

다중생체인식을 이용한 사용자 인증

이경희, 최우용, 지형근, 반성범, 정용화*

*한국전자통신연구원, 생체인식기술연구팀

Person Authentication using Multi-Modal Biometrics

Kyunghee Lee, Wooyong Choi, Hyungkeun Jee, Sungbum Pan, Yongwha Chung*

*Biometrics Research Team, ETRI

요약

생체인식 기술은 전통적인 비밀번호 방식 또는 토큰 방식보다 신뢰성 면에서 더 선호되지만, 환경의 영향에 매우 민감하여 성능의 한계가 있다. 이러한 단일 생체인식 기술의 한계를 극복하기 위하여 여러 종류의 생체 정보를 결합한 다중 생체인식(multimodal biometrics)에 관한 다양한 연구가 진행되고 있다. 본 논문에서는 다중 생체인식 기술을 간략히 소개하고, Support Vector Machines(SVM)을 이용하여 얼굴 및 음성 정보를 함께 이용한 다중 생체인식 실험으로 성능이 개선될 수 있음을 확인하였다.

I. 서론

사람의 신체적, 행동적 특성에 기반한 생체인식은 전통적인 비밀번호 방식 또는 토큰 방식보다 더 신뢰할 만하다. 보편성, 유일성, 영구성, 획득성을 갖는 신체적, 행동적인 특성은 자동 사용자 인증 시스템의 생체 특징으로 사용될 수 있다. 또한 생체인식 기반의 인증은 도난당하거나 망각될 우려가 없기 때문에 전통적인 방법에 비해 선호된다.

생체인식 기술을 이용한 자동 사용자 인증 시스템은 다양한 응용분야에서 중요한 역할을 하는데, 전형적인 응용 예로 액세스 통제, 감시 분야, 비디오 데이터베이스로부터의 얼굴검색 등이 있다. 최근에는 인터넷 뱅킹을 포함한 다양한 종류의 금융서비스, 인터넷을 통한 중요한 자료에 대한 정보보호, 안전한 전자정부 구축 등에 신뢰성을 제공하는 핵심기술로 부각되고 있다.

비교적 단순한 생체 특징으로는 지문, 얼굴, 음

성 등이 있고, 반면에 좀 더 정교한 특징으로는 홍채 패턴과 적외선 카메라로부터 얻어진 얼굴 열상 등이 있다. 이러한 단일 생체인식은 이상적인 환경에서는 비교적 신뢰할 만하지만, 실제 상황에서는 환경에 매우 민감하다. 예로 음성인식은 소음 환경에서는 급격하게 인식률이 떨어지며, 얼굴인식 또한 조명, 사용자의 포즈에 큰 영향을 받는다. 최근에는 이러한 한계를 극복하고자 여러 종류의 생체인식 기술을 함께 사용하여 상호보완성을 제공하는 다중 생체인식에 관한 관심이 고조되고 있다.

다중 생체인식 시스템은 다중 생체 정보의 구성과 정보 통합이 이루어지는 기준에 따라 다음과 같이 분류할 수 있다[1].

- i) 다중 센서(multiple sensors) : 예를 들면 지문 획득을 위한 광학, 초음파, 반도체 방식 등의 센서들이 가능하다
- ii) 다중 생체특징 시스템(multiple biometric system) : 지문, 얼굴과 같은 상이한 생체 특징들을 결합한다
- iii) 동일 생체특징의 다중 유닛(multiple units

of the same biometric) : 양쪽 눈의 홍채 특징, 양쪽 손의 특징, 10개 손가락 지문 등을 각각 결합한다

- iv) 동일 생체특징을 여러 번 획득(multiple instances of the same biometric) : 같은 손가락의 지문을 여러 번 획득하거나, 여러 개의 음성 샘플, 여러 장의 얼굴 영상을 결합한다
- v) 동일 입력 생체특징 신호에 대한 다중 표현과 매칭 알고리즘(multiple representation and matching algorithms for the same input biometric signal) : 예로 하나의 지문에 대하여 다른 특징 추출과 매칭 알고리즘을 결합한다

두 개 이상의 단일 생체인식 시스템의 통합은 여러 단계에서 일어날 수 있는데, 다음과 같이 세 가지 통합 단계로 나누어 볼 수 있다[2].

- i) 특징 추출 단계(feature extraction level)에서의 통합 : 각 센서로부터 얻어진 데이터들을 하나의 특징 벡터로 통합한다
- ii) 매칭 점수 단계(matching score level)에서의 통합 : 각각의 단일 생체 매칭 모듈로부터 출력된 매칭 점수들을 통합한다
- iii) 결정 단계(decision level)에서의 통합 : 각각의 단일 생체 결정 모듈로부터 출력된 수락 또는 기각 결정들을 통합한다

단일 생체인식 시스템의 구성에도 응용분야에 따라 어떤 생체 정보를 사용할 것인지에 대한 고려가 있어야 하지만, 다중 생체인식 시스템을 구성하기 위해서는 응용분야에 따른 생체 정보의 선택, 상호 보완성이 좋은 생체 정보들의 조합 선택, 생체 정보의 통합 단계 선택, 생체 정보 구성 시스템의 비용 대비 효과 분석 등에 대한 고려가 반드시 필요하다.

본 논문의 구성은 다음과 같다. II장에서는 본 논문에서 인증 시스템으로 사용한 Support Vector Machines(SVM)에 대하여 간략하게 살펴보고, III장에서는 SVM을 이용한 다중 생체인식 시스템의 구현과 초기 성능에 대하여 설명한 후, IV장에서 결론 및 향후과제에 대하여 언급한다.

II. Support Vector Machines

SVM(Support Vector Machines)은 V. Vapnik 과 그의 공동 연구자들에 의해 패턴 인식에서 효과적이고 일반화된 방법으로 최근에 제안되었다

[3]. SVM은 구조적 위험 최소화(structural risk minimization)의 개념에 기반한 최적의 선형 결정 평면(optimal linear decision surface)을 찾는 이진 분류 방법이다. 결정 평면은 학습 집합의 원소들의 가중화된 조합으로 이루어진다. 이러한 학습 집합의 원소들을 *support vectors* 라고 부르고, 이것은 두 클래스 간의 경계면을 나타낸다.

N 개의 학습 샘플들 (\mathbf{x}_i, y_i) 가 있다고 할 때, 여기서 $\mathbf{x}_i \in R^d$ 이고 y_i 는 관련된 레이블이다 ($y_i \in \{-1, 1\}$). 선형 분리 데이터를 가정할 경우, 최대 여백 분류의 목표는 *support vectors*의 거리가 최대화되는 초월면(hyperplane)에 의해 두 개의 클래스를 분리하는 것이다. 이러한 초월면은 최적 경계 초월면(OSH: optimal separating hyperplane)이라고 불린다. OSH는 식(1)과 같이 나타낼 수 있다.

$$f(\mathbf{x}) = \sum_{i=1}^N \alpha_i y_i (\mathbf{x}_i^T \mathbf{x}) + b \quad (1)$$

여기서 계수들 α_i 와 b 는 2차 프로그래밍(quadratic programming) 문제의 해이다[3]. *support vectors*에 대한 α_i 는 0이 아닌 값을 갖는다.

선형 분리가 불가능한 데이터인 경우에는, SVM은 입력벡터를 선형 초월면이 발견되는 고차원의 특징 공간 F 로 비선형 매핑을 할 수 있다. 다음의 식(2)와 같은 고차원 매핑은 비선형 SVM을 형성하는데 사용된다.

$$\phi: R^d \mapsto F \quad (2)$$

목표함수와 결정함수의 두 가지 모두 데이터 벡터 \mathbf{x} 의 내적에 의해 표현됨에 따라, 계산적으로 복잡한 매핑 $\phi(\cdot)$ 는 명시적으로 계산할 필요가 없다. Mercer의 조건을 만족하는 커널 함수(kernel function) $K(\mathbf{x}, \mathbf{x}_i)$ 는, $(\mathbf{x} \cdot \mathbf{x}_i)$ 대신에 사용되는 $(\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i))$ 를 다시 치환할 수 있다[3]. 따라서 결정 평면은 다음 식(3)과 같은 수식을 갖는다.

$$f(\mathbf{x}) = \sum_{i=1}^N \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (3)$$

다음의 커널 함수들이 SVM에 종종 사용되는데, 다항식 커널(polynomial kernels)은 다음의 식(4)와 같이 주어진다.

$$K(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}^T \mathbf{x}_i + 1)^p \quad (4)$$

또한 RBF 커널(Radial Basis Function kernels)은 식(5)와 같으며,

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{x} - \mathbf{x}_i\|^2\right) \quad (5)$$

tangent hyperbolic kernels 은 다음 식(6)과 같다.

$$K(\mathbf{x}, \mathbf{x}_i) = \tanh(\alpha \cdot \mathbf{x}^T \mathbf{x}_i + \beta) \quad (6)$$

III. SVM을 이용한 다중생체인식

1. SVM을 이용한 단일생체인식

1) SVM을 이용한 얼굴인식

얼굴에 의한 인증은 인간에게 보다 친숙하고 얼굴 영상 획득이 비강제성을 띤다는 점에서 사용하기 편리한 생체인식 기술이다. 그러나, 얼굴 영상은 얼굴의 3차원 방향 변화, 빛의 방향 및 세기 변화, 머리모양, 표정 변화에 따라 차이가 현저하므로, 자동 얼굴인식 개발은 해결하기 어려운 문제이다.

얼굴인식에 관한 기존의 연구들은 전체적인(holistic) 접근방법과 분석적인(analytic) 접근방법으로 크게 분류해 볼 수 있다[4]. 전체적인 접근 방법은 패턴의 전역적인 특성을 고려하는 것으로, 얼굴영상 영역을 정규 직교 기저벡터(orthonormal basis vectors)의 집합을 사용하여 표현하는 것이다. 현재 가장 널리 알려진 기저벡터는 Turk 등이 제안한 고유얼굴(eigenface)이다[5]. 분석적인 접근 방법은 눈, 코, 입 등과 같은 얼굴의 요소들을 얼굴영상으로부터 추출하여, 이러한 얼굴의 요소 특징들간의 지형적인 특성들의 차이를 인식에 사용하는 것이다[6]. 이러한 특징 기반 기법은 빠른 인식 속도와 적은 메모리를 요구한다는 장점을 가지는데 반해, 특징점 선택과 추출이 어렵다는 단점이 있다. 최근 패턴 분류에서 주목을 받고 있는 이진 분류기(binary classifiers)인 SVM을 얼굴인식 분야에 적용하는 연구가 활발히 진행되고 있다[7][8]. Phillips는 PCA 기반 얼굴인증 방법과 비교하여 SVM기반 얼굴인증 시스템이 더 좋은 성능을 나타냄을 보여 주었다[9].

본 논문에서는 얼굴인식 분야에서 널리 사용되고 있는 주요소 분석 (Principal Component Analysis :PCA)[5]을 통한 SVM 분류기를 이용하

여 얼굴 인증을 수행하였다. XM2VTS 멀티모달 데이터베이스[10]를 이용하여 임의로 추출한 42명에 대하여 실험하였다. 각 얼굴 영상에 대하여 50개의 PCA 계수를 추출하여 개인별로 SVM 인증기를 학습을 통하여 생성한 후, 테스트 데이터로 실험하였다. 식(5)의 RBF 커널을 이용하여 σ 값을 0.1 과 0.5 로 실험한 결과, 표 1에서와 같이 FAR(False Acceptance Rate), FRR(False Reject Rate) 모두 비교적 좋은 결과를 얻었다.

2) SVM을 이용한 화자인식

사람의 음성을 이용하는 화자인식은 비교적 값이 싸고 손쉽게 구할 수 있는 마이크를 사용하며 사용자의 거부감이 적다는 장점이 있다. 그러나 주위 소음에 민감하다는 단점이 있다. 전통적으로 현재까지 많이 사용되어온 화자인식 방법으로는 DTW(Dynamic Time Warping)[11], HMM(Hidden Markov Model)[12], VQ(Vector Quantization)[13], GMM(Gaussian Mixture Model)[14] 등이 있다. DTW와 HMM은 주로 문맥종속 시스템에 많이 쓰이고, VQ와 GMM은 문맥독립 시스템에 많이 쓰이는 방법이다. 또한 최근 SVM을 이용한 화자인식 연구들이 발표되고 있다[15].

음성 정보를 이용한 화자인식 실험에서도 얼굴인식 실험에서 사용한 XM2VTS 데이터베이스 내의 42명에 대하여 실험하였다. 각각의 음성 화일에서 24개의 특징을 추출하여 개인별로 SVM 인증기를 학습을 통하여 생성한 후 테스트하였다. 얼굴인식과 같은 방법으로 식(5)의 RBF 커널을 이용하여 σ 값을 0.1 과 0.5 로 실험하였다. 표 1에서 알 수 있듯이 음성만을 이용한 사용자 인증은 좋지 않은 결과를 얻었다.

표1. 얼굴+음성을 이용한 다중 생체인식 실험결과

	에러율	얼굴	음성	얼굴+음성
RBF $\sigma=0.1$	FAR	0.025	0.095	0.022
	FRR	0.105	0.395	0.070
RBF $\sigma=0.5$	FAR	0.009	0.051	0.008
	FRR	0.116	0.488	0.058

2. SVM을 이용한 다중생체인식

본 논문에서는 사용자에게 보다 친숙하고 사용하기 편리하지만, 단일 생체인식으로는 주위 환경

에 민감하게 영향을 받는 얼굴 생체정보와 음성 생체정보를 결합한 다중 생체인식 실험을 하였다. 얼굴특징 50개와 음성특징 24개를 특징 추출 단계에서 통합한 74개의 특징을 이용하여, 개인별 SVM 인증기를 생성한 후 인증 실험을 하였다. 얼굴과 화자 단일 생체인식에서 사용한 XM2VTS 데이터베이스의 42명의 데이터를 사용하였으며, RBF 커널을 이용하여 σ 값을 0.1 과 0.5 로 실험하였다. 실험 결과 표 1에서와 같이 FAR(False Acceptance Rate)이 2.2%, 0.8% 이고, FRR(False Reject Rate)이 7.0%, 5.8%로 얼굴, 화자 각각의 단일 생체인식에 비해 훨씬 개선된 결과를 얻었다. 특히 에러율이 높은 음성 생체특징을 얼굴 특징에 결합하여 사용하였는데도 불구하고, 비교적 인식률이 좋았던 얼굴만을 이용한 단일 생체인식보다도 더 좋은 결과를 얻었다.

IV. 결론 및 향후과제

본 논문에서는 다중 생체인식 시스템의 필요성과 그 구성에 대하여 간략히 소개하였다. 또한 SVM 분류기를 이용한 얼굴과 음성 생체정보 각각에 대한 단일 생체인식과, 이를 통합한 다중 생체인식에 대한 비교 실험을 통하여 다중 생체인식의 성능 개선을 확인하였다. 향후 보다 다양한 생체 정보의 통합과, 여러 통합 단계에서의 다중 생체인식에 관한 면밀한 연구가 요구된다.

참고문헌

[1] S. Prabhakar and A. Jain, "Decision-Level Fusion in Fingerprint Verification", *Pattern Recognition*, Vol. 35, pp.861-874, 2002.
 [2] A. Ross and A. Jain, "Information Fusion in Biometrics", *Pattern Recognition Letters*, Vol. 24, pp.2115-2125, 2003
 [3] V. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, 1998.
 [4] K. Lam and H. Yan, "An Analytic-to-Holistic Approach for Face Recognition based on a Single Frontal View", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 7, pp. 673-686, 1998
 [5] M. Turk and A. Pentland, "Face Recognition using Eigenfaces", *Proceedings of IEEE International Conference on Computer*

Vision and Pattern Recognition, pp. 586-591, 1991
 [6] R. Brunelli and T. Poggio, "Face Recognition :Features versus Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042-1052, 1993
 [7] G. Guo, S. Z. Li, and K. L. Chan, "Support vector machines for face recognition", *Image and Vision Computing* 19, pp. 631-638, 2001
 [8] K. Lee, Y. Chung, and H. Byun, "SVM-based face verification with feature set of small size", *IEE Electronics Letters*, Vol.38, No.15, pp.787-789, 2002
 [9] P. Jonathon Phillips, "Support Vector Machines Applied to Face Recognition", *Advances in Neural Information Processing Systems 11*, MIT Press, pp. 803-809, 1999
 [10] XM2VTS Multimodal Face Database, <http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb>
 [11] J. Campbell, "Speaker recognition: a tutorial", *Proc. of the IEEE*, Vol. 85, No.9, pp.1437-1462, 1997
 [12] Q. Li, B. Juang, C. Lee, Q. Zhou, and F. Soong, "Recent advancements in automatic speaker authentication", *IEEE Robotics and Automation Magazine*, pp.24-34, Mar. 1999
 [13] J. He, L. Liu, and G. Palm, "A New Codebook Training Algorithm for VQ-based Speaker Recognition", *Proc. ICASSP*, Vol. 2, pp.1091-1094, 1997
 [14] C. Alamo, F. Gil, C. Munilla, and L. Gomez, "Discriminative training of GMM for speaker identification", *Proc. ICASSP*, Vol. 1, pp.89-92, 1996
 [15] W. Choi, K. Lee, and Y. Chung, "Learning Support Vectors for Speaker Verification", *Proceedings of WISA*, pp.349-354, 2002