨 후의 음성에 대해 각각 스펙트럼을 비교한 그림이다. [그림 8]을 통해 잡음의 에너지가 매우 큰 저주파 영역은 잡음 저감 후에도 원 스펙트럼 모양이 다소 훼손된 것을 확인할 수 있으며 그 외 음성 우세 영역은 상당히 복원이 되는 것을 알 수 있다.

### V. 결론

본 논문에서는 잡음 우세 영역으로부터 음성 우세 영역 스펙트럼이 과차감되는 것을 방지하기 위해 변형된

스펙트럼 차감법을 제안하였다. 대역 제한된 백색 가우스 잡음과 실제 차량에 유입되는 잡음을 사용한 실험을 통해 제안하는 방법이 잡음 환경에서 스펙트럼 상의 원음성의 고조파 성분들을 어느 정도 유지하거나 복원할 수 있다는 것을 확인하였다. 제안한 가중치 함수는 함수 테이블로 간단히 구현할 수 있어 실시간 처리에도 매우 적합할 것으로 판단되며, 향후 음성 인식 단계와 연계한 실험을 계속하고자 한다.

### 참고 문헌

- [1] 김남수, "잡음 환경에서의 음성인식", Telecommunications Review, 제13권, 제5호, pp.650-661, 2003.
- [2] M. K. Hasan, S. Salahuddin, and M. R. Khan, "A Modified A Priori SNR for Speech Enhancement Using Spectral Subtraction Rules," IEEE Signal Processing Letters, Vol.11, No.4, pp.450-453, 2004(4).
- [3] A. Blin, S. Araki, and S. Makino, "Underdetermined Blind Separation of Convolutional Mixtures of Speech Using Time-Frequency Mask and Mixing Matrix Estimation," IEICE Transactions on Fundamentals of Electronics, Vol.E88-A, pp.1693-1700, 2005.
- [4] J. Jensen and J. Hansen, "Speech Enhancement Using a Constrained Iterative Sinusoidal Model," IEEE Transactions on Speech and Audio Processing, Vol.9, No.7, pp.731-740, 2001(10).
- [5] D. Ealey, H. Kellher, and D. Pearce, "Harmonic tunneling: tracking non-stationary noises during speech," Eurospeech, pp.437-440, 2001.
- [6] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by additive noise," Proceedings of the IEEE Conference on

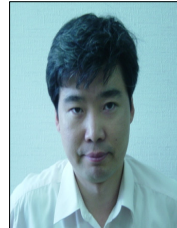
Acoustics, Speech, and Signal Processing, pp.208-211, 1979(4).

- [7] N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System," IEEE Transactions on Speech and Audio Processing, Vol.7, No.2, pp.126-137, 1999(3).
- [8] P. Lockwood and J. Boudy, "Experiments with a Nonlinear Spectral Subtractor(NSS), Hidden Markov Models and the pro-jection, for robust speech recognition in cars," Speech Communication, Vol.11, pp.215-228, 1992.
- [9] B. Jounghoon and K. Hanseok, "Spectral Subtraction Using Spectral Harmonics for Robust Speech Recognition in Car Environments," ICCS2003, LNCS Vol.2660, pp.1109-1116, 2003.
- [10] W. Hess, *Pitch Determination of Speech Signals*, Springer-Verlag Berlin Heidelberg New York Tokyo 1983.
- [11] L. Rabiner and R. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall 1978.
- [12] R. W. Aldhaheri and F. E. Al-Saadi, "Text-Independent Speaker Identification in Noisy Environment Using Singular Value Decomposition," ICICS-PCM, pp.1624-1627, 2003(12).

저 자 소 개

김 현 태(Hyun-Tae Kim)

정회원



- 1989년 2월 : 부산대학교 전자공학과(공학사)
- 1995년 2월 : 부산대학교 전자공학과(공학석사)
- 2000년 2월 : 부산대학교 전자공학과(공학박사)

▪ 2002년 3월 ~ 현재 : 동의대학교 멀티미디어공학과 부교수

<관심분야> : 음향신호처리, 멀티미디어통신

박 장 식(Jang-Sik Park)

정회원



- 1992년 2월 : 부산대학교 전자공학과(공학사)
- 1994년 2월 : 부산대학교 전자공학과(공학석사)
- 1999년 2월 : 부산대학교 전자공학과(공학박사)

▪ 1997년 3월 ~ 현재 : 동의과학대학 디지털정보전자과 교수

<관심분야> : 음성 및 음향신호처리, 멀티미디어통신, 입체음향