

음향학적 파라미터를 이용한 한국어 연결숫자인식의 성능개선

Performance Improvement of Korean Connected Digit Recognition Based on Acoustic Parameters

김 승 회*, 김 형 순*

(Seung Hi Kim*, Hyung Soon Kim*)

*이 연구는 부산대학교 기성희재단 학술연구조성비에 의한 연구입니다.

요 약

본 연구에서는 한국어 연결숫자인식에 있어서 모델간의 변별력을 향상시키기 위하여 음향학적 파라미터(Acoustic Parameter)를 사용하는 것을 제안한다. 제안된 방법은 음성학적 지식에 근거하여 적절한 주파수 대역별 에너지의 비의 로그값을 추가적인 특징 파라미터로 사용한다. 실험결과, 제안된 방법을 사용함으로써 기본 인식시스템에 비해 오류율이 최고 46% 정도 감소됨을 확인할 수 있었다. 그리고 채널보상 기술을 함께 적용함으로써 69% 정도의 오류를 감소시킬 수 있었다.

ABSTRACT

This paper proposes use of acoustic parameters to improve the discriminability among digit models in Korean connected digit recognition. The proposed method used the logarithmic values of energy ratio between the predetermined frequency bands as additional feature parameters, based on the acoustic-phonetic knowledge. The results of our experiment show that the proposed method reduced the error rate by 46% in comparison with the baseline system. And incorporation of channel compensation technique in the proposed method yielded error reduction of about 69%.

I. 서 론

음성인식의 한 분야인 음성입력에 의한 숫자인식은 음성 다이얼링 서비스나 은행업무자동화, 신용카드 번호입력 등과 같은 실생활의 여러 부분에서 활용될 수 있다. 숫자인식은 고립숫자인식과 연결숫자인식의 두 부류로 나눌 수 있다. 그 중에서 자연스럽게 연속적으로 발음한 숫자열을 인식하는 연결숫자인식은 독립적으로 발음된 하나의 숫자를 인식하는 고립숫자인식에 비해 몇 가지 문제점을 가진다[1]. 우선 연속적으로 발음되는 숫자열에서는 숫자들의 경계가 모호하고, 인접한 숫자들간의 상호 조음현상(coarticulation)에 의해 각 숫자들의 고유발음이 변하며, 발음된 숫자들의 길이를 모르는 경우도 많다. 그러나 실생활에서 사용자들은 자연스럽게 발음하는 것을 원하며, 따라서 사용자의 편의성을 고려한다면 연결숫자인식이 필수적으로 요구된다.

그 동안 연결숫자인식에 관한 많은 연구가 이루어져 왔으며, 한국어 숫자의 경우에도 최근 들어 활발한 연구들이 이루어지고 있다[2][3][4]. 그러나 한국어 숫자들은

우선 모두 단음절이고, 또한 숫자들간의 혼동가능성도 커서 영어 등 타언어권의 숫자인식에 비해 난이도가 높다. 한국어 숫자인식에서 오인식되는 경우들을 살펴보면, 몇몇 숫자쌍들로 오류분포가 집중되는 것을 알 수 있다. 따라서 이런 숫자들간의 변별력을 높일 수 있다면 인식시스템의 성능은 크게 개선될 수 있을 것이다. 그러나, 현재까지 이들 오인식되는 숫자들간의 변별력을 향상시키기 위한 연구들은 별로 이루어지지 못한 실정이다.

이 문제를 해결하기 위한 접근 방향은 특징 파라미터 추출과정에서의 접근방법과 숫자모델 훈련과정에서의 접근방법의 두 가지로 나누어 볼 수 있다.

우선 첫 번째로 특징 파라미터 추출과정에서 각 숫자들에 대한 사전 지식을 기반으로 하여 이들간의 변별력을 크게 할 수 있는 음향학적인 특징 파라미터를 추가적으로 추출하여 사용하는 방법을 생각해 볼 수 있다. 본 논문에서는 부가적인 특징 파라미터 추출이라는 접근방법에 의해 인식성능개선을 도모하였다.

두 번째 방법은 기존의 HMM 파라미터 추정방법으로 널리 사용되는 최대우도추정법(Maximum Likelihood Estimation) 대신에 훈련시 모델 파라미터 간의 차별성에 중점을 둔 Minimum Classification Error(MCE)나 Maximum Mutual

* 부산대학교 전자공학과
접수일자: 1999년 3월 16일

Information Estimation(MMIE) 방법 등을 사용하는 것이다. 1994년 Yves Normandin은 TIDIGITS 연결숫자 database에 대해 MMIE를 사용하여 최고 99.11%의 숫자열 인식 성능을 얻었다[5]. 본 논문에서는 두 번째 접근방법은 시도하지 않았다.

II. 음향학적 특징 파라미터

2.1 오인식되는 숫자쌍들에 대한 특징 분석

인식 실험 후의 결과를 토대로 오인식되는 양상들을 살펴보면 자주 오인식되는 숫자쌍들은 '일⇔이', '일⇔칠', '구⇔오', '오⇔공' 등이다. 먼저 '일⇔칠'의 오인식은 일반적으로 그 빈도가 그리 높지 않은 편이다. 다만 일부 화자에 있어서는 숫자열 첫 머리의 /칠/이 /일/로 오인식되는 빈도가 높게 나오기도 한다. /치/의 스펙트로그램은 그림 1에서 표시된 부분과 같이 전반적으로 백색 잡음과 유사한 형태를 띄나 3~4kHz 이상의 대역에 에너지가 많이 분포하는 특징을 가진다.

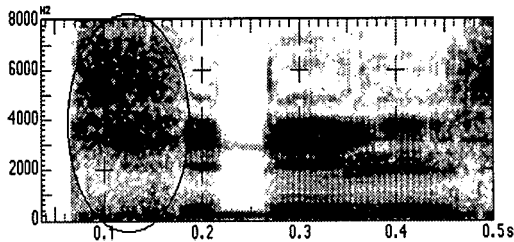


그림 1. 숫자열 /칠일/에 대한 스펙트로그램의 예
Fig.1. An example of spectrogram of the sound /chil-il/.

'구⇔오'의 경우에는 모음인 /우/와 /오/가 스펙트럼상으로 상당히 유사하다. 동일 화자내에서는 /우/와 /오/의 스펙트럼이 상대적인 차이를 보이는 경우도 있으나, 개개 화자간의 특성 차이를 극복할 만큼 음소간 특성의 편차가 크지 않기 때문에 /우/와 /오/가 단독으로 있을 경우에는 구별이 힘들다. 따라서 '구⇔오' 또는 '오⇔공'에서는 자음인 /기/이 인식에서 차지하는 비중을 높일 경우 구분이 보다 용이해 질 것으로 생각할 수 있다. 모음인 /우/나 /오/와 자음인 /기/의 스펙트럼 특성은 아주 다르다.

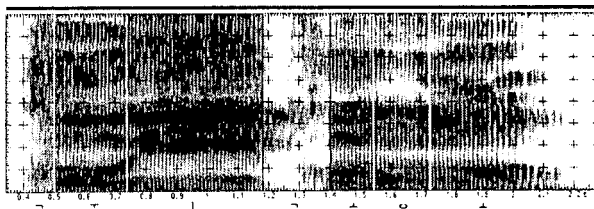


그림 2. 숫자열 /구이공/에 대한 스펙트로그램의 예
Fig. 2. An example of spectrogram of the sound /gu-i-gong-ol/.

/우/나 /오/에서는 높은 에너지가 분포하는 포먼트(formant) 주파수가 1kHz 이하와 1kHz~4kHz 대역에 뚜렷이 존재하는 것을 관찰할 수 있다. 그러나 /기/의 경우 뚜렷한 에너지 peak 열이 보이지 않으며 대체적으로 에너지가 분산되어 있는 편이다(그림 2 참조).

'일⇔이'는 화자에 상관없이 고루 발생하는 오인식 쌍이다. 실제로 연결숫자음성에서 해당구간만을 들었을 경우에는 사람도 이 둘을 구분하기 힘들다. 스펙트로그램을 보더라도 이 두 음절은 아주 유사한 특징을 가지며, 따라서 스펙트로그램을 이용한 방법으로는 이 둘의 구분이 힘들 것으로 예상된다. '일'이라는 음절에서 모음 /이/ 다음에 나오는 실측음 /리/의 주파수 패턴이 /이/와 큰 차이를 보이지 않으며, 음절 '이'에서 모음 /이/가 문맥의 영향을 받아 주파수 패턴이 변형되기 때문에 /일/과의 구분이 더욱 힘들다(그림 3 참조).

이상에서 살펴본 바와 같이 연결숫자인식에 있어서 스펙트로그램 상에서의 에너지의 분포 패턴에 관한 정보, 특히 자음과 모음간의 에너지 분포 패턴의 차이를 이용한다면 일의 경우를 제외한 오인식 숫자쌍들에 대해서는 이들간의 변별력 향상에 상당히 도움이 될 것으로 예상할 수 있다.

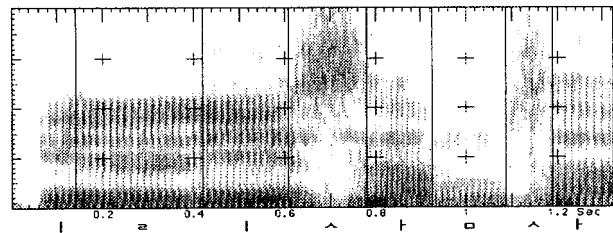


그림 3. 숫자열 /일이삼사/에 대한 스펙트로그램의 예
Fig. 3. An example of spectrogram of the sound /il-i-sam-sa/.

2.2 음향학적 특징 파라미터 (Acoustic Parameter)

음소들 그룹간의 구별에 있어서는 음운학적 자질론에서 출발한 음향학적 파라미터(이하 AP)를 사용하는 것이 캡스트럼에 비교하여 화자들간의 차이에 보다 둔감한 것으로 보고되었다[6]. Hansen[6]은 TIMIT 데이터베이스에서 각 음소들의 그룹들을 구분하기 위해, 주파수 밴드별 에너지에 기반을 둔 AP와, 유성음일 확률(voicing probability), 그리고 제1, 제2, 제3 포먼트(formant) 주파수 및 그 대역폭을 사용하였다. Hansen의 논문에서 에너지에 기반한 대부분의 파라미터들은 입력음성 중 가장 큰 값으로 정규화되었다. 그러나, 이 방법은 일단 음성이 다 입력된 다음에야 정규화의 기준이 되는 가장 큰 값을 구할 수 있으므로 실시간 처리에 문제가 있다. 그래서 본 논문에서는 정규화된 에너지를 사용하지 않는다. 대신에 다음과 같이 주파수 밴드별 에너지의 비에 기반을 둔 파라미터나 혹은 그들의 이동 평균(moving average)을 통해서 앞에서 언급한 에너지 분포 패턴에 관한 정보를 구한다.

$$E(t)_{AP} = \log \left(\frac{E_{A-\beta}(t)}{E_{\beta-A}(t)} \right) \quad (1)$$

$$E(t)_{AP,MA} = \frac{1}{5} \sum_{i=-2}^2 E(t-i)_{AP} \quad (2)$$

여기서 $E_{A-\beta}(t)$ 는 t 번째 프레임에서 주파수 f 에서 $f/2$ 까지의 에너지를 말한다. 숫자음에서 모음을 포함한 유성음의 경우 f_0 를 제외하고는 대부분 400Hz 근처에서 에너지값이 peak를 이룬다. 그리고 500Hz를 경계로 에너지가 급속히 감소하는 경향이 있으며, 4kHz 이상에서는 전체적으로 에너지가 작은 것을 관찰할 수 있다. f_0 의 경우에는 0~1kHz 대역에 강한 에너지가 나타난다. 무성음의 경우, 특히 /s/이나 /z/과 같은 마찰, 파찰지음들은 전반적으로 백색잡음과 유사한 형태이나, 4kHz 이하보다는 그 이상의 대역에서 높은 에너지를 가지는 경우가 많다. 본 논문에서는 이들 500Hz, 1kHz와 4kHz를 주파수 대역을 나누는 경계로 사용하였다.

AP를 사용하는 방법으로는 크게 훈련과정에서 모델에 AP를 포함하여 훈련하는 방법과 모델에는 포함시키지 않고 인식과정에서 AP를 이용하여 적절한 penalty를 부과하는 방법이 있다. 본 논문에서는 AP를 모델에 포함한 경우에 대해 실험하였다.

III. 채널 특성 및 화자 특성의 보상

HMM을 기반으로 하는 인식 시스템의 성능저하에 영향을 미치는 중요한 요소 중의 하나는 모델을 훈련할 때의 환경과 실제 인식이 작동되는 인식과정에서의 환경의 차이이다. 따라서 인식 시스템을 배경잡음이 존재하는 곳에서 사용하기 위해서는 배경잡음에 의한 영향을 보상해 주어야 한다. 음성신호에 영향을 미치는 배경잡음은 크게 음성신호에 가해지는 부가잡음과 채널잡음에 의한 음성신호의 왜곡으로 나눌 수 있다.

특히 채널잡음은 주파수 영역에서 clean speech에 곱해지는 선형필터로 모델링할 수 있으며 cepstrum 영역에서는 clean speech에 더해지는 형태로 표현된다[7]. 따라서 부가잡음의 영향을 무시할 수 있다면 왜곡된 음성신호는 cepstrum 영역에서 clean speech에 채널잡음을 더한 것이다.

이와 같은 채널의 영향을 보상하기 위한 방법으로는 cepstral mean subtraction(CMS), signal bias removal(SBR)[8], RASTA(Relative SpecTra)[9] 등 여러 방법들이 제안되었다. 이들 방법 중에서 CMS 및 RASTA 처리가 계산량에 비해 우수한 성능을 나타내는 것으로 알려져 있으며, 본 논문에서는 global CMS[10][11]와 local CMS[7], RASTA의 방법을 사용하여 채널왜곡 보상을 시도하였다. 음성신호의 장구간 스펙트럼이 화자의 특성에 영향을 받는다는 것은 잘 알려진 사실이며, 음성신호의 장구간 스펙트럼 특성은 시간에 대해 독립적이거나 느리게 변한다. TIDIGIT 데이터베이스에 대해 실험한 결과,

시간영역에서의 스펙트럼 파라미터열을 주파수영역으로 변환했을 때 0에서 1Hz까지의 분산은 숫자 그룹의 차이보다는 화자의 목소리에 훨씬 많은 영향을 받는 것으로 나타났다[12]. 결론적으로, 이 파라미터열의 저주파 영역의 요소들을 필터를 통해 감쇄시킴으로서 인식시스템은 화자에 보다 더 독립적인 시스템이 될 것이다. 대부분의 채널 왜곡 보상방법들은 위의 파라미터열에서 낮은 주파수 성분들을 제거하는 것이며, 따라서 clean speech에 대해서도 채널왜곡보상기법을 적용할 경우 화자특성보상에 의한 얼마간의 성능향상을 기대할 수 있다.

IV. 실험 및 결과

본 절에서는 연속 HMM을 이용하여 연결숫자인식 실험을 하며, 앞 절에서 설명한 방법들을 통해 인식성능의 개선정도를 평가하고자 한다. 우선 숫자사이의 변별력 향상을 위해 제안된 특징 파라미터를 사용하여 주위 잡음이 거의 없는 깨끗한 환경에 대해 인식실험을 수행한다. 다음으로 채널보상 알고리즘을 적용하여 실험하며, 이어서 제안된 음향학적 특징 파라미터와 채널보상 알고리즘을 같이 사용한 실험을 행한다.

실험에 사용하는 음성 특징 파라미터로는 기본적으로 12차 MFCC, delta MFCC, 그리고 delta 에너지를 사용하여 총 25차의 파라미터를 사용한다.

4.1 데이터 베이스 및 기본 인식 시스템

국어공학센터의 한국어 4연숫자 음성 DB 중에서 남성 화자들의 음성만을 사용하였다[13]. 음성은 방송부스에서 녹음되었으며, 16kHz로 sampling하고 16bits로 양자화되어 있다. 대부분 서울지역의 화자로 구성되어 있으며, 아나운서도 1명 포함되어 있다. 본 데이터 베이스에서 사용된 숫자들은 공, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구의 10개이다. 모델을 만들기 위한 훈련용 데이터 베이스로는 30명의 화자가 발성한 4130개의 4연숫자문장을 사용하며, 인식실험을 위한 데이터 베이스로는 훈련용 화자에 속하지 않은 9명의 화자가 발성한 1260개의 문장을 사용하였다.

실험에 사용한 숫자 모델링 방법으로 숫자에서 발생할 수 있는 음소를 기본으로 앞뒤의 문맥을 고려한 문맥중속 음소모델인 triphone을 기본 모델링 단위로 정의하였다. 제한된 데이터를 효율적으로 사용하기 위하여 tree-based clustering 기법을 적용하였으며, mixture 수를 증가시키면서 인식실험을 하였다. 음소당 상태수는 3개와 5개에 대해 검토하였으며, 상태수 5개일때의 성능이 보다 우수하여 이후 음소당 상태수 5개에 대해서만 실험을 진행하였다(표 1 참조). 다음에 나오는 표의 개별숫자에 대한 인식률은 대체(substitution) 오류만을 고려한 것으로, 개별숫자가 각각의 자리에서 정확하게 인식되지 않았으면 틀린 것으로 간주하였다.

표 1. 기본 인식시스템의 인식률(%)
Table 1. Recognition rate of baseline system(%)

Mixture 개수	숫자열	개별숫자
1	95.2	98.6
3	95.2	98.6
5	94.7	98.5
7	94.8	98.5
9	95.0	98.6
11	95.0	98.6
13	94.4	98.4

4.2 부가적인 특징 벡터 및 채널 보상 알고리즘을 적용한 실험

주파수 밴드는 0~500Hz(E1), 0.5~1kHz(E2), 1~4kHz(E3), 4~8kHz(E4)로 나누었으며, 음향학적 파라미터는 이들 주파수 밴드별 에너지의 비의 로그값을 사용하였다.

표2에서 E1부터 E4까지 한 가지씩만 사용된 것은 각각의 주파수 밴드 에너지를 전체 에너지로 나누어 최고값을 취한 경우이며, 복수개의 주파수 밴드가 사용되는 때에는 앞쪽 주파수 밴드 에너지들을 마지막 주파수 밴드 에너지로 정규화 시키도록 하였다. 실험결과 E1, E2, E3에 대해 log(E1/E3), log(E2/E3)의 값에 대한 이동평균 및 이들의 차분값을 부가적인 특징 벡터로 추가한 경우의 인식률이 가장 높았다(표 2참조).

표 2. 음향학적 파라미터(AP)를 사용한 실험의 인식률(%)
Table 2. Recognition rates of the experiments using acoustic parameters.

사용한 주파수 대역	최적 mixture 개수	숫자열	개별숫자
-	3	95.2	98.6
E1	3	95.4	98.7
E2	1	95.2	98.7
E3	1	94.7	98.5
E4	3	94.7	98.5
E1, E2, E3	5	97.0	99.2
E1, E2, E3(이동평균)	5	97.2	99.2
E1, E2, E3(이동평균)+차분값	3	97.4	99.3

표 3. 채널 보상 기술을 적용한 실험의 인식률(%)
Table 3. Recognition rate of the experiments using channel compensation techniques.

적용한 채널 보상기술	최적 mixture 개수	숫자열	개별숫자
-	3	95.2	98.6
GCMS	7	97.4	99.3
LCMS 150ms	5	97.4	99.3
LCMS 300ms	5	97.9	99.4
LCMS 450ms	7	97.8	99.4
RAST	3	97.5	99.3

실험은 각각의 경우에 대해 mixture 수를 1개에서 13개까지 늘려가면서 수행하였으며, 가장 높은 인식률을 나타낸 mixture 개수에 대한 결과만을 표에 나타내었다.

채널 보상 및 개개 화자 특성의 보상을 위해서 global CMS(GCMS), local CMS(LCMS), RASTA 방법을 적용하여 실험하였다. 300ms의 윈도우를 사용한 LCMS 방법의 결과가 가장 우수했으며 GCMS, RASTA에 대해서도 유사한 성능이 얻어졌다(표 3참조).

다음으로 음향학적 파라미터를 추가한 후에 채널 보상 알고리즘을 적용하여 실험을 하였다. 주파수 밴드 E1, E2, E3에 대한 음향학적 특징 파라미터를 추가하고 RASTA 방법을 적용하였을 때 98.5%의 가장 좋은 결과를 나타내었다. 표 4에는 부가적인 특징 벡터와 채널 보상 기술을 함께 적용한 여러 실험들 중에서 우수한 성능을 나타낸 실험들의 결과만을 나타내었다.

기본 인식시스템의 실험결과를 분석해 보면 전체 오류 중 '일⇔이'의 오류가 전체 오류의 37%, '구⇔오⇔공'의 오류가 전체의 48%를 차지하였다. 음향학적 파라미터를 도입해서 인식성능이 가장 우수한 경우 '일⇔이'의 오류는 개선정도가 미미하나, '구⇔오⇔공'의 오류는 68% 개선된 결과를 보이며, 채널 보상 알고리즘만을 적용했을 경우에는 '일⇔이'의 경우 41% 개선된 결과를 나타내었다. 음향학적 특징 파라미터를 추가한 뒤 채널 보상 알고리즘을 적용한 경우에는 기본 인식시스템의 결과에 대해 '일⇔이'가 55%, '구⇔오⇔공'은 86% 개선된 결과를 나타내었다. 즉, 음향학적 파라미터를 도입함으로써 '구⇔오⇔공'의 오류는 상당히 개선됨을 알 수 있다. 참고로 전화채널환경의 데이터 베이스에 대해서는 음향학적 파라미터를 도입하였을 때 인식률의 향상을 가져오지 못했다.

표 4. 부가적인 특징 벡터와 채널 보상 기술을 함께 적용한 실험의 인식률(%)

Table 4. Recognition rate of the experiments using additional APs and channel compensation techniques.

적용 알고리즘	최적 mixture 개수	숫자열	개별숫자
-	3	95.2	98.6
E1,E2,E3+RASTA	3	98.5	99.6
E1,E2,E3(이동평균)+LCMS300	9	98.3	99.6
E1,E2,E3(이동평균)+delta+RASTA	3	98.3	99.5
E1,E2,E3,E4+RASTA	5	98.5	99.6
E1,E2,E3,E4(이동평균)+RASTA	5	98.4	99.6
E1,E2,E3,E4+delta+RASTA	5	97.9	99.5

V. 결론

본 논문에서는 오인식이 빈번한 숫자쌍 모델들간의 변별력을 향상시킴으로써 연결숫자인식 시스템의 성능을

개선시키기 위해 부가적인 음향학적 특징 파라미터를 추출하여 사용하였다. Triphone을 기본 모델링 단위로 사용하였으며, 숫자간의 변별력 향상을 위해 부가적인 음향학적 파라미터로서 주파수 밴드별 에너지의 비의 로그값을 사용하였다. 실험결과 인식성능이 가장 우수한 경우 46% 정도 오류율이 감소됨을 확인할 수 있었다. 그리고 채널 및 화자특성의 차이를 보상하기 위해 몇 가지 채널 보상 기술을 추가적으로 적용함으로써 최고 69% 정도의 오류율 감소를 볼 수 있었다.

본 논문에서는 구현하지 않았으나, 서론에서 언급한 바와 같이 숫자간의 변별력 향상을 위해 훈련과정에서 MMIE나 MCE 기법을 적용하게 되면 추가적인 성능향상이 기대되며, 앞으로 이에 대한 연구가 계속 진행되어야 할 것으로 판단된다.

ACKNOWLEDGEMENT

이 연구는 국어공학센터에서 구축한 한국어 4연숫자 음성 DB를 사용하였습니다.

참 고 문 헌

1. L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall International, Inc., 1993.
2. 양태영 외 8인, "연결 숫자음 인식에서의 상태 및 단어 유지 확률을 이용한 지속시간 모델링," 제10회 신호처리 합동학술대회 논문집, pp. 313-316, 1997.
3. 이성권, 김순현, "Voice dialing system을 위한 연속음성인식," 제 11회 신호처리합동학술대회 논문집, pp. 685-688, 1998.
4. 김기성, 김승희, 김형순, 지민제, "한국어 연결숫자인식을 위한 숫자 모델링에 관한 연구," 제 15회 음성통신 및 신호처리 워크샵 논문집, pp.293-297, 1998.
5. Y. Normandin, R. Cardin and R. De Mori, "High-performance connected digit recognition using maximum mutual information estimation," *IEEE Trans. on Speech and Audio Processing*, vol.2, no.2, pp.299-311, Apr. 1994.
6. A. V. Hansen, "Acoustic-Phonetic Features used in Automatic Speech Recognition," Ph.D. thesis, Department of Communication Technology, Aalborg University, 1998.
7. A. E. Rosenberg, "C. H. Lee and F. K. Soong, Cepstral channel normalization techniques for HMM-based speaker verification," in *Proc. IEEE ICSLP*, pp.1835-1838, 1994.
8. M. G. Rahim and B. H. Juang, "Signal bias removal for robust telephone based speech recognition in adverse environments," in *Proc. IEEE ICASSP*, pp.445-448, 1994.
9. H. Hermansky, "RASTA processing of speech," *IEEE Trans. on Speech and Audio Processing*, vol.2, no.4, pp.578-589, Oct. 1994.
10. S. Furui, "Cepstral analysis technique for automatic speaker verification," *IEEE Trans. on ASSP*, vol.ASSP-29, pp.254-272, 1981.
11. C. Mokbel, P. Paches-Leal, D. Jovet and J. Monne, "Compensation of telephone line effects for robust speech recognition," in *Proc. IEEE ICSLP*, pp.987-990, 1994.
12. C. Nadeu, P. Paches-Leal and B. H. Juang, "Filtering the time sequences of spectral parameters for speech recognition," *Speech Communication*, vol.22, pp.315-332, June 1997.
13. Korean Speech Data Base CD-ROM, 국어공학센터, 1998.

▲김 승 희 (Seung Hi Kim)



1997년 : 부산대학교 전자공학과
학사

1999년 : 부산대학교 전자공학과
석사

1999년 ~ 현재 : 한국전자통신연구원
멀티모달 I/F 팀 연구원

※주관심분야: 음성인식, 합성,
음성신호처리

▲김 형 순 (Hyung Soon Kim)

현재 : 부산대학교 전자공학과 부교수
한국음향학회지 제17권 3호 참조