

다중 생체인식 기반의 모바일 인증 시스템 성능 개선

정강훈*, 김상훈**, 문현준*
*세종대학교 컴퓨터공학과
**한경대학교 정보제어공학과
e-mail:hmoon@sejong.ac.kr

Enhancement of Mobile Authentication System Performance based on Multimodal Biometrics

Kanghun Jeong*, Sanghoon Kim**, Hyeonjoon Moon*
*Dept of Computer Science and Engineering, Sejong University
**Dept of Information and Control, Hankyong National University

요 약

본 논문은 모바일 환경에서의 다중생체인식을 통한 개인인증 시스템을 제안한다. 다중생체인식을 위하여 얼굴인식과 화자인식을 선택하였으며, 시스템의 인식 시나리오는 다음을 따른다. 얼굴인식을 위하여 Modified census transform (MCT) 기반의 얼굴검출과 k-means 클러스터 분석 (cluster analysis) 알고리즘 기반의 눈 검출을 통해 얼굴영역 전처리를 수행하고, principal component analysis (PCA) 기반의 얼굴인증 시스템을 구현한다. 화자인식을 위하여 음성의 끝점 추출과 Mel frequency cepstral coefficient (MFCC) 특징을 추출하고, dynamic time warping (DTW) 기반의 화자 인증 시스템을 구현한다. 그리고 각각의 생체인식을 본 논문에서 제안된 방법을 기반으로 융합하여 인식률을 향상시킨다.

1. 서론

최근 통신 인프라의 발전에 따른 휴대폰의 변화 및 사용자들의 다양한 요구에 의하여 스마트폰이 출현하게 되었으며, 이렇게 대두된 스마트폰의 다양한 인터페이스 및 기능이 하나의 모바일 기기에서 제공됨으로 인하여 모바일 컨버전스 (convergence)가 우리의 실생활에서 실현되게 되었다. 이에 따른 스마트폰 기술 발전과 수요가 폭발적으로 성장하였으며, 스마트폰에서의 사용자 편의성을 고려한 개인 신원의 확인이나 인증 및 개인 정보보호의 중요성이 급속도로 증가하고 있다.

기존의 모바일 기기 및 일상생활에서 널리 사용되는 패스워드 등 사용자가 알고 있는 정보를 이용한 사용자 인증 방법은 망각, 분실 또는 도난의 이유 등으로 높은 보안 성능을 제공하지 못할 수 있는 문제가 있다. 이에 기존의 신원 확인 방법보다 더 안전하고 신뢰할 수 있는 사용자 인증 방법으로 신체의 고유 특성을 이용한 생체인식 기술이 부각되고 있다[1]. 그러나 높은 보안성을 요구하는 곳에서 단일생체 정보만으로는 인식 성능 면에서 한계가 존재한다. 따라서 다중생체인식을 통하여 둘 이상의 상호 보완적인 생체 정보를 동시에 이용하거나 다중 정합 알고리즘을 적절히 조합하는 기술이 각광받고 있다[2].

본 논문에서는 얼굴과 음성 등의 개별 생체인식 정보에 가중치를 적용한 융합 기법을 사용하여 문맥중속의 화

자인증 시스템과 얼굴인증 시스템을 모바일 환경에 적합하도록 수정 및 보완한 사용자인증 시스템을 제안한다.

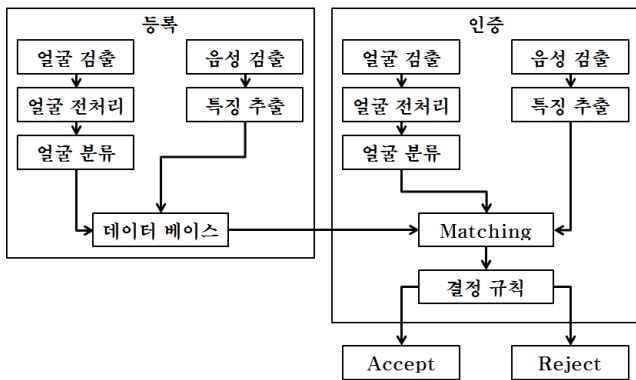
2. 얼굴인증 시스템

얼굴인식 기술은 사람들의 생체 특징 중 얼굴 특징 정보를 이용하여 각 사람의 신원을 인식하는 생체인식 기술 중 하나이다. 얼굴인식 기술은, 그 특성상 비공격적 (non-aggressive), 비강압적(non-intrusive) 인식 방법으로서, 인식 대상에게 보다 적은 거부감을 주는 사용자 친화적 생체인식 방법으로 알려져 있다. 본인인증을 위한 목적으로 생체정보를 획득 할 때, 영상 획득을 위한 추가적인 장비설치에 비용이 많이 들지 않으며, 사용자의 자연스러운 움직임으로부터 영상을 획득할 수 있기에 비접촉으로 자연스럽게 인식할 수 있는 장점이 있다. 본 논문에서는 modified census transform (MCT) 기반의 얼굴 검출(face detection)을 수행하고, 검출된 얼굴의 눈 위치를 기준으로 얼굴 영상을 정규화 하는 전처리(pre-processing) 과정을 거쳐서[3], principal component analysis (PCA)를 통해 얼굴인증을 시도하는 시스템을 구성한다[4].

3. 화자인증 시스템

화자인식 기술은 사람의 음성 정보를 이용하여 신원을 인식하는 기술로, 얼굴인식과 마찬가지로 인식 대상에게

적은 거부감을 주는 생체인식 방법이다. 본 논문에서는 주어진 문맥에 대한 발성을 생체특징으로 하는 문맥종속 화자인증 시스템을 구성하였으며, 전처리 과정과 인증 과정의 구조를 갖는다. 전처리 과정에서는 frame energy 및 zero-crossing rate (ZCR) 방식을 이용하여 음성구간과 묵음구간을 분리하는 끝점검출(end-point detection)과 인간의 청각 특성을 잘 반영하는 멜 캡스트럼 계수 (Mel-frequency cepstral coefficient; MFCC)를 특징 파라미터 (feature parameter)로 추출한다[5]. 인증은 추출된 특징 파라미터와 사전에 데이터베이스에 등록된 화자의 특징 파라미터를 비교하기 위해 동적시간간축 (dynamic time warping; DTW) 알고리즘을 이용하여 두 특징 사이의 패턴 유사도를 나타내는 거리 (distance)를 측정하고 그 값을 임계값과 비교하여 본인 여부를 확인한다[6].



(그림 1) 다중 생체인증 시스템

4. 단일 생체인증 실험

4.1 실험 환경

본 논문의 실험은 Samsung Galaxy S hoppin (CPU:1GHz, RAM:500MB)장치에서 수행되었으며, 운영체제는 Android 2.3.6이다. 실험에는 10명이 참가했으며 음성과 얼굴 정보를 사용하여 본인인증 실험과 타인인증 실험을 각 30회씩 수행하였으며, 각 실험은 다음 두 가지의 과정을 따른다. 첫 번째로 PCA 기반의 얼굴인증 시스템을 이용하여 얼굴 특징정보 추출을 통한 유사도를 얻어내고, 두 번째로 DTW 기반의 화자인증 시스템을 통한 입력패턴과 출력패턴의 유사도를 구한다. 각 생체인증 시스템의 유사도는 Euclidean distance로 측정한다. 실험은 조도가 170~220lx 이고 소음이 45~50dB인 조명의 밝기 및 주변소음의 환경 변화가 비교적 크지 않은 실내에서 진행되었다.

4.2 얼굴인증 실험

PCA 기반 얼굴인증 실험에서는 10명의 probe set (얼굴인증 시스템에 새로 입력되어 신원 확인의 주체가 되는 집합)을 대상으로 각각 30회씩 실험을 수행하였다. Android device의 camera로부터 800×480 크기의 영상을 입력받아서 얼굴검출 및 전처리를 거친 얼굴 영상을 고유얼굴에 투사시켜 고유얼굴에 대한 분포값 벡터를 계산한다. 이렇게 얻어진 probe들을 기존의 데이터베이스에 저장된 gallery와 비교하여 유사도 거리(distance)를 측정한다. 실험은 동일한 사람의 probe와 gallery를 사용한 본인인증 실험과 동일하지 않은 probe와 gallery를 사용하여 타인인증을 시도한 실험으로 나뉜다. 본인인증 실험에서 distance의 평균은 2803.8이며 표준편차는 595.7으로 나타났다, 타인인증 실험의 distance의 평균은 4959.1이고, 표준편차는 668.4으로 나타났다.

4.3 화자인증 실험

DTW 기반 화자인증 실험은 화자 10명에 대해서 주어진 단어를 30번씩 발음한 음성 데이터로 실험을 수행한다. 본 논문에서는 Android device의 microphone으로부터 8bit 모노 8000Hz 샘플링으로 녹음한 음향신호를 대상으로 하였으며, 녹음된 음향에서 음성을 추출하고 특징을 추출하여 기존에 데이터베이스에 등록되어있는 화자와 DTW 매칭을 통해 유사도 거리 (distance)를 측정한다. 데이터베이스에 등록되어있는 화자와 동일한 화자가 인증을 시도한 본인인증 실험에서 전체 distance의 평균은 139.43 이고, 표준편차는 19.93로 나타났으며, 타인인증 실험 결과 전체 distance의 평균은 218.07이고, 표준편차는 33.47로 나타났다.

5. 다중생체인식 융합 기법 및 성능 평가

5.1 고안된 다중생체인식의 결정 규칙

다음 식은 평균 μ 와 표준편차 σ 를 이용하여 가우시안 (Gaussian) 분포를 적용한 확률밀도함수 (probability density function; PDF)를 유도하는 식이다.

$$g_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

위의 식으로부터 생체인식의 본인인증 실험과 타인인증 실험에서 얻은 평균 μ 와 표준편차 σ 를 사용해서 가우시안 확률밀도함수를 적용하였다. [그림 3]은 얼굴인식의 본인인증 실험과 타인인증 실험결과의 가우시안 확률밀도

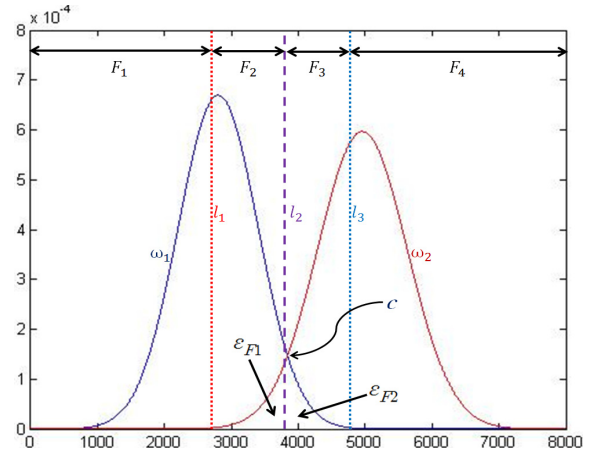
함수를 보여준다. ω_1 은 본인인증 실험의 확률밀도함수이고, ω_2 는 타인인증 실험의 확률밀도함수이다. ε_{F1} 영역의 넓이는 승인으로 결정했을 때 타인일 경우의 오류확률 (false acceptance ratio; FAR)이며, ε_{F2} 영역의 넓이는 거부로 결정했을 때 본인일 경우의 오류확률 (false rejection ratio; FRR)이다. 이 오류확률 ε_1 과 ε_2 는 다음 식을 통하여 구할 수 있다.

$$\varepsilon_1 = \int_{-\infty}^{D_{L2}} p(x|\omega_2)dx, \quad l_1 \leq D_{L2} < l_2$$

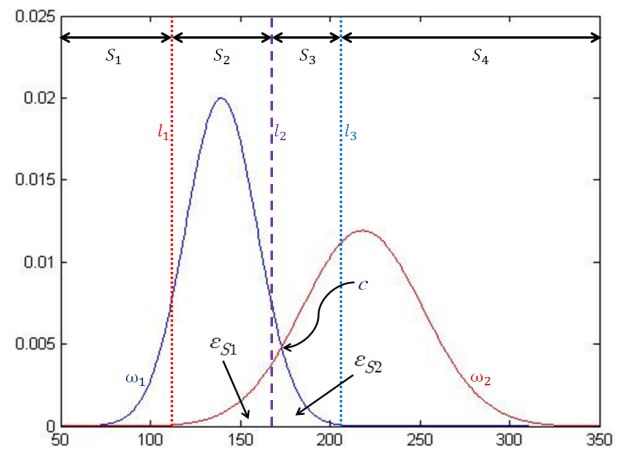
$$\varepsilon_2 = \begin{cases} 1 - \int_{-\infty}^c p(x|\omega_1)dx + \int_{-\infty}^c p(x|\omega_2)dx - \int_{-\infty}^{D_{L2}} p(x|\omega_2)dx, & \text{if } l_2 \leq D_{L2} < c \\ 1 - \int_{-\infty}^{D_{L2}} p(x|\omega_1)dx, & \text{else if } c \leq D_{L2} < l_3 \end{cases}$$

위의 식에서 D_{L2} 는 각 생체인식 시스템의 결과인 유사도를 나타내는 Euclidean distance이며, [그림 2]과 같이 ω_2 의 시작점과 ω_1 의 끝점 (확률이 0에 근접하는 지점)을 distance측과 직교하는 2개의 직선 l_1, l_3 으로 나타낼 수 있고, ε_{F1} 과 ε_{F2} 의 넓이가 동등하도록 나눌 수 있는 직선 l_2 을 얻을 수 있다. 여기서 l_2 는 승인/거부 (Accept/Reject) 결정에 있어서 최소 오류율 (equal error rate; EER)을 갖는 최적 결정 경계이다. l_1, l_2, l_3 세 개의 직선을 기준으로 F_1, F_2, F_3, F_4 의 네 영역으로 나눌 수 있으며, c 를 두 확률밀도함수의 교차점이라고 했을 때, 얼굴인식 실험 결과에서 l_1 은 2797.5, l_2 는 3832.9, l_3 는 4850.7, c 는 3940.7로 나타났다. 마찬가지로 [그림 3]에서는 화자인식의 본인인증 실험과 타인인증 실험의 가우시안 확률밀도함수를 보이며, ω_1 와 ω_2 는 각각 본인인증 실험의 확률밀도함수와 타인인증 실험의 확률밀도 함수이다. 얼굴인식 실험 결과와 같은 방법으로 S_1, S_2, S_3, S_4 의 네 영역으로 나눌 수 있으며, ε_{S1} 의 넓이와 ε_{S2} 이 넓이는 각각 화자인식에 대한 FAR과 FRR이다. 화자인식 실험의 결과는 l_1 은 109.70, l_2 는 168.06, l_3 는 208.92, c 는 173.06으로 나타났다.

다음 [그림 4]은 본 논문에서 제시하는 다중생체인식 융합 기법의 최종적인 결정 규칙으로, 얼굴인식과 화자인식의 분할된 영역의 조합 4×4 , 즉 16가지의 경우에 대한 결정규칙을 보여준다. 제시한 다중생체인식 시스템의 결정 규칙은 결정이 모호한 영역에서 오류확률을 기반으로 결정하기 때문에 하나의 인식기가 비교적 낮은 인식률을 보이더라도 전반적인 시스템의 인식성능을 유지할 수 있는 장점이 있다. 또한, FRR에 비해 위험한 FAR을 줄이고자 F_4 나 S_4 의 영역에 distance가 존재하면 거부로 결정한다.



(그림 2) 얼굴인식 실험 결과 (본인인증(좌)과 타인인증(우)의 PDF)



(그림 3) 화자인식 실험 결과(본인인증(좌)과 타인인증(우)의 PDF)

얼굴	F_1		F_2		F_3		F_4		
화자	S_{1-3}	S_4	S_{1-2}	S_3	S_4	S_1	S_2	S_{3-4}	S_{1-4}
결정 규칙 (Decision rule)	if $\varepsilon_{F1} \geq \varepsilon_{S2}$		else A		if $\varepsilon_{F2} < \varepsilon_{S1}$		else A		
A : Accept	A	R	A	R	R	A	R	R	R
R : Reject									

(그림 4) 제안된 다중생체인식의 결정 규칙

5.2 다중생체인식 실험 결과 및 분석

표 1은 단일생체인식 시스템과 다중생체인식 시스템의 실험결과를 보여준다. 다중생체인식 실험은 단일생체인식 시스템과 동일한 10명을 대상으로 하였으며 본인인증 실험과 타인인증 실험을 각 30회씩 수행하였다. 직렬 다중생체인식은 얼굴인식과 화자인식의 판단이 모두 수락일 때 승인해주는 AND 형태의 직렬 분류기로 구성된 시스템이며, 단일생체인식 시스템과 직렬 다중생체인식 시스템의 결정경계는 각각의 임계값 l_2 (EER)를 적용하여 판단하였다. 직렬 다중생체인식 시스템의 FAR은 0.9%로 가장 낮고, FRR이 12.1%로 가장 높으므로 직렬 다중생체인식 시스템을 실생활에 도입하여 사용하기에는 적합하지 않다.

표 1. 다중생체인식과 단일생체인식의 비교

	FAR	FRR	인식률
단일 얼굴 인식	4.6%	4.6%	95.4%
단일 화자 인식	6.7%	6.7%	93.3%
제안된 다중생체인식	1.8%	4.9%	95.1%

인증시스템의 가장 중요한 성능척도인 FAR이 단일생체인식 시스템에 비해서 크게 감소하였으므로, 제안된 다중생체인식 시스템은 단일생체인식 시스템에 비해서 신뢰할 수 있는 인증 시스템이다.

6. 결론

본 논문에서는 기존의 단일생체인식 시스템의 인식률 향상을 위해 생체인식 시스템에서 많이 사용되고 있는 얼굴인식과 화자인식을 이용한 다중생체인식 시스템을 제안하였다[7]. 제안된 다중생체인식 시스템과 단일생체인식 시스템의 인식률을 비교해보면 단일생체인식 시스템인 얼굴인식 시스템의 인식률은 95.4%, 화자인식은 93.3%로 각각 나타났으며, 다중생체인식 시스템의 본인 인식률은 95.1%로 나타났다. 하지만 보안감시 분야에서 중요한 성능척도인 FAR이 단일 생체인식 시스템의 얼굴인식에서는 4.6%, 화자인식은 6.7%인 반면, 다중생체인식 시스템은 1.8%로 감소하였다. 이는 사칭자의 보안 위협으로부터 개인정보를 보호할 수 있는 안정된 보안 시스템임을 입증한다. 그리고 제안된 다중생체인식 시스템의 얼굴인증과 화자인증의 결정 규칙을 통하여 하나의 생체인식이 조명 혹은 주변소음과 같은 생체인식에 장애가 되는 환경 변화의 요인에 의해 정확하지 않은 인식률을 보이더라도 단일 생체인식에 비해 안정적인 인식성능을 유지할 수 있는 장점이 있다.

Acknowledgement

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단 기초연구사업의 지원을 받아 수행된 연구임(2012-0006947).

참고문헌

- [1] A. K. Jain, "Biometrics Personal Identification in Networked Society," Kluwer Academic Publishers, 1999.
- [2] R. N. Rodrigues, L. L. Ling, V. Govindaraju, "Robustness of Multimodal Biometric Fusion Methods Against Spoof Attacks," Journal of Visual Languages & Computing, Volume 20, Issue 3, Pages 169-179, June 2009.
- [3] Z. Qian, D. Xu, "Automatic Eye Detection Using Intensity Filtering and K-means Clustering," Pattern Recognition Letters, Volume 31, Issue 12, Pages 1633-1640, 1 September 2010.
- [4] J. Ruiz-del-Solar, P. Navarrete, "Eigenspace-Based Face Recognition: A Comparative Study of Different Approaches," IEEE Trans. Systems, Man, and Cybernetics- Part C 35(3), 315-325, 2005.
- [5] R. Vergin, D. O'Shaughnessy, A. Farhat, "Generalized Mel Frequency Cepstral Coefficients for Large-Vocabulary Speaker-independent Continuous-Speech Recognition," Speech and Audio Processing, IEEE Transactions on, Vol. 7, No. 5, pp. 525-532, Sep 1999.
- [6] E. Keogh, C. A. Ratanamahatana, "Exact indexing of dynamic time warping," Knowledge and Information Systems, Volume 7, Issue 3, pp 358-386, March 2005.
- [7] <http://play.google.com>, "BiometricsVerification"