
다중생체인식 기법을 이용한 사용자 인식률 향상

금명환* · 이규원* · 이봉환**

Improvement of User Recognition Rate using Multi-modal Biometrics

Myung-Hwan Geum · Kyu-Won Lee · Bong-Hwan Lee*

요 약

단일 생체인식 시스템의 인식률을 높이는 것은 생체인식 방법마다 취약점이 있기 때문에 그 한계가 있기 마련이다. 얼굴 인식의 경우 조명과 같은 환경적 요인으로 인식률이 저하될 수 있으며, 화자 확인의 경우도 잡음과 같은 환경적 요인으로 인식률이 크게 저하될 수 있다. 따라서 두 가지 이상의 생체특징을 결합하여 다중 생체인식 시스템을 구현함으로써 그 취약점을 보완하는 추세에 있다. 본 논문에서는 얼굴 인식과 화자 확인 시스템을 결합하여 다중 생체인식 시스템을 구현하였고, 일반적인 가중치 합 알고리즘에 환경 변수를 적용하여 기존의 다중 생체인식 시스템보다 인식률을 향상시켰다. 본 시스템은 비밀키 기반의 애플릿으로 구현되어 있으므로 웹상의 사용자 인증을 필요로 하는 응용에 활용될 수 있다.

ABSTRACT

In general, it is known a single biometric-based personal authentication has limitation to improve recognition rate due to weakness of individual recognition scheme. The recognition rate of face recognition system can be reduced by environmental factor such as illumination, while speaker verification system does not perform well with added surrounding noise. In this paper, a multi-modal biometric system composed of face and voice recognition system is proposed in order to improve the performance of the individual authentication system. The proposed empirical weight sum rule based on the reliability of the individual authentication system is applied to improve the performance of multi-modal biometrics. Since the proposed system is implemented using JAVA applet with security function, it can be utilized in the field of user authentication on the generic Web.

키워드

Face recognition, Speaker verification, Multi-modal biometrics, User authentication

I. 서론

인터넷의 발전에 힘입어 통신망에 접속되어 있는 컴퓨터와 인간의 상호작용은 다양한 형태로 급속히 발전하고 있다. 원격에 위치한 컴퓨터 시스템에 접근하여 사용자 인증을 받기 위해서는 비밀번호 또는 인증서

(certificate)를 주로 사용하였으나 비밀번호를 기억하거나 인증서를 소지해야 하는 불편함이 따를 뿐 아니라 해킹으로 인한 위험이 점차 증가하는 추세에 있다. 이러한 문제를 해결하기 위하여 최근 사람이 가지고 있는 생체 정보를 이용한 사용자 인증 기술이 활발하게 연구되고 있다. 생체인식 기술은 하나의 생체 정보를 이용하는 단

* 대전대학교 정보통신공학과

** 교신저자

일 생체 인식 기술과 두 가지 이상의 생체 정보를 이용하는 다중 생체 인식(Multi-modal biometrics)으로 나누어진다. 단일 생체 정보만을 이용하여 하나의 생체 특징에 의한 인증 기술로는 정밀도가 부족하고 실용이 곤란했던 생체 인식을 다수 개의 생체 정보를 활용하여 시스템의 한계를 극복하고 요구되는 성능을 만족시키고 신뢰도를 높이는 다중 생체 인식 기술들이 제안되고 있다[1-3]. 향후 이런 인식 기술들은 PC 보안 제품, 전자상거래, 모바일 전자상거래 시장으로 확대될 것으로 예상된다. 특히 온라인상에서 활용될 수 있는 생체 인식 기술의 개발과 공개 키 기반 인증과의 연동을 통한 새로운 서비스 기술이 창출될 전망이다. 따라서 본 논문에서는 지문과 같은 접촉식인증 방식을 탈피하여 비접촉식인 얼굴과 음성 생체 정보를 이용한 다중 생체 인식 시스템을 제안한다. 또한 현재 널리 사용되고 있는 가중치 알고리즘에 인증 환경 변수를 적용하여 인증의 신뢰도를 향상시킨다.

제 2장에서는 제안한 다중 생체 인식시스템에 대하여 기술하고, 제 3장에서는 얼굴과 음성 데이터를 이용한 다중 생체인식시스템의 실험 결과를 분석한다. 제4장에서는 결론 및 향후 연구 방향을 제시한다.

II. 다중 생체 인식 시스템

1. 제안한 다중 생체인식시스템의 구성

본 논문에서 제안한 다중 생체 인식 시스템은 본 연구팀에서 개발한 PCA(Principal Component Analysis)를 이용한 얼굴인식 시스템[4]과 GMM-UBM(Gaussian Mixture Model - Universal Background Model)을 이용한 화자확인 시스템[5]을 결합한 시스템으로서, 일반적인 다중 생체 인식 방법 중 매칭 단계에서의 융합 방법[6]을 채택하고 있으며, 시스템의 구성도는 그림 1과 같다.

얼굴 인식 시스템은 PCA를 이용하여 최초 사용자가 입력한 얼굴 이미지와 음성 데이터에서 그 특징점을 추출하여 각각의 DB에 저장되어 있던 사용자 템플릿과의 매칭도 즉, 유사도를 평가하고 그 두 유사도를 조도값과 SNR (Signal-to-Noise Ratio)을 고려한 가중치 부여 알고리즘에 의해 최종 결정을 내리게 된다.

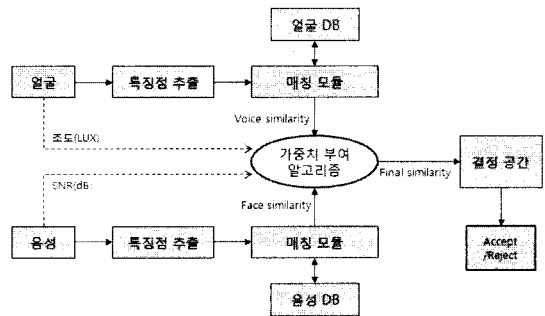


그림 1. 제안한 다중생체인식 시스템의 구성도
Fig. 1. Block Diagram of the Proposed Multi-modal Biometric System

2. 환경변수 및 가중치 부여 알고리즘

다중 생체 인식 시스템의 경우, 단일 생체 인식 시스템으로부터 나온 유사도를 어떤 방법으로 결합하는가에 따라 단일 생체인식 시스템보다 좋은 성능을 가지게 된다[7]. 본 시스템에서는 일반적인 가중치 부여 알고리즘에 인증 환경에 따른 환경변수를 적용하여 각 단일 시스템에 주어지는 일반적인 가중치에 대해서 신뢰도가 높으면 그 수치를 증가시키고, 신뢰도가 낮으면 그 수치를 감소시킴으로써 다중 생체인식시스템의 성능을 개선하도록 하였다. 식(1)은 본 연구에서 제안하는 가중치 부여 알고리즘으로서 각 단일 시스템에 환경변수 e 가 적용되었음을 알 수 있다.

$$e_v \cdot p \cdot S_v + e_f \cdot (1-p) \cdot S_f \quad (1)$$

여기서 e_v 는 화자 확인 시스템에 적용된 환경변수로서, 녹음된 음성 데이터의 SNR에 따라 그 값이 정해지며, SNR이 높을수록 e_v 값도 높아진다. e_f 는 얼굴 인식 시스템에 적용된 환경변수로서, 얼굴 촬영 당시 얼굴의 밝기, 즉 조도값(lux)에 따라 그 값이 정해지며, 조도값이 높을수록 큰 값을 갖게 되고 조도값이 낮을수록 작은 값을 가진다. p 는 본 시스템에 주어지는 가중치로서, 0과 1 사이의 실수값이며, 가장 최적의 수치는 여러 번의 실험 과정을 거쳐 정해진다. 그리고 S_v 와 S_f 는 각각 화자 확인 시스템과 얼굴 인식 시스템의 유사도이다.

3. 다중 생체 인식 시스템 구현

본 시스템은 그림 2에 나타난 바와 같이 온라인상에서도 안전하게 사용자의 얼굴과 음성으로써 인증을 받

을 수 있도록 클라이언트와 서버간의 모든 통신은 암호화 되도록 구현되었다[8-9]. 먼저 클라이언트에서는 사용자 얼굴과 음성 에 대한 생체 데이터를 획득하기 위해서 CCD 카메라와 마이크가 필요하며, 획득한 생체 데이터들은 서버로 암호화 되어 전송된다.

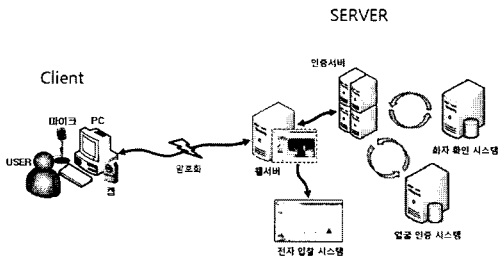


그림 2. 다중 생체인식시스템의 구현
Fig. 2. Implementation of Multi-modal Biometric System

여기서 암호화는 비밀키 기반 구조로 되어 있으며, Triple DES와 Diffie-Helman 키 교환을 이용하였다. 서버로 전송된 사용자의 생체 데이터들은 각각의 단일 인증 시스템에서 인증 절차가 이루어지고, 그 결과를 보고 인증서버에서 최종 결정을 내리게 된다. 최종 승인 결과가 결정되면 사용자는 웹서버에 구축되어 있는 타겟 응용인 전자 입찰 시스템에 로그인 된다.

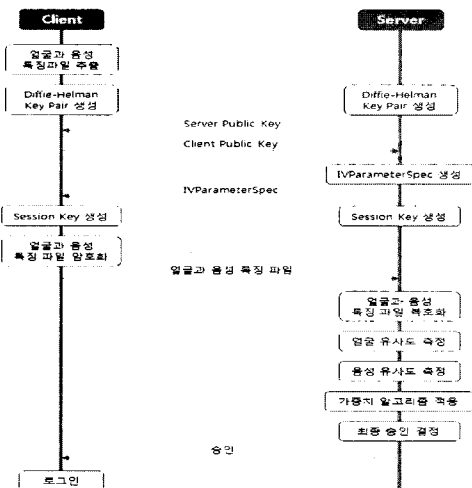


그림 3. 사용자 인증 과정
Fig. 3. User Authentication Process

그림 3은 얼굴과 음성 데이터를 이용한 사용자 인증 과정을 나타낸다.

그림 4는 사용자가 초기 인증을 위하여 웹서버에 접속 시 전자 입찰 시스템에 로그인 을 위해서 얼굴과 음성 데이터를 입력하기 위한 인터페이스이다.



그림 4. 얼굴과 음성 데이터 입력 인터페이스
Fig. 4. Face and Voice Data Input Interface

사용자는 얼굴 데이터를 입력하기 위하여 카메라를 바라보고 촬영 버튼을 누르면 얼굴 이미지가 저장된다. 이때 카메라와의 거리는 얼굴 크기가 화면상에 2/3정도 차지하는 것이 좋다. 이어서, REC 버튼을 누르고 지정된 문장을 발성 후 정지 버튼을 누르면 음성 파일이 저장되고 사용자 인증 절차가 시작된다.

인증 절차 후 사용자 승인이 결정되면 사용자는 타겟 응용으로 개발한 전자 입찰시스템에 로그인되어 전자 입찰을 진행할 수 있다. 그림 5는 전자 입찰시스템에 로그인 후 메인 화면이다.

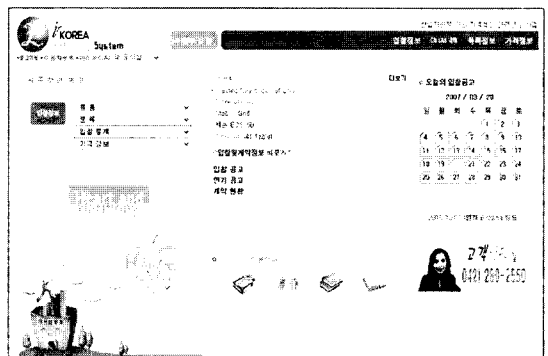


그림 5. 로그인 후 전자 입찰 시스템 메인 화면
Fig. 5. Screen shot of the E-bidding System

본 전자 입찰 시스템은 입찰 정보, 목록 정보, 가격 정보, 마이페이지, 물품 검색, 운전자 공지, 입찰 및 계약정보 바로가기, 오늘의 입찰 공고 및 이용 안내 기능이 제공되며, 사용자는 입찰을 할 경우 상단의 입찰 정보를 클릭하면 경매 입찰 할 수 있는 화면으로 이동하게 된다.

III. 실험 및 결과

1. 얼굴 DB

얼굴 인식은 카메라로 촬영된 얼굴 이미지에서 그 특징점을 찾아 인식하기 때문에 얼굴 이미지에서 눈, 코, 입, 얼굴 등 그 윤곽이 뚜렷이 드러나야 한다. 때문에 얼굴 인식에 가장 큰 영향을 끼치는 것은 얼굴에 비치는 조명의 밝기이다. 따라서, 조명의 밝기를 일반 사무 환경에서 접할 수 있는 한도 내에서 세 단계로 나누어 얼굴 이미지 DB를 구축하였다.

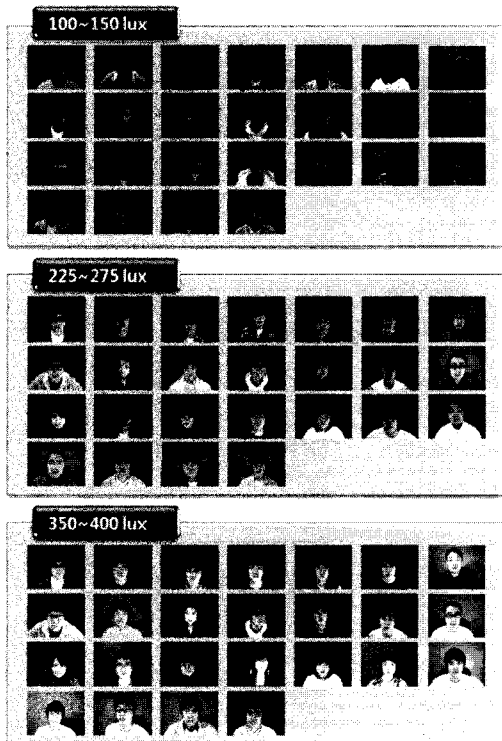


그림 6. 서로 다른 조도에서 촬영한 얼굴 이미지
Fig. 6. Face Image DB under Different Illumination Environment

일반적으로 사무실이나 공부방에 알맞는 밝기는 200~300LUX이기 때문에, 조도계를 이용하여 어두울 때 (100~150 LUX), 보통일 때(225~275 LUX), 밝을 때 (350~400 LUX)로 적절히 구분하여 각각의 밝기에서 촬영을 실시하였다. 촬영 회수는 일인당 각각의 밝기에서 10회씩 총 30회 촬영을 하고, 총 실험 인원은 25명으로 하였다. 따라서, 자기 인증 실험 회수는 750(25*30)회 할 수 있었고, 사칭자 인증 실험 회수는 각각의 밝기에서 600(25*24)회씩 실험하여 총 1800(600*3)회 실시하였다. 각 환경에서 촬영한 얼굴 이미지는 그림 6과 같다.

2. 음성 DB

본 논문의 실험에서 사용한 음성 DB는 ETRI 한국어 중가(中價) 마이크 화자 인식용 DB이다. 이는 ETRI의 음성정보 연구센터에서 구축한 한국어 화자인식용 음성 DB로, 사무실 PC 환경에서 중가의 마이크를 이용하여 250명 (기간별 : 주차 100명, 월차 100명, 3개월차 50명)의 화자가 발성한 2연 숫자, 4연 숫자, 문장으로 구성된 음성 DB이다. 문장 음성의 발생 목록은 개인정보와 관련된 10개의 질문과 3어절 이내로 구성된 단문 10개로 구성되며, 한 화자당 동일한 목록을 5회 발생하고, 녹음 간격에 따라 주차/월차/3개월차로 구분하여 4회 반복한 것이다.

녹음 환경은 SNR 25 dB 이상 확보가 가능한 조용한 사무실 환경으로, 펜티엄 III 이상의 데스크탑 PC에서 중가 마이크(모델명: Shenheiser MD425)를 통해 수집한 것이다. 사운드카드로는 Sound Blaster PCI 128 Digital을 사용하였고, FP-11 Pre AMP를 사용하여 적절한 에너지 레벨이 되도록 하였다. 음성 데이터는 16KHz/16Bit, Linear PCM, Intel byte-format으로 저장되었다.

본 논문에서 화자확인 실험은 주차 화자 50명을 대상으로 구성되었다. 그리고 훈련 시점과 1주일 차이의 음성인 주차화자의 1주차 음성을 사용하였다. 이 테스트 DB는 훈련과 독립적인 3개 단문(훈련에 사용되지 않은 단문)의 5회 음성을 사용하여 각 화자당 15 단문으로 하였다. 그리고 실험 대상 화자모델의 MAP 화자적용 시각 화자의 0주차(주차 화자의 음성중 기간별 첫 녹음 시점) 6개 단문의 5회 발생 음성을 사용하였는데, 화자당 60초(6000 프레임)로 제한하였다.

잡음 보상에 대한 실험을 위해 잡음 환경의 음성을 제작하였다. 화자확인 인식 시 사용된 테스트 음성(주차

화자 50명의 1주차 화자당 3단문의 5회 음성)에만 잡음을 추가하였다. 추가된 잡음은 균중 잡음으로 ITU-T 권고에 근거한 5dB, 10dB, 20dB의 SNR을 적용하였다.

사용된 모든 음성은 음성구간 검출을 위해 앞, 뒤의 묵음(전 200ms, 후 200ms)을 제거한 단문의 음성을 프레임 길이 25ms, 프레임 주기 10ms, Hamming window[10]를 사용하여 MFCC 13차를 추출하였다. 특징 추출은 HTK ver 3.3을 이용하였다.

3. 실험 방법

본 다중 생체인식시스템에서는 얼굴과 음성의 두 생체 데이터가 필요하므로, 위에서 설명한 얼굴 DB와 음성DB를 적절하게 일대일 매칭시켜 실험을 실시하였다. 실험 환경으로는 얼굴 DB에서의 세 가지 환경과 음성 DB에서의 세 가지 환경을 적절히 조합해 그림 7에서와 같이, 총 9 가지 실험 환경을 구성하였다.

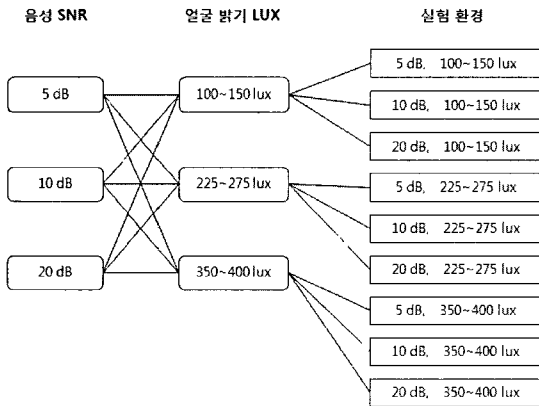


그림 7. 다중 생체인식시스템의 실험 환경
Fig. 7. Experimental Environment of the Multi-modal Biometric System

이와 같이 다양한 실험 환경을 꾸민 이유는 다중 생체인식 시스템에서는 각각의 단일 시스템이 서로 보완적인 관계에 있기 때문에 한쪽의 시스템이 좋지 않은 환경에 있을 경우 다른 한쪽의 시스템이 얼마나 보완해 줄 수 있는지를 분석하기 위해서다.

실험은 각 환경에서 자기 인증을 750회, 사칭자 인증을 7500회 실시하였으므로, 총 실험 회수는 자기 인증 6,750(750*9)회, 사칭자 인증 67,500(7,500*9)회가 된다.

4. 실험 결과 및 분석

표 1은 각 환경에서 실험을 통해 단일 인식시스템과 다중 생체인식시스템에서 환경변수를 적용했을 경우와 적용하지 않았을 경우의 EER(Equal Error Rate)[3]을 정리하여 나타낸 것이다.

환경변수 미적용과 환경변수 적용은 다중 생체 인식 시스템에서 사용한 가중치 부여 알고리즘에 환경변수를 적용한 경우와 적용하지 않은 경우를 나타낸다.

표 1. 각 실험환경에서의 EER(%)
Table 1. EER for Each Experimental Environment(%)

실험환경 (dB, lux)	얼굴인식	화자확인	환경변수 미적용	환경변수 적용
5, 100~150	14	28.66	9.91	-
5, 225~275	7.4	28.66	6.08	5.98
5, 350~400	9.35	28.66	7.91	7.31
10, 100~150	14	19.9	9.4	8.12
10, 225~275	7.4	19.9	4.84	4.64
10, 350~400	9.35	19.9	6.12	-
20, 100~150	14	11.2	7.25	5.91
20, 225~275	7.4	11.2	4.11	-
20, 350~400	9.35	11.2	4.3	4.1

얼굴인식의 경우 조도값(lux)에 따라 인식률이 다르고, 화자확인의 경우 SNR에 따라 그 인식률이 다르다는 것을 알 수 있다. 또한, 모든 환경에서 다중 생체인식시스템의 EER이 현저히 낮아졌고, 환경변수를 적용했을 경우에는 적용하지 않았을 때와 비교하여 EER이 더 감소함을 알았다. 표 1에서 측정된 각 시스템의 성능 차이를 요약하면 표 2와 같다.

표 2에서 보는 바와 같이, 본 연구에서 제안한 다중 생체인식시스템은 얼굴 인식 시스템 보다는 [20dB, 100~150lux]의 환경에서 최고 8.09%, 화자 확인 시스템 보다는 [5dB, 225~275lux]의 환경에서 최고 22.68%, 환경변수를 적용한 경우 적용하지 않았을 경우 보다는 [20dB, 100~150lux]의 환경에서 최고 1.34% 성능이 향상되었다.

표 2. 다중인식시스템의 성능 개선 비율(%)
Table 2. Recognition Rate Improvement of Multi-modal Biometric System

실험환경 (dB, lux)	얼굴인식 대비 개선율(%)	화자확인 대비 개선율(%)	환경변수 미적용 대비 개선율(%)
5, 100~150	4.09	18.35	-
5, 225~275	1.42	22.68	0.1
5, 350~400	1.51	20.82	0.6
10, 100~150	4.6	10.5	1.28
10, 225~275	2.59	15.09	0.2
10, 350~400	3.23	13.78	-
20, 100~150	8.09	5.29	1.34
20, 225~275	3.29	7.09	-
20, 350~400	5.20	7.05	0.2

그림 8은 표 2를 도식화 한 것으로 얼굴인식 대비 개선율은 [20dB, 100~150lux]의 환경에서, 화자확인 대비 개선율은 [5dB, 225~275lux]의 환경에서, 환경변수 미적용 대비 개선율은 [20dB, 100~150lux]의 환경에서 최고의 수치를 보였다.

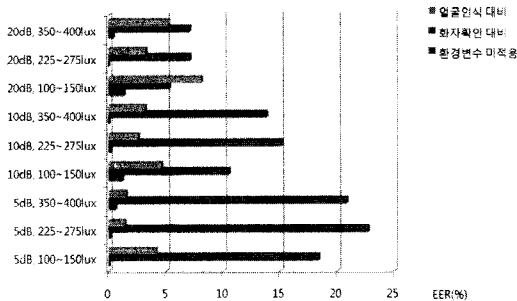


그림 8. 각 환경에서의 성능 개선 비율 요약(%)
Fig. 8. Summary of Recognition Rate Improvement

IV. 결론 및 향후 연구

본 논문에서는 얼굴 인식 시스템과 화자 확인 시스템을 가중치 부여 알고리즘으로 결합하여 단일 인증시스템과의 성능을 비교 분석하였다. 또한, 얼굴과 음성의 생체 정보를 입력할 때 생체 인식 시스템의 성능 저하의 요인이 될 수 있는 사용자 환경 중에서 얼굴 표면의 밝기와

녹음된 음성파일의 잡음 정도를 이용하여 그 환경변수를 정하여 인식율의 차이를 비교하였다.

구현한 다중인식시스템의 인식률은 얼굴 인식 시스템 보다는 [20dB, 100~150lux]의 환경에서 최고 8.09%, 화자 확인 시스템 보다는 [5dB, 225~275lux]의 환경에서 최고 22.68%, 환경변수를 적용하지 않았을 경우 보다는 [20dB, 100~150lux]의 환경에서 최고 1.34% 성능이 향상됨을 보였다. 또한, 그 인식률은 인증 환경에 따라서 크게 차이를 보였다. 따라서 각 시스템에 적용되는 가중치에 환경 조건에 따른 환경변수를 적용함으로써 인식률을 향상시킬 수 있음을 확인하였다. 하지만 환경변수를 적용하지 않았을 때보다 적용하였을 경우 그 인식률의 향상이 비교적 미세하여 좀 더 많은 환경 변수를 연구하여 적용해 보고, 환경변수의 범위 또한 다양하게 적용함으로써 좀 더 높은 성능 개선점을 찾아야 할 것이다.

감사의 글

본 논문은 산업자원부 지역혁신인력양성사업 및 정보통신연구진흥원의 대학 IT연구센터 지원사업(IITA-2008-C1090-0801-0014)의 연구결과로 수행되었음.

참고 문헌

- [1] M. Indovina, U. Uludag, R. Snelick, A. Mink, and A. Jain, "Multimodal Biometric Authentication Methods: A COTS Approach", Proc. MMUA 2003, Workshop on Multimodal User Authentication, Santa Barbara, CA, Dec. 11-12, 2003.
- [2] Arun Ross, Anil Jain, "Information fusion in biometrics", *Pattern Recognition letters*, Vol. 24, pp. 2115-2125, 2003.
- [3] D. Dessimoz, J. Richiardi, C. Champod, and A. Drygajlo, "Multimodal Biometrics for Identity Documents", *In Proc. of the 4th European Academy of Forensic Science*, June 13-16, Helsinki, Finland, 2006.
- [4] Kyu-Won Lee, "Realtime Face Tracking using Motion Analysis and Color Information," 한국 해양 정보통신 학회논문지 제 11권 5호, pp. 977-984, May. 2007.

- [5] Jae-Kil Choi, "A study on Histogram Reformation for performance improvement of speaker verification," Master's Thesis, Graduate School, Daejeon University, Oct. 2006.
- [6] Arun Ross, Anil Jain, "Information fusion in biometrics," Pattern Recognition letters, Vol. 24, pp. 2115-2125, 2003.
- [7] Hyo-Sup Kang, "An Empirical Study of Multi-Modal Biometrics using Face and Fingerprint," 한국정보과학회 학술발표논문집 제29권 제2호, pp. 622-624, 2002.
- [8] Sung-keun Yun, "A Multiple Signature Authentication System Based on BioAPI for WWW," 정보과학학회 논문지 제31권 제9호, Sep. 2004.
- [9] Biometrics Working Group, "Best Practices in Testing and Reporting Performance of Biometric Devices Version 2.0," Biometrics Working Group, Aug. 2002.
- [10] Oppenheim, A.V., and R.W. Schaffer, *Discrete-Time Signal Processing*, Prentice-hall, pp. 447-448, 1989.



이 봉 환(Bong-Hwan Lee)

1985년 서강대학교 전자공학과 졸업(학사)

1987년 연세대학교 대학원 전자공학과 졸업(석사)

1993년 Texas A&M 대학교 대학원 전기 및 컴퓨터공학과 졸업(박사)

현재 대전대학교 정보통신공학과 교수

※관심분야 : 그리드컴퓨팅, 유비쿼터스헬스케어, 네트워크보안 등

저자소개



금 명 환(Myung-Hwan Geum)

2006년 대전대학교 정보통신공학과 (학사)

2008년 대전대학교 정보통신공학과 (석사)

현재 (주)베넷 정보기술 연구원

※관심분야 : 그리드컴퓨팅, 유비쿼터스헬스케어, 네트워크보안 등



이 규 원(Kyu-Won Lee)

1986년 연세대학교 전자공학과(학사)

1988년 연세대학교 대학원 전자공학과(석사)

1998년 연세대학교 대학원 전자공학과 박사(박사)

현재 대전대학교 정보통신공학과 부교수

※관심분야 : 영상처리, 멀티미디어 등