

## 화자확인에서 일정한 결과를 얻기 위한 빠른 순시 확률비 테스트 방법

### Fast Sequential Probability Ratio Test Method to Obtain Consistent Results in Speaker Verification

김 은 영<sup>1)</sup> · 서 창 우<sup>2)</sup> · 전 성 채<sup>3)</sup>

Kim, Eunyoung · Seo, Changwoo · Jeon, Sungchae

#### ABSTRACT

A new version of sequential probability ratio test (SPRT) which has been investigated in utterance-length control is proposed to obtain uniform response results in speaker verification (SV). Although SPRTs can obtain fast responses in SV tests, differences in the performance may occur depending on the compositions of consonants and vowels in the sentences used. In this paper, a fast sequential probability ratio test (FSPRT) method that shows consistent performances at all times regardless of the compositions of vocalized sentences for SV will be proposed. In generating frames, the FSPRT will first conduct SV test processes with only generated frames without any overlapping and if the results do not satisfy discrimination criteria, the FSPRT will sequentially use frames applied with overlapping. With the progress of processes as such, the test will not be affected by the compositions of sentences for SV and thus fast response outcomes and even consistent performances can be obtained. Experimental results show that the FSPRT has better performance to the SPRT method while requiring less complexity with equal error rates (EER).

**Keywords:** speaker verification (SV), sequential probability ratio test (SPRT), frame overlap, Gaussian Mixture model (GMM).

#### 1. 서론

생체 신호처리(biometrics signal processing) 방법은 생리학적(physiological) 또는 행동 특성(behavioral characteristics)을 이용해서 사용자를 인증(authentication) 하는 유익하면서도 신뢰할 수 있는 접근 방법이다[1]. 개인의 독특한 특징을 이용하는 생체 인식 방법은 키나 패스워드와 같은 전통적인 방법보다 높은 보안(security)을 제공할 수 있다. 이러한 생체 식별 시스템(biometric identification system)은 추출된 특징에 따라서 세 가지로 분류할 수 있다[2]. 첫 번째는 정적 특징(static feature)으로 얼굴이나 지문과 같이 한번 추출된 특징이 변하지 않는 것을 말한다. 두 번째는 동적 특징(dynamic feature)으로 사람의 목소리나 사인과

같이 시간이나 환경에 따라서 변할 수 있는 경우이다. 마지막으로 화학-물리적(chemical-physical) 특성들을 들 수 있다. 이들 생체 인식 방법 중에서 사람의 목소리를 이용한 화자인식(speaker recognition: SR) 방법은 가장 자연스러움과 편리성 그리고 비용이 적게 요구되기 때문에 널리 연구되고 있다[3].

최근 목소리를 이용해서 사용자를 인증하는 SR 방법에서 화자확인(speaker verification SV) 방법이 활발하게 연구되어 왔다. 그러나 현재 연구되고 있는 SV 방법은 생성된 데이터에 대해서 모든 데이터를 입력 받아서 사용자의 유사도(likelihood)를 계산한 후에 결정 판별값(decision threshold)과 비교하는 방법이다. 이런 접근은 모바일 기기와 같은 낮은 프로세스 속도와 작은 메모리를 사용할 경우 많은 계산 시간을 요구하는 문제점이 있다[4].

이런 문제점을 해결하고 빠른 응답 결과를 얻기 위해서 순시 확률비 테스트(sequential probability ratio test: SPRT) 방법이 SV에 적용되었다[5, 6]. SPRT는 통계적인 순시 결정(statistical sequential decision) 방법으로 SV에서 빠른 응답과 성능 개선을 위한 효과적인 접근방법이다. SV에 SPRT를 적용할 경우 몇 가지의 잇점이 보고되었다. (i)이론적으로 확인 오류율(error rate)을 제어할 수 있고, (ii)주어진 기준 성능에서 요구되는 입력 샘플

1) 숭실대학교 key710503@ssu.ac.kr, 교신저자

2) 숭실대학교 cwseo@ssu.ac.kr

3) 숭실대학교 sarim@keri.re.kr

접수일자: 2010년 1월 29일

수정일자: 2010년 2월 15일

게재결정: 2010년 3월 5일

플(input sample)의 수가 어떤 다른 방법보다 작으며, (iii)화자 종속 적용 처리를 쉽게 실현할 수 있다[7, 8]. SPRT에서 결정 과정을 수행하기 위해서는 입력 샘플들이 확률 밀도 함수(probability density function: PDF)로부터 상호 독립적으로 동일하게 분포되어 있다(independent and identically distributed: *i.i.d.*)고 가정한다[9]. 그러나 SV에 SPRT를 적용하는 과정에는 사용되는 데이터에 따라서 두 가지의 문제점이 발생할 수 있다. 첫 번째는 이론적으로 발성 또는 문장 길이에 종속적이기 때문에 짧은 발성에서는 테스트 데이터의 부족에 의해서 제한된 성능을 가질 수 있다. 이런 문제를 해결하기 위해서 확장된 순시 확률비 테스트(extended sequential probability ratio test: ESPRT) 방법이 제안되었다[10]. ESPRT는 처음 확인 과정에서는 SPRT를 수행하고 그리고 만약 유사도 값이 결정 판별에 미치지 못할 경우에 프레임 시프트를 통해서 새로운 테스트 데이터를 생성하여 결정 과정을 반복적으로 수행하는 방법이다. 두 번째는 사용된 문장의 구성에서 자음과 모음의 조합에 의해서 SV의 성능에 영향을 미칠 수 있다. 즉 SPRT를 수행하는 첫 번째 단계의 테스트 데이터에 자음성분이 많이 포함될 경우 성능이 떨어질 수 밖에 없다는 것이다.

본 논문은 위의 두 번째 문제점을 해결하기 위한 방법으로 SV에 사용된 문장의 구성에 상관없이 항상 일정한 성능을 얻을 수 있는 효과적인 방법을 제안한다. 제안된 빠른 순시 확률비 테스트(fast sequential probability ratio test: FSPRT) 방법은 SV를 위한 프레임 생성에서 처음에는 오버랩 없이 생성된 프레임만으로 테스트 과정을 수행한 후 판별 기준에 미치지 못할 경우에 오버랩이 적용된 프레임을 순차적으로 사용하는 방법이다. 따라서 이렇게 진행할 경우 테스트 과정에서 사용된 문장의 자음과 모음의 구성에 상관없이 항상 일정한 성능 결과를 얻을 수 있을 뿐만 아니라 기존의 SPRT보다 더 빠른 응답 결과를 얻을 수 있다. 제안된 방법의 우수성을 확인하기 위해서 SPRT, ESPRT, 그리고 일반적인 방법을 SV의 등가오류율(equal error rate: EER)에서 비교 실험의 결과로 적용하였다.

논문은 다음과 같이 구성되었다. 2장에서는 통계적인 순시 결정 방법들을 설명하였다. 3장에서는 제안된 FSPRT 방법을 기술하였다. 그리고 4장과 5장에서는 실험 결과 및 결론을 서술하였다.

## 2. 통계적인 순시 결정 방법 (statistical sequential decision)

### 2.1 순시 확률비 테스트

(sequential probability ratio test: SPRT)

일반적으로 연속적인 특징벡터에는 상관성이 존재 하지만, 순시 확률비 테스트(SPRT)를 적용하는 SV에서는 상관성이 없는 즉 특징벡터들 사이에는 상호 독립이라고 가정한다. 입력 음

성으로부터 얻어진 특징벡터를  $1, 2, 3, \dots, F, \dots, T-1, T$  라 하고  $s(x_t)$  와  $b(x_t)$  를 Gaussian Mixture model(GMM)을 통해서 얻어진 화자(speaker)와 모든 사칭자(imposter)들의 확률밀도함수라 하자. 특징벡터  $x_t$  가 주어지고 그리고 상호 독립이라고 가정할 때, SPRT를 위한 유사도 비(likelihood ratio)는 다음과 같이 구할 수 있다[9, 11].

$$L_m(x_1, \Lambda, x_F) = \prod_{t=1}^F \frac{p_s(x_t)}{p_b(x_t)} \tag{1}$$

유사도 비를 이용한 결정 판별은 다음과 같이 나타낼 수 있다.

$$L_m(x_1, \Lambda, x_F) > Th_a \quad : \text{ 화자로 인증,} \tag{2}$$

$$L_m(x_1, \Lambda, x_F) < Th_r \quad : \text{ 사칭자로 거절} \tag{3}$$

여기서  $Th_a$  와  $Th_r$  는  $a > Th_r$  조건을 만족시키는 판별 값이다. 여기서 만약 유사도 값이  $Th_a$  와  $Th_r$  사이에 속한다면 또 다른 특징벡터  $x_{F+1}$  가 선택되고 그리고 결정 판별은 반복된다.

판별값  $Th_a$  와  $Th_r$  는 오거절율(false rejection rate: FRR)  $E_r$  과 오인증율(false acceptance rate: FAR)  $E_a$  사이에서 다음 관계를 가진다[5].

$$Th_a \leq \frac{1 - E_r}{E_a} \tag{4}$$

$$Th_r \geq \frac{E_r}{1 - E_a} \tag{5}$$

### 2.2 확장된 순시 확률비 테스트

(extended sequential probability ratio test: ESPRT)

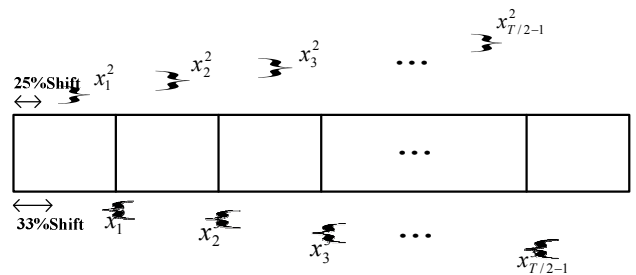


그림 1. 샘플 시프트를 이용한 확장된 순시 확률비 테스트 방법

Figure 1. ESPRT method using the sample shift

SPRT에서 T개의 전체 데이터에 대한 최종 유사도 비  $L_m(x_1, \Lambda, x_F, \Lambda, x_T)$  가  $Th_a$  와  $Th_r$  사이에 속할 경우 테스트 데이터의 부족에 의한 성능 제한을 받을 수 있다. 이런 문제를 해결하기 위해서 확장된 순시 확률비 테스트(ESPRT) 방법이 제안되었다. <그림 1>은 ESPRT에서 테스트 데이터를 위한 프레임의

시작점을 25%와 33%의 샘플 시프트를 사용하여 새로운 테스트 데이터를 생성하는 방법이다. 즉, 길이가  $1/2 - 1$  인 두 셋의 테스트 데이터  $x_2^2$  와  $x_3^3$  를 생성할 수 있다. 첫 번째 프레임에서 샘플 시프트 비  $S_i$  는 한 프레임에서의 샘플의 수  $N$  과 반복 횟수  $I$  사이의 관계에 의해서 결정된다[10].

$$S_i = \frac{N(i-1)}{2 \cdot i}, \quad i=1, \Lambda, I \quad (6)$$

여기서  $I = 2$  와  $3$  은 첫 번째 시작 프레임에서 각각 25%와 33%의 샘플 시프트를 나타낸 것이다. ESPRT를 수행하기 위한 첫 번째 단계부터 최종 테스트 데이터  $x_T$  까지는 SPRT와 동일한 과정으로 진행된다.

$$L_m(x_1, \Lambda, x_F, \Lambda, x_T) = \prod_{i=1}^T \frac{p_s(x_i)}{p_b(x_i)} \quad (7)$$

여기서,  $\Lambda, x_F, \Lambda, x_T$  이  $h_a$  보다 크거나  $h_r$  보다 작다면 테스트 과정은 종료되지만, 만약  $(x_1, \Lambda, x_F, \Lambda, x_T) < Th_a$  이면 <그림 1>에서와 같이 또 다른 특징벡터가 선택되고 그리고 다음 단계를 위한 결정에 적용된다.  $\geq 2$  경우에, 각 샘플 시프트에서 결정을 위한 유사도 비는 다음과 같이 결정된다.

$$L_m(x_1^i, \Lambda, x_{T-1}^i) = \prod_{i=1}^{T/2-1} \frac{p_s(x_i^i)}{p_b(x_i^i)} \quad (8)$$

만약  $\sum_{i=1}^I L_m(x_i^i)$  이  $h_a$  와  $h_r$  사이에 빠진다면, 최종 판별 값  $h_f$  이 결정되고 최종 결정 판별은 다음과 같이 이루어진다.

$$\frac{1}{I} \sum_{i=1}^I L_m(x_i^i) \geq Th_f \quad : \text{화자로 인증}, \quad (9)$$

$$\frac{1}{I} \sum_{i=1}^I L_m(x_i^i) < Th_f \quad : \text{사칭자로 거절} \quad (10)$$

이와 같은 방법으로 진행할 경우, 새로운 테스트 데이터를 생성해서 사용할 수 있기 때문에 일반적으로 성능향상을 기대할 수 있다.

### 3. 빠른 순시 확률비 테스트

(fast sequential probability ratio test: FSPRT)

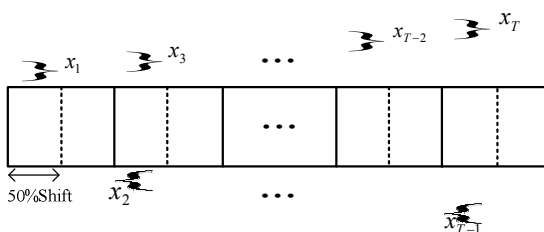


그림 2. 프레임 오버랩을 이용한 빠른 순시 확률비 테스트 방법.

Figure 2. FSPRT method using the frame overlap

본 장에서는 SPRT와 ESPRT에서 제기된 문제점을 해결하면서 균일한 응답 결과를 얻을 수 있는 빠른 순시 확률비 테스트 (FSPRT)를 제안한다. 제안된 FSPRT에서 첫 번째 판별을 위한 프레임은 <그림 2>의 실선으로 표시된 프레임간 오버랩(본 논문에서는 50% 적용) 적용 없이 생성된 프레임 =  $1, 3, 5, K, T - 2, T$  만을 다음과 같이 사용하였다.

$$L_m(x_1, x_3, \Lambda, x_{T-2}, x_T) = \prod_{i=1}^T \frac{p_s(x_i)}{p_b(x_i)}, \quad t=1, 3, \Lambda, T-2, T \quad (11)$$

여기서,  $x_3, \Lambda, x_{T-2}, x_T$  이  $h_a$  보다 크거나  $h_r$  보다 작다면 테스트는 종료되지만, 그렇지 않을 경우 <그림 2>의 점선으로 표시된 프레임에서 오버랩이 적용된 첫 번째 프레임  $x_2$  가 선택되고 판별은 다음과 같이 계속된다.

$$L_m(x_1, x_3, \Lambda, x_{T-2}, x_T, x_2) = L_m(x_1, x_3, \Lambda, x_{T-2}, x_T) \cdot \frac{p_s(x_2)}{p_b(x_2)} \quad (12)$$

만약  $x_3, \Lambda, x_{T-2}, x_T, x_2$  의 유사도 비가 판별 기준에 미치지 못하면, 새로운 프레임을 적용하여 계속 판별을 시도한다. 이때 최종 프레임까지 적용하여 판별이 결정되지 않으면, 일반적인 방법과 동일하게 최종 판별치  $h_f$  가 선택되고 다음과 같이 결정된다.

$$\dots, x_{T-2}, x_T, x_2, \Lambda, x_{T-1}) \geq Th_f \quad : \text{화자로 인증}, \quad (13)$$

$$\dots, x_{T-2}, x_T, x_2, \Lambda, x_{T-1}) < Th_f \quad : \text{사칭자로 거절} \quad (14)$$

이와 같은 방법으로 진행할 경우, FSPRT의 첫 번째 결정 판별에서 전체 프레임 개수의 반만을 사용하였지만 테스트 데이터를 위한 문장의 시작부터 끝까지 다 적용한 것과 같은 효과를 가질 수 있다.

### 4. 실험 결과

제안된 방법에 대한 우수성을 확인하기 위해서 제안된 (FSPRT), SPRT[9], ESPRT[10], 그리고 일반적인 방법[12]들을 SV에서 비교 실험하였다. 수집한 데이터는 한국어 고립 단어 숫자음 “영(Zero), 일(II), 이(Li), 삼(Sam), 사(Sa), 오(Oh), 육(Yook), 칠(Chil), 팔(Pal), 구(Gu)”와 발성이 짧은 연속음 문장인 “열려라 참깨(Open sesame)” 그리고 발성길이가 비교적 긴 문장인 “무궁화 꽃이 피었습니다(Roses of Sharon have blossomed)”를 사용하였다[13].

실험에 사용된 음성 데이터는 대학원 실험실에서 200명의 화자(남·여 각각 100명)로부터 획득하였고, 개인별 화자의 데이터는 주 단위로 3주 동안 15개(주 5문장)의 데이터를 추출하였다. 처음 2주 동안 수집한 10개의 데이터를 학습(training) 과정

에 사용하였고, 마지막 주에 수집한 나머지 5개의 데이터를 테스트(test) 과정에 사용하였다. 따라서 SV의 테스트에서 사용된 오거절율(FRR)과 오인증을(FAR)을 위한 문장은 각각 1, 000 (=200×5)와 199, 000 (=200×199×5)이다.

음성의 분석과정에서 샘플링 주파수는 11.025kHz, 16bit 분해능, 그리고 음성분석 프레임은 16ms를 적용하였다. 특징벡터는 12차 mel frequency cepstral coefficient(MFCC)[12]와 12차 델타 켈스트럼(delta cepstrum)[14]을 사용하였다. <그림 3>은 실험 데이터 “열려라 참깨”를 화자 식별(speaker identification: SI)에서 각 프레임의 오버랩 정도에 따른 실험 결과를 나타낸 것이다. 그림에서와 같이 오버랩이 0%와 25%를 적용한 경우보다 50%와 75%를 사용하였을 때가 훨씬 더 좋은 결과를 보였다. 특히 GMM에서의 혼합성분(mixture) 개수가 증가되었을 때, 오버랩 정도가 낮은 경우에는 입력 데이터의 부족에 의해서 성능이 줄어들고 있음을 알 수 있었다. 그리고 높은 중첩을 적용한 경우에는 성능이 일정하게 향상되었다. 따라서 본 논문을 위한 실험에서는 계산량과 성능을 고려해서 50% 중첩을 적용하였다.

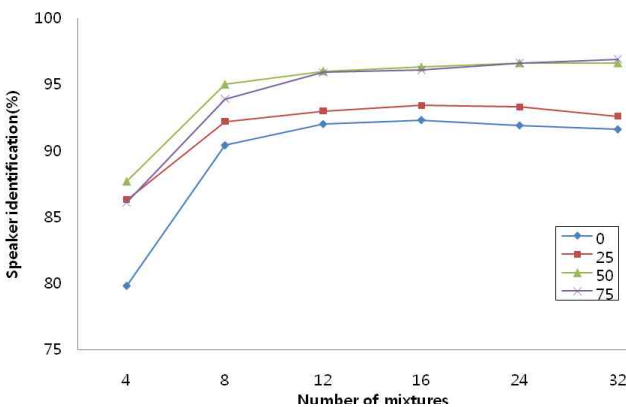


그림 3. 프레임 오버랩에 따른 화자식별 성능  
Figure 3. SI performance through frame overlap

실험에서 SV를 위한 사칭자 배경 모델(universal background model: UBM) [15, 16]을 위한 데이터는 학습과 테스트에 참여하지 않은 50명의 화자로부터 발생된 개인별 10개의 문장-독립(text-independent) 데이터 500(=50×10)를 이용하여 모델을 구성하였다. 사칭자 배경 모델은 250개의 혼합성분으로 구성되었고, 초기치 추정에는 벡터 양자화(vector quantization: VQ) 클러스터링을 이용하였다[17]. 결정 판별을 위한  $h_a$  와  $h_r$  는 로그 유사도(log likelihood) 값이 음수를 나타내기 때문에 각각 0.73과 1.28를 학습과정에서 구해진 로그 유사도를 기준으로 구한 최종 판별치  $h_f$  에 곱한 값을 적용하였다.

<표 1>은 실험 데이터 “무궁화 꽃이 피었습니다”를 이용하여 계산량을 비교한 것이다. 비교 실험은 각 방법에서 사용된 프레임 비를 일반적인 방법을 기준(100%)으로 비교하였다. 실험 결과로부터 SPRT와 FSPRT는 일반적인 방법보다 13.7%와 13.9% 작은 프레임을 요구하였다. 그렇지만, ESPRT의 두 방법

표 1. 사용된 프레임 비에 따른 계산량 비교 결과

Table 1. Comparison results of the computational complexity through used frame ratio

Methods	Frame ratio (%)
Conventional	-
SPRT	-13.7
ESPRT(I=2)	+6.2
ESPRT(I=3)	+8.3
FSPRT	-13.9

(I=2, 3)은 각각 6.2%와 8.3%가 증가하였다.

<그림 4>는 <표 1>에서 사용된 프레임 비를 기반으로 EER에서의 성능을 비교한 것이다. 그림에서와 같이 계산량이 증가한 ESPRT의 두 방법(I=3, 2)이 가장 좋은 성능을 보였다. 그리고 SPRT와 FSPRT는 사용된 프레임 비의 감소에 의해서 일반적인 방법보다 평균 0.2% 정도의 성능감소를 보였다.

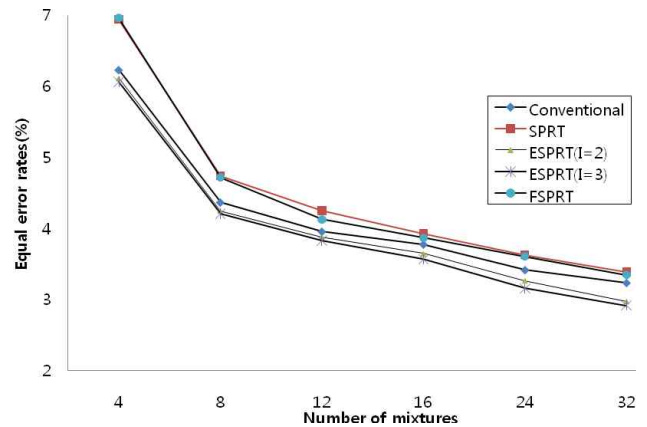


그림 4. 화자확인 성능의 비교 결과  
Figure 4. Comparison results of SV performance

마지막 실험에서는 제안된 방법의 우수성을 확인하기 위해서 발생 문장의 모음과 자음의 조합에 따른 성능 및 계산량을 비교하기 하였다. 실험에 사용된 데이터는 고립 단어 숫자음을 자음 성분이 많은 데이터 순{삼, 칠, 팔, 십, 영, 일, 사, 육, 이, 오}으로 나열하여 실험을 수행하였다. 실험 결과로부터 FSPRT, SPRT, 일반적인 방법, ESPRT(2), 그리고 ESPRT(3)는 순으로 프레임 사용이 증가되었다. 또한 이전 실험에 사용된 프레임 비와 비교했을 때, FSPRT는 0.8% 증가였지만 SPRT, ESPRT(2), 그리고 ESPRT(3)는 훨씬 더 많은 4.4%, 3.3%, 그리고 2.1%가 증가하였다.

표 2. 사용된 프레임 비에 따른 계산량 비교 결과

Table 2. Comparison results of the computational complexity through used frame ratio

Methods	Frame ratio (%)
Conventional	-
SPRT	-9.3

ESPRT(I=2)	+9.5
ESPRT(I=3)	+10.4
FSPRT	-13.1

<그림 5>는 <표 2>의 프레임 비를 기반으로 EER에서의 성능을 나타낸 것이다. 그림에서와 같이 가장 좋은 인식률을 보인 것은 사용된 프레임 비가 거의 10% 정도 증가한 ESPRT의 두 방법(I=2, 3)이지만, 자음 성분을 많이 포함한 데이터를 첫 번째 단계(시작 프레임 순으로 50% 적용)에 사용하였기 때문에 일반적인 방법과 거의 비슷한 결과를 보였다. FSPRT와 SPRT는 일반적인 방법보다 13.1%와 9.3% 정도 사용된 프레임 비의 감소에 의해서 각각 0.21%와 0.36% 낮은 결과를 보였다. 제안된 FSPRT와 SPRT를 비교했을 때, FSPRT는 사용된 프레임에서 SPRT보다 약 4%정도 작지만 인식률은 오히려 0.15% 정도 좋은 결과를 보였다. 실험 결과로부터 FSPRT가 SPRT와 ESPRT보다 항상 균일한 성능을 얻을 수 있는 것은 첫 번째 단계에서 데이터의 전구간에 걸친 프레임을 사용하였기 때문이다. 그리고 SPRT와 ESPRT는 첫 번째 단계에서 자음 성분이 많이 포함된 데이터가 사용되었기 때문에 성능 감쇠가 발생하였다.

따라서 제안된 방법이 SV를 위한 발성 문장의 구성에 영향을 받지 않고 항상 일정한 결과를 얻을 수 있는 효과적인 방법이라 할 수 있다.

### 5. 결론

본 논문에서는 화자확인(speaker verification: SV)에서 빠른 응답 결과를 얻을 수 있을 뿐만 아니라 발성 문장의 구성에 상관없이 항상 일정한 성능을 얻기 위한 효과적인 방법을 제안하였다. 제안된 빠른 순시 확률비 테스트(fast sequential probability ratio test: FSPRT)는 프레임 생성에서 처음에는 오버랩 없이 생성된 프레임만으로 테스트를 수행한 후 판별 기준에 미치지 못할 경우에 오버랩이 적용된 프레임을 순차적으로 사용하는 방법이다. 제안된 FSPRT와 SPRT를 한국어 고립단어 숫자음 데이터에 적용하여 비교했을 때, 요구되는 프레임이 4% 작으면서 향상된 결과를 보였다.

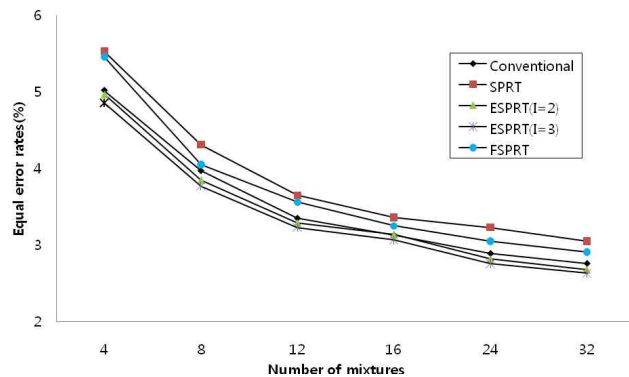


그림 5. 화자확인 성능의 비교 결과  
Figure 5. Comparison results of SV performance

### 참고문헌

- [1] Evans. N. W. D, Mason. J. S, Aukenthaler R. and Stapert R, "Assessment of speaker verification degradation due to packet loss in the context of wireless mobile devices", *COST275 Workshop on Biometrics over the Internet*, Rome, November 7-8, 2002.
- [2] Norgaard A, Pasic. E, "Requirements and biometric technology matching: two approaches from European projects", In: *Proc. of IEEE 39th Annual conference on security technology*, 2005.
- [3] Yang. Y, Wu, Z and Ma Z, "Real time user authentication system for PDA", In: *2004 IEEE International Conference on Systems, Man and Cybernetics(ICSMC 2004)*, pp.439-443, 2004.
- [4] Huang. X, Acero. A, Chelba. C, et al., "MiPad: a Multimodel Interaction Prototype", In: *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, pp.9-12, 2001.
- [5] Fukunaga. K, "Introduction to statistical pattern recognition", (Second edition), Academic Press, pp.110-119, 1990.
- [6] Therrien. C. W. "Decision Estimation and Classification: An Introduction to Pattern Recognition and related Topics", John Wiley & Sons, pp.38-41, 1989.
- [7] Harada. K, Noda. H., Kawaguchi. E, "Text-independent speaker verification using sequential decision", *Technical Report of IEICE*, PRMU97-81, 1997.
- [8] Lund. MA and Lee. C. C, "A robust sequential test for text-independent speaker verification", *J. Acoust. Soc. Am.*, Vol. 99, No.1, pp.609-621, Jan.1996.
- [9] Noda. H, Harada. K, Kawaguchi. E, "A context-dependent Sequential decision for speaker verification", *IEICE Trans. INF. & syst.* Vol. E82-D, No.10 Oct., 1999.
- [10] Seo. C, Sim. K, Kim. E, Ko. H, Lim. Y, "Speaker Verification System Using Extended Sequential Probability Ratio Test in PDA", *LNCS 5630, Springer*, pp.188-193, Jun., 2009.
- [11] Noda. H, Kawaguchi. E, "Adaptive speaker identification using Sequential Probability Ratio Test", *15th International Conference on Pattern Recognition*, Vol. 3, pp.262-265, 2000.
- [12] Reynolds. D, and Rose. R, "Robust text-independent speaker identification using Gaussian mixture speaker models", *IEEE Trans. on SAP*, Vol. 3, No.1, pp.72-82, 1995.
- [13] Seo. C, Zhao. M, Lim. Y, Jeon. S, "Effective Combination of Temporal Information and Linear Transformation of Feature Vector in Speaker Verification", *Jour. of Kor. Soc. Spee. Sci.*, Vol. 1, No.4, pp.127-132, 2009.

(서창우, 조미화, 임영환, 전성채, "화자확인에서 특징벡터의 순시 정보와 선형 변환의 효과적인 적용", *말소리와 음성과학*,

Vol. 1, No.4, pp.127-132, 2009.)

- [14] Young. S, Evermann. G, Kershaw. D, Moore. G, Odell. J, Ollason. D, Povey. D, Valtchev. V, and Woodland. P, *The HTK Book (for HTK version 3.2)*, Cambridge University Engineering Department, 2002.
- [15] Parthasarathy. S and Rosenberg. A. E, "General phrase speaker verification using sub-word background models and likelihood ratio scoring", *ICSLP-96*, 1996.
- [16] Matsui. T and Furui. S, "Likelihood normalization for speaker verification using a phoneme and speaker-independent model", *Speech Communication*, Vol. 17, pp.109-116, August, 1995.
- [17] Lind. Y, Buzo. A, and Gray. R. M, "An algorithm for vector quantizer design", *IEEE Trans. Commun.*, Vol. 28, pp.84-95, 1980.

• **김은영 (Kim, Eunyoung)**

숭실대학교 글로벌 미디어학부  
서울시 동작구 상도동 511번지  
Tel: 02-826-9872  
Email: key710503@ssu.ac.kr  
관심분야: 멀티미디어, 모바일 시스템  
2004~현재 글로벌 미디어학부 박사수료

• **서창우 (Seo, Changwoo)**, 교신저자

숭실대학교 글로벌 미디어학부  
서울시 동작구 상도동 511번지  
Tel: 02-820-0910 Fax: 02-826-9872  
Email: cwseo@ssu.ac.kr  
관심분야: 음성신호처리, 멀티미디어, 모바일 시스템  
2008~현재 글로벌 미디어학부 연구교수

• **전성채 (Jeon, Sungchae)**

한국전기연구원 융합기술연구단  
경기도 안산시 상록구 사1동 1271-19번지  
Tel: 031-8040-4151 Fax: 031-8040-4163  
Email: sarim@keri.re.kr  
관심분야: 신호처리, 방사선 센서, 방사선 신호검출용 ASIC  
설계  
2006~현재 전자의료기기연구센터 센터장