

유무선 전화를 통한 화자인식 알고리즘에 관한 연구

A Study on Speaker Recognition Algorithm Through Wire/Wireless Telephone

김 정 호*, 정 희 석*, 강 철 호*, 김 선 희**
(Jung-Ho Kim*, Hui-Seok Jung*, Chul-Ho Kang*, Sun-Hee Kim**)

*광운대학교 전자통신공학과, **광운대학교 SITRC
(접수일자: 2002년 10월 25일; 채택일자: 2003년 1월 16일)

본 논문에서는 방사 기저함수 (RBF: Radial Basis Function) 신경망을 이용하여 특징 파라미터를 사상시켜 화자인식의 성능을 개선하기 위한 알고리즘을 제안하였다. 동일한 화자의 유무선 전화의 벡터 영역이 서로 다르므로 제안한 화자확인 시스템은 유무선 학습 모델을 생성하기 위해서 먼저 음성인식을 통해 유무선 채널을 판별하고, 학습하지 않은 채널의 모델은 방사 기저함수 신경망을 이용하여 학습된 모델의 특징 벡터 (LPC-켄스트럼)를 사상하는 방법이다. 모의 실험 결과 기존의 켄스트럼 평균 차감법을 사용할 때보다 제안한 알고리즘을 적용했을 때의 인식율이 약 0.6%~10.5%의 성능 향상을 보여주었다.

핵심용어: 방사 기저함수 신경망, 평균 차감법, 화자인식

주요분야: 음성처리 분야 (2.5)

In this thesis, we propose the algorithm to improve the performance of speaker verification that is mapping feature parameters by using RBF neural network. There is a big difference between wire vector region and wireless one which comes from the same speaker. For wire/wireless speakers model production, speaker verification system should distinguish the wire/wireless channel that based on speech recognition system. And the feature vector of untrained channel models is mapped to the feature vector (LPC Cepstrum) of trained channel model by using RBF neural network. As a simulation result, the proposed algorithm makes 0.6%~10.5% performance improvement compared to conventional method such as cepstral mean subtraction.

Keywords: Radial basis function neural network, Cepstral mean subtraction, Speaker recognition

ASK subject classification: Speech signal processing (2.5)

I. 서론

최근에는 화자 인식시스템의 실용화가 늘어나면서 환경 변화에 강한 화자인식에 관한 연구가 활발하다. 잡음이 없거나 조용한 실험실 환경에서 우수한 성능을 나타내는 화자 인식시스템이 주위에 잡음이 노출된 실제 환경에서는 급격한 성능 저하가 발생한다. 또한 유무선 전화망을 통하는 경우, 사운드 카드나 음성을 받는 마이크가 다른 경우에 채널특성이 일정하지 않는 문제점이 발생하고, 지금까지의 연구 결과에 의하면 음성인식 또는 화자인식 실험에 있어서 치명적인 성능 저하를 가져온다 [1].

현제 전화망을 이용한 음성인식 또는 화자인식 서비

스가 제공되고 있는데 전화음성은 여러가지 요인들에 의하여 인식하기가 어렵다. 마이크나 유선 전화망 환경에서는 불일치 조건의 원인이 되는 주위의 배경잡음이나 전자기적인 충격에 의하여 발생하는 부가잡음 (additive noise)과 전화기의 마이크 특성과 전화선 및 교환기를 포함하는 채널의 특성에 의한 채널왜곡 (channel distortion)이 동시에 존재한다. 현재 서비스 중인 무선 전화망 환경에서는 8 kCELP, 13 kCELP, EVRC 방식이 있다. 3가지 방식은 유선전화망과 같은 잡음을 가질 뿐만 아니라, 음성 데이터를 낮은 전송율로 가변하여 전송하므로 화자의 특징 파라미터에 손실이 발생한다. 유무선 전화상의 화자인식에서 가장 큰 문제점은 동일한 채널에서의 인식율보다 서로 다른 채널 (예 학습:유선, 인식:무선)에 대한 인식율이 급격히 감소한다는 것이다. 일반적으로 유무선 의 서로 다른 채널에서 인식율은 마이크에 비해 크게 떨어진다. 따라서 본 연구에서는 유무선 전화환경에 강한

책임저자: 김정호 (anpaul@hanmail.net)
115-7(3) 서울시 노원구 월계동
광운대학교 전자통신학과
(전화: 02-940-5136; 팩스: 02-917-5136)

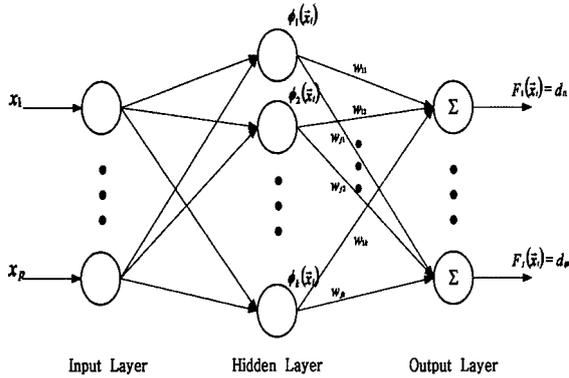


그림 1. 방사 기저함수 신경망의 기본구조
Fig. 1. Basis configuration of RBF neural network.

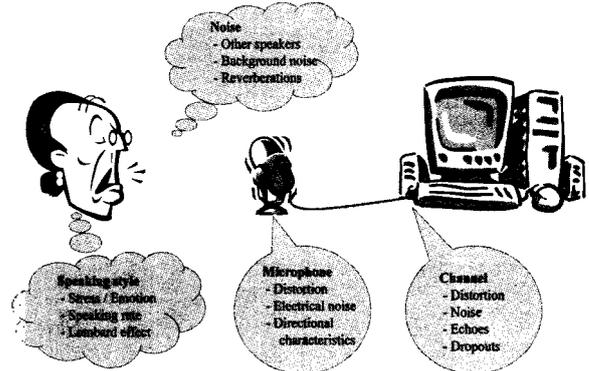


그림 2. 화자인식에서의 성능저하 요인
Fig. 2. Degradation factors of speaker recognition.

한 화자인식에 관한 알고리즘을 제안한다.

II. 방사 기저함수 신경망

방사 기저함수 (RBF: Radial Basis Function) 신경망은 그림 1과 같이 3개의 층으로 구성되어 있다. 입력이 주어지면 주어진 입력으로 계산된 은닉층 (hidden layer)의 출력과 우리가 원하는 출력값을 이용하여 가중치가 학습되어진다. 가중치를 학습시키는 방법으로는 선형 적응필터에서 사용하는 알고리즘을 그대로 사용한다[2-4].

방사 기저함수 신경망에서 입력과 출력간의 사상 (mapping)은 다음과 같은 두 단계로 나누어서 설명할 수 있다.

1. 입력층과 은닉층간의 비선형 변환 단계
2. 은닉층과 출력층간의 선형 변환 단계

비선형 변환은 방사 기저함수에 의해서 정의되며, 선형 변환은 가중치들에 의해서 정의된다. 방사 기저함수 신경망의 핵심은 은닉층에 있는 방사 기저함수들이며 일반적으로 가우시안 함수 (Gaussian Function)를 가장 많이 이용한다. 그림 1은 방사 기저함수 신경망의 기본구조를 나타낸다.

$$F_i(\vec{x}_i) = \sum_{k=1}^K w_{jk} \phi_k(\vec{x}_i) \quad (1)$$

여기서, $1 \leq i \leq N$, $1 \leq k \leq K$, $1 \leq j \leq J$ 가우시안 방사 기저함수는 식 (2)와 같이 정의된다.

$$F_i(\vec{x}_i) = \sum_{k=1}^K w_{jk} \exp\left(-\frac{1}{2\sigma_k^2} \|\vec{x}_i - \vec{\mu}_k\|^2\right) \quad (2)$$

σ_k 는 가우시안 함수의 표준편차이고, $\|\vec{x}_i - \vec{\mu}_k\|$

는 입력 \vec{x} 와 중심점 $\vec{\mu}_k$ 간의 유클리드 거리이다.

가우시안 방사 기저함수 신경망에서 학습해야 할 파라미터로는 표준편차, 방사 기저함수들의 중심점 (center), 가중치 (weight)가 있다. 학습과정에서 중심점과 가중치가 수렴하도록 반복적으로 수행한다. 입력과 사상시키려는 출력 벡터가 낮은 차수인 경우에 중심점을 임의로 설정해도 되지만 높은 차수인 경우에 임의로 설정하면 수렴을 하지 못하고 발산하는 문제점이 있다.

III. 잡음 처리 방법

3.1. 전화망의 잡음요인 및 채널 왜곡 보상종류

유무선 전화망의 음성인식이나 화자확인에서 발생하는 잡음 원인은 크게 두 가지로 부가잡음 (additive noise)과 채널왜곡 (channel distortion)으로 구분할 수 있다. 전화기를 이용하는 경우 화자의 음성이 전화기 가까이에서 발음을 하게 되므로 부가잡음보다는 채널 왜곡의 영향을 더 받는다[5]. 그림 2는 화자인식에서 성능을 저하시키는 여러 가지 요인을 나타내고 있다.

채널 특성을 알 수 없는 전화망에서의 채널왜곡 보상 방법으로는 캡스트럼 영역에서의 채널 바이어스를 제거하는 방법과 시간적인 정보를 이용하는 방법이 있다.

3.2. 캡스트럼 영역에서의 채널 바이어스 제거

채널왜곡은 캡스트럼 영역에서 바이어스로 표시되며 이를 추정하여 제거함으로써 채널보상을 할 수 있다. 대표적인 채널 보상 방법으로는 캡스트럼 평균 차감법 (CMS: Cepstral Mean Subtraction)[6]과 신호 바이어스 제거 (SBR: Signal Bias Removal)[7]이다.

3.3. 시간적인 정보를 이용하는 방법

일반적으로 음성은 일정한 범위의 속도로 변화하는데 대하여 채널왜곡은 시간적으로 일정하거나 천천히 변화한다. 시간적으로 천천히 변화하는 정보를 제거함으로써 채널의 영향을 받지 않는 특징 벡터를 얻을 수 있다. 상대적 스펙트럼 (RASTA: RelATive SpecTrAl) 처리[8]가 가장 많이 이용된다.

하지만 SBR과 RASTA-PLP 방법들은 음성인식에서 좋은 성능을 보여 주지만, 화자인식에서는 화자의 특성을 감소시키는 문제로 화자인식에 적합하지 않다.

IV. 제안한 화자인식 알고리즘의 구조

채널왜곡에 강인한 채널보상 알고리즘도 동일한 화자로서 서로 다른 채널에서 코드북 영역이 다른 경우에 채널 보상을 하지 못한다. 학습과 인식이 동일한 채널에서 이루어졌을 때 실험한 결과는 서로 다른 채널에서 유무선 화자를 확인하는 실험을 한 경우와 비교할 때 급격한 인식율의 감소가 발생한다. 서로 다른 채널에 대한 인식율의 급격한 감소의 원인은 여러 가지가 있지만, 가장 큰 원인으로 유선과 무선의 코드북이 서로 다른 영역을 가진다는 것이 있다. 음성인식 시스템에서 유무선 데이터를 이용해 유선과 무선을 인식하는 실험을 했을 경우 LPC-킵스트림과 RASTA-PLP에서 약 90% 이상 유무선을 판별할 수 있었다.

학습단계에서 화자가 사용하는 채널이 유선인지 무선인지 판단할 수 없는 상황에서 인식 실험을 했을 경우 학습과 다른 채널일 때 인식율의 감소가 발생하므로, 학습 단계에서 먼저 음성인식 시스템을 이용하여 어느 채널로 학습하는지 유무선을 판별하고 학습하지 않는 채널에 대한 모델을 동시에 만들 경우 유무선 대한 인식율을 향상시킬 수 있다.

4.1. 음성의 유무선 채널 판별

학습단계에서 유선과 무선을 판별하기 위해서 음성인식 시스템을 사용하였다. 음성인식의 학습과정에 대한 모델은 오프라인으로 유선과 무선 각각의 코드북은 화자 확인의 월드모델(world model)의 유무선 공통 코드북으로 동일하게 생성하였고, 이산 HMM (Discrete Hidden Markov)을 이용하여 유선과 무선의 음성 모델을 만들었다. 인식과정에서 유무선의 비터비 계산(Viterbi scoring) 결과를 비교하여 화자확인 시스템에서 음성인식의 비터비 계산이 큰 쪽은 기존의 학습 알고리즘을 이용하고, 비

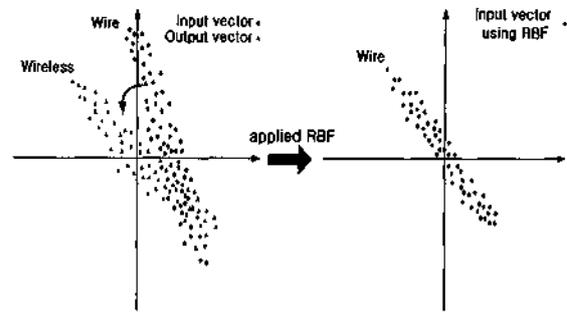


그림 3. 유선벡터를 무선벡터 영역으로 사상
Fig. 3. Mapping wire telephone vector to wireless telephone vector.

터비 계산이 작은 쪽은 제안한 알고리즘을 이용하여 학습하게 된다.

4.2. 화자확인 시스템의 학습단계

화자확인 시스템의 학습단계에서는 어떤 채널로 학습을 하더라도 동시에 유무선 모델을 생성하게 된다. 예로 유선전화로 학습을 할 경우 유선을 제외한 무선에 대한 학습 모델이 필요하게 된다. 무선전화에 대한 모델은 그림 3과 같이 유선벡터를 무선벡터 영역과 비슷한 벡터로 사상(mapping)시켜 만들 수 있다. 사전에 구성한 유무선 학습 모델을 이용하여 화자확인 시스템에 적용하였다.

학습단계에서 음성인식으로 유무선을 판별하고, 판별된 채널에 대해서는 일반적으로 사용하는 화자확인 방법으로 모델을 구성한다. 학습하지 않은 채널은 학습하지 않은 채널의 공통 코드북을 이용하여 출력벡터를 설정하고, 음성인식에서 판별한 채널의 특징벡터를 입력벡터로 하고, 방사 기저함수 신경망을 이용하여 출력벡터로 사상시켜 유무선 전화에 대한 모델을 동시에 구성할 수 있다. 그림 3은 유선 벡터를 무선 벡터 영역으로 사상하는 방법을 간략하게 보여주고 있다.

그림 4는 제안한 방법으로 입력벡터와 방사 기저함수 신경망을 적용한 혼합구조이다. 학습 채널로 유선전화를 사용하고, 학습하지 않은 채널을 무선벡터 영역으로 사상하는 방법을 예로 들었다. 방사 기저함수 신경망을 적용하기 위해 입력 벡터를 유선전화로 얻어진 LPC-킵스트림 벡터로 정하고, 사전에 구성한 무선 공통 코드북을 이용하여 출력 벡터를 얻는다. 방사 기저함수 중심점은 학습시 3번 발음한 유선 LPC-킵스트림 값을 클러스터링 알고리즘(clustering algorithm)을 이용해서 초기화시킨다. 최적화된 가중치와 중심점을 다시 입력 벡터인 유선 LPC-킵스트림과 방사 기저함수 신경망에 적용하면 무선 영역으로 사상된 벡터를 얻을 수 있다.

4.3. 화자확인 시스템의 인식단계

화자확인 시스템의 인식단계에서는 미리 준비된 월드 모델과 학습에서 생성된 화자의 개인 모델간의 계산에

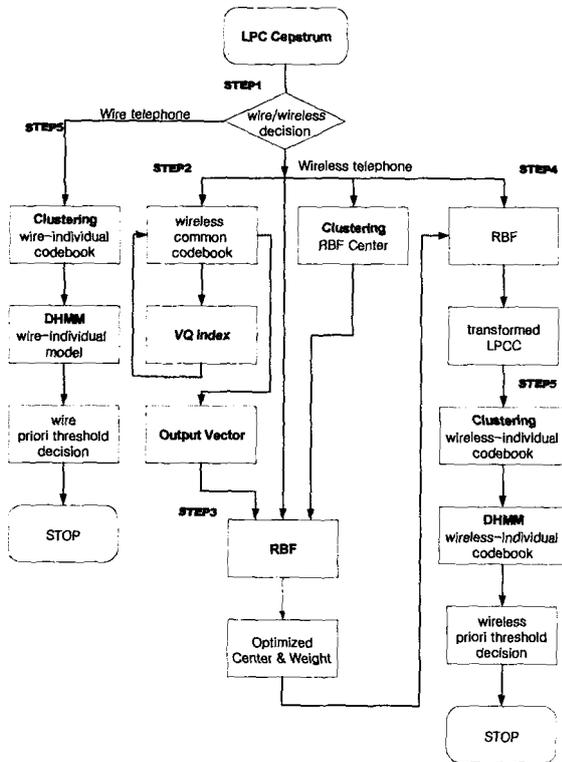


그림 4. 방사 기저함수를 사용하여 제안한 학습단계 구성도
Fig. 4. The proposed training block diagram by RBF.

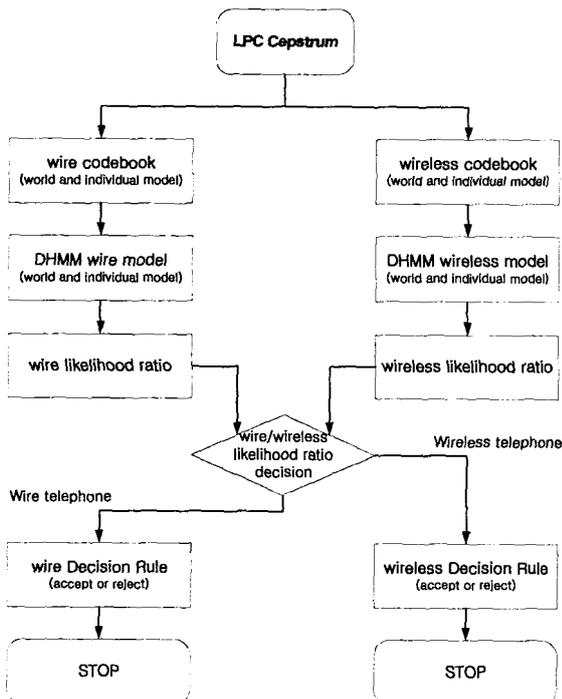


그림 5. 인식단계 구성도
Fig. 5. Recognition block diagram.

근거한 일정한 결정규칙에 승인 및 거부를 판단하여 본인 여부를 결정하는 역할을 하게 된다. 유무선 두 가지 방식으로 학습된 모델에 대해 본인 여부를 확인하기 위해서는 유무선 각각에 대한 우도비를 가지게 된다. 유선과 무선에 대한 우도비 크기를 비교해서 우도비가 큰 값을 선택하는 결정규칙에 의해 본인 여부를 판단하게 된다. 학습 과정과 같이 음성인식을 이용하지 않아도 인식단계에서 발음한 채널에 대한 동일한 채널의 우도비가 상대적으로 크기 때문에 음성인식 과정을 적용하지 않아도 우도비 크기에 의해 유무선을 판별할 수가 있다.

V. 실험 환경 및 결과

화자확인 시스템의 전화음성 데이터는 Dialogic Board를 이용하여 유선전화는 실험실 환경에서 발음하였고, 무선전화는 실제환경에서 발음하여 Vox 파일로 저장하였다. 유무선 각각의 공통 코드북과 월드모델은 오프라인으로 20~30대의 50명의 남성 화자로부터 5번 발음한 데이터로 사전에 구성하였다. 실험에 사용된 단어는 ("안녕하세요")로 50명의 남성 화자에 의해 3일차, 회당 10회씩 발음하였다.

실험에 사용된 음성 데이터는 8 kHz 8 bit vox 파일을 화자확인 시스템에 적용하기 위해 8 kHz 8 bit wav 파일로 변환하였다. 현재 서비스 중인 무선 vocoder 방식과 같은 조건으로 음성의 한 프레임은 240으로 하여 80 frame씩 이동하면서 해밍 윈도우 (hamming window)을 취한 후 14차 LPC-켄스트럼 특징 파라미터로 구성하였다.

실험은 음성인식 시스템과 화자확인 시스템의 학습/인식 단계로 구분할 수 있다. 음성인식 시스템의 학습단계는 오프라인으로 각각의 유무선 코드북은 128개의 코드워드로 구성하였고, 이산 HMM의 상태수는 5개로 하고, Baum-Welch 재추정의 수렴조건 임계값은 0.005로 설정하였다. 인식단계는 비터비 알고리즘을 이용하여 유무선을 판별하였다. 화자확인 시스템의 학습단계는 공통 코드북보다 적은 개인 코드북을 생성하고, 이산 HMM 과정은 음성인식과 같은 조건으로 개인 문턱치를 결정한다. 방사 기저함수 신경망의 입력 벡터와 출력 벡터는 14차 LPC-켄스트럼을 사용하고, 방사 기저함수 중심점의 수는 클러스터링 알고리즘의 기본 조건인 입력 프레임의 10분의 1로 설정하였다. 화자확인 시스템의 인식단계는 유무선 각각에 개인 코드북과 공통 코드북을 합성한 코드북을 통해 VQ과정을 수행하고, 이산 HMM 모델을 통해서

표 1. 유무선 음성인식의 인식율

Table 1. Recognition rate of speech recognition between wire telephone and wireless telephone.

Channel	wire	wireless
LPC Cepstrum	721/800= 90.1%	702/800= 87.75%
RASTA-PLP	745/800= 93.1%	722/800= 90.2%

표 2. 동일한 유선 채널에서 기존의 방법과 제안한 방법의 인식율 비교

Table 2. Comparison of the recognizing rate between conventional and proposed method on the wire telephone line at same channel.

wire(training) → wire(recognition)				
Cepstral Mean Subtraction				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	37/350	207/24500	98/500	243/24500
%	10.6	0.85	19.6	0.99
total error rate	FR: 15.81%, FA: 0.92%			
proposed method				
50 users	after training		after 3 days	
	FR	FA	FR	FA
합계	37/350	527/24500	93/500	575/24500
%	10.6	2.15	18.6	2.35
total error rate	FR: 15.2%, FA: 2.25%			

인인모델과 월드모델의 비터버 계산으로 얻어진 우도비는 길정규칙에 따라서 인증 여부를 판단하게 된다[9,10].

그리고 펜티엄 41.6 GHz, 램 256 MB를 기준으로 실험의 계산량 비교하였다. 학습시 3번 발음한 후 모델 생성시점은 cepstrum 차감법 경우 10.2초이고, 제안한 채널 보상 방법은 32.4초가 소요되었다. 인식시 cepstrum 차감법 경우 2.4초이고, 제안한 채널 보상 방법은 4.9초가 소요되었다.

표 1은 음성인식 실험에 의해서 유무선을 판별한 실험 결과이다. 실험 결과에 의해서 유무선 전화음성은 LPC-cepstrum과 RASTA-PLP에 약 90% 이상 채널을 구별할 수 있음을 실험 결과로 보여주고 있다.

표 2, 3는 동일한 채널에서의 인식 실험 결과이다. 제안한 채널보상 방법이 cepstrum 평균 차감법(CMS)보다 오거부율(FR)이 감소하였지만, 오수락율(FA)은 증가하였다. 오수락율의 경우에 유무선의 우도비를 비교하여 큰 우도비를 선택하는데 우도비가 유무선을 잘못 판별하여 cepstrum 평균 차감법에 비해 오수락율의 오류율이 증가하였다. 무선환경에서 인식 실험 결과 유선보다 여

표 3. 동일한 무선 채널에서 기존의 방법과 제안한 방법의 인식율 비교

Table 3. Comparison of the recognizing rate between conventional and proposed method on the wireless telephone line at same channel.

wireless(training) → wireless(recognition)				
Cepstral Mean Subtraction				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	58/350	174/24500	147/500	189/24500
%	16.6	0.71	29.4	0.77
채널 인식율	FR: 24.12%, FA: 0.74%			
proposed method				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	55/350	473/24500	141/500	495/24500
%	15.7	1.93	28.2	2.02
total error rate	FR: 23%, FA: 1.97%			

표 4. 서로 다른 채널에서 기존의 방법과 제안한 방법의 인식율 비교

Table 4. Comparison of the recognizing rate between conventional and proposed method at different channel.

wire(training) → wireless(recognition)				
Cepstral Mean Subtraction				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	207/500	178/24500	243/500	183/24500
%	41.4	0.73	48.6	0.75
total error rate	FR: 45%, FA: 0.74%			
proposed method				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	164/500	714/24500	181/500	703/24500
%	32.8	2.91	36.2	2.86
total error rate	FR: 34.5%, FA: 2.9%			

러 가지 환경 조건에 큰 인식율 차이를 보이고 있다. 표 4, 5는 서로 다른 채널에서의 인식 실험 결과이다. 표 2, 3 보다 급격하게 인식율 감소가 발생한다. 유무선의 벡터가 서로 다른 영역에 있으므로 동일한 채널에서 실험 한 경우보다 인식율이 급격하게 감소한다. 제안한 방법은 cepstrum 평균 차감법보다 오거부율(FR)이 10.3~10.5% 감소했지만, 오수락율(FA)도 2.09~2.16% 증가하였다.

표 5. 서로 다른 채널에서 기존의 방법과 제안한 방법의 인식율 비교

Table 5. Comparison of the recognizing rate between conventional and proposed method at different channel.

wireless(training) → wire(recognition)				
Cepstral Mean Subtraction				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	222/500	185/24500	239/500	208/24500
%	44.4	0.755	47.8	0.849
total error rate	FR: 46.1%, FA: 0.8%			
wireless(training) → wire(recognition)				
Cepstral Mean Subtraction				
50 users	after training		after 3 days	
	FR	FA	FR	FA
total	171/500	706/24500	187/500	709/24500
%	34.2	2.88	37.4	2.89
total error rate	FR: 35.8%, FA: 2.89%			

VI. 결론

실제 화자확인 시스템에서 불일치조건 (Mis-match condition)은 인식율의 급격한 하락을 가져오게 된다. 특히, 전화환경에서의 유·무선 채널의 불일치는 화자확인 시스템에서 주변 환경 잡음이나 화자내 변이와 함께 가장 큰 오인식 요인이 된다.

따라서, 화자의 특징 벡터가 잡음 환경에 변화에 적응하고 잡음에 강인한 채널보상에 대한 연구가 매우 중요하다. 본 논문에서는 학습과정에서 유선과 무선 모델을 동시에 구성하는 방법을 제안하였다.

모의실험 결과, 서로 다른 채널에서 기존의 채널 보상 방법인 캡스트럼 평균 차감법을 사용한 경우, 학습이 유선이고 인식이 무선인 경우 55%, 학습이 무선이고 인식이 유선인 경우 53.9%의 본인 승인율을 나타낸 반면, 본 연구에서 제안한 RBF를 이용하여 특징 벡터를 사상하는 알고리즘을 사용하였을 경우 본인 승인율이 캡스트럼 평균 차감법에 비해 10.3~10.5%로 크게 향상된 것을 알 수 있다.

참고 문헌

- 진세훈, 이재희, 강철호, "화자확인 시스템을 위한 적응적 모델 갱신과 사전 문턱치 결정에 관한 연구." 한국음향학회지, 19 (5), 20-26, 2000.
- S. Haykin, *Neural Networks*, Prentice-Hall International Inc, 1999.
- M. W. Mak and S. Y. Kung, "Estimation of elliptical basis function parameters by the EM algorithm with application to speaker verification," *IEEE Trans. on Neural Networks*, 11 (4), 961-969, July 2000.
- M. W. Mak and S. Y. Kung, "Robust speaker verification over the telephone by feature recuperation," *Intelligent Multimedia, Video and Speech Processing*, 2001, Proceedings of 2001 International Symposium, 433-436, 2001.
- 전원석, 신원호, 양태영, 김원구, 윤대희, "음성 인식을 위한 전화망에서의 잡음 처리," 한국음향학회지, 17 (1), 4-8, 1998.
- R. J. Mammone, X. Ziang, R. P. Ramachandran, "Robust speaker recognition—a feature-based approach," *IEEE Signal Processing Mag.*, 58-71, September 1996.
- M. G. Rahim and B. H. Juang, "Signal bias removal by maximum likelihood estimation for robust telephone speech Recognition," *IEEE, Trans. Speech & Audio Processing*, 4 (1), 19-30, 1996.
- H. Hermansky and N. Morgan, "RASTA processing of speech," *IEEE Trans. Speech & Audio Processing*, 2 (4), 578-589, 1994.
- 위진우, 강철호, "잡음환경에 강인한 HMM 기반 화자확인 시스템," 한국음향학회지, 20 (7), 67-75, 2001.
- L. Rabiner, B.-H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall International Inc, 1993.

저자 약력

● 김 정 호 (Jung-Ho Kim)



2000년 2월: 대전대학교 통신공학과 (공학사)
2002년 8월: 광운대학교 전자통신공학과 (석사)
2002년 8월~현재: 한국파워보이스 근무
* 주관심분야: 화자인식, 잡음처리

● 정 화 석 (Heui-Seok Jung)

한국음향학회지 제18권 제3호 참조

● 강 철 호 (Chul-Ho Kang)

한국음향학회지 제20권 제7호 참조

● 김 선 희 (Sun-Hee Kim)

1985년 2월: 연세대학교 불어불문학과 (문학사)
1986년 8월: 파리 7대학 (석사)
1988년 2월: 파리 고등사회과학대학원 (EHESS) (인어학 박사)
현재: 광운대학교 객원교수