

홍채와 음성을 이용한 고도의 개인확인시스템

Development of Advanced Personal Identification System Using Iris Image and Speech Signal

이대종*, 고현주**, 객근창**, 전명근**

Dae-Jong Lee, Hyoun-Joo Go, Keun-Chang Kwak, Myung-Geun Chun

* 한국전기안전공사부설 전기안전연구원

** 충북대학교 전기전자 및 컴퓨터공학부 컴퓨터정보통신연구소

* Electrical safety research Institute

** Chungbuk National University

School of Electrical and Computer Engineering

요 약

본 논문에서는 홍채인식과 화자식별 방법을 결합한 고도의 개인확인 시스템을 제안한다. 제안된 방법은 홍채인식과 화자식별의 장점만을 부각시켜 최종적으로 결정법칙에 의해 화자를 인식하는 구조로 되어 있으며, 특히 음성에 노이즈가 첨가된 환경하에서도 우수한 성능을 보일 수 있도록 시스템을 구성하였다. 제안된 알고리즘의 성능을 검증하기 위하여 실험한 결과, 보안의 정도가 높은 상급(High)에 해당하는 인식률은 홍채만을 이용한 경우보다 56.7%, 음성을 이용한 화자식별 방법만을 사용한 경우보다 10% 정도 인식률이 향상되었다. 또한 음성에 노이즈를 첨가한 경우에도 보안의 정도가 높은 상급(High)에 해당하는 인식률은 홍채만 이용한 경우보다 인식률이 30%, 노이즈가 첨가된 음성만 이용한 경우보다 60% 정도 향상된 인식률을 보였다.

Abstract

This proposes a new algorithm for advanced personal identification system using iris pattern and speech signal. Since the proposed algorithm adopts a fusion scheme to take advantage of iris recognition and speaker identification, it shows robustness for noisy environments. For evaluating the performance of the proposed scheme, we compare it with the iris pattern recognition and speaker identification respectively. In the experiments, the proposed method showed more 56.7% improvements than the iris recognition method and more 10% improvements than the speaker identification method for high quality security level. Also, in noisy environments, the proposed method showed more 30% improvements than the iris recognition method and more 60% improvements than the speaker identification method for high quality security level.

Key Words : 강인한 화자인식(Robust speaker recognition), 홍채패턴(Iris Pattern), 정보보호(information security), 생체인식(biometrics)

1. 서 론

정보화 사회를 맞이하여 기업은 물론 개인도 질 높은 정보를 축적하고 있으며, 이러한 축적된 정보는 네트워크를 통하여 손쉽게 전파 될 수 있을 뿐더러 효과적으로 공유될 수 있는 환경이 이미 구축되어 있다. 더욱이 인터넷의 발달로 촉발된 전자상거래의 규모는 해마다 기하급수적으로 늘어나고 있어, 현대인은 네트워크화된 세상에 살고 있다고 해도 과언이 아니다. 이런 정보의 교환과 전자상거래에 있어서 인가된 대상자와 그렇지 않은 사람을 구별하고, 정보를 보호하거나

안전한 상거래를 이룩하는 일은 어느 때보다도 중요한 문제로 부각되고 있다.

시스템 보안을 위한 방법 중 패스워드에 의한 개인 확인과 인증은 가장 많은 운영체제에서 지원되고 있는 구현이 용이하고 별도의 특수 장치가 필요 없는 시스템 보안의 방법이다. 그러나 개인이 아무리 자신의 패스워드 관리를 잘 한다 하여도 그것이 타인에게 누설 될 수 있으므로 이러한 패스워드에 의한 시스템 보안을 대체 할 수 있는 새로운 개인확인 및 인증 방법이 시급히 요구되고 있는 상황이다[1]. 기존의 단순한 형태의 개인확인 및 검증방법의 한계를 극복하여, 절도나 누출에 의하여 도용 될 수 없으며 변경되거나 분실할 위험성도 없는 새로운 형태의 신분검증 방법에 대한 연구 분야가 바로 바이오메트릭스(Biometrics)분야이다. 생체인식기법이라고 번역되는 바이오메트릭스는 "자동화된 방법으로 특정 개인의 특성을 검증하거나 신분을 인식하기 위해, 측정 가능한 물리적 특성 또는 개인의 생체학적 특성을 연구하는

접수일자 : 2002년 10월 9일

완료일자 : 2003년 4월 22일

본 연구는 한국과학재단 목적기초연구(R01-2002-000-00315-0) 지원으로 수행되었음.

학문"으로 정의될 수 있다. 위와 같은 바이오메트릭스 기술에서 대상으로 하고 있는 인간의 생체학적 특성은 크게 지문이나 얼굴, 홍채, 음성인식, 서명인식 등이 사용되고 있다 [2-4].

바이오메트릭스 기술에서 사용하고 있는 대상들 중 먼저, 지문(Fingerprint)에 대해 살펴보자. 각 개인의 지문은 땀샘이 용기되어 일정한 흐름을 형성한 것으로, 그 모양이 개인마다 서로 다를 뿐만 아니라 태어날 때의 모습 그대로 평생 동안 변하지 않음이 증명되어 실생활에 사용되어져 왔다. 인식에 필요한 개인별 특징추출에 사용되는 방법으로는 주파수 공간에서의 Fourier, Wavelet 변환, 또는 신경회로망이나 퍼지논리에 의한 것 등을 들 수 있다. 그러나, 지문에는 기준 좌표축이 존재하지 않으므로 임의로 회전되어 채취된 지문에 대한 처리가 어려우며, 또한 지문은 신체의 일부로서 유연성을 가지고 있어 채취할 때마다 그 모양이 달라 보이며, 지문은 손상될 수 있어 고정능의 개인 인증시스템에는 그 문제점이 지적되고 있다[5].

한편, 얼굴영상인식의 문제에서는 많은 응용분야와 다양한 연구에도 불구하고 만족할 만한 신뢰도를 보장받지 못하고 있다. 왜냐하면 정지영상의 경우는 영상으로부터의 얼굴분리가 어렵고, 동영상을 이용한 경우는 시간정보를 이용하기 때문에 얼굴의 분리는 가능하여 정지영상을 이용한 경우의 문제가 없으나, 영상의 질이 몹시 저하된다는 문제가 따른다. 또한 노화에 따른 얼굴의 변형이나 머리카락의 길이, 표정의 변화, 주위 조명에 많은 영향 등으로 인한 특징 추출의 문제로 인해 아직까지 개선할 사항이 많은 상황이다[6].

최근에 사람의 눈동자에 존재하는 홍채(Iris) 패턴에 의한 개인 인식 및 인증에 대한 연구가 관심을 끌고 있다. 홍채는 사람이 태어난 후 얼마간의 세월이 흐른 뒤 그 모양이 변하지 않은 특성을 지니면서, 사람마다 독립적인 특징을 가지므로 신원인식의 시스템으로서 높은 신뢰도를 지닐 수 있다. 그러나, 홍채인식에 의한 생체인식 시스템 역시, 그것이 가지는 영상처리적인 특징으로 말미암아 조영이나 안경으로 인한 빛의 반사능의 감박임으로 인한 불안정한 영상 등으로 인하여 그 인식율이 저하받고 있다[7].

인간이 가지는 또 다른 생체 특성으로 음성이 있다. 음성은 허용성의 차원에서 인간에게 가장 자연스럽게 취득될 수 있을 뿐만 아니라, 전화나 통신망 등을 통하여 원격리에서도 쉽게 보내어 질 수 있기 때문에 음성인식과 화자인식 분야에서 많이 연구되어 왔다. 특히, 음성을 이용한 생체인식 기법은 그 어떤 생체 특징 보다도 향후의 네트워크 및 유무선 통신망의 발달에 발 맞추어 그 활용분야가 넓어지리라 기대되고 있다. 그러나 음성을 이용한 화자 인식기법에 있어서도 화자의 기본이나 말하는 속도의 변화, 그리고 녹음에 사용된 매체의 특성변화 등으로 고 신뢰도의 화자인식 시스템 구성에 어려움을 겪고 있다[8].

위에서 살펴 본 바와 같이 사람이 가지는 생체 특징은 생체인식의 관점에서 볼 때, 각각 그 장점과 단점을 가지고 있다. 따라서 고 신뢰도의 생체인식 시스템을 구현하여 여러개의 생체 특징을 고려하는 다중 생체 인식 기법에 관한 연구가 최근 들어 시도되고 있다. 이에 대한 연구로는 먼저 얼굴과 음성인식을 결합한 형태[9,10]와 얼굴인식과 지문인식을 결합한 형태[11] 등이 있지만 아직 초기 연구단계에 불과할 뿐 현재까지도 뚜렷할 만한 연구결과는 미미한 실정이다.

따라서, 본 논문에서는 지금까지 알려진 생체인식 기법 중 가장 신뢰도가 큰 홍채패턴 인식과 가장 보편적으로 널리 쓰일 수 있는 음성에 기반을 둔 화자식별 기법을 활용하여 새

로운 고신뢰도의 다중 생체인식 시스템을 제안하고자 한다. 우선, 홍채패턴 인식을 위해 검출된 1차원 홍채패턴에 대하여 Discrete Fourier Transform(DFT) 계수 추출과정과 데이터 길이 정규화과정을 거친 후 correlation의 계수 값을 사용하여 홍채를 인식하는 기법[4]을 적용하였다. 다음으로, 음성정보를 이용하여 화자식별을 위해 우수한 다해상도 특성을 지니고 있는 웨이블릿 서브밴드 필터뱅크에 기반을 둔 화자식별기법을 적용하였다[12]. 마지막 단계는 독립적인 홍채인식기와 화자식별기의 융합단계로서, 이를 위해 본 논문에서는 홍채인식과 화자식별의 장점만을 부각시킨 최종 결정법칙을 제안 및 적용하여 노이즈가 존재하는 환경 하에서도 우수한 인식성능을 보일 수 있도록 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 사용된 방법으로, 논문 [12]에서 제안한 홍채인식 알고리즘과 논문 [22]에서 제안한 음성에 기반을 둔 화자식별 알고리즘을 설명한 후, 마지막으로 본 논문에서 제안한 홍채와 음성을 결합한 다중생체 인식 알고리즘을 설명한다. 3장에서는 실험을 통해 제안한 방법의 타당성을 검토하고, 마지막으로 4장에서 결론을 맺는다.

2. 홍채와 음성을 결합한 다중생체 인식 알고리즘

2.1 홍채인식 알고리즘[4]

그림 1에서는 본 논문에서 이용된 홍채인식 알고리즘의 인식과정을 나타냈다. 논문 [4]에서 제안된 홍채인식 알고리즘의 인식과정을 간략히 설명하면, 우선 영상전처리 과정에서는 획득된 영상에서 인식할 부분인 홍채, 즉 동공의 초점 및 동공의 윤곽선을 검출하는 단계이다.

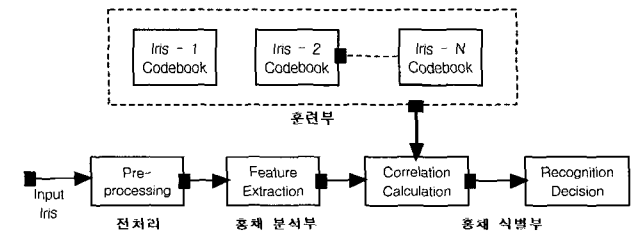


그림 1. DFT 계수를 이용한 홍채인식 알고리즘
Fig. 1. Iris recognition algorithm using DFT coefficients

홍채 분석부에서는 검출된 홍채의 특징점 즉 DFT계수를 구하고 훈련부에서는 인식하고자 하는 홍채를 각각 대표하는 코드북을 만드는 과정이다. 마지막 단계인 인식부에서는 입력된 홍채와 미리 작성된 코드북과의 상관도를 계산하여 상관도가 가장 높은 코드북을 인식하고자 하는 홍채로서 인식하는 단계이다. 이에 대하여 자세히 알아보면 다음과 같다.

홍채를 이용한 인증시스템을 구축하기 위하여 우선 카메라로부터 홍채 영상을 취득한다. 그림 2에 LG전자에서 상용 중인 홍채인식기의 카메라를 통하여 취득한 홍채영상의 일례를 보이고 있다. 홍채인식을 하기 위해 사용되는 홍채패턴은 대부분 눈동자의 검은 부분인 동공의 반지름을 기준으로 하여 1.3배, 1.4배, 1.5배인 부분에 집중되어 있다[4]. 본 논문에서는 동공 반지름의 1.3배인 원에 위치한 일차원 홍채패턴을 이용하고자 한다. 이와 같은 일차원 홍채패턴을 얻기 위한 동공의 반지름은 동공의 초점과 동공의 외부 윤곽선을 이용

하여 구해진다. 반지름을 구하기 위한 동공추출은 동공의 밝기가 배경보다 검다라는 점에 착안하여 자동경계값 검출방법으로 홍채영상의 동공만을 추출하였다. 그런 다음 추출한 동공의 x-축에서의 히스토그램과 y-축에서의 히스토그램의 중심 값이 만나는 점을 이용하여 동공의 초점을 찾아낸다. 이때, 중위수 필터(Median Filter)를 이용하여 잡음을 제거하고 윤곽선(Edge)을 유지할 수 있게 하였다. 또한 동공의 경계를 단단하게 하기 위하여 8이웃화소인 성분 명칭화(Component labeling) 기법을 사용하였다.

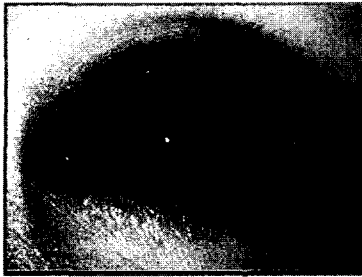


그림 2. 홍채 영상
Fig. 2. Iris image

이에 대한 검증으로는 초점으로부터 좌우대칭 여부를 확인하는 방법을 사용하였으며, 이렇게 찾아낸 동공의 초점으로부터 동공의 반지름(r)을 알 수 있고, 이로부터 동공의 밖으로 반지름 길이의 30% 큰 원을 그릴 수 있다. 이렇게 찾아낸 원으로부터 데이터를 취득할 수 있는데, 이때 원의 반지름 $R=r(1+0.3)$ 이 되고 전체 데이터의 개수는 $R \times \cos 45^\circ \times 8$ 이 된다.

위와 같은 영상 전처리 과정을 통해 얻어진 1차원의 밝기 데이터는 동공의 크기에 따라 밝기 데이터의 길이가 모두 다르다. 이는 동일한 사람이라 할지라도 조명에 따라 동공의 크기가 변하기 때문이며, 이러한 문제점을 해결하기 위하여 본 논문에서는 최근접 보간법을 이용한 resize 기법으로 비교대상의 밝기데이터와 길이를 맞추었다. 이렇게 얻은 데이터에 대해서 DFT계수를 얻을 수 있다. 홍채영상의 특징패턴으로 사용된 DFT계수는 영상의 회전, 이동, 크기에 불변하고, 데이터의 시작점이 다른 경우에도 계수 값이 같으므로 회전에 무관하며, 또한 영상 전체의 밝기를 높이거나 낮출 때에도 계수 값이 변하지 않아 조명등에 의한 영상 밝기 변화에도 무관한 특성을 보인다.

위에서와 같은 방법을 이용하여 각 사람의 눈을 대표하는 코드북을 만들 수 있다. 코드북은 한사람에 대해 세번의 영상으로 이루어진 것으로, 세 개의 밝기데이터 길이를 같게 하고, 평균을 구한 후 DFT계수를 획득하는 방법을 사용하였다. 최종 홍채인식은 코드북(codebook)내에 저장된 기준 데이터와 인식하고자 하는 홍채 데이터간의 공분산 행렬에 기반을 둔 correlation의 계수값에 의한 유사도를 구하여 유사도 값이 가장 높은 코드북을 인식 홍채로 인식한다. 이때 correlation은 공분산 행렬을 이용한 것으로 최소값 0에서 최대값 1 사이의 값으로 표현된다.

2.2 음성을 이용한 화자식별 알고리즘[12]

본 논문에서 사용된 화자식별 알고리즘은 입력음성을 웨이블릿 변환에 의해 각각의 주파수 대역별로 분할한 후 벡터양자화 알고리즘을 이용하여 인식하는 구조로 되어 있다. 벡터양자화를 이용한 일반적인 화자식별 시스템은 그림 3에서 보

는 바와 같이 음성 검출부, 음성 분석부, 훈련부, 인식부로 구성되어 있다. 화자식별 시스템을 구성하는 각 부분을 간단히 설명하면, 우선 음성 검출부에서 음성의 시작점과 끝점을 검출하며, 음성 분석부에서는 각 단어들의 음성에 대한 분석을 거쳐서 최종적으로 멜주파수 켈스트럼계수인 특징벡터를 구하는 부분이다. 훈련부에서는 음성 분석부에서 얻어진 특징벡터들을 K-means 클러스터링 기법을 이용하여 각각의 음성을 대표하는 코드북을 만드는 과정이다. 인식부에서는 식 (1)에서 보인 바와 같이 유클리디안 방법을 이용하여 입력된 화자데이터를 훈련과정에서 만들어둔 각 단어에 대한 코드북과 비교하여 벡터 공간상의 거리가 가장 가까운 것을 선택하고 선택된 코드북에 해당하는 화자를 인식으로 결정하는 부분이다[13].

$$d(x, y) = (x - y) W^{-1} (x - y)^T, \quad W = I \quad (1)$$

여기서, x : 기준 코드북내의 특징벡터
 y : 입력음성의 특징벡터

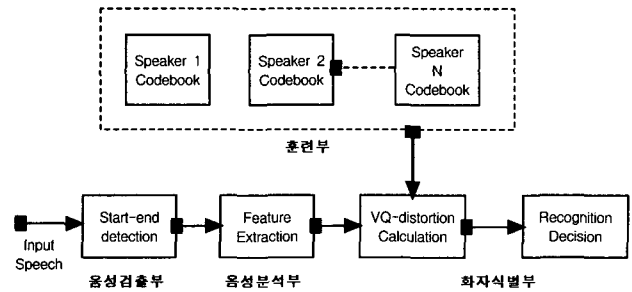


그림 3. 벡터양자화에 의한 화자식별
Fig. 3. Speaker identification based on vector quantization

위와 다르게 그림 4에서는 논문 [12]에서 제안된 웨이블릿 필터뱅크를 이용한 화자식별기의 인식과정을 보여 주고 있다. 여기서, 논문 [12]와 동일하게 웨이블릿 필터뱅크의 구조는 주파수 대역을 균등하게 분할하는 방식인 웨이블릿 패킷구조 방식을 사용하였으며, 화자인식을 위해 사용된 필터는 4개의 필터뱅크로 구성되었고 이중 다양한 실험에서 가장 낮은 인식률을 보인 대역을 제외하였다. 그리고 가장 널리 사용되는 Daubechies 함수를 이용하여 음성신호를 해석하였다.

각각의 필터에서 출력되는 화자의 음성신호는 음성 분석부에서 특징벡터를 계산한 후 기존의 벡터양자화 기법과 동일하게 K-means 알고리즘을 이용하여 독립적인 코드북을 미리 만들어 놓는다. 인식과정에서는 인식하고자 하는 화자의 음성신호가 입력되면 웨이블릿 변환을 하여 주파수별로 음성신호를 분할한 후 음성 분석부에서 각각의 주파수 대역에 대한 특징벡터를 계산한다. 이렇듯 음성 분석부에서 계산된 특징벡터는 미리 벡크별로 만들어 놓은 코드북과의 거리를 계산한 후 독립적인 화자 인식률을 산출한다. 최종적으로는 각각의 대역별에서 출력된 화자인식률을 이용하여 최종 인식률을 결정하기 위한 다중밴드에서의 의사결정방법을 결정해 주어야 한다. 본 논문에서도 논문 [12]와 동일하게 그림 5와 같은 의사결정법칙을 이용하였다.

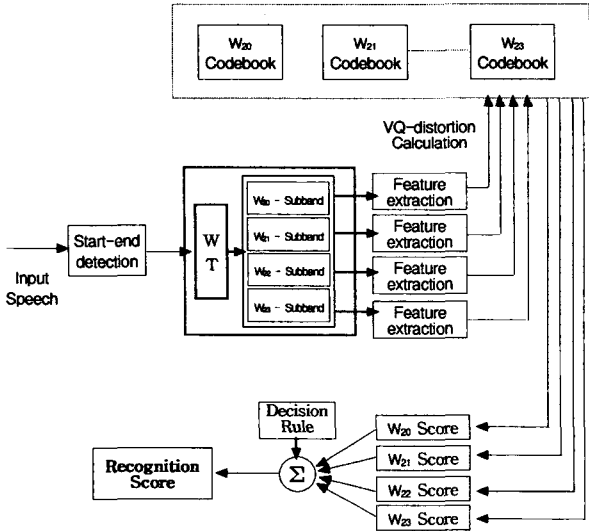


그림 4. 웨이블릿 필터뱅크에 기반을 둔 강인한 화자식별
Fig. 4. Robust speaker identification based on the wavelet filter banks

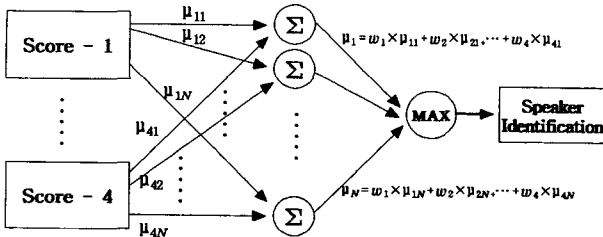


그림 5. 다중밴드에서의 의사결정 방법
Fig. 5. Multiple band decision-making method

그림 5에서 보는 바와 같이 각 대역별에서 산출된 인식률은 음성신호를 프레임으로 나누고 각각의 프레임에서 얻어진 특징벡터와 코드북과의 거리계산에 의하여 산출되기 때문에 어느 특정 단어에 대한 정보만을 가진 것이 아니라 인식하고자 하는 대상에 대한 소속도를 가지고 있다. 이러한 소속도를 μ_{ij} 로 표현했으며, 여기서 μ_{ij} 는 서브밴드 i 에서 단어 j 의 소속도를 나타낸다. 마지막으로 최종인식 단계에서는 각각의 소속도를 합산하여 나타낸 μ_k (k 의 단어가 인식될 소속도)를 구할 수 있으며, 이 중에서 소속도가 가장 높은 화자를 인식대상으로 선정한다. 본 논문에서는 각 밴드별의 인식률에 따른 가중치 w 를 [1 1 1 0]으로 설정하여 최고주파 성분에 대한 영향을 무시하였다.

2.3 홍채와 음성을 결합한 다중생체 인식 알고리즘

그림 6에서는 본 논문에서 제안하는 홍채와 음성을 이용한 고도의 개인확인 시스템 구성도를 나타냈다. 그림 6에 나타낸 바와 같이 홍채인식과 화자식별 결과에 의해 유사도가 가장 높은 값을 선택한 후 최종적으로 표 1과 같은 결정법칙에 의하여 정보보호의 중요도에 따라 임계치를 개별적으로 설정한 개인확인 시스템을 구축하였다. 여기서, 홍채인식과 화자식별 결합에 의한 고도의 개인확인 시스템 성능평가를 위해 인식정도를 정보보호의 중요도에 따라 상급(High), 중급(Middle), 하급(Low)으로 구분하였다. 또한 홍채인 경우

유사도가 0.6 미만, 음성인 경우 유사도가 0.5미만인 경우에는 거절(reject)하는 구조로 되어 있다.

표 1. 최종결정법칙
Table 1. Decision-making rule

구분	최종결정법칙	
승락 (Accept)	상급 (High)	홍채인식에 의한 유사도 ≥ 0.8 또는 음성인식에 의한 유사도 ≥ 0.8
	중급 (Middle)	홍채인식에 의한 유사도 ≥ 0.7 또는 음성인식에 의한 유사도 ≥ 0.7
	하급 (Low)	홍채인식에 의한 유사도 ≥ 0.6 또는 음성인식에 의한 유사도 ≥ 0.5
거절 (Reject)	홍채인식에 의한 유사도 < 0.6 또는 음성인식에 의한 유사도 < 0.5	

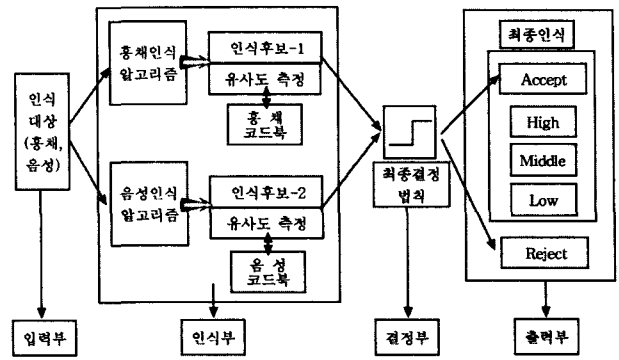


그림 6. 홍채와 음성을 이용한 고도의 개인확인 시스템
Fig. 6. Advanced personal identification system using iris and speech.

3. 실험결과 및 고찰

우선 홍채영상의 취득을 위해 LG전자에서 상용중인 홍채 인식기의 카메라를 사용하였다. 다섯 사람의 눈에 대한 홍채 패턴을 한 사람에 대해 다섯 번 촬영하여 제시된 알고리즘을 MATLAB으로 구현하여 실험하였다. 한사람에 대한 다섯 개의 영상에 대해 3개의 홍채는 기준데이터, 즉 학습데이터 (training data)로 사용하였고, 나머지 2개의 홍채는 검증데이터 (checking data)로 사용하였다.

화자의 음성 데이터는 잠음이 최대인 상태에서 녹음된 우리말 숫자음 /영/에서 실험하였다. 홍채에 사용된 화자인 남성화자 2명과 여성화자 3명이 각각 9회씩 발음한 음성신호 중 3개는 학습, 또는 기준패턴을 만들기 위해서 사용하였고, 나머지 6개는 화자인식실험을 위해서 사용하였다. 녹음된 음성데이터의 샘플링 주파수는 11.025kHz이며, 기준패턴인 코드북의 사이즈는 16으로 정하였다. 일반적으로 음성신호의 특징파라미터는 약 20ms 구간에서 음성신호가 정상 (stationary)이라는 가정아래 20ms의 프레임 단위로 구하여 지는데, 본 논문에서도 20ms의 Hamming window를 사용하였고, 프레임 양 끝단의 신호정보를 보충하기 위하여 10ms 씩 중첩을 시켜서 윈도우를 이동시켰다. 이렇게 Hamming window를 사용하여 원 신호를 프레임 단위로 분할한 후 각

표 2. 홍채인식 결과

Table 2. Personal identification system using iris image

Person (Biometric)	Training data				Test data			
	Accept			Reject	Accept			Reject
	High	Middle	Low		High	Middle	Low	
A (Iris)	3	3	3	0	1	2	2	0
B (Iris)	2	3	3	0	0	0	2	0
C (Iris)	1	2	3	0	0	2	2	0
D (Iris)	3	3	3	0	2	2	2	0
E (Iris)	1	2	3	0	0	1	2	0
Total	10 (66.7%)	13 (86.7%)	15 (100%)	0 (0%)	3 (30%)	7 (70%)	10 (100%)	0 (0%)

표 3. 음성에 의한 화자식별 결과

Table 3. Personal identification system using speech signal

Person (Biometric)	Training data				Test data			
	Accept			Reject	Accept			Reject
	High	Middle	Low		High	Middle	Low	
A (Speech)	3	0	0	0	1	4	6	0
B (Speech)	3	0	0	0	6	6	6	0
C (Speech)	3	0	0	0	4	6	6	0
D (Speech)	3	0	0	0	6	6	6	0
E (Speech)	3	0	0	0	6	6	6	0
Total	15 (100%)	0 (0%)	0 (0%)	0 (0%)	23 (76.7%)	28 (93.3%)	30 (100%)	0 (0%)

표 4. 노이즈가 첨가된 음성에 의한 화자식별 결과

Table 4. Personal identification system using noisy speech signal

Person (Biometric)	Training data				Test data			
	Accept			Reject	Accept			Reject
	High	Middle	Low		High	Middle	Low	
A (Speech)	0	1	3	0	0	0	5	1
B (Speech)	3	3	3	0	6	6	6	0
C (Speech)	0	0	3	0	0	0	3	3
D (Speech)	0	1	3	0	0	1	4	2
E (Speech)	3	3	3	0	3	5	6	0
Total	6 (40.0%)	8 (53.3%)	15 (100%)	0 (0%)	9 (30.0%)	12 (40.0%)	24 (80.0%)	6 (20.0%)

각의 프레임에 포함된 데이터에서 13차의 펠레스트럼 계수를 구하였다. 벡터 양자화 과정에서 음성의 시작점과 끝점을 정확하게 검출하는 것은 매우 중요한데, 본 논문에서는 Rabiner와 Sambur에 의해 제안된 단시간 평균에너지 (Short-time average energy)와 단시간 영교차율 (Short-time zero crossing rate)을 이용한 알고리즘을 사용하였다[13].

본 논문에서 제안한 홍채와 음성을 이용한 고도의 개인확인 시스템 성능을 위해 사용된 데이터 중 학습데이터의 검증을 위해 홍채 1개당 대응하는 화자음성데이터 1개를 적용하여 총 15개의 데이터를 사용하였고, 검증데이터로는 1개의 홍채에 대하여 순차적으로 화자음성 3개를 사용하여 총 사용된 데이터의 수는 [(홍채데이터: 2) × (화자음성데이터: 3) × 5명] 30개의 데이터를 사용하였다.

표 2에서는 홍채인식결과를 보이고 있다. 표 2에서 보는 바와 같이 검증데이터인 경우 보안정도가 높은 상급(High)에 해당하는 값은 총 학습데이터 15개 중 10개로 66.7%의 점유율을 차지하고 있으며, 중급(Middle)인 경우 86.7%, 하급(Low)인 경우는 100%로 나타났고 거절(Reject)인 경우는 0%로 나타났다. 검증데이터인 경우 상급(High)에 해당하는 값은 총 검증데이터 10개 중 3개로 30%의 점유율을 차지하고 있으며, 중급(Middle)인 경우 70%, 하급(Low)인 경우는 100%로 나타났고 거절(Reject)인 경우는 0%로 나타났다. 표 2에서 보는 바와 같이 학습데이터와 검증데이터 모두 거절률은 0%이지만 검증데이터인 경우 확실한 보안이 요구되는(유사도가 80%이상)인 경우는 30%에 지나지 않음을 알 수 있다.

표 3에서는 음성에 의한 화자식별 결과를 보이고 있다. 표 3에서 보는 바와 같이 검증데이터인 경우 보안의 정도가 고정밀급(High)에 해당하는 값은 총 학습데이터 15개 중 15개로 100%의 점유율을 차지하고 있으며, 중급(Middle), 하급(Low), 거절(Reject)들은 모두 0%로 나타났다. 검증데이터인 경우 고정밀급(High)에 해당하는 값은 총 검증데이터 30개 중 23개로 76.7%의 점유율을 차지하고 있으며, 중급(Middle)인 경우 93.3%, 하급(Low)인 경우는 100%로 나타났고 거절(Reject)인 경우는 0%로 나타났다. 표 3에서 보는 바와 같이 적용 알고리즘에 차이가 다소 발생하겠지만 일반적으로 홍채보다 음성에 의한 인식이 보안의 정도가 높은 시스템에 보다 우수한 결과를 보인다고 할 수 있으나, 음성에 의한 화자식별인 경우 노이즈가 존재하는 경우 인식이 매우 저조해지는 문제점을 수반한다.

표 4에서는 음성에 일반적인 차량소음이 첨가된 경우의 화자식별 인식을 보이고 있다. 표 4에서 알 수 있는 바와 같이 노이즈가 첨가됨에 따라 학습데이터인 경우 보안의 정도가 높은 상급(High)이 40%, 중급(Middle)이 53.3%, 하급(Low)이 100%로 나타났고, 검증데이터인 경우 30%, 40%, 80%로 그리고 거절률(Reject rate)도 20%로 나타나 제안된 알고리즘을 사용한 경우에도 어느 정도 노이즈에 민감한 특성을 보이고 있다. 일반적인 벡터양자화(VQ)기법만을 쓴 경우는 거절율이 60%이상을 보였다.

표 5에서는 홍채인식과 음성에 의한 화자식별 방법을 결합한 개인확인 시스템인 경우 인식결과를 보이고 있다. 표 5에서 보는 바와 같이 검증데이터인 경우 보안의 정도가 높은 상급(High)에 해당하는 값은 총 학습데이터 15개 중 15개로 100%의 점유율을 차지하고 있다. 검증데이터인 경우 상급(High)에 해당하는 값은 총 검증데이터 30개 중 26개로 86.7%의 점유율을 차지하고 있으며, 중급(Middle)인 경우 100%, 하급(Low)인 경우와 거절(Reject)인 경우는 모두 0%로 나타났다. 표 5에서 보는 바와 같이 보안이 정도가 높은 상급(High)에서의 인식을 86.7%는 홍채만 사용한 경우보다 56.7%, 음성을 이용한 화자식별 방법만을 사용한 경우보다 10% 정도 인식이 향상됨을 알 수 있다. 표 6에서는 음성에 노이즈를 첨가한 경우 제안된 방법의 인식결과를 보이고 있다. 검증데이터인 경우 보안의 정도가 높은 상급(High)에서의 인식을 60.0%로 나타나 홍채만 이용하여 인식하는 경우보다 인식이 30%, 노이즈가 첨가된 음성만 이용한 경우보다 60% 정도 인식이 향상되었다. 따라서, 본 논문에서 제안된 다중생체인식 기법은 특히 노이즈가 존재하는 환경하에서 우수한 결과를 보이고 있음을 알 수 있다.

표 5. 홍채와 음성을 결합한 개인확인시스템에 의한 인식결과
Table 5. Personal identification system using iris image and speech signal

Person (Biometric)	Training data				Test data			
	Accept			Reject	Accept			Reject
	High	Middle	Low		High	Middle	Low	
A(Iris+Speech)	3	0	0	0	4	6	0	0
B(Iris+Speech)	3	0	0	0	6	6	0	0
C(Iris+Speech)	3	0	0	0	4	6	0	0
D(Iris+Speech)	3	0	0	0	6	6	0	0
E(Iris+Speech)	3	0	0	0	6	6	0	0
Total	15 (100%)	0 (0%)	0 (0%)	0 (0%)	26 (86.7%)	30 (100%)	0 (0%)	0 (0%)

표 6. 홍채와 음성(노이즈 첨가)을 결합한 개인확인시스템에 의한 인식결과

Table 5. Personal identification system using iris image and noisy speech signal

Person (Biometric)	Training data				Test data			
	Accept			Reject	Accept			Reject
	High	Middle	Low		High	Middle	Low	
A(Iris+Speech)	3	3	3	0	3	6	6	0
B(Iris+Speech)	3	3	3	0	6	6	6	0
C(Iris+Speech)	1	2	3	0	0	6	6	0
D(Iris+Speech)	3	3	3	0	6	6	6	0
E(Iris+Speech)	3	3	3	0	3	5	6	0
Total	13 (86.8%)	14 (93.3%)	15 (100%)	0 (0%)	18 (60.0%)	29 (96.7%)	30 (100%)	0 (0%)

4. 결 론

본 논문에서는 홍채와 음성의 다중 생체정보를 이용한 새로운 개인확인시스템을 제안하였다. 제안된 다중 생체인식기법은 홍채인식과 화자식별의 장점만을 부각시켜 최종적으로 결정법칙에 의해 개인의 신분을 확인하는 구조로 설계함으로써 보안의 중요성이 높고 고신뢰성이 요구되는 장소 및 시스템에 적용 가능하도록 하였다.

제안된 알고리즘의 성능을 검증하기 위하여 실험한 결과, 보안의 정도가 높은 상급(High)에 해당하는 인식률은 홍채만을 이용한 경우보다 56.7%, 음성을 이용한 화자식별 방법만을 사용한 경우보다 10% 정도 인식률이 향상됨을 알 수 있었다. 또한 음성에 노이즈를 첨가한 경우에도 보안의 정도가 상급(High)에 해당하는 인식률은 홍채만 이용한 경우보다 인식률이 30%, 노이즈가 첨가된 음성만 이용한 경우보다 60% 정도 인식률이 향상되었다. 따라서, 음성에 노이즈가 첨가한 경우에도 제안된 다중생체인식 기법이 우수한 결과를 보이고 있음을 알 수 있다. 이로부터 본 연구결과는 정보보호의 중요성이 매우 높은 장소 또는 시스템에 적용 가능하리라 기대되며, 이를 위해 많은 데이터와 더불어 제 3자에 의한 성능평가도 행해져야 할 것으로 생각된다.

참고문헌

[1] 특허청, "2001 신기술동향 조사 보고서 - 생체인식시

스템 - ", 2001.

[2] A. Jain, R. Bolle, and S. Pankanti, "Biometrics : Personal Identification in Networked Society", Kluwer Academic Publish, 1999.

[3] 한국전자통신연구원, "생체측정시스템", 2000.

[4] 고현주, 이상원, 전명근, "개인확인 및 인증 알고리즘을 위한 홍채패턴인식", 한국정보처리학회, 제8-C권 제5호, pp. 499-506, 2001.

[5] J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression," IEEE Transaction on Acoustics, Speech and Signal Processing, Vol. 7, pp. 1169-1179, 1988.

[6] Pong C. Yuen, J. H. Lai, "Face Representation Using Independent Component Analysis", The Journal of the Pattern Recognition, Vol. 35, pp. 1247-1257, 2002.

[7] W. W. Boles and B. Boashash, "A Human Identification Technique Using Images of the Iris and Wavelet Transform", IEEE Transaction on Signal Processing, Vol. 46, No. 4, pp.1185-1188, 1998.

[8] J. P. Gampbell, Jr., "Speaker Recognition : a Tutorial", Proceedings of the IEEE, Vol. 85, pp. 1436-1462, 1997.

[9] S. Ben-Yacoub and et al., "Fusion of Face and Speech Data for Person Identity Verification", IEEE Transactions on Neural Networks, Vol. 10, No.5, pp. 1065- 1074, 1999.

[10] U. Dieckmann and et al, "SESAM: A Biometric Person Identification System Using Sensor Fusion", Pattern Recognition Letters, Vol. 18, pp. 827-833, 1997.

[11] L. Hong and A. Jain, "Integrating Faces and Fingerprints for Personal Identification", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol 20, No. 12, pp. 1295-1298, 1998.

[12] 이대중, "웨이블렛 필터뱅크에 기반을 둔 소음에 강한 음성인식시스템", 박사학위논문, 충북대학교, 2002. 8.

[13] L. R. Rabiner and B. H. Juang, Fundamentals of speech recognition, Prentice-Hall, 1993.

저 자 소 개



이대중(Dae Jong Lee)

1995년 : 충북대학교 전기공학과(학사)
 1997년 : 충북대학교 전기공학과(공학석사)
 2002년 : 충북대학교 전기공학과(공학박사)
 2000년~현재 : 한국전기안전공사
 전기안전연구원

관심분야 : 음성신호처리, 서명인식, 다중생체인식



고현주(Hyoun Joo Go)

1999년 : 한밭대학교 제어계측공학과(학사)
2002년 : 충북대학교 제어계측공학과
(공학석사)
2002년~현재 : 충북대학교 제어계측공학과
박사과정

관심분야 : Biometrics, 감정인식, Computer vision



곽근창(Keun Chang Kwak)

1996년 : 충북대학교 전기공학과(학사)
1998년 : 충북대학교 전기공학과(공학석사)
2002년 : 충북대학교 전기공학과(공학박사)
2003~현재 : Dept. Electrical and
Computer engineering, University of
Alberta, Post-doc

관심분야 : 계산지능, 생체인식



전명근(Myung Geun Chun)

1987년 : 부산대학교 전자공학과(학사)
1989년 : 한국과학기술원 전기 및 전자공학
과(공학석사)
1993년 : 한국과학기술원 전기 및 전자공학
과(공학박사)
1993년~1996년 : 삼성전자 자동화연구소
전입연구원

2000년~2001년 : University of Alberta 방문교수

1996년~현재 : 충북대학교 전기전자 및 컴퓨터공학부 교수

관심분야 : Biometrics, 감정인식, 음성신호처리, 얼굴인식