

논문 2005-42SC-4-5

# 모듈화된 얼굴인식 시스템을 이용한 성능 시험에 관한 연구

## (A Study on the Performance Evaluation based on Modular Face Recognition System)

홍 태 화\*, 문 현 준\*\*, 신 용 녀\*\*\*, 이 동 근\*\*\*, 김 재 성\*\*\*

(Tae-Hwa Hong, Hyeon-Joon Moon, Yong-Nyuo Shin, Dong-Geun Lee, and Jae-Sung Kim)

### 요 약

생체인식 기술 중 변별력과 활용성, 편리성이 뛰어난 얼굴인식 기술은 출입통제나 금융관련 업무 처리와 같이 보안관련 응용분야에서 필요성이 급속도로 요구되고 있다. 따라서 얼굴인식 알고리즘의 발전과 더불어 현 기술의 상태를 파악하고 발전 방향을 제시하기 위한 성능 시험 방법에 대한 연구 또한 중대한 이슈로 부각되고 있다. 본 연구에서는 얼굴인식 시스템의 성능 시험을 위한 프로토콜의 설계 기준을 제시하고 XM2VTS 데이터베이스를 사용하여 PCA를 기반으로 한 인식 시스템을 디자인하여 Identification 시나리오와 Verification 시나리오 상에서 성능 시험 결과를 제시한다.

### Abstract

Face recognition out of biometrics is considerable interesting due to high performance and accessibility in applications to security such as access control and banking service. Therefore, a study on the protocol of the performance test is an important issue to understand the art-of state and to show a direction in future works, in addition to developing algorithms. We present a design criterion for the performance test protocol of face recognition system and show the result of experiment executed on identification and verification scenario based on PCA algorithm and XM2VTS DB

**Keywords :** Face Recognition System, Performance test Protocol, PCA, Identification scenario, Verification scenario

## I. 서 론

생체인식 기술은 주요시설의 출입통제, 금융관련 업무 처리와 같이 보안이 필수적인 응용분야에 있어서 필요성이 급속도로 증대되고 있다. 특히 최근에는 생체인식 기술로서 변별 능력이 탁월하고 활용성 및 편리성이

뛰어난 얼굴인식 기술이 크게 부각되고 있다<sup>[1]</sup>.

얼굴인식(face recognition) 기술이란 입력된 영상에 존재하는 사람의 얼굴을 검출하여 데이터베이스에 저장된 얼굴 영상들과 비교하여 그 신원(identity)을 인증(verification) 또는 인식(identification) 하는 기술을 일컫는다. 얼굴인식 기술의 장점은 다른 생체인식기술인 지문인식, 홍채인식, 정맥인식 등과는 달리 자신의 신체 일부를 인식장치에 직접 접촉이 불필요 하므로 다른 방식에 비해 생체정보의 획득방법에서 강제성이 적으므로 다른 생체인식 기술이 사용자에게 줄 수 있는 거부감과 불편함이 존재하지 않는다. 이러한 특징들을 바탕으로 최근 10여 년 동안 얼굴인식에 대한 연구는 정지된 얼굴영상(still image)에서부터 비디오영상을 통한 얼굴인식에 이르기까지 광범위한 분야에서 적극적으로 진행되어 왔다<sup>[2]</sup>. 90년대 초반 Turk와 Pentland에 의해 PCA

\* 학생회원, 연세대학교 전기전자공학과  
(Dept. of Electrical Electronic Engineering, Yonsei University)

\*\* 정회원, 세종대학교 소프트웨어공학과  
(School of Computer Engineering Department of Computer Software, Sejong University)

\*\*\* 정회원, 한국정보보호진흥원  
(Korea Information Security Agency)

※ 본 연구는 한국정보보호진흥원의 생체인식시스템 시험기술 표준화 연구비로 지원됨

접수일자: 2005년1월7일, 수정완료일: 2005년7월1일

를 이용한 eigenfaces<sup>[3]</sup> 기법이 알려졌고 얼굴영상의 클래스간의 유사도를 최대화하고 동일 클래스 내에 있는 영상들의 유사도를 최소화시키는 fisherface<sup>[4]</sup> 기법이 제시되었다. 그러나 이러한 통계적 기법들은 선형성에 의존하기 때문에 한계성을 드러냈고 비선형인 얼굴영상의 변화요인들을 모델링하기 위해 커널을 이용한 KPCA 방법들이 제시되었다.

인간 스스로에 있어서 얼굴인식 메커니즘은 그다지 어려운 문제가 아니지만 컴퓨터 시스템을 이용한 얼굴인식은 얼굴 자체의 상호 유사성과 외부의 환경이나 내재적인 요소로 인해 다양한 변화를 갖는 얼굴의 특성으로 인해 결코 쉬운 일이 아니기 때문에 얼굴인식은 컴퓨터 비전, 패턴인식 연구 분야에서 가장 도전적인 과제 중 하나이다<sup>[1]</sup>.

얼굴인식 기술은 두 가지 주요 카테고리를 통해 점점 발전해 나가고 있다. 하나는 전처리, 얼굴검출, 형태추출, 분류기법 등 얼굴인식관련 문제에 대한 신뢰적인 해답을 제시하는 알고리즘의 발전이고 또 하나는 다양한 시나리오와 영상 데이터의 설정을 통한 성능시험 방법의 발전으로서 시스템을 수행할 때 발생하는 다양한 요구조건들을 알고리즘이 얼마나 잘 대처하는 지를 측정하는 것이다<sup>[6]</sup>. 특히 후자의 경우는 얼굴인식에서 최근 새롭게 대두되고 있는 관심분야로서 성능평가(performance evaluation)와 학습기반예측(predictive learning)과의 상호작용을 보다 명확히 제시하려는 것이다. 즉 주어진 영상정보에 대한 성능을 평가하고 이를 바탕으로 미래에 주어질 영상정보들에 대한 성능을 예측할 수 있어야 한다는 것이다. 통계적 학습이론들은 미래 영상정보들을 예측할 때 발생할 수 있는 위험요소들을 추정해 낼 수 있는 기법들을 제공한다. 특히 모델의 성능은 분류기(classifier)의 복잡성이나 테스트 집합에 대한 트레이닝 집합의 상대적인 영상의 수나 질에 많은 영향을 받게 된다.

실질적으로 대부분의 얼굴인식 관련 연구에서는 다양한 알고리즘을 제안하고 있고 이들은 상대적으로 적은 영상데이터를 통해 보통 95%이상의 인식률을 보고하고 있는 실정이다. 그러나 이러한 알고리즘들은 서로 다른 가정과 영상데이터, 평가방법을 채택하고 있기 때문에 직접적인 비교가 불가능하다. 따라서 표준화된 얼굴영상 DB와 얼굴인식 시스템의 경쟁력을 비교, 시험할 수 있는 성능평가 방법의 제시는 그 필요성이 절실하다고 할 수 있다. 따라서 특정한 응용 환경에서 얼굴인식 시스템을 활용하기 위해서는 어떠한 인식 모델을

이용하는 시스템을 도입하는 것이 효과적인지, 또는 같은 모델을 이용하더라도 어떠한 분류기를 이용하는 것이 바람직할지 결정하는 것이 중요한 문제가 된다. 얼굴인식 시스템의 성능을 객관적으로 신뢰성 있게 시험하는 기술은 이러한 문제에 대한 해결책을 제시하려는 의도에서 연구되기 시작하였고 개발자 입장에서도 개발된 기술에 대한 객관적 평가로 인한 신뢰성과 시장성 확보 면에서 매우 중요하다고 할 수 있다.

본 연구에서는 얼굴인식 시스템의 성능 시험을 위한 설계 기준을 제시하고 PCA를 기반으로 한 얼굴인식 시스템을 이용하여 인식 및 인증의 두 가지 성능 시험 시나리오에 대한 시험 결과를 제시한다. II장에서는 성능 시험 프로토콜의 설계 기준을 제시하고, III장에서는 얼굴인식 시스템의 모듈별 구성을 설계하고 PCA를 벤치마크 알고리즘으로서 설명한다. IV장에서는 성능 시험 시나리오를 제시하고 V장에서는 다양한 성능 시험 방법과 PCA를 이용한 성능 시험 결과를 제시한다.

## II. 성능 시험 프로토콜 설계

### 1. 성능 시험 계획 단계

성능시험 프로토콜을 계획하는 단계는 전체 성능시험 과정의 가장 핵심적인 단계이다. 성능시험의 절차 및 방법에 대한 객관적이고 명확한 규정이 설정되어야 보다 정확하고 신속한 성능시험이 수행될 수 있다. 따라서 얼굴인식 시스템과 성능 분석 시나리오와 같은 기술적인 부분과 별도로 피시험자의 자격 기준, 시험 전 고려사항, 피시험자 관리 기준 등을 미리 설정하여 기술적 성능 시험의 성공 가능성을 높일 수 있다.

성능시험 프로토콜은 공정성 기준에 의거하여 설계되어야 하고 객관적이고 엄정한 피시험자의 자격 기준을 설정하여 독창적이고 상업적으로 이용가치가 충분한 시스템의 참가를 유도해 낼 수 있어야 한다. 또한 피시험자는 자신의 시스템을 성능 시험 전까지 최적화시키고 성능 시험 목적을 명확하게 정의하여야 한다<sup>[4]</sup>.

얼굴인식 시스템 시험은 목표에 따라 기술 시험, 시나리오 시험, 운영 시험의 세 가지 큰 형태 중의 하나로 분류할 수 있다<sup>[6][7]</sup>. 기술 시험(technology evaluation)의 목표는 단일한 얼굴인식 기술에서 경쟁 관계에 있는 알고리즘들을 비교하는 것이다. 알고리즘 시험은 일반적인 카메라에 의해 수집된 표준화된 데이터베이스에서 수행된다. 그러나 이 같은 데이터베이스에 대한 성능도 데이터가 수집된 환경과 데이터에 등록된 사람들의 인

적 구성(population)에 따라 변화하므로, 시험할 알고리즘에 대해서 너무 쉽지도 너무 어렵지도 않은 데이터베이스를 구축하는 것이 중요하다. 알고리즘 개발과 튜닝을 위한 샘플 데이터는 시험 전에 분배되지만, 실제 시험은 사전에 공개되지 않았던 데이터에 대해 이루어져야 한다. 시험은 일반적으로 오프라인으로 수행하며 그 결과는 재현할 수 있어야 한다. 한편 이러한 알고리즘 시험을 위해 기준이 되는 벤치마크 알고리즘의 선정 또한 중요하다.

시나리오 시험(scenario evaluation)의 목표는 프로토타입 또는 모의실험에서 전반적인 시스템 성능을 결정하는 것이다. 즉, 실제 응용을 모델링한 환경에서 완전한 시스템을 시험한다. 각 시스템이 독자적인 획득용 카메라를 사용함에 따라, 같은 환경에서 같은 구성원에 대한 데이터를 수집하도록 하는 것이 중요하다. 각 시스템의 데이터 저장 용량에 따라 시험은 오프라인 또는 온라인 형태로 수행할 수 있다. 시험 결과의 재현은 모델링된 시나리오가 조절될 수 있는 한도 안에서 가능하다.

운영 시험(operational evaluation)의 목적은 특정한 대상 구성원을 가진 특정한 환경에서 완전한 시스템의 성능을 결정하는 것이다. 시험 시스템의 데이터 저장 용량에 따라 오프라인 시험은 불가능한 경우도 있다. 운영 환경간의 명시되지 않는 차이 때문에 (얼굴인식 시스템의 경우에 조명, 포즈, 표정 등의 변화) 운영 시험의 결과는 일반적으로 재현이 불가능한 경우가 많다.

## 2. 성능 시험 진행 단계

성능시험 수행 전의 계획이 설정되면 피시험자를 모집하고 성능시험에 대한 프로토콜 공고, 입출력 파일 형식 설계, 사전 테스트를 거쳐 실질적인 성능시험을

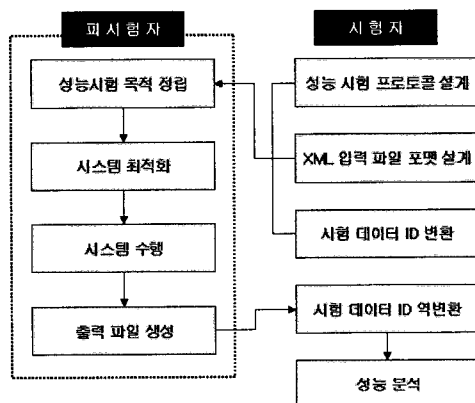


그림 1. 성능 시험 수행 절차  
Fig. 1. The procedure of performance test.

수행하게 된다. 실질적인 성능시험은 시험자가 피시험자에게 입력파일(target 및 query의 정보 파일, 영상저장 폴더를 담은 USB 드라이브 또는 하드 디스크 등)을 제공하면서 시작된다. 이 때 각 영상파일들은 성능시험의 공정성을 위해 변환된 ID를 사용하여 파일명이 재작성 된다. 물론 이러한 파일명은 출력파일의 생성 후 다시 원 ID를 포함하는 파일명으로 복구하는 과정을 거치게 된다.

## III. 모듈화된 얼굴인식 시스템

본 연구에서는 성능 시험의 효과적 수행을 위해 얼굴인식 시스템의 모듈화를 구현한다. 모듈화된 얼굴인식 시스템은 크게 전처리, 특징추출, 인식(인식의 결과는 성능 시험의 데이터가 됨)의 3가지 모듈로 구분할 수 있다. 이를 통해 성능 시험의 객관성을 유지할 수 있다.

모듈화된 얼굴인식 시스템을 구현하기 위한 벤치마크 알고리즘으로써 PCA 기반 인식 알고리즘을 이용한다. PCA는 2차 모멘트의 통계적 정보를 이용한 선형 인식 알고리즘으로서 가장 널리 알려져 있고 이로 인해 대다수의 논문들이 PCA를 벤치마크하여 논문의 독창성 및 우수성을 주장한다.

### 1. Principal Component Analysis (PCA)

PCA는 샘플 데이터로부터 최적화된 특징을 추출해 내기 위한 통계적 차원 축소 기법이다<sup>[8]</sup>. 1990년에 Kirby와 Sirovich는 인간의 얼굴을 분석하고 표현하는데 PCA를 이용하였고 1991년에 Turk와 Pentland<sup>[3]</sup>는 얼굴을 인식하는 과정에서 PCA를 이용하였다. PCA 기반 얼굴 인식 시스템에 있어서 입력은 동일한 해상도를 갖는 N장의 샘플 얼굴 영상이 된다. PCA를 수행할 때 샘플 영상들의 해상도가  $n \times m$  이라면 각 영상은 벡터공간  $\mathbb{R}^{n \times m}$ 에 속한 하나의 벡터로서 표현된다. 이 때 총 N개의 벡터  $t_1, \dots, t_N$ 가 생성되는데 이들을 정규화시키면 N-1개의 고유벡터,  $e_1, \dots, e_{N-1}$ 와 고유값,  $\lambda_1, \dots, \lambda_{N-1}$ 을 계산할 수 있다. 이러한 고유벡터들을 특히 고유얼굴(eigenfaces)이라 부른다. 이 때 생성된 고유벡터들을 정규화시키고 이들을 고유값의 크기가 감소하는 방향 ( $\lambda_k > \lambda_{k+1}$ )으로 정렬한다. k번째 고유값,  $\lambda_k$ 는 k번째 고유벡터 위로 샘플 영상들을 투영했을 때의 분산값을 의미하게 되고 정렬 수행에 의해 높은 순위(k는 k+1보다 높은 순위가 됨)에 있는 고유벡터일수록 샘플 영상

간의 더 많은 변화를 내포하게 된다. 따라서 낮은 순위의 고유벡터일수록 영상간의 작은 변화를 내포하고 이들은 노이즈로 간주되어 보통 계산에서 제외한다. 만약  $M \leq N-1$  인  $M$ 을 선택한다면 정렬된 순서에 의거하여  $M$ 개의 고유벡터가 얻어지고 이들로 구성된 공간에 얼굴 영상들을 투영하면  $M$ 차원의 공간이 생성되는데 이를 얼굴 공간(face space)이라 부른다. 따라서 이 때 투영된 얼굴은  $M$ 차원 벡터공간상에 존재하는 하나의 벡터가 된다.

만약 인식단계를 위한 갤러리(gallery : 얼굴인식 시스템에 미리 공개되어 정보가 데이터베이스에 저장된 영상 집합)와 프로브(probe : 얼굴인식 시스템에 새로 입력되어 신원 확인의 주체가 되는 집합) 집합이 선정되면 각각을 얼굴 공간에 투영시켜 해당되는 벡터를 얻어 낼 수 있고 이들은 가지고 유사도 측정(similarity measurement)을 수행한다.  $G$ 개의 갤러리 영상이 선정되어 이들을 얼굴 공간에 투영시키면  $g_1, \dots, g_G$ 와 같은  $G$ 개의 벡터가 생성되고 이 때 하나의 프로브가 선택되어 얼굴 공간에 투영했을 때 생성된 벡터가  $p_i$ 라 하면 이 프로브 생성 벡터는  $G$ 개의 각 갤러리 생성 벡터와 유사도를 측정할 수 있게 된다. 두 영상,  $p_i$ 와  $g_j$ 의 유사도,  $s_i(j)$ 을 측정하기 위해서 다양한 방법이 존재하나 일반적으로 식 (1)과 같이 유클리디안 거리(euclidean distance)를 이용한다.

$$\begin{aligned}
 p_i &= [p_i^1 \ p_i^2 \ \dots \ p_i^M] \\
 g_j &= [g_j^1 \ g_j^2 \ \dots \ g_j^M] \\
 s_i(j) &= \sqrt{\sum_{r=1}^M (p_i^r - g_j^r)^2} \tag{1}
 \end{aligned}$$

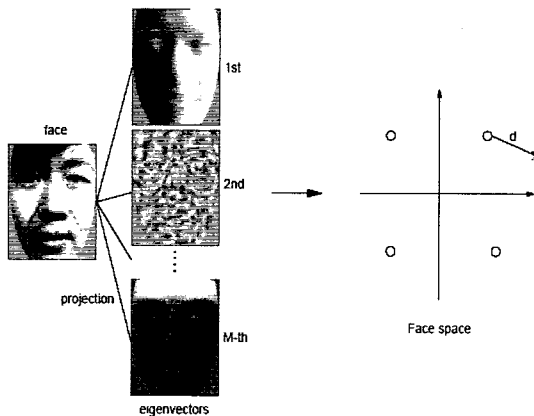


그림 2. PCA를 통해 생성된 얼굴 공간  
 Fig. 2. Facespace generated by PCA.

$s_i(j)$ 이 최소값을 가진다면  $p_i$ 에  $g_j$ 의 신원을 부여하게 되는 것이다. 그림 2는 PCA를 통해 형성된 고유얼굴과 이를 통해 형성된 고유공간을 보여주고 있다.

### 2. 시스템 구성

PCA 기반 얼굴인식 시스템은 그림 3과 같이 전처리(preprocessing), 특징추출(feature extraction), 인식(recognition)의 3가지 모듈로 구성된다<sup>[1]</sup>. 먼저 전처리 모듈은 입력된 샘플 영상들을 특징 추출이 용이하도록 정규화하고 통계적, 기하학적으로 미리 가공하는 것이다. 이러한 모듈의 목적은 인식성능에 영향을 줄 수 있는 여러 변화 요인들을 감소시키거나 제거함으로써 입력 샘플 영상을 보다 표준화된 영상으로 변형시키는 것이다. 전처리 모듈은 크게 3단계로 수행되는데 먼저 기하학적(geometrical) 처리 단계로서 회전이동(rotation), 평행이동(translation), 크기변화(scaling)를 통해 얼굴에서 두 눈의 중심을 기준 위치로 이동시킨다. 두 번째는 마스킹 단계로서 옷이나 머리카락과 같은 얼굴이외의 부분을 얼굴로부터 분리하고 또한 배경으로부터 얼굴을 분리시키는 과정이다. 마지막으로 히스토그램 평활화(histogram equalization) 과정을 거치는데 이는 픽셀의 명도값을 평균적으로 고르게 분포시킴으로써 인식 성능에 중대한 영향을 미치는 조명 성분의 영향을 감소시킨다.

두 번째 모듈은 샘플 영상으로부터 고유값과 고유벡터를 계산하여 주성분을 추출하고 이를 바탕으로 하나의 얼굴 공간을 생성시킨다. 이 때 각 갤러리 영상들을 생성된 얼굴 공간에 투영시켜 해당되는 벡터를 발생시키게 된다. 마지막 인식 모듈은 크게 두 가지 단계로 분리할 수 있는데 그것은 입력된 프로브 영상을 얼굴 공간에 투영시켜 벡터를 생성하고 이 벡터를 미리 발생된 갤러리의 각 벡터와 유사도를 비교하여 결정조건에 최적적으로 만족되는 갤러리 영상에 신원을 부여하는 결정 과정이 포함된다.

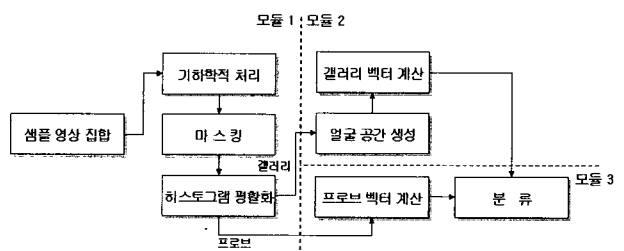


그림 3. PCA 기반 얼굴인식 시스템의 블록선도  
 Fig. 3. Block diagram of the PCA-based face recognition system.

### IV. 성능 시험 시나리오

#### 1. 인식(Identification) 시나리오

인식 시나리오는 ‘그는 누구인가? (Who is he ?)’ 라는 질문으로부터 시작한다. 즉 프로브 집합으로부터 하나의 영상이 새롭게 입력되면 그 영상의 신원을 갤러리 집합에 속한 영상들과 적합한 판단기준에 의거하여 서로 비교함으로써 가장 잘 부합되는 갤러리 영상의 ID를 부여하는 방식이다. 성능 분석은 최고순위 결정과 CMC(Cummulative Match Characteristic) 곡선의 두 가지 방법을 통해 수행된다<sup>[2][9]</sup>.

인식 시나리오에 대한 성능 시험은 매우 간단하다.  $P$ 를 프로브 집합이라 하고,  $|P| = N$ 을 프로브 집합의 크기라 하자. 마찬가지로  $G$ 가 갤러리 집합이고 갤러리 집합의 크기를  $|G| = M$ 라 하면,  $P = \{p_1, p_2, \dots, p_N\}$ ,  $G = \{g_1, g_2, \dots, g_M\}$ 가 된다. 이 때 유사도를 판단하는  $s_i(k)$ ,  $p_i \in P, g_k \in G$ 를 정의할 수 있다.  $s_i(k)$ 가 작은값을 갖을수록 두 영상의 유사성은 증가한다. 또한 두 영상이 동일하다면 당연히  $s_i(k)=0$ 이 된다. 인식 시나리오 상에서 성능을 시험하는 과정은 아래와 같다.

(1) 각 프로브 영상,  $p_i \in P$ 에 대해 모든 갤러리 영상과의 유사성 스코어 계산한다.

$$s_i(1), s_i(2), \dots, s_i(M)$$

(2) 크기 순으로 유사성 스코어를 정렬한다.

(3) 인식여부를 판정한다. 만약 프로브  $p_i$ 에 대해 일치하는 갤러리의 영상의 인덱스를  $id(i)$ 라 한다면  $p_i$ 와  $g_{id(i)}$ 는 동일인의 영상일 것이다. 이 때 두 가지의 경우에 대해 성능을 평가할 수 있는데 첫 번째는 최고 순위 결정 방법이고 두 번째는 R순위 결정 방법이다. 최고 순위 결정 방법은  $s_i(id(i))$ 이 가장 최소값을 갖는지에 대한 판단이다. 즉 최소값을 갖는다면 프로브가 정확하게 인식되는 것이다. 두 번째의 경우는 적당한 순위,  $R < M$ 을 선택하여 정렬된 순서에 대해 R번째 안에  $s_i(id(i))$ 의 값이 있는지를 판단하는 것이다. 이 때 R 순위까지의 누적된 인식률을 순위에 따른 분포로 나타내어 성능을 평가하는 데 이것을 CMC 곡선이라 한다<sup>[2][10]</sup>.

$$\sum_{i=1}^k f_i = R_k R_{|G|} = |P| \tag{2}$$

$$CMC_k = \frac{R_k}{|P|} CMC_{|G|} = 1 \tag{3}$$

$k$  : 순위 ( $k \leq G$ )

$f_k$  : 유사도가  $k$  순위로 매치된 프로브 영상의 수

$R_k$  : 유사도가  $k$  순위 안에 든 모든 프로브 영상의 수

$CMC_k$  : 모든 프로브에 대한  $k$  순위 까지의 누적 분포율

#### 2. 인증(Verification) 시나리오

인증 시나리오는 ‘그가 A인가? (Is he A ?)’라는 질문으로부터 시작한다. 즉 시스템에 접근한 어떤 사람이 자신의 신원 정보와 새로운 얼굴 영상을 제공하면 그 신원 정보에 따라 기존의 데이터베이스에 저장된 해당 영상과 새로 입력된 영상의 유사도를 비교하여 인증 확인 여부를 결정하는 것이다.

시스템적으로 인증 시나리오에서는 프로브 영상( $p$ ) 안에 있는 사람이 갤러리 영상( $g$ ) 안에 있는 사람과 동일한 지 여부를 판단한다. (만약 실제적으로  $p$ 와  $q$ 가 동일한 사람이면,  $p=q$ , 그렇지 않으면  $p \neq q$ 라고 표현한다고 하자) 이 때 시스템의 성능은 2가지 통계적 확률을 가지고 평가할 수 있는데 하나는 옳게 신원 인증을 수행할 확률, 즉 실제적으로  $p=q$ 일 때 시스템이  $p=q$ 라고 판단할 확률이고, 두 번째는 잘못 신원 인증을 수행할 확률, 즉 실제적으로  $p \neq q$ 일 때 시스템  $p=q$ 라고 판단할 확률이다. 전자를 VR(verification rate,  $P_v$ )이라 하고 후자를 FAR(false alarm rate,  $P_F$ )이라 한다<sup>[2][11]</sup>.

각각 하나의 프로브( $p_i$ )와 갤러리( $g_j$ )가 주어진다면, 프로브의 신원을 인증하거나 거부하는 결정은 유사도를 나타내는 식,  $s_i(j)$ 에 의해 행해진다. 어떠한 임계값,  $c$ 에 대해 스코어가  $c$ 보다 작거나 같다면( $s_i(j) \leq c$ ) 신원을 인증하게 되고, 반대로 스코어가  $c$ 보다 크다면( $s_i(j) > c$ ) 신원은 거부된다. 일반적으로 이러한 임계값을 시스템에 적절하게 선정한 후 그것의 최소값부터 최대값까지 변화시킴으로써 PV와 PF의 조합을 계산할 수 있는데 이러한 조합을 그래프에 나타낸 것이 바로 ROC(receive operating characteristic) 곡선이다.

위에서 설명했듯이 시스템은 한 사람에 대한 ROC 곡선을 계산할 수 있다. 그러나 실질적으로 다수의 사람들로 이루어진 어떤 집단에 대한 ROC 곡선이 필요하다. 이를 위해 갤러리는 사람당 하나의 영상만을 보유하게 함으로써 라운드-로빈 방법을 사용하여 집단에 대한 ROC 곡선을 계산한다. 이 과정은 3가지 단계를 통해 수행된다<sup>[2]</sup>.

(1) 프로브 집합을 2개의 비집합(disjoint)된 부분집합으로 분리하는 것이다. 갤러리 집합에서 하나의 영상,  $g_j \in G$ 이 주어지면 프로브 집합은  $D_j, F_j$ 로 분리되는

때 이 때  $D_j$ 는  $g_j$ 와 동일한 프로브 집합 내의 모든 영상들로 이루어진 부분집합이고  $F_j$ 는  $g_j$ 와 다른 프로브 집합 내의 모든 영상들로 이루어진 부분집합이다.

(2) 주어진 임계값과 선택된 갤러리 영상,  $g_j$ 에 대해 VR,  $P_V^{c,j}$ 과 FAR,  $P_F^{c,j}$ 을 계산하는 것이다.

$$P_V^{c,j} = \frac{|D_j^c|}{|D_j|}, \quad P_F^{c,j} = \frac{|F_j^c|}{|F_j|} \quad (4)$$

$D_j^c: (s_i(j) \leq c, \text{ given } p_i \in D_j)$  을 만족하는  $p_i$ 의 수

$F_j^c: (s_i(j) \leq c, \text{ given } p_i \in F_j)$  을 만족하는  $p_i$ 의 수

(3) 주어진 임계값에 대한 전체적인(모든 갤러리 영상에 대한) VR,  $P_V^c$ 과 FAR,  $P_F^c$ 을 계산한다. 이 때 이들은  $P_V^{c,j}$ 와  $P_F^{c,j}$ 의 가중치가 부여된 평균으로 계산된다.

$$P_V^c = \frac{1}{|G|} \sum_{j=1}^{|G|} \frac{|D_j|}{\frac{1}{|G|} \sum_j |D_j|} P_V^{c,j} = \frac{1}{\sum_j |D_j|} \sum_{j=1}^{|G|} |D_j| \cdot P_V^{c,j} \quad (5)$$

$$P_F^c = \frac{1}{|G|} \sum_{j=1}^{|G|} \frac{|F_j|}{\frac{1}{|G|} \sum_j |F_j|} P_F^{c,j} = \frac{1}{\sum_j |F_j|} \sum_{j=1}^{|G|} |F_j| \cdot P_F^{c,j} \quad (6)$$

최종적으로 임계값을 최소에서 최대로 변화시키면서 전체 ROC 곡선을 그릴 수 있다. ROC 곡선이 그려지면 EER(error equal rate)을 정의할 수 있는데 이것은 FRR(false reject rate,  $1-P_V^c$ )와 FAR이 같아지는 점을 의미한다. 성능 면에서 ERR이 작을수록 좋은 시스템이라고 할 수 있다.

인증 시나리오에서는 갤러리와 프로브의 크기가 영향을 미칠 수 있는데 갤러리는 반드시 사람당 한 장의 영상을 포함하고 있어야 하고 프로브는 사람당 한 장 이상의 영상을 포함하고 있어야 하며 프로브가 갤러리에서 속한 모든 사람들의 영상을 가지고 있을 필요는 없다. 그러나 프로브에 속한 모든 사람에 대해 갤러리는 그들의 영상을 가지고 있어야만 한다.

## V. 성능 시험

### 1. 데이터베이스

본 연구에서는 성능 시험을 위해 XM2VTS (eXtended MultiModal Verification for Teleservices and Security applications) 데이터베이스를 이용하였다. XM2VTS<sup>[12]</sup> 데이터베이스는 영국 Surrey 대학을 중심으로 유럽 각

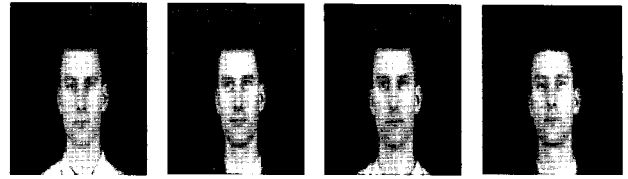


그림 4. XM2VTS의 샘플 영상

Fig. 4. Examples of face images in XM2VTS database.

표 1. 시나리오에 따른 실험 데이터 구성

Table 1. Configuration of experimental data to test scenarios.

	인식 시나리오	인증 시나리오
해상도	130×150 gray scale raw 영상	
ground truth	양 눈의 좌표	
training	100명에 대한 세션 1, 4 영상	100명에 대한 세션 1 영상
gallery	100명에 대한 세션 2 영상	100명에 대한 세션 1 영상
probe	100명에 대한 세션 3 영상	100명에 대한 세션 2, 3, 4 영상

국가들이 참여하여 구축되었고 그림 4에서처럼 한달에 1회씩 총 4회에 걸쳐 295명에 대해 4장의 정면영상을 보유하고 있다. 파일의 확장자는 ppm을 취하고 각 영상의 해상도는 720×576이며, 양쪽 눈에 대한 ground truth 정보를 가지고 있다.

본 연구에서 성능시험을 위해 XM2VTS 데이터베이스를 표 1과 같은 환경을 구축하였다.

### 2. 성능 시험 카테고리

본 연구에서는 전처리 모듈과 얼굴 공간 생성 모듈, 유사도 계산 모듈에 대한 성능시험을 수행하였다. 전처리 모듈에서는 히스토그램 평활화의 수행여부에 따른 성능 시험을 수행하였고 얼굴 공간 생성 모듈에서는 공간의 차원 선택(즉, 고유벡터의 축소)에 따른 성능 시험을 수행하였으며 유사도 계산 모듈에서는 다양한 유사도 계산 공식에 따른 인식 결과를 분석하였다.

#### 가. 히스토그램 평활화의 수행방법에 따른 성능 시험

전처리 과정은 인식 성능에 매우 큰 영향을 미치는 중요한 단계이다. 픽셀값의 정보에 의존하는 알고리즘들에 있어서 대비효과를 증가시키는 히스토그램 평활화 과정은 매우 중요하다 따라서 마스킹 처리를 기준으로 히스토그램 평활화 과정은 마스킹 전, 마스킹 후, 마스킹 전후 모두, 둘 다 하지 않을 경우의 4가지에 대한 성

능 시험을 수행하였다.

나. 고유벡터의 수에 따른 성능 시험

작은 고유값을 갖는 고유벡터들은 특징 추출을 위한 샘플 영상들 사이의 적은 변화를 내포하기 때문에 일종의 잡음으로서 간주될 수 있다. 따라서 이러한 벡터들은 인식 시스템의 성능 면에 있어서 기여도가 낮아서 이들을 제거함으로써 시스템의 성능을 향상시킬 수 있다. 총 199개의 정렬된 고유벡터 중 선택된 n개의 고유벡터가  $e_1, e_2, \dots, e_n$ 와 같다면 n의 값을 변화시켜 그에 따른 성능의 변화를 확인하였다.

다. n순위 고유벡터 제거에 따른 성능 시험

큰 고유값을 갖는 상위 고유벡터들은 샘플 영상들 간의 많은 변화와 차이를 내포하게 된다. 만약 이러한 벡터들이 조명 성분에 대한 정보를 주로 내포한다면 이들을 제거함으로써 성능의 향상을 도모할 수 있게 된다. 본 연구에서는 고유벡터의 수를 199개로 고정했을 때 최상위 고유벡터 3개를 제거하면서 성능의 변화를 확인하였다. ( $e_1, e_2, \dots, e_{199}$ 일 때  $e_1, e_2, e_3$ )

라. 분류기(classifier)에 따른 성능 시험

유사도를 측정하는 수단은 다양한 방법들이 제공되고 있다. 본 연구에서는 FERET<sup>[13]</sup>에서 사용한 분류기들을 사용하여 성능 시험을 수행하였다.

3. 성능 시험 결과

그림 5, 6, 7, 8은 각 성능 시험에 대한 CMC 곡선을 보여주고 있고 표 2, 3, 4, 5 (표에서 A, B, C는 각각 top 매치율, 10위까지의 누적매치율, 누적매치율이 1위 되는 순위를 나타낸다.) 는 top 매치율과 10위까지의 누적 매치율, 인증 시나리오에 대한 EER 값을 보여주고 있다. 히스토그램 평활화의 수행 시험의 경우 마스크를 처리하기 전에 히스토그램 평활화를 수행한 경우에 보다 높은 성능을 보였지만 히스토그램 평활화를 마스크 후에 처리했을 경우 히스토그램 평활화를 하지 않는 것보다 안 좋은 성능을 보였다.

얼굴공간 차원의 축소 시험에서는 100% 고유벡터를 사용하는 것보다는 80% 정도에서 보다 좋은 성능을 보였지만 축소율을 증가 시킬수록(적용되는 고유벡터의 수가 감소할수록) 성능은 낮아짐을 알 수 있다.

최대 고유벡터 제거의 경우에는 고유벡터를 제거

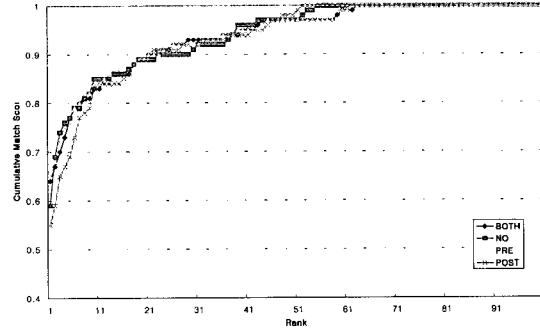


그림 5. 히스토그램 평활화 수행에 따른 CMC 곡선  
Fig. 5. CMC curve to the execution of histogram equalization.

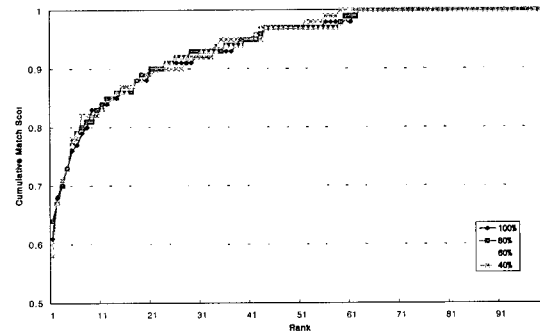


그림 6. 얼굴공간의 차원 축소(고유벡터의 수)에 따른 CMC 곡선  
Fig. 6. CMC curve to the rate of high order eigenvectors.

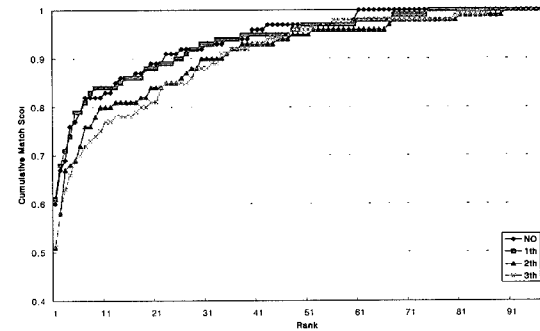


그림 7. 최상위 고유벡터 제거에 따른 CMC 곡선  
Fig. 7. CMC curve to the elimination of the nth highest eigenvector.

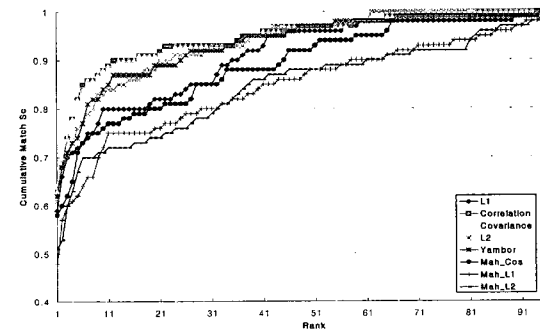


그림 8. 분류기에 따른 CMC 곡선  
Fig. 8. CMC curve to classifiers.

표 2. 히스토그램 평활화 수행에 따른 성능 시험 결과

Table 2. Performance test result on the execution of histogram equalization.

구분	A	B	C	EER
수행 안함	0.59	0.85	55	0.10
마스킹 전 수행	0.65	0.84	62	0.11
마스킹 후 수행	0.55	0.83	52	0.10
마스킹 전후 수행	0.64	0.83	63	0.12

표 3. 얼굴공간의 차원 축소(고유벡터의 수)에 따른 성능 시험 결과

Table 3. Performance test result on the rate of high order eigenvectors.

구분	A	B	C	EER
100%	0.61	0.83	62	0.12
80%	0.64	0.83	63	0.11
60%	0.60	0.82	61	0.11
40%	0.58	0.82	59	0.11

표 4. n순위 고유벡터 제거에 따른 성능 시험 결과

Table 4. Performance test result on the elimination of the nth highest eigenvector.

구분	A	B	C	EER
제거 안 함	0.60	0.82	61	0.12
1th 고유벡터 제거	0.61	0.84	75	0.11
2th 고유벡터 제거	0.51	0.80	90	0.14
3th 고유벡터 제거	0.51	0.75	87	0.14

표 5. 분류기에 따른 성능 시험 결과

Table 5. Performance test result on the variation of classifiers.

구분	A	B	C	EER
L1	0.59	0.80	95	0.16
L2	0.61	0.83	62	0.12
Correlation	0.63	0.88	83	0.08
Covariance	0.63	0.88	81	0.08
Cosine+Mahalanobis	0.58	0.76	95	0.13
L1+Mahalanobis	0.48	0.72	99	0.08
L2+Mahalanobis	0.51	0.71	98	0.08
Angle+Mahalanobis	0.62	0.84	93	0.09

하지 않았을 경우와 1번 고유벡터를 제거했을 경우에서 좋은 성능을 나타내었고 EER의 값도 낮은 수준을 유지하였다.

분류기의 선택에 따른 시험에서는 covariance와 correlation을 사용한 경우가 top 매치율과 EER 평가에서 좋은 성능을 나타내었고 Mahalanobis 거리에 L1와 L2를 합성한 분류기의 경우는 낮은 성능을 나타내었다.

#### 4. 성능 분석 메뉴

모듈화된 얼굴인식 시스템의 성능 시험을 위해 별도의 분석 도구를 개발하였다. 분석 메뉴화면은 그림9, 10

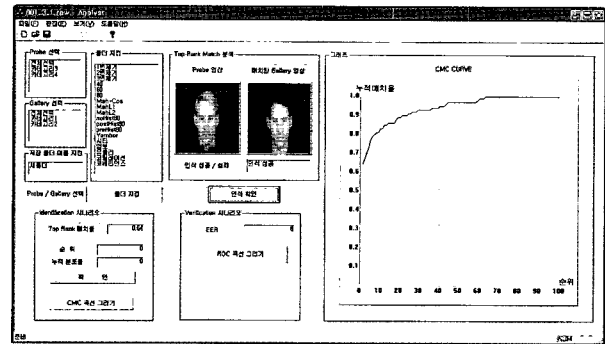


그림 9. 인식 시나리오를 수행한 성능 분석 툴의 화면  
Fig. 9. User interface of performance analysis tool showing the result by the identification scenario.

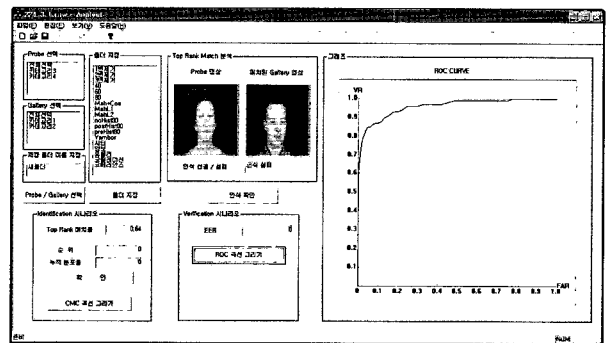


그림 10. 인증 시나리오를 수행한 성능 분석 툴의 화면  
Fig. 10. User interface of performance analysis tool showing the result by the verification scenario.

에서처럼 인식 시나리오와 인증 시나리오 상에서의 성능 시험을 분석할 수 있도록 구성되어 있다. 구성 메뉴는 아래와 같다.

- (1) probe / gallery 선택
- (2) 저장 폴더명 설정
- (3) 분석 폴더 지정
- (4) top rank 매치율 분석
- (5) top rank 매치율 출력
- (6) 해당순위 누적 매치율 출력
- (7) EER 출력



## (8) CMC 곡선 / ROC 곡선 출력

## VI. 결 론

본 연구에서는 성능 시험 프로토콜 설계 기준을 제시하였고 성능 시험의 객관성을 확보하기 위해 모듈화된 얼굴인식 시스템을 제시하였다. PCA를 기반으로 한 인식 알고리즘을 벤치마크 알고리즘으로 이용하여 인식 시나리오와 인증 시나리오의 성능 비교 실험을 수행하였으며 성능 시험 분석을 위한 도구를 개발하였다. 성능 시험 프로토콜은 객관적이고 공정한 기준이 요구되기 때문에 FERET이나 FRVT2000 및 FRVT2002의 적용기준 사례와 국외의 생체인식 포럼과 같은 표준화 그룹의 동향을 바탕으로 국내의 현 기술 수준을 파악하고 이를 평가할 수 있는 표준화 작업이 매우 중요하다.

성능 시험 결과는 일반적으로 얼굴 영상의 특징(포즈, 조명, 표정 등)에 대한 갤러리와 프로브의 설정에 따라 성능의 차이가 나타나는데 본 연구에서는 정면영상만을 다루고 세션의 변화만이 존재하는 XM2VTS 데이터베이스를 사용하였다. 본 연구에서 볼 수 있듯이 성능에 영향을 주는 요소는 데이터베이스 영상의 질, 수, 특성뿐만 아니라 전체 인식시스템의 수행과정에서 각 모듈 단계에서의 운영방식 또한 다양한 성능 결과를 나타내는 주요 요소임을 알 수 있었다. 앞으로 국내의 얼굴인식 및 생체인식 분야에 대한 기술 발전을 위해 성능 시험에 대한 체계적인 표준화가 절실하며 학문적으로는 인식 알고리즘의 개발과 더불어 다양한 얼굴 인식 응용분야를 객관적으로 평가할 수 있는 시나리오의 개발과 성능 시험 카테고리에 대한 연구 또한 매우 중요한 이슈 중 하나이다.

## 참 고 문 헌

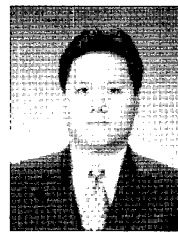
- [1] H. Moon and P. J. Phillips, "Computational and performance aspects of projection-based face recognition algorithms", *Perception*, vol. 30, pp. 303-321, 2001.
- [2] H. Moon, "Performance Evaluation Methodology for Face Recognition Algorithms", Ph.D thesis, Dept. of Computer Science and Engineering, SUNY Buffalo, 1999.
- [3] M. Turk and A. Pentland, "Eigenfaces for recognition", *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.
- [4] P. Belhumieur, J. Hespanha and D. Krieman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 711-720, 1997.
- [5] S. Gutta, J. Huang, D. Singh, I. Shah, B. Takacs, and H. Wechsler, "Benchmark studies on face recognition", In M. Bichsel, editor, *International Workshop on Automatic Face and Gesture Recognition*, 1995.
- [6] "Best Practices in Testing and Reporting Performance of Biometric Devices", NPL Report CMSC, version 2.01, 2002.
- [7] "Face Recognition Vendor Test 2002", Evaluation Reports, 2003.
- [8] I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.
- [9] H. Moon and J. Kim, "Biometrics identification and verification using projection-based face recognition system", *Proceedings of WISA 2003*, pp. 380-394, 2003.
- [10] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1090-1104, 2000.
- [11] H. Moon, "Biometrics Person Authentication using Projection-based Face Recognition System in Verification Scenario", *Proceedings of ICBA 2004*, pp. 207-213, 2004.
- [12] K. Messer, J. Matas, J. Kittler, J. Luetin, and G. Maitre, "XM2VTSDB: The Extended MeVTS Database," *Proc. of International Conference on Audio- and Video-Based Person Authentication*, pp. 72-77, 1999.
- [13] P. J. Phillips, P. Rauss, and S. Der, "FERET (face recognition technology) recognition algorithm development and test report", *Technical Report ARL-TR-995*, U.S. Army Research Laboratory, 1996.

## 저 자 소 개



홍 태 화(학생회원)  
1997년 2월 연세대학교  
전기공학과 학사  
1999년 2월 동대학원  
전기전자공학과 석사  
1999년 3월~현재 동대학원  
박사과정

<주관심분야 : 컴퓨터 비전, 얼굴인식, 패턴인식, 영상처리>



문 현 준(정회원)  
1990년 2월 고려대학교  
전자공학과 학사  
1992년 6월 SUNY Buffalo  
컴퓨터공학 석사  
1999년 9월 SUNY Buffalo  
컴퓨터공학 박사

1999년 11월~2003년 2월 Viisage Technology  
Senior Researcher

2004년 3월~현재 세종대학교 컴퓨터공학부  
부교수

<주관심분야 : 컴퓨터 비전, 얼굴인식, 패턴인식, 영상처리>



신 용 녀(정회원)  
1999년 2월 숭실대학교  
컴퓨터과학과 학사  
2001년 9월 고려대학교  
컴퓨터과 학과 석사  
2002년 1월~현재 한국정보보호  
진흥원 기술표준팀 연구원

<주관심분야 : 생체인식, 정형기법, 정보보호>



이 동 근(정회원)  
2001년 2월 경북대학교  
컴퓨터과학과 학사  
2003년 2월 동대학원  
컴퓨터과학과 석사  
2003년 1월~현재 한국정보보호  
진흥원 기술표준팀 연구원

<주관심분야 : 네트워크, 무선인터넷, 정보보호>



김 재 성(정회원)  
1986년 2월 인하대학교  
전산학과 학사  
1989년 2월 동대학원  
전산학과 석사  
1996년 7월~현재 한국정보보호  
진흥원 기술표준팀 팀장

<주관심분야 : 생체인식, 인지과학, 정보보호>