

# 국내 생체 인식 시스템 성능 평가를 위한 표준안 연구\*

문지현\*\*, 김학일\*\*

## A Study on Standards for Performance Evaluation of Biometric Recognition Systems

Ji-hyun Moon\*\*, Hak-il Kim\*\*

### 요 약

최근 생체 인증 제품이 물리적 보안뿐만 아니라 논리적 접근 제어를 위한 제품으로서 활성화되고 있다. 그러나 국내에서는 이러한 제품의 성능평가 환경이 전무한 상태이며 평가 기준 또한 마련되어 있지 않은 실정이다. 향후 생체 인식 시스템의 사용 확대가 예상됨에 따라 국내 실정에 맞는 생체인식 시스템의 성능시험 평가 기준 및 평가 방법론의 개발이 절실하다 하겠다. 본 논문은 국외 생체인식시스템 평가 기관들의 성능평가에 관한 정량적, 정성적 평가기준 및 평가 방법을 분석하여 그 결과를 토대로 국내 생체인식시스템의 성능평가 방법 및 기준 모델을 정립하는 것을 목적으로 하여, 센서, 알고리즘, 그리고 응용 시스템 전반에 관한 평가 시나리오 및 평가 항목을 정의하고 이를 위한 평가 표준안을 제안하였으며, 이를 바탕으로 4 가지 지문 입력 센서들에 대한 성능을 비교하여 정리하였다.

### ABSTRACT

The purpose of this research is to establish Korean standards of best practice and common criteria for testing and evaluating the performance of biometric recognition systems. First of all, research activities in leading countries such as USA, Great Britain, and Germany have been studied and analyzed. Then, the standards of best practice and common criteria are proposed in the aspects of sensors, algorithms, and application systems. The results of this work will be utilized for test and evaluation of commercial biometric products by KISA and provided to private industries for their own evaluation of products.

**keyword** : *Biometrics, Standards, Performance, Evaluation, Recognition*

### 1. 서 론

생체 인식이란 계측 가능한 개인의 신체 또는 습관적 행위 특성을 추출하여, 기 등록된 내용과의 동일 여부를 확인함으로써 개인 신분을 확인하는 기술을 말한다. 생체 인식 기능을 이용한 자동화된 개인

인증 시스템은 높은 보안성과 사용의 편리성을 장점으로 지금까지 활발한 연구가 계속되고 있다.

최근 시장에 참여하는 생체 인식 시스템 제품의 종류가 늘어가는 가운데, 공정한 기술적 평가를 시도하여 소비자에게 정확한 선택 기준을 제공하고, 개발자로 하여금 신속한 문제 도출로 신뢰도와 완성도를

\* 본 연구는 정보통신부 정보통신선도기반기술개발사업 과제(2001-S-093)로 수행되었습니다.

\*\* 인하대학교 자동화공학과 신호 및 정보 공학 연구실(g1991095@inhavision.inha.ac.kr, hikim@inha.ac.kr)

높이는데 기여하고자 하는 노력이 절실한 시점이다.

지금까지 있어왔던 성능 평가는 개발자 자신이 제품의 설계 목표를 완수하였는지를 확인하기 위한 개별적인 노력이었다. 따라서 서로의 주장에 대한 객관적인 비교가 불가능하였으며, 이는 소비자에게 생체 인식 시스템에 대한 많은 지식을 요구함과 동시에 시스템 활용에 대한 정확한 이해를 방해하는 요소가 되어 왔다. 그러나, 국내에서는 이러한 제품의 객관적인 성능 평가 환경이 전무한 상태이며, 평가 기준 또한 마련되어 있지 않은 실정이다. 향후 생체 인식 시스템의 사용 확대가 예상됨에 따라 국내 실정에 맞는 생체 인식 시스템 성능 시험 평가 기준 및 평가 방법론 개발이 요구된다. 또한 현재 국외 선진국의 평가 기관에서 시행하고 있는 생체 인식 시스템의 성능 평가 사례분석을 통한 국내 생체 인식 시스템 성능 평가 모델 정립이 절실히 요구된다.

본 연구는 이러한 국외 생체 인식 시스템 평가 기관들의 성능 평가에 관련된 정량적, 정성적 평가 기준 및 평가 방법을 분석하여 그 결과를 토대로 국내 생체 인식 시스템의 성능 평가 방법 및 기준 모델을 정립하고자 한다. 이를 위해 국외 여러 나라들의 연구 및 평가 모델, 시험 관련 표준 등에 대한 사례를 분석하여 국내의 생체 인식 시스템을 위한 성능 평가 모델 확립의 기반으로 삼는다. 본 연구는 생체 인식 시스템의 기술적 성능 평가에 그 내용을 한정

하며, 시스템의 보안에 관련된 사항들은 평가의 범위에서 제외된다.

## II. 성능 평가 관련 국외 사례 분석

외국의 연구 현황은 90년대 초반 국립 연구 기관들을 중심으로 시작되어 최근에는 미국, 영국, 독일 등에서 활발하게 이루어지고 있다. 미국의 Sandia 연구소에서는 100명에 가까운 지원자들을 시스템 사용자 집단으로 하여 실험하였으며<sup>(1)</sup>, 영국의 Biometric Working Group에서 자체 정립한 평가 표준을 바탕으로 선택된 생체 인식 시스템의 성능을 평가하였다<sup>(2)</sup>. 독일에서는 "BioIS"라는 생체 인식 장비 성능 평가를 위한 연구 프로젝트를 수행하여 이를 위해 실험할 생체 인식 장비에 따른 테스트 절차를 정의한 표준을 제시하였다<sup>(3)</sup>. 또한 최근에 들어 특정 생체 정보를 이용한 인증 시스템의 인증 성능을 평가, 비교해 보려는 노력이 활발해지고 있으며, 대표적인 예로는 지문 인식 알고리즘의 성능을 비교, 평가한 Fingerprint Verification Competition 2000과 얼굴 인식 알고리즘의 성능을 비교, 평가한 FERET Program 및 Face Recognition Vendor Test 2000이 있다<sup>(4,5)</sup>. [표 1]은 위에서 설명한 국외 연구 사례를 정리한 것이며, [표 2]는 [표 1]에서의 연구에서 사용된 평가 항목들을 정리한 것이다.

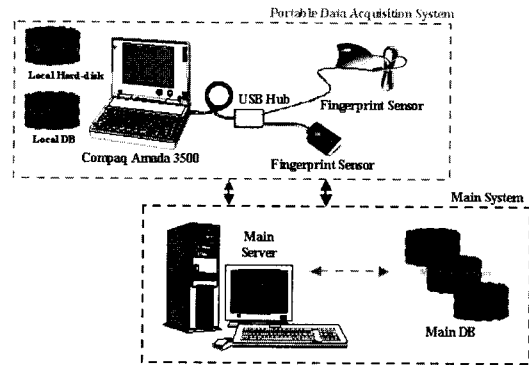
[표 1] 성능 평가에 관한 국외 연구 사례 정리

	주관	목적	대상
Biometric Product Testing (2000년 3월-12월)	영국 National Physical Laboratory, CESG(Communications Electronics Security Group)	현 기술로 획득할 수 있는 인식 시스템의 성능수준을 측정, 평가에의 참여도 및 평가 활동의 증진을 위한 방법론을 연구	지문, 얼굴, 장문, 홍채, 혈관 패턴, 음성 인식 시스템 (총 7개 회사의 7개 시스템 평가)
BioIS Project (1999년 3월-8월)	독일 Fraunhofer 대학, 독일 Federal Criminal Investigation Office, German Information Security Agency	생체 인식 시스템이 가져야 될 최소 요구 사항을 평가하며, 시스템 특성의 신빙성 있는 평가를 위한 표준을 정의	지문, 얼굴, 장문, 홍채, 서명, 음성 인식 시스템 (총 8개 회사의 12개 시스템 평가)
FERET (1993년-1998년)	미국 DoD Counterdrug Technology Development Program Office	보안, 지식, 법적인 개인 인증에 도움이 될 자동화된 얼굴 인식 기술의 개발 및 그 성능의 평가	얼굴 인식 알고리즘 (총 6개 학교/회사 참가)
FRVT2000 (2000년 5월-6월)	미국 DoD Counterdrug Technology Development Program Office, National Institute of Justice, DARPA	얼굴 인식 시스템의 기술적인 평가를 시도하고, 평가 방법에 대한 교육적인 효과 유도	얼굴 인식 알고리즘 (총 5개 회사 참가)
FVC2000 (2000년 6월-9월)	이탈리아 볼로냐 대학, 미국 미시간 주립대학, 미국 NBTC	현재의 기술 수준에 대한 보다 정확한 이해를 돕고 앞으로의 기술이 나아갈 방향을 제시하기 위한 공통의 기초를 정립함	지문 인식 알고리즘 (총 11개의 학교/회사 참가)

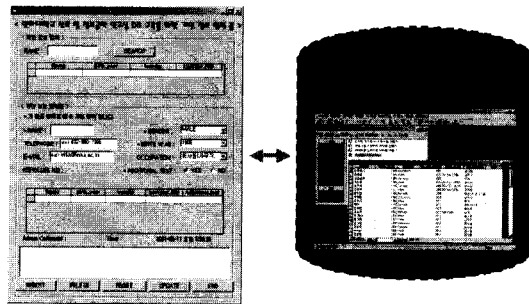
[표 2] 국외 평가 사례에서 사용된 평가 항목들

	Biometric Product Testing	BioS Project	FERET	FRVT 2000	FVC 2000
FAR/FRR	√	√	√		√
FMR/FNMR	√				
EER/AEER			√		√
ROC curve			√		√
Average ROC			√		
CMC curve				√	
Fail to Enroll	√	√			√
Fail to Acquire	√				
Fail to Match					√
User throughput	√	√			√
Algorithm throughput	√	√			√
Verification probability			√		
Upper-bound curve			√		
Environmental test				√	
Product usability test				√	

이미지는 로컬 저장장치에 파일 형태로 저장되었다. 개발된 메인 시스템은 [그림 4]에서 보는 바와



(그림 1) 성능 평가를 위한 시스템의 전체 구성도



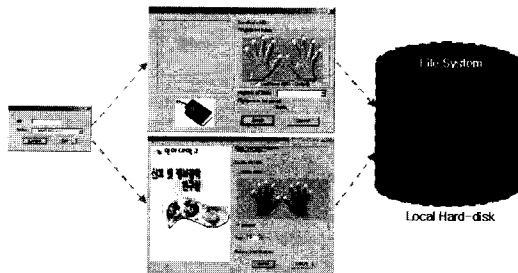
(그림 2) 개인 정보 저장을 위한 사용자 인터페이스

### III. 성능 평가 관련 국내 연구 현황

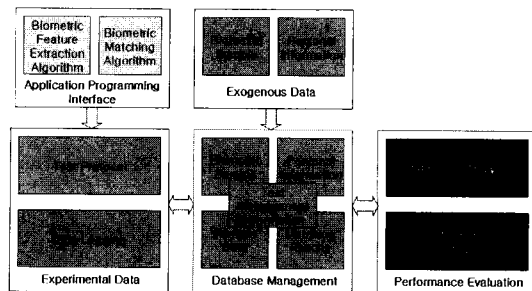
생체 인식 시스템의 성능을 평가하는데 있어서는 다양한 오류 원인을 분석하고 이들의 영향을 평가하며, 개선부문에 대한 효과를 가시화 할 수 있는 기능을 자동화하는 것이 필요하다. 이와 같은 필요성을 바탕으로 실제 구현된 '생체 인식 시스템 성능 평가를 위한 플랫폼'을 소개한다.<sup>[6]</sup>

생체 인식 시스템 성능 평가 플랫폼은 [그림 1]과 같이 크게 생체 자료 수집을 위한 이동식 정보 수집 시스템과 수집된 자료를 기반으로 알고리즘에 대한 기능적 성능 평가를 수행하는 메인 시스템으로 구분된다. 이동식 정보 수집 시스템은 [그림 2], [그림 3]과 같은 인터페이스를 이용해 이동식 PC에 저장된다.

실험적인 평가를 위해 시도된 지문 이미지 획득을 위한 정보 수집 시스템은 Pentium III 300MHz, 128MB 메모리를 장착한 노트북 컴퓨터를 기본으로 하고 있으며, 자료 저장을 위한 DBMS로 MS Access를 사용하였다. 각각의 지문 입력 센서는 USB 타입의 하드웨어 인터페이스를 사용하는 것이 선택되어 USB 허브를 이용해 연결되었으며, 획득된 지문



(그림 3) 지문 이미지 수집을 위한 사용자 인터페이스



(그림 4) 성능 평가를 위한 기능 단위별 시스템 구성도

[표 3] 기능 특성에 따른 데이터베이스 테이블의 분류

Category	Class	Items	Sub-Items
Data	Endogenous data	Gender, Data of Birth, Biometric sample	
	Exogenous data	Name or ID, Job, Sample acquisition date, Quality Index, Sample Sequence, Environmental condition	
	Devices	Typology of identification methods, Model, Specifications, Data format	
	Algorithms	Algorithm, Developer, Version, Template format	Feature extraction, Matching, Classification, etc.
Result	Intermediate result	Items	Intermediate result formats
	Algorithm result	Template, Classify result, Matching similarity, etc.	Result format
	Experimental result		
Log	Experimental log	Date, Operator, Experiment conditions, Job size	Counts of success and failure
	Statistics	Database size, Record size, Experiment log, etc.	

같은 기능들을 갖추도록 설계되었다. 플랫폼은 33개의 데이터베이스 테이블과 밀접하게 연동되며 운용되는데, 본 플랫폼을 이루는 데이터베이스는 그 기능 성격에 따라 [표 3]과 같이 정보(Data), 결과(Result), 로그(Log)의 세 가지 영역으로 분류할 수 있다. 정보를 다루는 데이터베이스 테이블들은 등록자, 장비, 알고리즘과 같이 평가에 필요한 기본적인 정보들을 저장, 관리한다. 결과를 다루는 테이블들은 각 실험에 따른 최종 결과뿐만 아니라 평가의 중간 단계에서 생길 수 있는 임시 결과까지 다양한 정보를 관리할 수 있도록 하였다. 로그를 다루는 테이블들은 전체 실험에 대한 작업 진행 사항 및 작업 결과를 저장하고 관리한다.

## V. 성능 평가 표준안

본 논문은 국외 생체 인식 시스템 평가 기관들의 성능평가에 관한 정량적, 정성적 평가기준 및 평가 방법을 분석하여 그 결과를 토대로 국내 생체 인식 시스템의 성능평가 방법 및 기준 모델을 정립하는 것에 그 목적이 있다. 따라서 다음과 같이 센서, 알고리즘 및 응용 시스템 전반에 관한 평가 시나리오 및 평가 항목을 정의하고 이를 위한 올바른 평가 표준안을 제안함으로써 평가 결과의 신빙성 확보 및 실험의 공정성을 확립하는 기반을 마련한다.

### 4.1 성능 평가 시나리오

국내 생체 인증 시스템을 평가하기 위한 테스트 공간 구성은 영국의 BWG에서 공포한 "Best Practice"<sup>(7)</sup> 문서에 제안된 방식을 충분히 고려하여 국내 실정에 맞게 재설정 되었다. 일반적인 사항들을 우선적으로 설정하여 환경 구성에 유연성을 줌과 동시에 테스트

시의 상황에 따른 최적의 환경 설정을 유도함으로써 실험의 공정성을 확보한다.

#### 4.1.1 지원자 집단

“생체 정보 측정 및 등록, 인증에 이르는 모든 과정에 있어 자신의 생체 정보를 제공하고 기 정립된 평가 절차를 충실히 이행할 것을 보장하는 자원자들의 집합”을 지원자 집단이라고 한다.

지원자 집단은 자원자들 중에서 선별된 후 결정되는데, 일반적으로 알고리즘의 성능 평가 시에는 통계적 자료를 통해 알 수 있는 대한민국 성인의 연령별 인구 분포 현황을 고려하여 선별한다.

시스템 관점에서의 시나리오 평가를 목적으로 하는 경우에는 시스템을 사용할 사용자 모집단에 그 분포가 가깝도록 지원자 집단을 선별한다. 계산 가능한 일반식을 구하는 것은 어려우나, 지원자 집단의 크기는 필요한 비교 횟수를 얻기에 충분한 수로 결정한다. 영국 BWG의 사례에서는 약 200명 정도를 선별하였는데 전원을 해당 연구소 관련 연구원들로 하였고, 선별된 200명의 연령 및 성별 분포는 연구소 구성 인원과 그 비를 맞추었다. 또한, 본 연구에 관하여 인하대학교를 중심으로 시범 수행된 지문 이미지 데이터베이스 구축 작업에 이용된 지원자 집단도 약 200명 정도였다. 그런데 본 작업은 지원자 집단의 선별 없이 이루어져 성별 및 직업 분포가 대한민국 인구 분포 현황을 정확히 고려하지 못하고 있다는 점에서 영국의 사례와 다소 차이가 있다. 그러나 인구 분포와는 별도로 유년층에서 노년층에 이르는 폭넓은 지원자 집단을 확보하였다는 점에서 의의가 있으며, 직접 지원자들을 방문하여 작업하였기 때문에 차후에 있을 지문 데이터베이스 구축 작업을 위한 기반을 확고히 하였다는 점에서 또한 의의를 찾을 수 있다.

4.1.2 환경

테스트를 위한 공간은 보통의 사무실용 방을 사용하는 것으로 한다. 광원은 일반적으로 많이 사용되는 형광등을 사용하는데 되도록 균일한 밝기를 오래 유지할 수 있는 것으로 선택하고, 창이 있는 경우 차양을 이용해 일조량의 영향을 최소화시킨다. 얼굴 인식과 같이 조명의 영향을 받는 시스템의 경우 자연광의 유무가 인식 성능에 영향을 미치므로 시간에 따라 변화하는 자연광의 완벽한 차단을 위한 노력이 필요한 것이다. 음성 또는 화자 인식 시스템의 경우 주변의 잡음이 성능에 영향을 줄 수 있으나 실제 사용자의 환경과 유사하게 구성되도록 하기 위해 음성 또는 화자 인식 시스템만을 위한 방음 시설은 따로 갖추지 않는 것으로 한다. 공간의 위치 및 테스트 상황에 따라 사무실용 방 전체를 방음 처리하는 것은 고려할 만 하다. 생체 인식 시스템을 설치할 시에는 제공 회사에서 제안하는 최적의 설정치를 사용하여 설치함을 원칙으로 한다.

4.1.3 등록 및 확인

조명, 온도 등 자연적인 환경의 변화에 민감한 생체 정보가 있기 때문에 어떤 장비로 어떤 생체 정보를 먼저 측정하느냐가 인식 성능에 영향을 미칠 수 있다. 이를 순서 효과(Order Effect)라고 하는데, 시스템 설치 시 고려되어야 함은 물론이고, 측정 시에도 지원자들의 수와 측정 시간 간격 등을 고려하여 특정 시스템에서의 측정 순서를 조정해 주는 관리자의 작업이 필요해진다. 그리고, 1차 생체 정보 등록 및 확인 후 2차 시도까지 적어도 3개월의 시간 간격을 둔다. 이는 생체 정보를 통해 얻어지는 템플릿들의 노화(aging)를 고려하기 위한 것이며, 평가 시에는 적어도 6 개월 이상의 시간 간격을 가진 템플릿들을 비교하도록 한다.

4.1.4 테스트

테스트는 지원자 집단을 대상으로 평가 대상이 되는 생체 인식 시스템에 대한 정보를 공개하고 알려주는 것에서 시작한다. 이는 지원자 집단에 속한 대부분의 자원자가 기기의 사용에 익숙하지 않음으로 해서 오는 오류의 가능성을 조금이나마 줄여 오류 관련 측정치의 정확성을 높이는 데 그 역할을 한다. 또한 항상 시스템 주변에는 자원자의 도움 요청에 응할 수 있는 관리자 혹은 도우미를 두어 자원자들의 불편을 덜어줄 수 있도록 해야 한다. 등록과 이를 통해 얻어지는 템플릿의 확인 및 인증은 같은 시스템

을 통해 이루어져야 하며, 처리 과정에서 얻을 수 있는 모든 정보를 저장할 수 있도록 해야 한다.

V. 성능 평가 항목 및 성능 평가 방법

4.1 시스템 공통

시스템을 평가함에 있어서는 크게 하드웨어 시스템과 소프트웨어 시스템의 두 부분을 고려한다. 하드웨어 시스템에 관한 평가 항목은 [표 4]와 같이 정리해 볼 수 있으며, 소프트웨어 시스템에 관한 평가 항목은 [표 5]와 같이 정리할 수 있다.

[표 4]에서의 평가 항목 중 해상도(Resolution)는 가로, 세로의 방향으로 모두 측정하는데 단위는 dpi(dot per inch)로 한다. 측정은 평판 스캐너를 이용해 측정된 일정 부분의 단위 영역을 장비별로 취득한 이미

[표 4] 하드웨어 시스템을 위한 평가 항목 정리

	하드웨어 시스템 평가 항목
시스템 요구사항	<ul style="list-style-type: none"> <li>▫ Processor(CPU, MHz)</li> <li>▫ Size of memory(MB)</li> <li>▫ Local storage(KB)</li> <li>▫ Additional hardware</li> <li>▫ Non-standard hardware</li> </ul>
생체 정보 관련	<ul style="list-style-type: none"> <li>▫ Resolution(dpi)</li> <li>▫ Biometric data quality</li> </ul>
단위 센서 관련	<ul style="list-style-type: none"> <li>▫ Biometric data size</li> <li>▫ Sensor Area (width(mm), height(mm))</li> <li>▫ Module Size (width(mm), height(mm), depth(mm))</li> <li>▫ Operational environment (temperature(°C), humidity(%), illuminance(LUX), noise(dB))</li> <li>▫ Power consumption(mW)</li> <li>▫ API standard compliant</li> <li>▫ Device interface</li> </ul>

[표 5] 소프트웨어 시스템을 위한 평가 항목 정리

	소프트웨어 시스템 평가 항목
시스템 요구 사항	<ul style="list-style-type: none"> <li>▫ Operating system</li> <li>▫ DBMS</li> <li>▫ Additional software</li> <li>▫ Non-standard software</li> </ul>
알고리즘 관련	<ul style="list-style-type: none"> <li>▫ Size of biometric template</li> <li>▫ Size of program to be installed</li> <li>▫ Processing time</li> </ul>
단위 센서 관련	<ul style="list-style-type: none"> <li>▫ User interface</li> <li>▫ Maintenance</li> </ul>

지들의 그것과 비교하는 방법을 사용해 수행될 수 있다. 또한 이미지 품질(Biometric data quality)은 이미지들의 히스토그램을 통해 대상을 비교해 볼 수 있다.

4.2 인식 기술

생체 인식 시스템의 기술 성능을 평가할 수 있는 정량화 가능한 평가 항목의 정립은 생체 인식 시스템 평가 표준안 개발의 주요 목적이며, 많은 실험과 연구를 통해 지속적인 수정과 보완이 필요한 평가의 핵심이라 할 수 있다. 제안하는 인식 기술 관련 평가 항목은 [표 6]과 같이 세 가지로 나누어 정리해 볼 수 있다.

[표 6] 인식 기술의 성능 평가를 위한 평가 항목 정리

인식 기술 평가 항목	
일반	<ul style="list-style-type: none"> <li>◦ FMR/FNMR</li> <li>◦ EER</li> <li>◦ ROC curve</li> <li>◦ Failure to Acquire rate/Failure to Enroll rate</li> <li>◦ Enrollment &amp; Matching throughput</li> </ul>
권장	<ul style="list-style-type: none"> <li>◦ Average ROC curve</li> <li>◦ Upper-bound on Algorithm Performance</li> <li>◦ Requirement of minimum processing resource</li> </ul>
제안	<ul style="list-style-type: none"> <li>◦ Transaction speed</li> <li>◦ Penetration rate</li> <li>◦ Binning error rate</li> </ul>

4.2.1 일반 항목

- False Match Rate(FMR) : 타인의 시도(NI, Number of Imposter Recognition Attempts)에 대해 임계값(t) 이상으로 나타나는 정합 점수(sim, Imposter Matching Score)를 가진 타인의 비율.

$$FMR(t) = \frac{card\{s_{im} | s_{im} \geq t\}}{NI}$$

- False Non-Match Rate(FNMR) : 본인의 시도(NG, Number of Genuine Recognition Attempts)에 대한 어떤 임계값(t) 이하를 보이는 정합 점수(sgm, Genuine Matching Score)를 가진 본인과 본인의 등록 거부 및 비교 거부 회수의 합(REJNG)의 비율.

$$FNMR(t) = \frac{card(s_{gm} | s_{gm} < t) + REJ_{NG}}{NG}$$

- Equal Error Rate(EER) : FMR과 FNMR의 누적 합이 동일할 때의 Error Rate.
- Receiver Operating Characteristic Curve (ROC Curve) : FNMR을 FMR에 대한 함수 곡선으로 표현한 것
- Failure to Acquire rate(FTA) : 생체 정보를 획득하기 위한 입력기의 성능을 평가하는 항목으로서 기기 불안정 또는 사용의 불편 정도를 평가한 것
- Failure to Enroll rate(FTE) : 인식 알고리즘이 정의한 특징량을 추출하지 못하는 경우에 대한 평가 항목. (이러한 경우의 사용자는 등록 및 사용이 불가능하게 된다. 이 비율을 제외하고 등록이 가능한 사용자에 대한 FMR/FNMR만을 논의하는 것은 성능 평가에 있어 그 인식률의 신빙성을 떨어뜨릴 수 있는 오류를 범할 수 있다.)
- False Acceptance Rate(FAR) : 전체 시스템에 대한 정합 오류의 양을 측정하는 것으로서, n번의 정합 시도에 대해, 등록되어 있지 않은 사용자가 등록된 사용자로 오인되어 접근이 허용될 확률

$$FAR = (1 - FTA)(1 - (1 - FMR)^n)$$

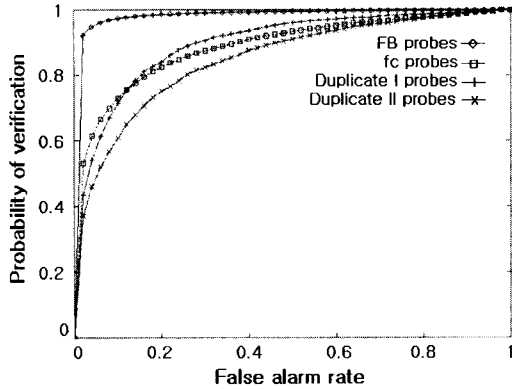
- False Reject Rate(FRR) : FAR과 마찬가지로 전체 시스템에 대한 평가 항목으로, n번의 정합 시도에 대해, 등록되어 있는 사용자의 접근이 거부될 확률

$$FRR = (1 - FTA)FNMR^n + FTA$$

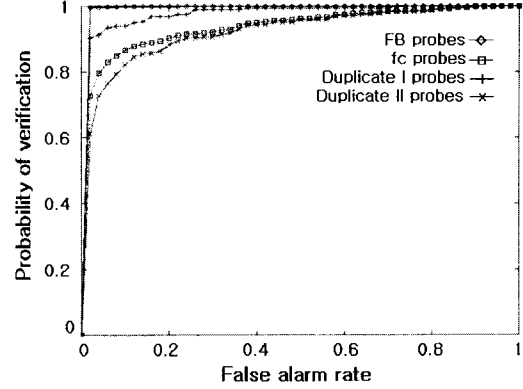
- Enrollment & Matching throughput : 등록과 정합 과정에 있어서의 알고리즘이 소모하는 시간 각각을 소프트웨어 측면에서 측정

4.2.2 권장 항목

- Average ROC curve : 실험에 사용된 다수 집합의 생체 정보 데이터베이스에 대해 실험한 알고리즘 전부의 ROC 결과들을 평균하여 나타낸 곡선, 현 시점에서의 구현 가능한 최적의 성능을 가능해 볼 수 있는 척도로써 사용됨(FERET<sup>(5)</sup>에서 사용된 평가 항목으로, 그 결과의 예는 [그림 5]와 같다.)



(그림 5) FERET에서 사용된 Average ROC curves의 예



(그림 6) FERET에서 사용된 Upper-bound ROC curves의 예

- Upper-bound on algorithm performance : 실험에 사용된 다수 집합의 생체 정보 데이터베이스에 대해 실험 대상 알고리즘들 중 가장 우수한 성능을 보인 알고리즘을 해당 데이터베이스의 성능 상한이라 정의하고 그 ROC를 곡선으로 나타냄. 성능 상한은 최소 EER을 가지는 알고리즘으로 선택 (이 항목 역시 FERET<sup>(5)</sup>에서 사용되었으며, 그 결과의 예는 (그림 6)과 같다.)
- Requirement of minimum processing resource : 인식 알고리즘을 실험할 시 필요한 최소 자원량을 측정함. 자원을 대표하는 것으로는 최소 요구 메모리, 데이터 처리에 필요한 최소

요구 저장 공간 등이 있음

4.2.3 제안 항목

- Transaction speed : 단일 비교에 요구되는 최소 시간을 측정함. System throughput(S)은 이 Transaction speed의 역으로 표현될 수 있음

$$S = \frac{1}{T_{transaction}}$$

$$T_{transaction} = T_{collect} + T_{compute}$$

(표 7) 지문 입력 센서 비교표

하드웨어 평가항목	SWEEPEE03	CCD sensormodule	HFDU01A	Precise100SC
Sensor manufacturer	ATMEL	DELSY	NITGEN	VERIDICOM
Sensor type	Thermal Array	CCD	Optical	Capacitive
Resolution	500	508	500	500
Biometric data quality				
Biometric data size		256x288	248x292	300x300
Sensor area	14, 0.4	9.22, 13.82		15, 15
Module size	17.3, 1.7			61, 92, 19
Temperature	0~70	-40~85	-40~60	0~50
Humidity				
Illumination		0~30000		
Noise				
Power consumption	20	< 150		< 200
API standard compliant				BioAPI
Device interface	USB	Parallel	USB	Parallel, USB
소프트웨어 평가항목	SWEEPEE03	CCD sensormodule	HFDU01A	Precise100SC
Operating system			Win9X, ME, NT, 2000	Win9X, ME, NT, 2000, XP
Size of biometric template			400	variable
Processing time			0.08, 0.08	< 5, < 1
User interface				
Maintenance				
FRR			1% after three trials	< 0.1%
FAR			0.30%	< 0.001%

- Penetration rate : 전체 데이터베이스와 입력된 하나의 생체 정보를 비교하는 경우의 탐색 속도

$$P = \frac{E(\text{number of comparison})}{\text{Number of templates in Database}}$$

- Binning error rate : 실험할 데이터베이스를 올바르게 선택하지 못했을 경우 생길 수 있는 에러율

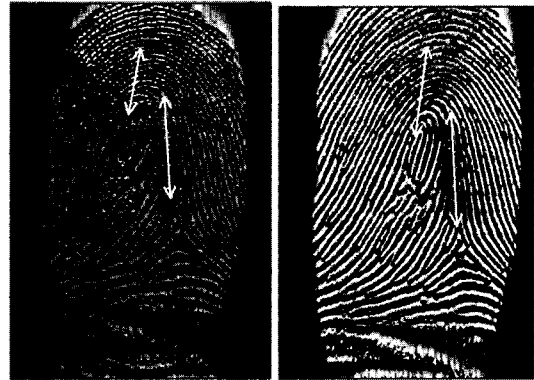
Ⅶ. 지문 입력 센서의 성능 평가 정리

정의된 평가 항목들을 바탕으로 다음 4 가지의 지문 입력 센서에 대해 성능을 비교해 보았다. 비교 결과는 [표 7]과 같다. [표 7]을 보면, 하드웨어 시스템 평가를 위해 정의된 평가 항목들 중 일부는 비교 대상에서 제외되었는데 이는 각 센서 단위가 아닌 인식 시스템 전체에 대한 평가 항목에 관한 내용들이기 때문이며, 소프트웨어 시스템 평가를 위한 평가 항목들 중 일부가 제외된 것은 이에 따른 자체 평가를 할 수 없는 경우로 그 이유를 설명할 수 있다. 대부분의 센서 제공 업체는 자사의 제품에 최적화된 인식 알고리즘을 SDK의 형태로 함께 제공한다. SDK로 제공된 알고리즘의 결과는 인식에 있어서의 확인(True) 또는 확인불가(False)로 주어지게 되며, 이 경우 인식 성능의 정량적인 비교가 불가능하므로 ROC curve와 같은 가시적인 결과를 산출해 내기가 어렵다. 따라서 전체 시스템이 아닌 이와 같은 단위 센서별 평가는 표준으로 제안한 평가 항목들을 완벽하게 적용하여 비교할 수 없음을 알 수 있다.

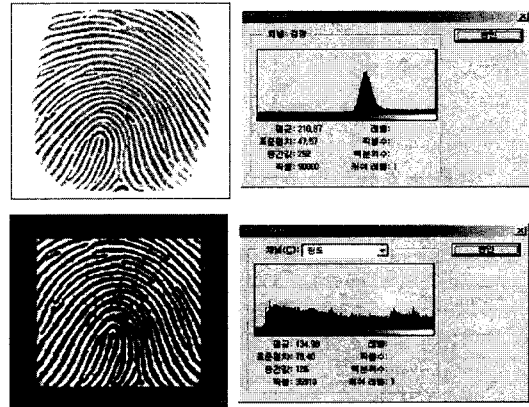
센서를 통한 지문 영상의 해상도 평가 결과는 [표 8]과 같이 나타난다. 이는 [그림 7]에서와 같은 수직 방향 또는 수평 방향의 비교 기준을 선정하는 작업을 통해 준비된다. 비교 대상이 될 비교 기준은 지문 영상 상에 나타나는 특징점들 중 가장 쉽게 찾을 수 있는 몇 개의 특징점을 선별하여 수직선 또는 수평선을 그리는 것으로 결정된다. 각 센서로 획득한 지문 영상 상에서 비교 기준을 찾고 이들의 간격을 평판 스캐너를 사용해 얻은 비교 기준의 간격과 비교하면 지문 영상의 해상도를 측정해 볼 수 있다. 표 8은 센서에서 제공하는 지문 영상의 원래 크기(original)와 평판 스캐너에서 얻은 값과의 비교 결과(증감비율)를 각 센서에 대해 보여주고 있으며, 이 비교 결과를 원래의 이미지에 적용해 원 영상의 크기를 보정한 결과(re-sampling)도 나타내고 있다.

[표 8] 지문 영상 해상도 평가를 위한 측정 결과표

	원 이미지	증감비율	리샘플링
NITGEN	248	0.76%	250
	292	2.36%	299
VERIDICOM	300	-2.23%	293
	300	-0.34%	299
DELSY	256	-28.46%	183
	288	-4.93%	274
ATMEL	312	-6.98%	290
	452	-1.82%	444



[그림 7] 지문 영상 해상도 비교



[그림 8] 지문 영상 품질 평가를 위한 히스토그램 비교 (위: Veridicom 센서 이미지, 아래: AuthenTec 센서 이미지)

지문 영상의 품질 비교를 위해 히스토그램을 비교한 결과는 [그림 8]과 같다. 히스토그램은 영상 처리를 위한 전문 프로그램을 사용해 쉽게 얻을 수 있으며, 이를 통해 영상의 히스토그램이 센서 특성에 따라 다양한 분포도로 나타내어짐을 눈으로 확인할 수 있다.



**Ⅷ. 결 론**

본 논문을 통해 제안한 평가 항목은 기존의 국외 사례들을 바탕으로 하여 국내 개발의 생체 인증 시스템 성능 평가를 위해 적합하다 사료되는 항목들을 정리한 것이다. 또한 단위 센서 및 생체 정보의 특성을 파악하기 위한 평가 항목을 추가로 제안함으로써 본 연구는 기존의 것과는 다른 독창성을 가지게 되었다.

평가가 업체 주관적인 관점에서 이루어지면 특정 시스템의 유리한 성능만이 지나치게 부각되어 제품을 선택, 사용하게 될 사용자의 현명한 판단을 흐리는 결과를 보이게 될 것이다. 따라서 사용자들에게 해당 시스템에 대한 객관적이고 신빙성 있는 성능의 비교 기준을 제시해 주기 위해서는 모든 시스템에 공통으로 적용될 수 있음과 동시에 누구나 그 실효성을 인정해 줄 수 있는 평가 항목의 선정은 중요하다.

평가 항목은 공신력 있는 평가 기관의 평가 작업과 함께 결과의 신뢰도를 높이는데 기여하게 된다. 본 논문에서 제안하는 방법론 및 평가 항목은 한국 정보 보호 진흥원(KISA) 평가 인증 사업단과의 공동 연구를 통해 도출되었으며, 현재 KISA 내에 구축 중인 생체 인증 기술 실험실에 평가 플랫폼, 평가 시나리오 및 평가 항목 등을 활용할 예정에 있다.

지금까지의 국외 생체 인식 관련 평가 기관들의 평가 기준 및 평가 방법에 대한 사례들을 분석해 본 결과, 단일 생체 인식 기술에 대한 인식 에러율과

평균적인 인식 시간들을 주요 평가 항목으로 하고 있음을 알 수 있었다. 그러나 향후 생체 인식 기술의 발전 방향은 다중 생체 인식 기술을 이용한 시스템으로 나아가는 추세이기 때문에 다중 생체 인식 기술에 대한 인식 에러율의 평가 방법의 정립이 필요하다.

**참 고 문 헌**

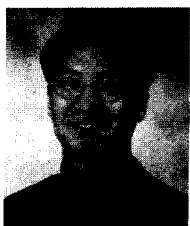
- [1] Sandia National Laboratories, <http://www.sandia.gov/>
- [2] CESG/BWG Biometric Test Programme, <http://www.cesg.gov.uk/biometrics/pBiometricTestReportpt1.pdf>
- [3] BioIS Project, [http://www.igd.fhg.de/igd-a8/projects/biois/biois\\_de.html](http://www.igd.fhg.de/igd-a8/projects/biois/biois_de.html)
- [4] Fingerprint Verification Competition 2000, <http://bias.csr.unibo.it/fvc2000>
- [5] Face Recognition Vendor Test 2000, <http://www.dodcounterdrug.com/facialrecognition/>
- [6] 문지현, 김학일, 류춘우, 이응봉, 전성욱, "생체 인식시스템 성능 평가를 위한 연구", 정보과학회지, 제19권, 제7호, pp. 60~70, 2001.7.
- [7] "Best Practice in Testing and Reporting Performance of Biometric Devices", <http://www.cesg.gov.uk/biometrics/pdfs/BestPractice.pdf>

**〈 著 者 紹 介 〉**



**문 지 현 (Ji-hyun Moon) 학생회원**

1999년 2월 : 인하대학교 항공우주 자동화공학군 자동화 공학 전공  
 1999년 3월~현재 : 인하대학교 자동화공학과 석박사 통합과정  
 2001년 1월~7월 : ETRI 생체인식기술연구팀 위촉연구원  
 <관심분야> 생체인식, 패턴인식, 정보보호



**김 학 일 (Hak-il Kim) 정회원**

1983년 : 서울 대학교 공과대학 제어계측공학과 졸업  
 1985년 : (미) Purdue Univ. 전기/컴퓨터 공학과 석사 학위 취득  
 1990년 : (미) Purdue Univ. 전기/컴퓨터 공학과 박사 학위 취득  
 1990년~현재 : 인하대학교 정보통신공학부 부교수  
 1998년 1999년 : (캐) 사이먼 프레이저 대학교 시스템 공학과 방문교수  
 2001년~현재 : 한국 생체 인식 협의회 성능 평가 분과 위원장  
 <관심분야> 생체인식, 영상처리, 로봇비전