

음성 인식을 위한 전화망에서의 잡음처리

Noise Processing for Speech Recognition in the Telephone Line

전 원 석*, 신 원 호*, 양 태 영*, 김 원 구**, 윤 대 회*

(Won-Suk Jun*, Won-Ho Shin*, Tae-Young Yang*, Weon-Goo Kim**, Dae-Hee Youn*)

※본 연구는 '96 정보통신연구관리단 대학기초연구지원사업(과제관리번호:96033-IT2-11)에 의해 이루어진 것입니다.

요 약

본 논문에서는 다양한 전화선 채널을 통하여 수집된 음성 데이터에 포함된 잡음 및 채널 왜곡을 제거하여 음성인식 시스템의 성능을 향상시키는 방법에 관하여 연구하였다.

전화선을 통과한 음성에 포함된 채널 잡음 및 왜곡을 제거하는 방법으로는 음성신호를 보상하는 방법으로 CMS(Cepstral Mean Subtraction), SBR(Signal Bias Removal)과 SM(Stochastic Matching)의 성능을 비교 평가하였다.

잡음제거 방식의 성능을 평가를 위하여 음소 단위의 반연속 HMM을 이용한 화자독립 단독음 인식을 수행하였다. 인식 실험 결과, 멜 켈스트립을 사용한 경우에 CMS가 가장 우수한 성능을 내었고 다음으로 SM과 SBR 순으로 나타났다. 또한 특징벡터를 주변 잡음에 강인하게 하는 가중함수(RPS, BPL)를 사용한 켈스트립 계수와 잡음제거 방식을 함께 사용한 경우에 인식 성능이 더욱 향상되었다.

ABSTRACT

In this paper, we study methods that improve the performance of speech recognition system by subtracting the channel noise included in the speech data collected through the various telephone channels.

The performance of CMS(Cepstral Mean Subtraction), SBR(Signal Bias Removal) and SM(Stochastic Matching) for subtracting the channel noise and distortion are compared.

In order to evaluate the performance of the noise subtraction methods, speaker independent isolated word recognition experiments were conducted using phoneme-based semi-continuous HMMs. Experiment results show that CMS showed the best performance and the next is SM and SBR when mel cepstral coefficients were used as feature vectors. When Noise subtraction methods were used with cepstral weight functions(RPS, BPL) enable cepstral coefficients robust to the environmental noise, better performance enhancement could be achieved.

I. 서 론

잡음이 섞인 음성 신호로부터 잡음을 제거하는 잡음 제거(noise subtraction) 기술은 음성 신호처리의 중요한 연구 분야중 하나이다. 특히 최근에는 음성 인식 시스템의 실용화가 늘어나면서 잡음 환경에서의 음성 인식에 관한 연구가 활발히 진행되고 있다. 잡음 및 주위 환경 변화에 적용할 수 있는 음성 인식 시스템 구현은 실용적인 음성 인식 시스템을 개발하는데 있어서 고려해야할 중요한 문제들 중의 하나이다.

잡음에 강인한 음성 인식에 관한 연구는 그동안 실험실 환경에서 연구되어온 음성 인식 시스템에 대한 실용 가능성을 평가하기 위하여 다양한 잡음 환경에서 인식 평가가 이루어지고 있으며 이러한 평가를 바탕으로 인식 성능을 개선시킬 수 있는 여러 가지 방법들이 제안되어져 왔다 [1-5]. 그러나 이러한 방법들은 음성인식 시스템이 사용되어질 환경(예로, 자동차, 비행기, 전화선 등)을 대상으로 개발되어진 것이므로 다른 형태의 잡음 환경에서도 우수한 성능을 나타낸다고 볼 수는 없다. 특히 전화선을 통한 음성정보 서비스를 수행하는 경우, 사용자의 환경에서 발생하는 잡음의 형태는 주변잡음, 채널에 의한 신호왜곡, 화자의 발음변화 등 매우 다양하게 된다. 따라서 이러한 잡음이 존재하는 환경에서 음성인식 시스템을 사용하기 위해서는 다양한 종류의 잡음과 환경 변화에 대한 광

*연세대학교 전자공학과

**군산대학교 전기공학과

접수일자:1997년 7월 30일

범위한 실험과 연구가 요구되어지고 있다[6-13].

본 논문에서는 전화 채널을 통한 음성신호에 대하여 음성인식 시스템 성능 개선을 위한 효율적인 전처리 기술의 개발을 목표로 하였다. 이러한 목표를 위하여 전화망을 통하여 구축한 음성 인식 시스템 및 데이터베이스를 기반으로 잡음 및 채널 왜곡에 강인한 잡음처리 방법을 적용하여 인식 실험을 수행하였다. 실험에서는 신호 편의 제거(Signal Bias Removal)[16], 확률적 매칭(SM:Stochastic Matching)[17]과 켈스트랄 평균 차감법(CMS:Cepstral Mean Subtraction)[18]들을 음소 단위의 반연속(semi-continuous) HMM을 이용한 음성인식 시스템에 적용하여 화자독립 단독음 인식을 수행하였다.

인식 실험 결과, 잡음제거 방법들은 모두 인식 성능이 향상되었다 특히, CMS가 가장 우수한 성능을 내었고 다음으로 SM, SBR 순으로 나타났다. 또한 특징벡터를 주변 잡음에 강인하게 하는 가중함수(RPS, BPL)를 사용한 켈스트럼 계수와 잡음제거 방식(SBR, SM)을 함께 사용한 경우에 인식 성능이 더욱 향상되었다. 그러나 CMS의 경우에는 오히려 인식 성능이 저하되는 결과를 얻었다.

본 논문의 구성은 2장에서 전화망 및 잡음 환경에 적용하기 위한 잡음 처리 기술에 대하여 소개하고 3장에서 이에 따른 실험 및 결과를 고찰하였고 4장에서 결론을 맺었다.

II. 잡음 처리 기술

HMM을 사용한 음성인식 시스템을 사용할 경우 학습 과정과 인식과정에서의 음향적인 불일치는 인식시스템의 성능을 크게 저하시킨다. 특히 전화망의 경우에는 주변잡음, 채널 왜곡, 마이크로폰 왜곡 등으로 인식 성능이 크게 저하된다. 따라서 잡음에 강인한 음성인식 시스템은 이러한 환경적인 차이점을 보상하여 주면서 인식을 수행하는 것이다.

전화선을 통한 음성신호는 채널특성에 의하여 신호에 편의(bias)가 발생되는데 이러한 음성신호의 특징벡터열 $Y = (y_1, \dots, y_T)$, 학습된 HMM Λ 와 추정된 편의 \hat{b} 가 주어졌을 때, 특징 벡터 영역에서의 편의 제거는 다음과 같이 이루어진다.

$$\hat{x}_t = y_t - \hat{b} \quad (1)$$

이러한 편의를 제거하는 방법으로는 신호 편의 제거(SBR), 통계적 매칭(SM) 방법과 켈스트랄 평균 차감법(CMS)등이 있다.

2.1 신호 편의 제거(SBR : Signal Bias Removal)

신호 편의 제거(SBR) 방법은 여러 가지 환경에 의하여 오염된 입력 신호에서 음성 신호와 바이어스를 분리하여, 이를 제거함으로써 채널왜곡이나 잡음에 의한 영향 등을 효과적으로 제거해 준다[16]. 그림 1은 신호 편의 제

거 방법을 사용한 음성 인식 시스템의 구조를 나타낸다. 그림에서 알 수 있듯이 신호 편의 제거는 전처리 단계인 특징벡터를 구하는 단계에서 수행될 수 있다. 신호 편의 제거 방법에서는 입력 음성과 학습 모델간의 음향적인 불일치를 가우시안 코드북을 사용하여 최소화 시킨다. 따라서 코드북은 학습 데이터 또는 모델 파라미터로부터 M 개의 코드워드 $(\mu_i), i = 1, \dots, M$ 와 공분산 행렬 $(\Sigma_i), i = 1, \dots, M$ 을 포함하여야 한다. 따라서 음향적인 불일치는 입력음성과 코드북간의 관계로부터 구해질 수 있다.

신호 편의 제거에서의 편의는 다음과 같이 정의된 최우도(maximum likelihood) 추정에 의하여 구해질 수 있다.

$$p(Y|B, \Lambda) = \prod_{t=1}^T \max_i p(y_t - b | \mu_i), b = (b_1, \dots, b_D)' \quad (2)$$

여기서

$$p(y_t - b | \mu_i) = K \exp \left\{ -\frac{1}{2} [(y_t - b) - \mu_i]' \Sigma_i^{-1} [(y_t - b) - \mu_i] \right\} \quad (3)$$

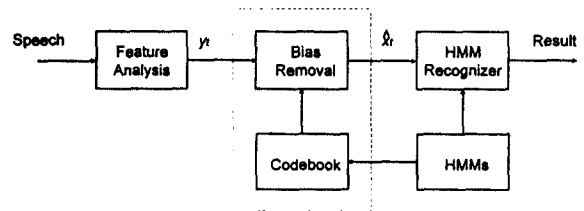


그림 1. 신호 편의 제거 방법을 갖는 음성인식 시스템의 구조

이고 K 는 b 와 무관한 정규화 상수이다. 초기 편의 추정값 \hat{b} 가 주어지면 식 (3)을 최대화하는 코드워드를 찾기 위하여 nearest neighbor 탐색을 수행한다. 이렇게 선택된 코드워드 z_t 는 다음과 같이 정의한다.

$$z_t = \mu_{i^*} \quad (4)$$

여기서

$$i^* = \arg \max_i p(y_t - b | \mu_i) \quad (5)$$

따라서 우도(likelihood) 함수는

$$p(Y|B, \Lambda) = K \exp \left\{ -\frac{1}{2} \sum_{t=1}^T [(y_t - b) - z_t]' \Sigma_i^{-1} [(y_t - b) - z_t] \right\} \quad (6)$$

이고 이를 최대화하는 편의 \hat{b} 는 다음과 같이 얻어진다.

$$\hat{b} = \frac{1}{T} \sum_{t=1}^T (y_t - z_t) \quad (7)$$

이를 이용하여 입력신호 y_t 로부터 편의가 제거된 $\hat{x}_t = y_t - \hat{b}$ 를 얻을 수 있다. 이러한 편의 추정 및 제거 과정은 수렴할 때까지 반복적으로 수행된다.

2.2 통계적 매칭(SM: Stochastic Matching)

통계적 매칭에서는 입력 음성 신호와 모델 파라미터간의 불일치 문제를 다룬다. 이러한 과정은 그림 2와 같이 두단계의 인식 과정을 필요로 하게 된다. 첫 번째 단계는 문자열 W 를 만들어 내는 일반적인 디코딩 단계일 뿐만 아니라 통계적 매칭의 편의를 추정하기 위한 상태열을 찾는 것이다. 우선, LTR(left-to-right) 형태의 N 개의 상태를 갖는 연속 분포(continuous distribution) HMM을 가정하고 상태 n 에서의 관찰 확률을 평균 μ_{nm} , 공분산 $C_{n,m}$ 과 가중 $W_{n,m}$ 을 갖는 M 개의 가우시안 분포의 혼합형태(mixture)로 정의하고 편의 $\hat{b} = (\hat{b}_1, \dots, \hat{b}_D)$ 를 EM 알고리즘을 사용하여 추정하면 다음과 같다[17].

$$\hat{b}_i = \frac{\sum_{t=1}^T \sum_{n=1}^N \sum_{m=1}^M \gamma_t(n, m) (y_{t,i} - \mu_{n,m,i}) / \sigma_{n,m,i}^2}{\sum_{t=1}^T \sum_{n=1}^N \sum_{m=1}^M \gamma_t(n, m) / \sigma_{n,m,i}^2}, \quad i=1, \dots, D \quad (8)$$

여기서 $\gamma_t(n, m) = p(Y_t = n, c_t = m | B, A)$ 이고 $C_{n,m} = \text{diag}(\sigma_{n,m,1}^2, \dots, \sigma_{n,m,D}^2)$ 를 가정하였다. 이렇게 추정된 편의를 사용하여 편의가 제거된 신호 $\hat{x}_t = y_t - \hat{b}$ 를 얻는다.

두 번째 단계에서는 편의가 제거된 신호 \hat{x}_t 가 새로운 인식 결과를 계산하기 위하여 다시 인식 시스템을 통과된다. 신호 편의 제거 방법과 마찬가지로 이러한 과정은 수렴할때까지 몇 차례 반복될 수 있다.

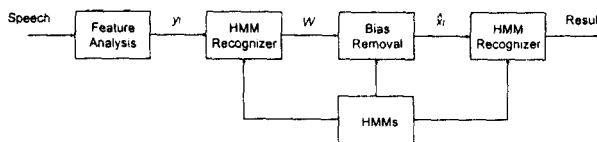


그림 2. 통계적 매칭 방법을 사용한 음성인식 시스템의 구조

2.3 켈스트랄 평균 차감법(CMS: Cepstral Mean Subtraction)

켈스트랄 평균 차감법은 진체 구간에 대하여 켈스트랄의 평균을 구하고, 이를 차감하여 채널의 효과를 제거하는 방법이다[18]. 전화망을 통한 음성 신호는 채널의 필터링 영향에 의해 선형 왜곡이 일어난다. 이를 간단히 표현하면, $Y(z) = S(z)G(z)$ 으로 표현되는데, 여기서 $S(z)$ 는 순수한 음성신호, $G(z)$ 는 전화 채널, $Y(z)$ 는 필터링된 음성을 의미한다.

이를 로그 영역으로 나타내면,

$$\log Y(z) = \log S(z) + \log G(z) \quad (9)$$

즉, 채널의 영향은 순수한 음성의 켈스트랄에 대해 부가적인 성분으로 나타나게 된다. 이 때, 순수한 음성의 켈스트랄에 대해 장구간 평균이 0이라고 가정하면, 채널 켈스트랄의 추정치는 필터링된 음성의 켈스트랄들을 평균하여 구할 수 있다. 채널 효과를 보상하기 위해서는 추정된 채널 켈스트랄을 제거하는 것이 켈스트랄 평균 차감법이다.

$$\hat{b} = \frac{1}{T} \sum_{t=1}^T y_t \quad (10)$$

그러므로, 채널의 영향이 보상된 켈스트랄은 다음과 같이 구해진다.

$$\hat{x}_t = y_t - \hat{b}, \quad t=1, \dots, T \quad (11)$$

III. 실험 및 결과

잡음제거 방법들의 성능을 평가하기 위하여 음소단위의 반연속 HMM[2]을 사용하여 화자 독립 단독음 인식 실험을 수행하였다.

3.1 데이터 베이스

데이터 베이스는 단축 다이얼을 위한 50단어를 선정하여 구성하였다. 학습 데이터는 20-30대 남성 화자 25명과 여성 화자 25명(총 50명)이 각 단어를 3회씩 발음한(50 단어 * 50명 * 3회 = 7500개) 음성으로 구성되었고, 인식 데이터는 학습 데이터에 포함되지 않은 20-30대 남성 화자 7명과 여성 화자 7명(총 14명)이 각 단어를 3회씩 발음한(50 단어 * 14명 * 3회 = 2100개) 음성으로 구성되었다. 서울 여러지역의 각 가정에서 전화를 통하여 발음된 각 음성은 DAT(Digital Audio Tape)에 저장되었다.

표 1. 데이터 베이스

Table 1. Data base

할아버지, 할머니, 아버지, 어머니, 아들, 딸, 형, 언니, 누나, 오빠, 동생, 장인어른, 장모님, 사돈, 삼촌, 고모, 이모, 사촌, 아저씨, 아주머니, 동창회, 친구, 회사, 은행, 보험회사, 증권회사, 전화국, 우체국, 사청, 동사무소, 파출소, 세무서, 경찰서, 철도역, 고속버스터미널, 도서관, 교회, 성당, 집, 수리센터, 중국집, 가스집, 찰집, 셔유집, 비디오, 세탁소, 병원, 가게, 학원, 서점
--

3.2 음성 분석 및 인식 시스템 구성

HMM을 이용한 화자 독립 단독음 인식 시스템의 구성은 다음과 같다. 음성 신호는 8kHz, 16비트로 표본화된다. 표본화된 음성 신호는 $1-0.95z^{-1}$ 의 전달 함수를 갖는 프

리엠퍼시스(pre-emphasis) 필터를 통과하고 32ms(256샘플)의 크기를 갖는 해밍 창을 사용하여 16ms씩 이동하면서 특징 벡터를 구한다. 실험에 사용된 특징 벡터는 멜 캡스트럼 계수(MFCC: Mel-Frequency Cepstral Coefficient), 차동 멜 캡스트럼 계수, 차동 에너지를 사용하였다.

멜 캡스트럼 계수는 1024샘플로 FFT를 수행하고 멜 밴드로 변환한 후 코사인 역변환하여 12차의 캡스트럼 계수를 구하였다. 또한 현재 프레임을 기준으로 전후 32ms의 멜 캡스트럼 차를 이용하여 차동 캡스트럼 계수를 구하였다. 에너지 파라미터는 전후16ms의 에너지 차와 에너지 차의 전후 32ms의 차를 구하여 2차의 에너지 벡터를 구하였다. 이러한 특징 벡터는 LBG 알고리즘을 사용하여 256개 코드를 갖는 3개의 코드북을 구하는데 사용되었고 이렇게 구해진 코드북은 반연속 HMM의 가우시안 분포의 초기치로 사용되었다.

각 음소 모델은 음성인 경우 5개, 묵음인 경우 3개의 상태수를 갖고 각 상태마다 2개의 천이를 갖는 LTR(left to right) 모델을 사용하였다. 각 단어의 모델은 각 단어의 음절수에 비례하도록 가변적인 상태갯수(음절수*4개)를 갖도록 하였다. 또한 관찰열의 발생 확률을 구할 때 사용되는 분포의 갯수 F 는 캡스트럼, 차동 캡스트럼, 차동 에너지에 각각 4, 4, 2개씩을 사용하였다. 반연속 HMM의 학습은 Baum-Welch 알고리즘을 사용하였고 인식시에는 Viterbi 알고리즘을 사용하였다.

3.3 잡음제거 방법에 따른 성능 평가

멜 캡스트럼을 사용하여 신호 편의 제거(SBR), 통계적 매칭(SM)과 캡스트럴 평균 차감법(CMS) 방법들의 성능을 평가하였다. 이때 특징벡터에는 가중함수를 사용하지 않았고 코드북의 크기는 256개를 사용하였다. 표 2에서 보면 잡음제거 방법을 사용하지 않은 경우보다 사용한 경우가 4.28%에서 9.95%의 인식 성능 향상을 보였다. 성능 향상의 정도는 CMS, SM, SBR 순으로 나타났다. 물론, 인식 성능면에서는 CMS가 가장 우수한 성능을 나타냈으나 CMS는 학습 데이터에도 CMS를 적용하여 인식 시스템을 재학습시켜야 하지만 SBR은 학습데이터에는 사용하지 않고 인식 데이터에만 사용할 수 있는 장점이 있다. SM도 이와 마찬가지로 많은 계산량과 복잡성을 요구한다.

표 2. 잡음제거 방법에 따른 인식 성능 평가
Table 2. Recognition rates for noise subtraction methods

잡음 제거 방법	인식률(%)
1) 사용하지 않음	88.67
2) SBR	92.95
3) SM	94.67
4) CMS	98.62

3.4 잡음제거 방법과 가중함수에 따른 성능 평가

특징벡터를 주변 잡음에 강인하게 하는 특징이 있다. 이러한 것은 낮은 차수의 캡스트럼 계수보다 높은 차수의 계수가 잡음에 강인하다는 것을 나타낸다. 따라서 잡음에 강인하다고 알려진 가중함수로 $w_i = i, i = 1, \dots, D$ 를 사용하는 Root Power Sum(RPS)[14]과 $w_i = 1 + 0.5L \sin(\pi_i/L), i = 1, \dots, D, (L$ 은 상수)를 사용하는 Band Pass Lifter(BPL)[15]를 사용하였고, 가중함수를 사용하지 않는 경우(EUC)와 비교하였다. 표 3에서 잡음제거 방법을 사용하지 않은 경우, 가중함수를 사용하면 인식 성능이 크게 향상되었다. 이러한 결과는 잡음이 존재하는 경우 가중함수를 사용한 특징벡터가 잡음에 강인하다는 결과와 일치한다. 또한 잡음제거 방식으로 SBR과 SM을 사용하는 경우도 가중함수에 따라 인식성능이 크게 향상되었다. 그러나 CMS의 경우에는 가중함수를 사용한 경우 오히려 인식 성능의 저하를 나타내었다. 이러한 이유는 각 단어의 평균을 편의로 계산하여 제거하는 방법을 사용하는 CMS가 높은 차수의 캡스트럼 계수의 변별력을 저하시키기 때문이다.

그러므로 SBR이나 SM의 경우에는 RPS의 가중함수와 함께 사용하는 경우가 우수한 성능을 내는 한편 CMS는 가중함수 없이 사용할 때 가장 우수한 성능을 나타냈다. 특히 SBR의 경우에는 전처리 과정에서 수행될 수 있고 SM보다 복잡성도 적은 장점이 있다. 그러나 가장 높은 인식 성능은 가중함수를 사용하지 않고 CMS를 사용한 경우이다.

표 3. 잡음제거 방법과 가중함수에 따른 성능 평가
Table 3. Recognition rates for noise subtraction methods and weight functions

잡음제거 방법	가중함수		
	EUC	RPS	BPL
1) 사용하지 않음	88.67	96.71	95.71
2) SBR	92.95	97.48	97.14
3) SM	94.67	97.48	97.14
4) CMS	98.62	96.57	94.52

IV. 결 론

본 논문에서는 전화 채널을 통한 음성 인식 시스템 성능 개선을 위한 효율적인 전처리 기술을 비교, 연구하였다. 전화망을 통하여 구축한 음성 인식 시스템 및 데이터 베이스를 기반으로 음소 단위의 반연속 HMM을 이용한 음성인식 시스템에 적용하여 화자독립 단독음 인식을 수행하였다.

잡음 및 채널 왜곡에 강인한 잡음처리 방법을 적용한

인식 실험에서는 캡스트럴 평균 차감법(CMS), 신호 변의 제거(SBR)과 확률적 매칭(SM) 방법들을 사용하였다.

잡음 처리 기술을 이용하여 실험한 결과 다음과 같은 결론을 얻을 수 있었다. 벨 캡스트럴을 사용하여 SBR, SM과 CMS 방법들의 성능 평가에서 CMS가 가장 우수한 성능을 내었고 다음으로 SM, SBR 순으로 나타났다. 또한 특징벡터를 주변 잡음에 강인하게 하는 가중함수(RPS, BPL)를 사용한 캡스트럴 계수와 잡음제거 방식을 함께 사용한 경우에 인식 성능이 더욱 향상된다. 그러나 CMS의 경우에는 오히려 인식 성능이 저하되는 결과를 얻었다.

이러한 세가지 방법을 비교하여 보면, CMS는 가장 우수한 인식 성능을 보이고 구현도 간단하지만 학습과 인식시에 모두 적용하여야 하는 단점이 있다. SBR의 경우에는 인식시에만 적용할 수 있고 구현도 비교적 간단하고 가중함수를 적용하면 비교적 우수한 성능을 나타내는 것을 알 수 있다. SM은 SBR과 마찬가지로 인식시에만 적용할 수 있으나 계산량이 많은 단점이 있다.

참 고 문 헌

1. A. Acero and R. M. Stern, "Environmental Robustness in Automatic Speech Recognition," in *Proc. ICASSP*, pp. 849-852, April 1990.
2. L. R. Rabiner, B. H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall Inc., 1993.
3. H. Hermansky, B. A. Hanson and H. Wakita, "Perceptually Based Linear Predictive Analysis of Speech," in *Proc. ICASSP*, pp. 509-512, March 1985.
4. B. A. Hanson and H. Wakita, "Spectral Slope Distance Measure with Linear Prediction Analysis for Word Recognition in Noise," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-35, No. 7, pp. 968-973, July 1987.
5. B. H. Juang, L. R. Rabiner and J. G. Wilpon, "On the Use of Bandpass Lifting in Speech Recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-35, No. 7, pp. 947-954, July 1987.
6. S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-27, No 2, pp. 113-120, April 1979.
7. H. Hermansky, N. Morgan, H. G. Hirsch, "Recognition of Speech in Additive and Convolutional Noise based RASTA Spectral Processing", in *Proc. ICASSP*, pp. 83-86, 1993.
8. J. Kochler, N. Morgan, H. Hermansky, H. G. Hirsch, G. Tong, "Integrating RASTA-PLP into Speech Recognition", in *Proc. ICASSP*, pp. 421-424, 1994.
9. H. Hermansky, N. Morgan, A. Bayya, P. Kohn, "Compensation for the Effect of the Communication Channel in Auditory-Like Analysis of Speech(RASTA-PLP)", in *Proc. EUROSPEECH*, vol. 3, pp. 1367-1370, Sep. 1991.
10. P. Alexandre, P. Lockwood, "Root Cepstral Analysis: A Unified View. Application to Speech Processing in Car Noise Environments", *Speech Communication*, vol. 12, no. 3, pp. 277-288, 1993.
11. P. Lockwood, P. Alexandre, "Root Adaptive Homomorphic Deconvolution Schemes for Speech Recognition in Noise", in *Proc. ICASSP*, pp. 441-444, 1994.
12. T. Usagawa, M. Iwata, M. Ebata, "Speech Parameter Extraction in Noisy Environment using A Masking Model", in *Proc. ICASSP*, pp. 81-84, 1994.
13. J. A. N. Flores, S. J. Young, "Continuous Speech Recognition on Noise using Spectral Subtraction and HMM Adaptation", in *Proc. ICASSP*, pp. 409-412, 1994.
14. A. H. Gray and Jr., J. D. Markel, "Distance Measures for Speech Processing," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-24, No. 5, pp. 380-391, Oct. 1976.
15. N. Nocerino, F. K. Soong, L. R. Rabiner and D. H. Klatt, "Comparative Study of Several Distance Measures for Speech Recognition," in *Proc. ICASSP*, pp. 25-28, March 1985.
16. Aaron E. Rosenberg, Chin-Hui Lee, Frank K. Soong, "Cepstral Channel Normalization Techniques for HMM-Based Speaker Verification", in *Proc. ICSLP*, pp. 1835-1838, 1994.
17. Mazin. G. Rahim, Bing-Hwang Juang, "Signal Bias Removal by Maximum Likelihood Estimation for Robust Telephone Speech Recognition", *IEEE Trans. Speech & Audio Processing*, vol. 4, No. 1, pp. 19-30, 1996.
18. A. Sankar and C. H. Lee, "A maximum-likelihood approach to stochastic matching for robust speech recognition", *IEEE Trans. Speech & Audio Processing*, vol. 4, No. 3, pp. 190-202, 1996.

▲전 원 석(Won Suk Jun)

1971년 6월 16일생



1996년 8월:연세대학교 전자공학과 졸업(공학사)

1996년 9월~현재:연세대학교 대학원 전자공학과 석사과정

※주관심분야: 음성인식, 잡음처리

▲신 원 호(Won Ho Shin)

한국음향학회지 1996년 15권 5호 참조

현재:연세대학교 대학원 전자공학과 박사과정

▲양 태 영(Tae Young Yang)

한국음향학회지 1996년 15권 5호 참조

현재:연세대학교 대학원 전자공학과 박사과정

▲김 원 구(Weon Goo Kim)

한국음향학회지 1996년 13권 1호 참조

현재:군산대학교 전기공학과 조교수

▲윤 대 회(Dae Hee Youn)

한국음향학회지 1996년 13권 1호 참조

현재:연세대학교 전자공학과 교수