

# 화자 확인 시스템을 위한 적응적 모델 갱신과 사전 문턱치 결정에 관한 연구

## A Study on Adaptive Model Updating and a Priori Threshold Decision for Speaker Verification System

진 세 훈\*, 이 재 회\*, 강 철 호\*  
(Se Hoon Chin\*, Jae Hee Lee\*, Chul Ho Kang\*)

### 요 약

화자 확인시스템에서 화자의 장기간 음성 변동에 대처하기 위해서는 작은 양의 데이터로써 화자 확인을 위한 HMM(hidden Markov model) 파라미터 갱신과 사전 문턱치 결정이 중요한 요소이다. 본 연구에서는 화자내 변이(intra-speaker variation)에 적용하는 모델 갱신방법과 이에 따른 문턱치 적응에 관한 방법을 제안한다. 제안하는 방법은 분기간 화자내 변이로 발생할 수 있는 오인식율을 Baum-Welch re-estimation을 통해 현재 화자 모델 파라미터에 새로운 음성 데이터를 적용시킴으로써 감소시킨다. 본 논문에서 제안하는 사전 문턱치 결정 방법은 기존의 월드 모델(world model) 방법과 군중 모델(cohort model) 방법의 하이브리드 형태로써 실험적으로 결정된다. 실험에 의해 모델 갱신을 하지 않은 경우보다 제안하는 모델 갱신방법의 화자 인식율이 우수함을 확인하였다. 또한, 사후 문턱치 결정에 의한 인식율과 제안한 사전 문턱치 결정에 의한 인식율의 차이가 근소함을 확인하였다.

핵심용어: HMM, 모델 갱신, 사전 문턱치, 월드 모델 방법, 군중 모델 방법  
투고분야: 음성처리 분야(2.5)

### ABSTRACT

In speaker verification system the HMM(hidden Markov model) parameter updating using small amount of data and the priori threshold decision are crucial factor for dealing with long-term variability in people voices. In the paper we present the speaker model updating technique which can be adaptable to the session-to-session intra-speaker variability and the priori threshold determining technique. The proposed technique decreases verification error rates which the session-to-session intra-speaker variability can bring by adapting new speech data to speaker model parameter through Baum Welch re-estimation. And in this study the proposed priori threshold determining technique is decided by a hybrid score measurement which combines the world model based technique and the cohort model based technique together. The results show that the proposed technique can lead a better performance and the difference of performance is small between the posteriori threshold decision based approach and the proposed priori threshold decision based approach.

Key Words: HMM, Model updating, Priori threshold, World model, Cohort model based technique

### I. 서 론

사회의 정보화가 급속히 진행됨에 따라 통신망을 통한 사용자의 대규모 데이터베이스에 대한 접근과 정보의 검색, 갱신, 수정이 빈번해지고 있다. 이에 따라 정보의 보안 문제가 심각해지고 있다. 따라서 사용자의 본인 여부를 판단하는 개인 확인 수단이 필수적이다. 기존의 카드, 도장,

신분증 등은 도난이나 위조에 매우 취약하고, 정보의 접근이 전화나 통신망들을 이용하여 원격지에서 이루어지는 경우, 개인 확인은 더욱 어려워진다. 이에 비해 화자 인식은 음성에 포함되어 있는 화자 정보를 추출하여 개인을 확인하는 기술로서 사칭자(impostor)에 대한 대처, 처리시간, 원격지 확인 등 여러 측면에서 가장 효과적인 기술중 하나이다. 화자 인식(Speaker Recognition)은 인식 대상에 따라 발성화자를 구분해내는 화자 식별(Speaker Identification)과 본인 여부를 판단하는 화자 확인(Speaker Verification)의 두 분야로 나눌 수 있다.

\* 광운대학교 전자통신공학과  
접수일자: 2000년 3월 1일

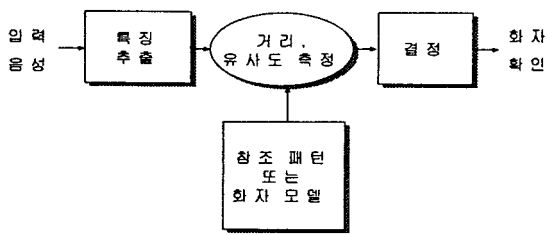


그림 1. 화자 확인 시스템의 블록도  
Fig. 1. Block diagram of the speaker verification system.

화자 확인의 일반적인 블록도를 그림 1에 나타내었다. 음성이 입력되면 그것의 특징(feature)을 뽑아 참조패턴이나 모델과의 거리나 유사도를 측정하여 그 결과 값에 따라 화자 확인을 행한다.

화자 확인 시스템은 매우 긴 시간동안 개별 화자의 신원을 확인해야 한다. 그러나, 일정기간동안 발생된 음성 데이터만으로 만들어진 화자 모델은 시간이 지남에 따라 변화하는 화자의 음성 정보 때문에 화자의 신원을 잘 구별할 수 없게된다. 따라서, 모델 갱신이 필요하다.

Furui는 틀(template)기반 화자 확인에서 최근의 음성정보를 이용하여 참조 틀(template)을 갱신하는 방법을 발표했다<sup>[1]</sup>. 또한, HMM기반의 화자 확인에서 참조 모델을 갱신하는 몇몇 연구들이 있었다<sup>[2][3]</sup>. 만약 각 화자의 많은 양의 음성정보가 저장되어있고 참조 모델이 새로운 음성 정보와 함께 다시 만들어진다면 매우 좋을 것이다. 하지만, 요구되는 메모리와 계산량이 매우 많으므로 작은 양의 음성정보만으로 화자 모델 파라미터를 갱신하는 방법이 요구된다.

화자 확인 시스템의 애러는 오거절(false rejection)과 오수락(false acceptance)의 두 가지 애러로 나눌 수 있는데 이 두 가지 애러는 서로 상반(trade-off)관계에 있다. 즉, 오수락(FA) 애러율을 줄이기 위해 문턱치를 높이면 오거절(FR) 애러율이 증가하고 반대로 오거절(FR) 애러율을 줄이기 위해 문턱치를 너무 낮추면 오수락(FA) 애러율이 증가하게 된다. 그러므로 시스템 구성시 이러한 관계를 잘 고려하여 적절한 문턱치를 가질 수 있도록 신중히 설계하여야 한다<sup>[4]</sup>.

## II. 기존의 우도비 정규화 기법

HMM(Hidden Markov Model)을 이용한 방법은 현재 음성인식 분야에서 널리 사용되고 있는 확률모델을 화자 인식에 적용하는 방법이다<sup>[5][6]</sup>. HMM은 학습 기능을 이용하여 화자 내의 변이를 흡수할 수 있으며, 입력 패턴의 비선형 정합을 수행하는 특성이 있다. 화자 확인에서 월드 모델(world model)기반 정규화 기법과 군중 모델(cohort model)기반 정규화 기법은 가장 많이 사용되는 정규화 기법이다<sup>[7][8]</sup>. 확인은 주어진 샘플과 주장되는 신원에 대해 확인 시스템이 수락 또는 거절의 결정을 내려주는 과정이다. 그 시스템은 주장되는 모든 신원에 대한

정보를 가지고 있어야 한다. V를 확인 과정이라 하고 I를 요구되는 신원 그리고 K가 정보라 하자. 입력 샘플 S에 대해 확인 과정은 다음과 같이 정의된다.

$$V : (S, I, K) \rightarrow (0, 1) \quad (1)$$

이 확인 과정은 식 (2)와 같이 미리 학습된 참조 패턴  $T_s$ 와 입력 샘플의 유사도 측정 M과 미리 정해진 문턱치  $\theta$ 로 이루어져 있다.

$$V : (S, I, K) = \begin{cases} 0 & M(S, I, T_s) < \theta \\ 1 & M(S, I, T_s) \geq \theta \end{cases} \quad (2)$$

그러나, 문턱치  $\theta$ 는 참조 패턴  $T_s$ 와의 비터비 계산(viterbi scoring) 방법에 따라서 달라지므로 정규화 할 필요가 있다. 기존의 월드 모델(world model)방법은 문장 종속 화자 독립 단어모델을 이용한다. 즉, 월드 모델(world model)과 테스트샘플간의 비는 화자 S의 틀(template)  $P(O|M(S, W))$ 의 우도비(Likelihood ratio)를 정규화하는데 사용된다.  $M(S_{world}, W)$ 를 단어 W에 대한 월드 모델(world model)이라 하면 정규화는 다음의 우도비(Likelihood ratio)를 가져온다.

$$R_{world} = \frac{P(O|M(S_i, W))}{P(O|M(S_{world}, W))} \quad (3)$$

그리고, 경쟁 기반 측정 기법으로써 군중 모델(cohort model)기법이 있다<sup>[9]</sup>. 이 방법에서 화자 S의 틀(template)  $P(O|M(S, W))$ 의 우도비를 정규화하기 위해서 식 (4)와 같이 신원이 요구되는 화자 모델과의 확률값과 가장 확률값이 큰 다른 화자 모델과의 비로써 정의한다.

$$R_{cohort} = \frac{P(O|M(S_i, W))}{\max_{j \neq i} P(O|M(S_j, W))} \quad (4)$$

## III. 제안하는 적응적 모델 갱신 기법과 사전 문턱치 결정 기법

### 3.1. 제안한 적응적 모델 갱신 기법

사람의 목소리는 시간이 지남에 따라 변화한다. 그렇다면 어떻게 이 변화에 대처할 것인가가 화자확인 시스템 성능 평가에 중요한 요소가 된다. 실제 온라인 환경에서 고객에게 모델 갱신을 위해서 많은 수의 발음을 요구할 수 없으므로 적은 수의 음성 발음으로 모델 갱신을 할 수 있어야 한다. 본 연구에서는 각 화자에 대한 현재 이산, 반연속 HMM파라미터를 이용해서 제한된 데이터 양으로써 화자 모델을 갱신한다. 여기서 이산 HMM의 경우, 공통의 코드북으로는 화자의 변이를 충분히 흡수할 수 없었기에 공통 코드북에 각 화자의 데이터를 흡수시켜 갱신하여 개인별 코드북 벡터를 생성하여 실험하였다.

본 연구에서는 3회에 걸쳐 화자 모델을 갱신하며 실험을 하였다. 식 (5)부터 (9)까지는 초기, 천이, 관측 파라미터와 평균, 분산의 재추정식이다.

$$\pi_i = \gamma_1(i) \tag{5}$$

$$a_{ij} = \frac{\sum_{t=1}^T \gamma_t(i, j)}{\sum_{t=1}^T \gamma_t(i)} \tag{6}$$

$$b_j(k) = \frac{\sum_{t=1}^T \xi_t(j, k)}{\sum_{t=1}^T \gamma_t(j)} \tag{7}$$

$$\mu_j = \frac{\sum_{t=1}^T \xi_t(j) o_t}{\sum_{t=1}^T \xi_t(j)} \tag{8}$$

$$\sigma_j^2 = \frac{\sum_{t=1}^T \xi_t(j) (o_t - \mu_j) (o_t - \mu_j)^2}{\sum_{t=1}^T \xi_t(j)} \tag{9}$$

윗 식에서의 중간 변수  $\gamma$ 와  $\xi$ 는 다음의 식(10), (11), (12), (13)과 같이 표현된다.

$$\gamma_t(i, j) = P(s_t = i, s_{t+1} = j | O, \lambda) \tag{10}$$

$$\gamma_t(i) = P(s_t = i | O, \lambda) \tag{11}$$

$$\xi_t(i, k) = P(s_t = i, o_t = v_k | O, \lambda) \tag{12}$$

$$\xi_t(k) = P(o_t = v_k | O, \lambda) \tag{13}$$

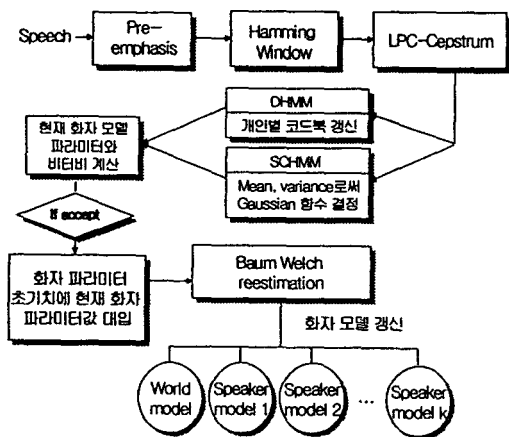


그림 2. 화자 모델 갱신 기법의 구성도  
Fig. 2. Block diagram of the speaker model updating technique.

이러한 화자 HMM 파라미터를 재추정하는데 있어서 파라미터 초기치를 현재 화자 HMM 파라미터로 선택하고 각 분기마다 새로 얻은 소량의 데이터를 비터비 계산(Viterbi scoring)에 의하여 인식되었을 경우에만 Baum-Welch re-estimation을 통해 적용시킴으로써 갱신된 모델을 얻는다. 그림 2는 화자 모델 갱신기법의 구성도를 보여준다.

3.2. 제안한 사전 문턱치 결정 기법

화자확인 시스템을 구현시에는 화자 확인을 위한 사전 문턱치를 결정하는 방법은 중요한 문제가 된다. 보통 화자 확인 연구에서 문턱치는 EER(equal error rate)을 갖도록 사후에 결정된다. 그러나, 실제환경에서 문턱치는 사후에 결정될 수 없으므로 화자 확인 전에 문턱치를 결정해서 보유하고 있어야 한다. 기존의 우도비(Likelihood ratio) 정규화 기법(월드 모델, 군중 모델기반)은 이론적으로 다른 배경에서 해석된다. 이러한 두 기법은 모델 갱신시 결합되면 더 좋은 화자 확인 성능을 내는 사전 문턱치를 결정할 수 있다. 예를 들어, 실험에서 한 화자 akm의 2 분기 모델 갱신을 위한 우도비 계산에서 월드 모델과의 우도비는 -24.693279, 자신의 화자 모델과의 우도비는 -8.609664가 나온다. 그리고, 군중 모델(타화자 모델)과의 우도비중 가장 큰 우도비는 -13.942381이다. 그러면 월드 모델 방법의 문턱치는 0.350082이 되고, 군중 모델방법의 문턱치는 0.617517이 된다. akm의 사후 문턱치(0.55)와는 차이가 있음을 알 수 있다. 따라서, 본 논문에서 제안하는 방식은 화자 모델 갱신 과정에서 분기별로 얻어진 소량의 학습 데이터를 비터비 계산(Viterbi scoring)을 통해서 월드 모델(world model)기법과 군중 모델(cohort model)기법의 하이브리드 형태의 사전 문턱치를 결정한다. 실험적으로 제안된 실험식(14)와 같이 모델 갱신에 쓰여진 데이터로 비터비 계산(Viterbi scoring)에 의한 우도비를 구한 후 평균을 구함으로써 사전 문턱치 R을 구하고, 구한 문턱치에 의해 확인 실험을 행하였다.

$$R = \frac{\sum_n R_{world}}{n} + \frac{\sum_n R_{hybrid}}{n^2} \tag{14}$$

$$\sum_n R_{hybrid} = \sum_n \text{MIN}_{j \neq i} P(O_M(S_{world, i}, W)) \tag{15}$$

여기서, n은 모델 갱신에 필요한 발성 횟수(본 논문에서는 5)이고,  $\sum_n R_{hybrid}$ 는 식 (15)와 같이 식 (3)과 식 (4)의 혼합 형태로서 월드 모델(world model)기반 우도비중 가장 큰 확률값을 갖는 다른 화자의 우도비이다. 본 논문에서 군중(cohort) 크기는 7이다. 그림 3에 제안하는 사전 문턱치 결정방식의 구성도를 나타내었다.

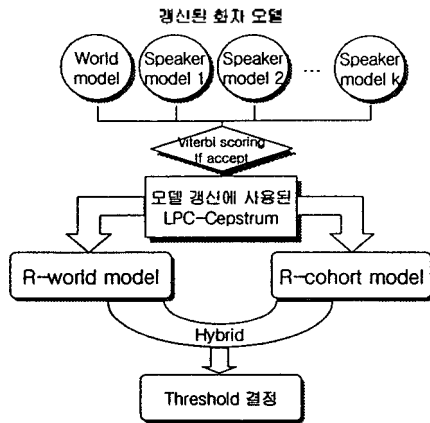


그림 3. 사전 문턱치 결정 기법의 구성도  
Fig. 3. Block diagram of the priori threshold decision.

#### IV. 실험 환경 및 결과

본 논문에서 사용된 단어는 13명의 남성화자에 의해 6개월 동안 분기당 30번씩 발음된 한 단어("안녕하세요")인데, 7명의 화자는 시스템의 고객으로써 쓰여졌고, 6명의 화자는 사칭자(impostor)로써 사용되었다. 실험은 7명의 고객이 발음한 10번의 단어로 HMM 학습 알고리즘에 의해 각각의 화자 모델을 만들었고, 7명의 고객 발음에 대해 고객들간의 오거부(FR)실험과 오수락(FA)실험, 그리고 6명의 사칭자 발음의 오수락(FA)실험을 분기별로 각각 행하였다. 또한, 월드 모델(world model)을 만드는 데에는 인식실험에 참여한 13명의 7명의 남성화자가 10번씩 발음한 200단어로 단어 모델을 만들었다. 모델 갱신에 필요한 데이터는 각 분기 초반에 수락된 5개의 데이터를 사용하여 해당 분기 모델 갱신에 사용하였다.

실험에 사용된 음성 데이터의 샘플링 주파수는 11.025kHz이고 256포인트(Point)의 음성 샘플을 한프레임(Frame)으로 하여 128포인트씩 shift하여 중첩되도록 해밍창(Hamming Windows)을 취한후 14차 LPC-켄스트럼을 구하였다. 이렇게 얻어진 14차 LPC-켄스트럼을 이용하여, 이산 HMM의 경우 코드북의 성능이 매우 중요하므로 k-means 알고리즘을 이용해 개인별 코드북 벡터를, 반연속 HMM의 경우는 한국어에 존재하는 모든 자음과 모음을 k-means 알고리즘을 이용해 64개의 코드워드 구성된 코드북 벡터를 생성하였다. 본연구에서는 실제 온라인 상에서 시스템 구현을 가능케 하기 위해 HMM기법 중 계산량이 적어 현실성이 있는 이산 HMM과 반연속 HMM알고리즘에 대해서만 실험하였다. HMM상태수(State number)는 이산 HMM과 반연속 HMM 모두 동일하게 6개로 하고, left-to-right모델을 적용하였다. 인식 실험에서는 로그함수(log function)로 계산 속도를 줄인 변형 비터비 알고리즘(Alternative Viterbi algorithm)이 사용되었다.

표 1에는 한 화자에 대한 이산 HMM과 반연속 HMM에서 기존 방법에 의한 사전 문턱치와 제안한 방법의 사전 문턱치, 그리고 사후 문턱치를 비교하였다. 실제 환경에서는 화자 모델 갱신시 갱신용 음성 정보와 사전 문턱치 결정을 위한 음성 정보를 따로 요구할 수 없으므로, 모델 갱신에 사용된 음성 정보를 다시 사전 문턱치 결정을 위해 사용하면 기존의 방법들은 사후 문턱치와는 상당한 차이가 있음을 알 수 있다.

표 1. 기존의 사전 문턱치 결정방법과 제안한 사전 문턱치 결정방법의 비교

Table 1. Comparison of the priori threshold between conventional and proposed technique.

		world model방법	cohort model방법	제안한 방법	사후 문턱치
11월	이산	0.453332	0.6623	0.69701	0.68
	반연속	0.350082	0.617517	0.556595	0.55
1월	이산	0.470321	0.612203	0.710256	0.72
	반연속	0.321107	0.620789	0.50721	0.54
2월	이산	0.402155	0.650843	0.694523	0.70
	반연속	0.40562	0.610439	0.54612	0.58

그림 4와 5에서는 모델 갱신 유무 및 사전, 사후 문턱치를 사용한 이산 HMM과 반연속 HMM의 오거부율을 비교하였다. 그림 4와 그림 5에서는 제안한 모델 갱신 방식과 사후 문턱치를 사용했을 때 이산 HMM의 경우, 오거부율이 월별로 19~21%, 반연속 HMM의 경우, 7~15% 가량 감소함을 볼 수 있다. 또한, 모델 갱신과 동시에 제안한 사전 문턱치 결정 방식을 사용했을 때 이산 HMM의 경우, 월별로 오거부율이 11~21%, 반연속 HMM의 경우, 14~15%가량 감소함을 알 수 있다.

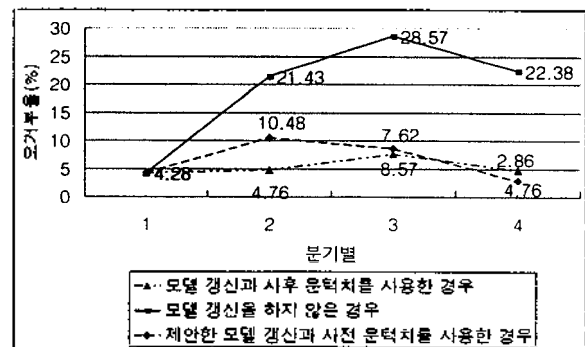


그림 4. 이산 HMM에서 모델 갱신 유무와 사전 사후 문턱치에 따른 오거부율 비교

Fig. 4. Comparison of the false rejection rate between conventional and proposed method with DHMM.

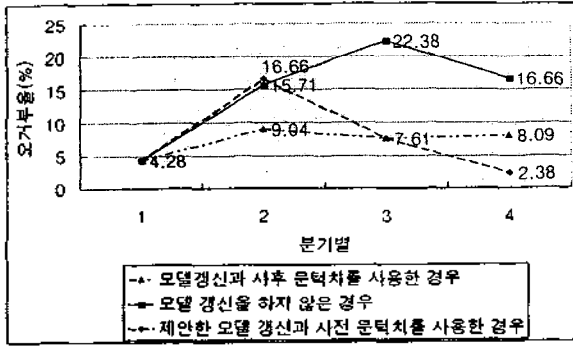


그림 5. 반연속 HMM에서 모델 갱신 유무와 사전 사후 문턱치에 따른 오거부율 비교

Fig. 5. Comparison of the false rejection rate between conventional and proposed method with SCHMM.

그림 6과 7에서는 모델 갱신 유무 및 사전, 사후 문턱치를 사용한 이산 HMM과 반연속 HMM의 오수락율을 비교하였다.

그림 6과 그림 7에서 보듯이 제안한 모델 갱신 방식과 사후 문턱치를 사용했을 때 이산 HMM의 경우, 오수락율이 월별로 1.9~2.8%, 반연속 HMM의 경우, 1.3~2.2% 가량 감소함을 볼 수 있다. 또한, 모델 갱신과 동시에 제안한 사전 문턱치 결정 방식을 사용했을 때 이산 HMM의 경우, 월별로 오수락율이 2.3~3.4% 감소, 반연속 HMM의 경우, 오히려 2%가량 증가함을 알 수 있다.

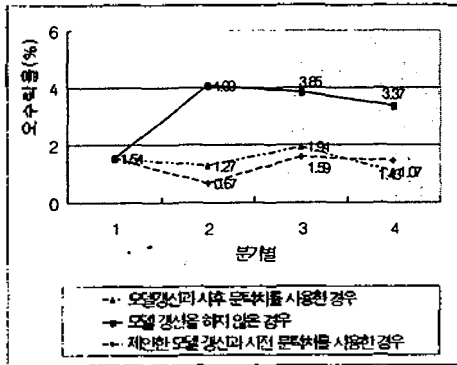


그림 6. 이산 HMM에서 모델 갱신 유무와 사전 사후 문턱치에 의한 오수락율 비교

Fig. 6. Comparison of the false acceptance rate between conventional and proposed method with DHMM.

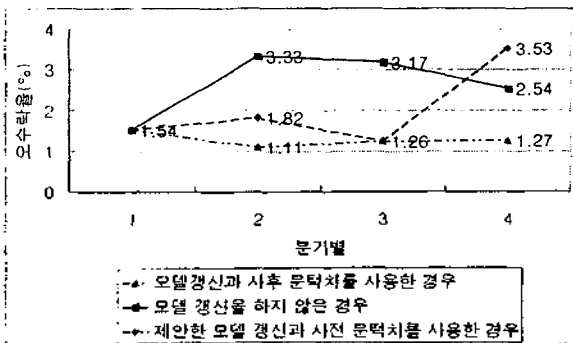


그림 7. 반연속 HMM에서 모델 갱신 유무와 사전 사후 문턱치에 따른 오거부율 비교

Fig. 7. Comparison of the false rejection rate between conventional and proposed method with SCHMM.

표 2, 3, 4에는 이산 HMM의 경우 월별 화자 오인식율을 그리고 표 5, 6, 7에는 반연속 HMM의 경우 월별 화자 오인식율을 각각 나타내었다.

표에서 보듯이 모델 갱신을 통해 월별로 발생하는 화자 내 변이에 적응하여 화자별 오인식율이 감소하고, 제안한 사전 문턱치에 의해 측정된 오인식율을 또한, 이산 HMM과 반연속 HMM 두가지 인식 알고리즘 모두에서 사후 문턱치에 의해 측정된 오인식율과 비슷한 성능을 보임을 확인할 수 있었다.

표 2. 이산 HMM에서 모델 갱신이 없는 월별 화자 오인식율  
Table 2. Verification error rates without model updating in DHMM

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	3/180+1/180	2/180+1/180	2/180+1/180
	FR	0/30	10/30	15/30	10/30
myj	FA	2/180+0/180	2/180+0/180	3/180+0/180	4/180+0/180
	FR	6/30	2/30	4/30	3/30
jjm	FA	0/180+7/180	4/180+7/180	5/180+7/180	1/180+7/180
	FR	0/30	14/30	20/30	12/30
jsd	FA	2/180+16/180	30/180+16/180	23/180+16/180	19/180+16/180
	FR	0/30	11/30	8/30	5/30
jsh	FA	1/180+1/180	5/180+1/180	12/180+1/180	10/180+1/180
	FR	3/30	4/30	5/30	6/30
kds	FA	0/180+6/180	4/180+6/180	3/180+6/180	2/180+6/180
	FR	0/30	3/30	5/30	5/30
kwk	FA	0/180+7/180	17/180+7/180	11/180+7/180	9/180+7/180
	FR	0/30	1/30	3/30	6/30
총합계	FA	1.71% (43/2520)	4.09% (103/2520)	3.85% (97/2520)	3.37% (85/2520)
	FR	4.28% (9/210)	21.43% (45/210)	28.57% (60/210)	22.38% (47/210)

표 3. 이산 HMM에서 제안한 모델 갱신과 사후문턱치 방법에 의한 월별 화자 오인식율

Table 3. Verification error rates when using proposed model updating and posteriori threshold in DHMM.

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	3/180+6/180	2/180+1/180	0/180+3/180
	FR	0/30	2/30	5/30	0/30
myj	FA	2/180+0/180	4/180+0/180	0/180+0/180	0/180+0/180
	FR	6/30	3/30	0/30	2/30
jjm	FA	0/180+7/180	3/180+6/180	4/180+7/180	1/180+3/180
	FR	0/30	1/30	6/30	0/30
jsd	FA	2/180+16/180	3/180+1/180	6/180+8/180	0/180+9/180
	FR	0/30	0/30	0/30	0/30
jsh	FA	1/180+1/180	0/180+1/180	5/180+1/180	2/180+1/180
	FR	3/30	0/30	5/30	5/30
kds	FA	0/180+6/180	0/180+2/180	0/180+5/180	0/180+6/180
	FR	0/30	2/30	0/30	0/30
kwk	FA	0/180+7/180	0/180+3/180	3/180+7/180	0/180+2/180
	FR	0/30	2/30	0/30	3/30
총합계	FA	1.54% (39/2520)	1.27% (32/2520)	1.94% (49/2520)	1.07% (27/2520)
	FR	4.28% (9/210)	4.76% (10/210)	7.62% (16/210)	4.76% (10/210)

표 4. 이산 HMM에서 제안한 모델 갱신과 사전 문턱치 기법에 의한 월별 화자 오인식율

Table 4. Verification error rates when using proposed model updating and priori threshold in DHMM.

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	2/180+1/180	2/180+1/180	0/180+3/180
	FR	0/30	5/30	5/30	0/30
myj	FA	2/180+0/180	4/180+0/180	0/180+1/180	0/180+0/180
	FR	6/30	3/30	0/30	2/30
ijm	FA	0/180+7/180	5/180+2/180	4/180+7/180	0/180+6/180
	FR	0/30	2/30	6/30	0/30
jsd	FA	2/180+16/180	1/180+0/180	6/180+8/180	0/180+9/180
	FR	0/30	3/30	0/30	0/30
jsh	FA	1/180+1/180	0/180+0/180	0/180+1/180	0/180+5/180
	FR	3/30	1/30	5/30	2/30
kds	FA	0/180+6/180	0/180+2/180	0/180+6/180	0/180+6/180
	FR	0/30	2/30	0/30	0/30
kwk	FA	0/180+7/180	0/180+0/180	0/180+4/180	0/180+7/180
	FR	0/30	6/30	2/30	2/30
총합계	FA	1.54% (39/2520)	0.67% (17/2520)	1.59% (40/2520)	1.43% (36/2520)
	FR	4.28% (9/210)	10.48% (22/210)	8.57% (18/210)	2.86% (6/210)

표 5. 반연속 HMM에서 모델 갱신 없는 월별 화자 오인식율  
Table 5. Verification error rates without model updating in SCHMM.

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	8/180+0/180	0/180+0/180	1/180+1/180
	FR	0/30	7/30	0/30	4/30
myj	FA	0/180+0/180	5/180+1/180	0/180+0/180	0/180+5/180
	FR	3/30	1/30	0/30	0/30
ijm	FA	0/180+7/180	0/180+2/180	2/180+1/180	0/180+0/180
	FR	1/30	3/30	5/30	1/30
jsd	FA	0/180+16/180	8/180+0/180	5/180+1/180	11/180+0/180
	FR	0/30	2/30	3/30	11/30
jsh	FA	1/180+1/180	1/180+0/180	3/180+16/180	1/180+1/180
	FR	1/30	2/30	1/30	1/30
kds	FA	0/180+6/180	0/180+0/180	0/180+1/180	0/180+0/180
	FR	4/30	0/30	6/30	0/30
kwk	FA	0/180+7/180	0/180+3/180	0/180+3/180	0/180+12/180
	FR	0/30	4/30	1/30	0/30
총합계	FA	1.54% (39/2520)	1.11% (28/2520)	1.26% (32/2520)	1.26% (32/2520)
	FR	4.28% (9/210)	9.04% (19/210)	7.61% (16/210)	8.09% (17/210)

표 6. 반연속 HMM에서 제안한 모델 갱신과 사후 문턱치로 측정된 화자 오인식율

Table 6. Verification error rates when using proposed model updating and posteriori threshold in SCHMM.

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	9/180+1/180	2/180+1/180	2/180+1/180
	FR	0/30	10/30	15/30	0/30
myj	FA	0/180+0/180	0/180+0/180	0/180+0/180	4/180+0/180
	FR	3/30	0/30	0/30	15/30
ijm	FA	0/180+7/180	0/180+7/180	5/180+7/180	1/180+7/180
	FR	1/30	17/30	23/30	12/30
jsd	FA	0/180+16/180	31/180+16/180	23/180+16/180	19/180+16/180
	FR	0/30	0/30	0/30	0/30
jsh	FA	1/180+1/180	5/180+1/180	12/180+1/180	0/180+1/180
	FR	1/30	1/30	3/30	5/30
kds	FA	0/180+6/180	0/180+6/180	0/180+6/180	0/180+6/180
	FR	4/30	3/30	6/30	0/30
kwk	FA	0/180+7/180	1/180+7/180	0/180+7/180	0/180+7/180
	FR	0/30	2/30	0/30	3/30
총합계	FA	1.54% (39/2520)	3.33% (84/2520)	3.17% (80/2520)	2.54% (64/2520)
	FR	4.28% (9/210)	15.71% (33/210)	22.38% (47/210)	16.66% (35/210)

표 7. 반연속 HMM에서 제안한 모델 갱신과 사전 문턱치 방법으로 측정된 화자 오인식율

Table 7. Verification error rates when using proposed model updating and priori threshold in SCHMM.

화자	월	9월	11월	1월	2월
akm	FA	0/180+1/180	8/180+0/180	1/180+0/180	21/180+13/180
	FR	0/30	6/30	0/30	0/30
myj	FA	0/180+0/180	3/180+1/180	3/180+1/180	0/180+8/180
	FR	3/30	3/30	0/30	0/30
ijm	FA	0/180+7/180	0/180+1/180	1/180+1/180	0/180+0/180
	FR	1/30	6/30	5/30	1/30
jsd	FA	0/180+16/180	1/180+0/180	19/180+5/180	22/180+1/180
	FR	0/30	9/30	0/30	2/30
jsh	FA	1/180+1/180	0/180+0/180	0/180+1/180	1/180+0/180
	FR	1/30	10/30	2/30	2/30
kds	FA	0/180+6/180	4/180+18/180	1/180+5/180	0/180+0/180
	FR	4/30	0/30	4/30	0/30
kwk	FA	0/180+7/180	0/180+10/180	0/180+2/180	3/180+20/180
	FR	0/30	1/30	1/30	0/30
총합계	FA	1.54% (39/2520)	1.82% (46/2520)	1.26% (32/2520)	3.53% (89/2520)
	FR	4.28% (9/210)	16.66% (35/210)	7.61% (16/210)	2.38% (5/210)

## V. 결론 및 고찰

실제 화자 확인 시스템에서는 장기간 음성의 변이에 적응을 위한 모델 갱신과 사전 문턱치가 매우 중요하다. 본 논문에서는 모델 갱신을 위해 현재 HMM파라미터와 분기별 소량의 데이터로써 화자 모델을 갱신하는 방법과 동시에 사전 문턱치를 결정하는 방법을 제안하였다. 실험 결과, 이산HMM에서는 모델 갱신을 하지 않았을 경우보다 제안하는 방법을 적용하였을 경우 19~21%의 오거부를 향상과 1.9~2.8%의 오수락을 향상을 확인하였고, 제안한 사전 문턱치 결정방법에 의한 인식율은 11~21%의 오거부를 향상과 2.3~3.4%의 오수락을 향상을 확인하였다. 그리고, 반연속HMM에서는 모델 갱신을 하지 않았을 경우보다 제안하는 방법을 적용하였을 경우가 7~15%의 화자 오거부와 1.2~2.2%의 화자 오수락을 향상을 가져오음을 확인하였고, 제안한 사전 문턱치 결정방법에 의한 인식율은 14~15%가량의 오거부를 향상, 약 2%의 오수락을 증가를 확인하였다.

제안하는 방법은 화자 확인 시스템에서 문제시되는 화자 내의 변이에 적응하고 사전 문턱치를 결정함으로써 실질적인 화자 확인 시스템 구현에 적용될 수 있을 것으로 생각된다.

## 참고 문헌

1. S. Furui, "Cepstral analysis technique for automatic speaker verification," IEEE Trans. ASSP-29, 2, pp.254-272, 1981.
2. A. Setlur and T. Jacobs, "Results of a speaker verification service trial using HMM models," Proc. Eurospeech, pp. 1-53-56, 1995.
3. Tomoko Matsui, Kiyooki Aikawa, "Robust model for speaker verification against session-dependent utterance variation," Proc. ICASSP, pp. 1-117-120, 1998.
4. Sadaoki Furui, "Recent advances in speaker recognition," Pattern Recognition Letters 18, 1997.
5. Lawrence Rabiner, Biing-Hwang Juang, *Fundamentals of Speech Recognition*, Prentice Hall International Inc, pp.321-389, 1993.
6. X.D.Huang, "Semi-continuous hidden markov models for speech recognition," Ph.D. thesis, Department of Electrical Engineering, University of Edinburgh, 1989.
7. Fredouille, C., Bonastre, Merlin, T. "Similarity normalization method based on World model and a posteriori probability for speaker verification", Eurospeech pp. II-983-986, 1999.
8. Yong Gu, Trevor Thomas, "A hybrid score measurement for HMM-based speaker verification," Proc. ICASSP, vol. 1-317-320 1999.
9. Toshihiro Isobe, Jun-ichi Takahashi. "A new cohort normalization using local acoustic information for speaker verification," Proc. ICASSP, pp. II-841-844, 1999.

### ▲ 진 세 훈(Se Hoon Chin)



1999년 2월 : 광운대학교 전자통신공  
학과 졸업(학사)

1999년 3월~현재 : 광운대학교 전자  
통신공학과(석사과정)

※ 주 관심분야 : 패턴 인식, 음성신  
호처리

### ▲ 이재희(Jae Hee Lee)

한국음향회지 제 19권 4호 참조

### ▲ 강철호(Chul Ho Kang)

한국음향회지 제 17권 8호 참조