

# 얼굴영상을 이용한 지능형 로봇의 개인식별 및 사용자 추적 성능평가 방법

## Performance Evaluation Method of User Identification and User Tracking for Intelligent Robots Using Face Images

김 대 진<sup>1</sup>, 박 광 현<sup>2</sup>, 홍 지 만<sup>3</sup>, 정 영 숙<sup>4</sup>, 최 병 욱<sup>†</sup>

Dae-Jin Kim<sup>1</sup>, Kwang-Hyun Park<sup>2</sup>, Jiman Hong<sup>3</sup>,  
Young Sook Jeong<sup>4</sup>, Byoung-Wook Choi<sup>†</sup>

**Abstract** In this paper, we deal with the performance evaluation method of user identification and user tracking for intelligent robots using face images. This paper shows general approaches for standard evaluation methods to improve intelligent robot systems as well as their algorithms. The evaluation methods proposed in this paper can be combined with the evaluation methods for detection algorithms of face region and facial components to measure the overall performance of face recognition in intelligent robots.

**Keywords:** User Identification, User Tracking, Performance Evaluation, Face Image, Intelligent Robots

### 1. 서론

본 논문은 일반 가정 및 사무실 환경에서 지능형 로봇의 개인식별(User Identification) 및 사용자 추적기술의 성능평가 방법을 다룬다. 얼굴인식(Face Recognition) 기술의 기본적인 성능 특징을 규정하기 위하여는 얼굴영상을 이용한 개인식별 및 사용자 추적기술의 성능을 측정하기 위한 방법이 규정되어야 한다. 따라서 본 논문의 목적은 성능적인 측면에서의 측정 방법을 제시하는 것이다. 측정 방법에는 사용되는 알고리즘 자체의 특징을 규정하지는 않는다. 또한 본 논문에서 설명하는 시험에서 부대 장치나 표식을 설치하는 것은 시험에 방해가 되지 않는 한 설치할 수 있으며, 얼굴영역 및 얼굴요소 검출기술의 성능평가 방

법과 함께 전체 얼굴인식기술의 성능을 평가하는데 이용할 수 있다.

### 2. 일반 시험 조건

본 논문에서 얼굴인식은 센서에 의해 주어진 데이터 집합으로부터 얼굴영역 검출, 얼굴요소 검출, 개인식별을 수행하는 일련의 과정을 의미하며, 개인식별은 센서 데이터 집합 내에서 추출된 얼굴영역 및 얼굴요소를 이용하여 사용자 개개의 신원을 식별하는 일련의 과정을 의미한다. 본 논문에서 제시하는 시험조건의 규정 일부를 구성하기 위하여 조도 기준인 KS A 3011<sup>[1]</sup>, JPEG 국제표준 ISO 10918-1<sup>[2]</sup>, MPEG-1 국제표준 ISO/IEC-11172<sup>[3]</sup>, MPEG-2 국제표준 ISO/IEC-13818<sup>[4]</sup> 등을 인용하며, 이러한 인용 규격은 최신판을 적용한다.

#### 2.1. 조명 조건

특별히 다른 조건이 없다면, 실내 환경을 상정하여 조도의 범위는 200 50 LUX (KS A 3011 규정<sup>[1]</sup>), 형광등 조명의 규격은 20W~40W/60Hz, 인공조명의 개수는 최소 2초 이상의 동일 규격 조명으로 한다. 형광등 조명은 일반 가

Received: 18.May.2009, Reviewed: 21.July, Accepted: 29.July

※ 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (IITA-2008-C1090-0803-0006)

<sup>†</sup> 교신저자: 서울산업대학교, 전기공학과 교수(bwchoi@snut.ac.kr)

<sup>1</sup> NanoScience Technology Center, University of Central Florida, 전문연구원(dkim@mail.ucf.edu)

<sup>2</sup> 광운대학교 정보제어공학과 조교수(akaii@kw.ac.kr)

<sup>3</sup> 숭실대학교 컴퓨터학부 조교수(jiman@ssu.ac.kr)

<sup>4</sup> 한국전자통신연구원 기술사업화본부 기술진흥팀 책임연구원(yjsjeong@etri.re.kr)

정 및 실험실 환경에 따라 변화될 수 있으며, 인공조명은 방향성에 따른 균일성을 확보하는 범위 내에서 그 수를 변화시킬 수 있다. 측정 범위 내의 조도 조건은 반복성 및 재현성이 확보되어야 하며, 시험 중에 변화가 없도록 주의하여야 한다.

**2.2. 배경 조건**

특별히 다른 조건이 없다면, 단일배경의 경우 JND(Just Noticeable Difference) 기준 95% 내에서의 픽셀밝기 변화를 갖는 동일색으로 구성하여야 하며, 주 대상자 이외의 사용자 혹은 인쇄된 얼굴은 포함하지 않는다. 복합배경의 경우에는 JND 기준 95%를 초과하는 픽셀밝기 변화를 갖는 동일색 혹은 두가지 이상의 색으로 구성하며, 주 대상자 이외의 사용자 및 인쇄된 얼굴이 포함된다. 여기서, JND는 최소 인지 가능한 픽셀 밝기의 변화로 정의되며, 유의미한 JND를 얻기 위해서는 다수 픽셀에 대한 통계학적 분석을 통해 가우시안(Gaussian) 확률분포함수를 얻는 방법이 일반적이다. 측정 범위 내의 배경 조건은 반복성 및 재현성이 확보되어야 하며, 시험 중에 변화가 없도록 주의하여야 한다.

**2.3. 얼굴데이터 구성**

시험을 위한 얼굴데이터는 통계적인 수치의 표현을 위해 최소 100개 이상의 샘플을 수집하여야 하며, 3D 얼굴데이터의 획득도 가능하나 일반적으로는 2D 얼굴데이터를 기준으로 한 측정 방법을 따른다. 다양한 시험을 위해 표 1~4와 같은 기준으로 데이터를 수집한다. 응용분야에 따라 표1~4는 서로 혼용될수 있다.

**3. 얼굴식별 알고리즘 성능 시험**

**3.1. 시험 조건**

얼굴식별의 성능평가를 위해, 얼굴영역 및 얼굴요소를 포함한다수의 정지영상 및 동영상이 주어지고, 피험자에 관한 표준영상 및 표준특징지 정보가전문가에 의해 주어 진다고 가정한다. 시험에 사용되는 영상은 전문가의 육안 에 의해 각 개인별 얼굴요소가 문제없이 식별이 되는 데이터 집합을상정한다. 또한, 데이터베이스에 등록되어 있지 않은 피험자에 대해서는 거부율(Rejection Rate)을 측정하여 성능평가의 지표로삼는다.

**3.2. 시험 방법**

개인식별 성능평가를 위해 다음 절차를 따른다.

(가) 시험 영상의 취득

표 1~4에 기술한 내용을 참고하여 시험 영상을취득한다.

표 1. 시험을 위한 얼굴데이터 구성 (일반항목)

| 항 목          | 세부 규격  | 유의 사항  |
|--------------|--|--|
| 성별           | 남/여 각 50%의 비율로 수집.   | 응용 목적에 따라 성별 구분이 필요한 경우에 유용함.  |
| 피부색          | Fitzpatrick Scale <sup>[5]</sup> 기준 2-6단계로 구성.각 단계별로 유사한%를 유지하되, 국가별 환경을 고려. 동양권의 경우, 3단계 피부색을 고려하여10%-70%-20% 로 구성. | 인공선텐 등에 의해 변화된피부색을 갖는 개인은 제외함.   |
| 연령           | 20세-50세 범위 내로 제한. 20대/30대/40대로 구분하여 각 40%-30%-30% 로 구성.  | 유아기/청소년기/노년기의 얼굴 제외.- 성장으로 인한 급격한 변화.<br>- 얼굴인식에 필요한 요소 발달 미비.<br>- 인식이 어려운 주름 배제. |
| 필요한 데이터 집합 수 | 100명 이상의 피험자 모집.   | 통계적 의미를 갖기에 충분한수를 확보할 것.   |
| 데이터 수집기간     | 1주 이상의 간격으로 1회당 10장 이상, 10회 이상 수집하여 개인별 총 100장 이상의 데이터 수집  | 의상/헤어스타일/액세서리/개인별 특성을반영.   |

표 2. 시험을 위한 얼굴데이터 구성 (변화요인)

| 항 목     | 세부 규격  | 유의 사항   |
|---------|--|---|
| 포즈 변화   | 정면/좌측/우측 및 상단/하단을 향하는 얼굴데이터를 수집. 각 방향에 대해 최소 15 간격으로 변화하는 얼굴 데이터 수집. 각 포즈변화에 대하여 3장 이상 수집.                         | 정적얼굴요소(눈/눈썹/코/입/귀)가 반드시 각 하나 이상 포함되어야 함. 최대 변화각도는 좌/우측 90, 상/하단 45로 한정함. 총 19가지의 포즈변화 가능. |
| 조명 변화   | 2.1절에 의거하여 동일한 밝기 (200 50 LUX - KS A 3011 규정) 를 갖는 인공 조명을 정면/좌측/우측/상단/하단에 최소 15 간격으로 하나씩 배치. 각 포즈변화에 대하여 3장 이상 수집. | 실험실에 미리 부착된 조명은 고려하지 않음. 총 5 가지의 조명변화 가능.   |
| 센서와의 거리 | 최소 40cm에서 최대 4m로 한정함. 거리간간격은 최소 30cm로 고정함. 고정된 거리별로 3장 이상 수집.  | 센서 규격에 따라 측정간격의 조정이 가능. 열 센서의 경우 비교적 세밀한 범위에서의 데이터 수집이 가능.                                |
| 얼굴의 수   | 각 개인의 정적얼굴요소를 모두 포함하는 경우에만하여 최대 2명까지 허용.   | 데이터 비율은 균등하게 수집. (50%-50%, 2명의 경우)  |

(나) 시험 영상의 전처리(크기변환 및 컬러공간변환)

이미 구성된 시험 영상 데이터집합을 사용하는 경우, 표 3 및 표 4에 주어진 표준 크기로 변환한다. 시험 영상이 컬러로 주어진 경우, 밝기변화에 무관한 HSV/HIS 컬러공간으로 변환한다<sup>6)</sup>.

(다) 특징 선정

개인식별을 위한 특징은 크게 국부적 특징(Local Feature)과 전역적 특징(Global Feature)으로 나뉘어진다. 국부적 특징은 각 얼굴요소 혹은 얼굴영역 내의 일부분에 대한 기하학적, 통계학적 특징으로 표현되며, 인간의시각시스템을 모사한 가보필터(Gabor Filter) 특징<sup>7)</sup>과 SIFT(Scale Invariant Feature Transform) 특징<sup>8)</sup>이 사용된다. 전역적 특징은 대부분통계학적 특징을 의미하며, PCA(Principal Component Analysis)<sup>9)</sup>, ICA(Independent Component Analysis)<sup>10)</sup>, DFT(Discrete Fourier Transform)<sup>11)</sup>, DWT(Discrete Wavelet Transform)<sup>12)</sup> 등이 사용된다. MIT의 Pentland 등에 의해 제안된 고유얼굴(Eigenface)은 PCA를 응용한 얼굴관련 특징으로써 가장 널리 사용된다<sup>13)</sup>. 본 논문에서는 국부적 특징으로 가보필터 특징을, 전역적 특징으로 고유얼굴 특징을 사용한다.

표 3. 시험을 위한 얼굴데이터 구성 (정지영상 규격)

| 항 목    | 세부 규격                                    | 유의 사항                            |
|--------|--|----------------------------------|
| 데이터 크기 | 320x240/640x480 픽셀                       | 구입한 센서의스펙을 따르되, 필요시 영상크기변환을 할 것. |
| 데이터 타입 | 흑백: 8비트, 1채널<br>컬러: 8비트, 3채널, RGB, YCbCr | 3채널 이상의 정보는 고려하지 않음.             |
| 데이터 포맷 | JPEG (ISO 10918-1 참조 <sup>2)</sup> )     | -                                |

표 4. 시험을 위한 얼굴데이터 구성 (동영상 규격)

| 항 목    | 세부 규격  | 유의 사항   |
|--------|--|---|
| 데이터 크기 | 352x288/320x240/720x576 픽셀   | 구입한 센서의스펙을 따르되, 필요시 영상크기변환을 할 것.                                |
| 데이터 타입 | 흑백: 8비트, 1채널<br>컬러: 8비트, 3채널, RGB, YCbCr   | 3채널 이상의 정보는 고려하지 않음.  |
| 데이터 포맷 | MPEG-1 (ISO/IEC-11172 참조 <sup>3)</sup> )<br>MPEG-2 (ISO/IEC-13818 참조 <sup>4)</sup> ) | 동영상의 경우, MPEG-1/MPEG-2 포맷을 표준으로 채택하며, I/P/B 프레임중에서 I 프레임만을 사용함. |

(라) 분류기 선정

개인식별을 위한 분류기는 (다)항에서 선정된 여러가지 특징들을 이용하여 데이터베이스에 속한 개인에 대한 식별, 혹은 데이터베이스에 속하지않은 개인에 대한 거부 결과를 얻는 것이 목적이다. 소수 개인을 위해 활용되는 이진화분류기, 선형 분류기 이외에도 다수의 비선형 분류기법이 존재하며, 베이스 정리가 그 근간을 이루고있다. 최근 각광을 받고 있는SVM (Support Vector Machine)은 특징벡터를 저차원에서 고차원으로 변환하여 선형 분류기의 형태로 변환한 것으로 커널(Kernel)의 선정에 따라 매우 좋은 성능을 보인다<sup>14)</sup>. RBF(Radial Basis Function) 커널이 SVM에서 사용되는 대표적인 커널이다. 온라인 개인식별 및 적용형 시스템 구성을 위해서 필수적인 높은 거부율은 개인식별의 성공률보다도 높은 관심을받는 항목이다. 본 논문에서는 SVM 분류기를 사용하는 것을 상정한다.

(마) 표준영상/데이터베이스 비교

3.3절에서 설명할 성능 측정 지수를이용하여, 시험 영상과 검증된데이터베이스를 비교한다. 표준 데이터베이스로 CMU에서 제안된 PIE DB<sup>15)</sup>, Cohn-Kanade DB<sup>16)</sup>, AT&T DB<sup>17)</sup>, Yale DB A/B<sup>18)</sup> 등은 광범위한 얼굴표정, 포즈, 인종변화를 포함하고 있어 객관적인 성능 비교에 많이 활용된다.

3.3. 개인식별 성능 측정

개인식별의 성능측정은 익히 알려진 다수의 성능지수를 이용하여 표현된다<sup>19)</sup>. 본 논문에서는 이들을 그 성격에 따라 표 5와 같이 분류하였다.

(가) 개인식별의 유일성

개인식별의 유일성은 정확도(Precision)로 나타낸다. 정확도는 참긍정(True Positive)으로 분류된 샘플의 수( $n_{TP}$ )와 거짓긍정(False Positive)으로 분류된 샘플의수( $n_{FP}$ )의 비로 나타낸다. 정확도는 개인식별 분류기의 정확함(Exactness) 혹은 정밀함(Fidelity)을 의미한다. 본 논문에서의 정확도(Precision)는 일반적으로 알려진 Accuracy와는 다르다. Accuracy가 분류 실험에서 나타나는 모든 경우의수

표 5. 개인식별 성능평가 지수

| 항 목             | 세부 규격               |
|-----------------|---------------------|
| 개인식별의 유일성       | 정확도 ( $\leq 1$ )    |
| 개인식별의 신뢰성       | 리콜값 ( $\leq 1$ )    |
| 개인식별의 우수성       | ROC 특성곡선의 면적 혹은 F 값 |
| 개인식별의 재연성       | 내부 공분산              |
| 비학습된 개인에 대한 거부성 | 거부율 ( $\leq 1$ )    |

$(n_{TP}+n_{FP}+n_{TN}+n_{FN})$ 에 대한 성공률, 즉

$accuracy = (n_{TP} + n_{TN}) / n_{TP} + n_{FP} + n_{TN} + n_{FN}$ <sup>[20]</sup>을 의미한다면, 정확도는 식(1)과 같이 긍정으로 분류된 데이터들 중에서 참긍정으로 분류된 샘플에 대한 수치를 의미한다. 여기서,  $n_{TN}$ 은 참부정(True Negative),  $n_{FN}$ 은 거짓부정(False Negative)으로 분류된 샘플의 수를 각각 의미한다.

$$precision = \frac{n_{TP}}{n_{TP} + n_{FP}} \quad (1)$$

(나) 개인식별의 신뢰성

개인식별의 신뢰성은 리콜(Recall) 혹은 신뢰율(Reliability)로 나타낸다. 리콜값은 참긍정으로 분류된 샘플의 수( $n_{TP}$ )와 거짓부정으로 분류된 샘플의 수( $n_{FN}$ )의 비로 나타낸다. 본 논문에서의 신뢰도는 특정 개인에게서 추출된 모든 데이터 중 올바르게 찾아진 데이터의 비(식(2) 참조)를 나타내며, 이것은 개인식별기의 기능적 완전성(Completeness)을 의미한다.

$$recall = \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (2)$$

(다) 개인식별의 우수성

개인식별의 우수성은 식(1)과 식(2)로 표현되는 정확도와 리콜값의 가중적 하모닉 평균(weighted harmonic mean)을 혼용한  $F$  값<sup>[21]</sup> 혹은 특정 파라미터의 변화에 따른 정확

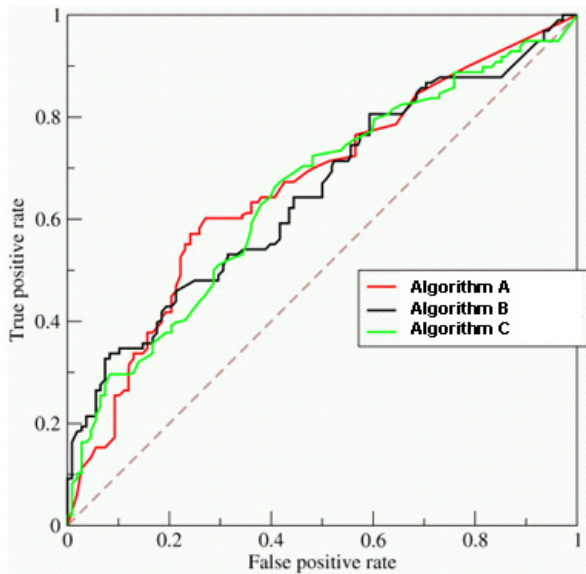


그림 1. ROC 특성곡선<sup>[22]</sup>

도와 리콜값의 그래프적 표현인 ROC (Receiver operating characteristic) 특성곡선의 면적으로 표현된다.  $F$  값은 정확도와 리콜값 사이의 비율( $\beta$ )에 따라  $F_\beta$  값으로 일반화할 수 있으며, 식(3)과 같이 정의된다.

$$F_\beta = \frac{(1 + \beta^2)(precision \cdot recall)}{\beta^2(precision + recall)} \quad (3)$$

그림 1은 전형적인 ROC 특성곡선을 이용하여 세가지 서로 다른 알고리즘의 비교를 수행한 것을 보여준다. 면적의 크기에 따라 A, C, B의 순으로 성능이 감소한다. ROC 특성곡선은 민감도(Sensitivity; True Positive Rate)와 특이도(Specificity; 1-False Positive Rate)의 조합으로 설명된다<sup>[23]</sup>. ROC 특성곡선을 이용하면 광범위한 변화에 대한 시스템의 성능을 효과적으로 비교할 수 있어 최근에는 기계학습, 데이터마이닝 등에서 매우 중요한 평가지수로 사용된다.

(라) 개인식별의 재연성

개인식별의 재연성은 분류기의 분류결과를 출력공간 상에 도사하였을 때 얼마나 근접해 있는가를 의미하는 수치이다. 통계적으로 이러한 재연성은 공분산(Covariance)의 비율로 표현되며 식(4)와 같이 정의된다.

$$R^i = \frac{\sum A(i)}{\sum_{j=1}^N \sum A(j)} \quad (4)$$

여기서,  $A(i)$ 는  $i$ 번째 개인에 대한 특징벡터 집합을 의미한다. 특징벡터집합  $A$ 에 대한 공분산은  $\Sigma_A = E(A - E(A))$ 와 같이 계산된다. 즉, 각 개인별로 계산된 공분산은 전체 인원내 대해 정규화(Normalize)되어 [0-1]에 해당하는 값으로 변환된다.

(마) 비학습된 개인에 대한 거부성 [24]

비학습된 개인에 대한 거부성은 전체 오답수( $n_{TN}+n_{FN}$ )에 대해 얼마나 정확하게 오답( $n_{TN}$ )이 분류되었나를 의미하는 거부율 (Negative Predictive Value)로 표기한다.

$$NPV = \frac{n_{TN}}{n_{TN} + n_{FN}} \quad (5)$$

### 4. 로봇의 얼굴인식 성능 시험

3절에서 살펴본개인식별 성능평가 방법은 특정 플랫폼에 한정되지 않는 일반적인 알고리즘의 성능평가에 활용

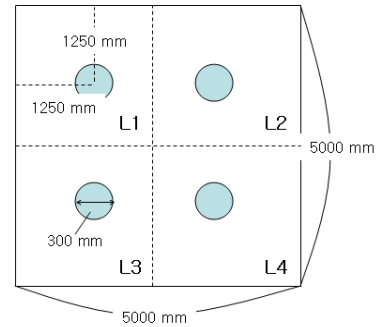
된다. 반면에 로봇과 같은 특정 플랫폼에서의 얼굴인식 알고리즘의 성능평가를 위해서는 구체적인 응용분야로서 얼굴인식을 통한 사용자 추적을 우선으로 꼽는다. 본 절에서는 얼굴인식 알고리즘의 로봇에 대한 응용으로써 사용자 추적 시스템의 성능평가 방법에 대해 기술한다.

#### 4.1. 시험 조건

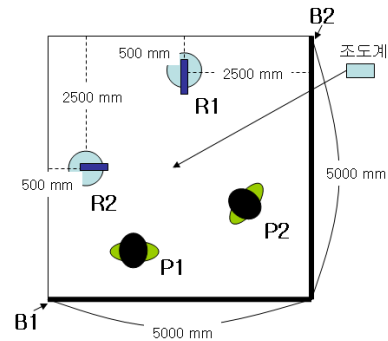
로봇의 얼굴인식 성능을 시험하기 위하여, 다음의 각 조건을 명시한다.

- (가) 시험 공간은 폭/넓이 각 5m 내외의 밀폐된 공간으로 한정한다. 이때, 높이는 3m 내외로 한정한다.
- (나) 시험 공간의 배경 조건은 단일색 벽면과 복합색 벽면으로 구성한다. 이 때, 색상은 국제표준 먼셀(Munsell) 컬러표준 [25]을 따르며, 단일색은 흰색(N9), 복합색은 적색(5R), 청색(5PB), 녹색(5G)으로 구성한다.
- (다) 시험 공간의 조명 조건은 2.1절을 참조하여 4기의 20W/60Hz 원형 형광등을 천정에 부착하고, 총 15가지의 조합에 의한 조명 조건을 시험한다. 즉, 모든 형광등을 끈 경우를 제외한 모든 조합을 상정한다. 각 형광등은 시험 공간의 천정 부위에 균일한 간격으로 배치한다. 이때, 형광등은 원형을 사용하여 방위에 따른 균등한 조명효과를 갖도록 한다.
- (라) 시험 공간 내에 배치하는 로봇은 고정형플랫폼(본체가 워크스테이션 형태로고정되며 카메라의 팬틸트줌(Pan-Tilt-Zoom) 운동이 가능한 형태) 혹은 이동형 플랫폼으로 구성되나, 얼굴인식 본연의 성능평가를 위하여 팬틸트줌 기능 및 이동형 플랫폼의 움직임에 의한 효과는 고려하지 않는다.
- (마) 시험 공간 내에는 두 명의 사용자가 자유롭게 활동할 수 있으나, 성능평가를 위해서는 다음 사항을 고려한다. 사용자는 표 1에 기술한 내용을 따라 로봇에 부착된 센서의 측정 범위 내에 위치하며, 각 사용자는 4기의 형광등에 의해 나뉘어진 구획 내에 위치한다. 두 명의 사용자는 하나의 구획 내에 다른 사용자의 얼굴을 모두 가리지 않는 한 함께 위치할 수 있으며, 사용자의 머리 위에 위치한 형광등은 반드시 켜져야 한다.
- (사) 시험 공간 내에 배치된 로봇에 부착되는 센서는 단일/복합 센서를 혼용할 수 있다.
- (아) 시험 공간 내에서 센서의 측정 범위 및 사용자의 움직임을 크게 제한하지 않는 선에서 기타 장애물/가구의 설치가 가능하다.
- (자) 시험 시간은 10분 내외로 제한한다.

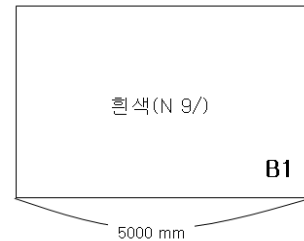
그림 2는 로봇의 얼굴인식 성능시험을 위한 시험장 규격을 나타낸다. 그림 2(가)에 도시한 것처럼, 천정에는 직경 300mm 내외의 원형 형광등이 균등하게 4기 부착되며, 전



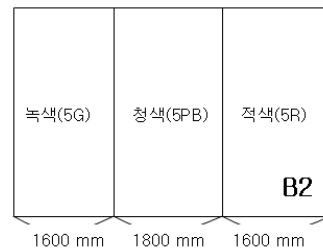
(가) 천정



(나) 바닥 및 로봇/사용자 배치



(다) 단일색 배경



(라) 복합색 배경

그림 2. 로봇의 얼굴인식 성능시험을 위한 시험장 규격

체 시험장을 총 4개의 구역(L1-L4)으로 나누어 조명조건을 구성한다. 그림 2(나)는 바닥 및 로봇, 사용자의 배치를 나타내며, 단일색배경 시험시 로봇은 R1 위치에서 단일색배경 B1 (그림 2(다) 참조)을 바라보고 위치하며, 복합색배경 시험시 로봇은 R2 위치에서 복합색배경 B2 (그림 2(라) 참조)를 바라보고 위치한다. 조명조건을 기록하기 위한 조도계는 시험장의 중앙에 위치하며, 두 명의 사용자(P1, P2)는 얼굴인식이 가능한범위 내에서 자유로운 움직임을 취한다. 복합색 배경에서 칭색은 세가지원색 중 가장 피부색과 잘 구분되므로 중앙에 위치하고 가장 넓은 범위로 구성한다.

### 4.2. 시험 방법

본 논문에서의 사용자 추적 시험은로봇의 위치변화에 의한 효과를 고려하지 않는다. 즉, 사용자 추적 동안에 이루어지는 로봇의 위치변화는 성과는 무관하게 이루어질 수 있으며본 논문에서는 사용자 추적 본연의기능에 의거한 시험을 수행하는 것을 전제로 한다. 사용자 추적의 성능시험은 다음과 같은 절차로수행된다.

#### (가) 시험 환경 선정

사용할 센서를 선정하고, 4.1.(나) 및 4.1.(다) 항을 참조하여 본 시험에서 사용할 조명 및 배경 조건을 선정한다. 이때, 사용자의 위치 및 겹침 조합에 따라 조명 조건이중속적으로 변화한다.

#### (나) 시험 영상 취득 및 전처리

표 1에 기술한내용을 참고하여 시험 영상을취득한다. 시험 영상의 전처리를 위해서는 3.2.(나)항에 명기된 방법을 이용한다.

#### (다) 특징 선정

3.2.(다)항에 기술한 내용을 바탕으로 특징을선정한다.

#### (라) 분류기 선정

3.2.(라)항에 기술한 내용을 바탕으로 분류기를 선정한다. 이때, 개인식별에 관한 분류기는 두 사용자를 서로 구분하기 위해 사용된다.

#### (마) 성능 평가 및 기록

4.3절에 소개되는 성능평가 지수를 이용하여 최소 10회 이상 수행된 시험 성능을 기록한다.

### 4.3. 로봇의 얼굴인식 성능 시험

표 6에 명기한 성능평가 지수는 실제 로봇에 탑재된 얼굴인식 기술의평가에 활용된다.

#### (가) 사용자 추적 성공률

사용자 추적의 성공률은 각 사용자에 대한 얼굴영역 검출, 얼굴요소 검출 및 개인식별 성능의 조합으로 표현된다. 단, 얼굴요소 검출 성능은 얼굴영역 검출 및 개인식별 성능에 비해 적은 가중치를 갖는다. 개인식별 분류기가 국부적 특징을 포함하지 않는 경우에는 얼굴요소 검출 성능에 대한 가중치는 0으로 둔다.

사용자추적 성공률 =

$$w_0 \times \text{얼굴영역추출 성능} + w_1 \times \text{얼굴요소추출 성능} + w_2 \times \text{개인식별 성능} \quad (6)$$

#### (나) 사용자추적 속도

동작속도는 ms 혹은fps (Hz) 단위로 기록하며 영상획득, 전처리, 얼굴영역 추출, 얼굴요소 추출, 특징추출, 개인식별로 구성되는 세부단계별 분석이필요하다.

#### (다) 사용자 추적 시작조건

최초 알고리즘의 동작시 사용자 혹은 시스템 운영자에 의한 수동 사용자 추적이 이루어질 것인지, 특정벡터(특히, 움직임 정보) 혹은 기타 특정센서를 이용하여 자동 사용자 추적이이루어질 것인지를 명시하고, 이를 구체적인 시작 조건으로 기술한다. 로봇 시스템의 자율성(Autonomy) 및 지능성(Intelligence) 측정의한 척도가 될 수 있으며, 본 논문에서는 광류(Optical Flow)<sup>[26]</sup>를 이용한 자동 사용자 추적을 시작조건으로 상정한다(광류 기반의 사용자 얼굴 추적은 [27] [28] [29] [30] 등에서 보고되었다).

#### (라) 사용자 추적 오류회복 조건

주어진 시험 환경에 따라 부분적으로 사용자 추적이 실패할 수 있다. 이러한 실패의 원인과해결책을 기술한 것이 오류회복 조건이며, 오류회복 조건은 가능한상세히 구체적인 방법론과 특정벡터를 통해 기술되어야 한다. 오류회복은 자동 혹은 수동으로 이루어질 수 있으며, 정량적으로는 오류회복 성공률을 %로 표현한다. 일반적으로 사용되는 방법은 실패하기 직전의얼굴영역 위치 변화를 모델링하고 이를 이용한 예측(Prediction) 기법이 널리 쓰인다. 본

표 6. 로봇의 얼굴인식 성능평가 지수

| 항 목         | 세부 규격                              |
|-------------|------------------------------------|
| 얼굴인식 성공률    | 얼굴영역/얼굴요소/개인식별 성능의 가중치합으로 표현.      |
| 얼굴인식 속도     | ms/FPS(Hz) (각 단계별 소요시간 기입)         |
| 얼굴인식 시작조건   | 자동 광류 정보를 이용한 성공률 기입.              |
| 얼굴인식 오류회복조건 | 자동 칼만필터에 의한 성공률 기입.                |
| 겹침/조명변화     | 다른 사용자, 가구에의한 겹침. 조명조합 및 평균 조도 기입. |

논문에서는 칼만필터(Kalman Filter)<sup>[31]</sup>를 이용한 사용자 추적 오류회복 기법을 상정한다 (칼만필터 및 그 변형을 이용한 사용자 추적 오류회복은 [32] [33] [34] [35] 등에서 보고되었다).

(마) 겹침/조명변화

시험 시간 내에 취득된 영상은 겹침이 없는 영상, 다른 사용자와의 겹침 영상, 기타 시험공간 내에 배치된 가구에 의한 겹침이 있는 영상으로 분류된다. 여기서, 겹침이 있는 영상을 제외한 경우에 한하여 사용자추적 오류회복 조건이 측정된다. 조명변화는 4.2.(가)항에 의해 정의된형광등의 선택적 조합, 조도계를 이용한 시험 공간내의 평균 조도를 이용해서 나타낸다. 이때, 조도계는 시험 공간의 중심에위치하여 균등한 측정을 수행한다.

5. 결론

본 논문에서는 지능형로봇을 위해 필수적인 요소인얼굴 인식에 관한 요소 중 개인식별 및 사용자 추적에대한 표준 평가방법을 제안하였다. 제안된 방법은 지능형 로봇에서의 얼굴인식 응용을 위한 플랫폼의 평가에 활용할 수 있다. 표준화된 평가 방법은 알고리즘의 발전과 함께 상업화에 중요하게 이용될 수 있을 것이다.

본 논문에서 제시한 개인식별 및 사용자 추적을 포함한 얼굴인식 성능의평가서는 표 7 및 표 8과 같이 정리할 수 있다.

표 7. 개인식별 성능 평가서 예시

| 실험 환경         |           |                       |           |
|---------------|-----------|-----------------------|-----------|
| 일련번호          | -         |                       |           |
| 일시            | 년 월 일     |                       |           |
| 장소            | 연구소 제연구동  |                       |           |
| 사용데이터 및 처리시스템 |           |                       |           |
| 데이터 규격        | 데이터의 크기   | 640x480 픽셀            |           |
|               | 데이터의 수    | 1,500장 (개인별 100-150장) |           |
|               | 데이터의 포맷   | 8bit 흑백 JPEG          |           |
|               | 데이터의 저장용량 | 440M Byte             |           |
| 데이터 처리시스템     | CPU의 종류   | Intel Core2Duo 3GHz   |           |
|               | 물리메모리     | 4GB DDR2 RAM          |           |
| 평가 항목         |           |                       |           |
| 유일성           | 0.78 0.02 | 우수성                   | 0.81 0.02 |
| 신뢰성           | 0.85 0.03 | 재연성                   | 0.65 0.08 |
| 거부성           | 0.05 0.01 | -                     | -         |

| 평가 결과 |  |
|-------|--|
| B     |  |

표 8. 로봇의 얼굴인식 성능 평가서 예시

| 실험 환경         |           |                                    |               |                                     |         |           |  |
|---------------|-----------|------------------------------------|---------------|-------------------------------------|---------|-----------|--|
| 일련번호          | -         |                                    |               |                                     |         |           |  |
| 일시            | 년 월 일     |                                    |               |                                     |         |           |  |
| 장소            | 연구소 제연구동  |                                    |               |                                     |         |           |  |
| 조명            | 조도        | 230 10 LUX                         |               |                                     |         |           |  |
|               | 광원        | 형광등 4기 / 20W / 60Hz                |               |                                     |         |           |  |
|               | 조명변화      |                                    | L1            | L2                                  | L3      | L4        |  |
|               |           | 1                                  | on            | off                                 | on      | off       |  |
|               |           | 2                                  | off           | on                                  | on      | on        |  |
| 3             |           | off                                | on            | off                                 | Off     |           |  |
| 4             | on        | on                                 | off           | on                                  |         |           |  |
| 배경            | 단일색 (N 9) |                                    |               |                                     |         |           |  |
| 온도            | 상온 (26 2) |                                    |               |                                     |         |           |  |
| 로봇 및 센서 시스템   |           |                                    |               |                                     |         |           |  |
| 로봇            | 플랫폼       | 이동형 PatrolBot (MobileRobots, Inc.) |               |                                     |         |           |  |
|               | 크기        | 59cm(폭) x 48cm(넓이) x 38cm(높이)      |               |                                     |         |           |  |
|               | 동작시간      | 3.5시간                              |               |                                     |         |           |  |
| 센서            | 종류        | 비전                                 |               |                                     |         |           |  |
|               | 규격        | NTSC / 8bit, 3채널 / RGB             |               |                                     |         |           |  |
|               | 동작속도      | 25 fps / 40 ms                     |               |                                     |         |           |  |
|               | 하드웨어규격    | IEEE 1394 / PCI                    |               |                                     |         |           |  |
| 사용데이터 및 처리시스템 |           |                                    |               |                                     |         |           |  |
| 데이터 규격        | 데이터의 크기   | 640x480 픽셀                         |               |                                     |         |           |  |
|               | 데이터의 수    | 1,500장 (개인별 100-150장)              |               |                                     |         |           |  |
|               | 데이터의 포맷   | 8bit 흑백 JPEG                       |               |                                     |         |           |  |
|               | 데이터의 저장용량 | 440M Byte                          |               |                                     |         |           |  |
| 데이터 처리시스템     | CPU의 종류   | Intel Core2Duo 3GHz                |               |                                     |         |           |  |
|               | 물리메모리     | 4GB DDR2 RAM                       |               |                                     |         |           |  |
| 평가 항목         |           |                                    |               |                                     |         |           |  |
| 사용자 추적 성공률    | 총합        | 95.2%                              | 사용자 추적 회복 성공률 | 78.5% (전체 영상의 48%에 대해서만 측정 겹침정보 참조) |         |           |  |
|               | 얼굴영역      | 96.1%                              |               | 다른 사용자와의 겹침                         | 35%     |           |  |
|               | 얼굴요소      | -                                  |               |                                     | 가구와의 겹침 | 13%       |  |
|               | 개인식별      | 94.3%                              |               |                                     |         | 겹침 없음 52% |  |
| 사용자 추적 시작 성공률 | 총합        | 355ms (3.8Hz)                      | 겹침            | 다른 사용자와의 겹침 35%                     |         |           |  |
|               | 영상취득      | 40ms                               |               | 가구와의 겹침 13%                         |         |           |  |
|               | 전처리       | 5ms                                |               | 겹침 없음 52%                           |         |           |  |
|               | 특징추출      | 250ms                              |               |                                     |         |           |  |
|               | 얼굴영역 추출   | 30ms                               |               |                                     |         |           |  |
|               | 개인식별      | 25ms                               |               |                                     |         |           |  |
| 사용자 추적 시작 성공률 | 89.5%     | 조명변화                               | 1             | 10%                                 | 230 LUX |           |  |
|               |           |                                    | 2             | 27%                                 | 235 LUX |           |  |
|               |           |                                    | 3             | 15%                                 | 225 LUX |           |  |
|               |           |                                    | 4             | 48%                                 | 240 LUX |           |  |
| 평가 결과         |           |                                    |               |                                     |         |           |  |
| A             |           |                                    |               |                                     |         |           |  |

---



---

**참고문헌**


---



---

- [1] KS A 3011 조도 기준
- [2] ISO 10918-1 JPEG 국제표준
- [3] ISO/IEC-11172 MPEG-1 국제표준
- [4] ISO/IEC-13818 MPEG-2 국제표준
- [5] T.B. Fitzpatrick, "Soleil et peau", *J. Med Esthet*, 1975.
- [6] Conversion between RGB and HSV/HSL, [http://en.wikipedia.org/wiki/HSL\\_and\\_HSV#Conversion\\_from\\_HSL\\_to\\_RGB](http://en.wikipedia.org/wiki/HSL_and_HSV#Conversion_from_HSL_to_RGB), 2009.
- [7] J.G. Daugman, "Two dimensional spectral analysis of cortical receptive field profile", *Vision Research* **20**: 847-856, 1980.
- [8] D. Lowe, "Object recognition from local scale-invariant features", *Proc. of the ICCV*, pp.1150-1157, 1999.
- [9] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space", *Philosophical Magazine* **2**(6): 559-572, 1901.
- [10] P. Comon, "Independent Component Analysis: a new concept?", *Signal Processing* **36**(3): 287-314, 1994.
- [11] A. Oppenheim *et al.*, *Discrete-time signal processing*, Prentice Hall, 1999.
- [12] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
- [13] M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuroscience* **3**(1): 71-86, 1991.
- [14] J. Christopher, "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery* **2**: 121-167, 1998.
- [15] Homepage of CMU PIE Database, [http://www.ri.cmu.edu/research\\_project\\_detail.html?project\\_id=418&menu\\_id=261](http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261), 2009.
- [16] Homepage of CMU Cohn-Kanade Facial Expression/Face Database, [http://vasc.ri.cmu.edu/idb/html/face/facial\\_expression/index.html](http://vasc.ri.cmu.edu/idb/html/face/facial_expression/index.html), 2009.
- [17] Homepage of AT&T Face Database, <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>, 2009.
- [18] Homepage of Yale A/B Face Database, <http://cvc.yale.edu/>, 2009.
- [19] J. Makhoul, "Performance measures for information extraction", *Proc.of DARPA Broadcast News Workshop*, 1999.
- [20] J.R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, University Science Books, pp.128-129, 1999.
- [21] V. Rijsbergen, *C.V.: Information Retrieval*, 2nd Edition, 1979.
- [22] Homepage of Wikipedia - Receiver operating characteristic, [http://en.wikipedia.org/wiki/Receiver\\_operating\\_characteristic](http://en.wikipedia.org/wiki/Receiver_operating_characteristic), 2009.
- [23] J.A. Hanley and B.J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases", *Radiology* **148**(3): 839-843, 1983.
- [24] D.G. Altman and J.M. Bland, "Diagnostic tests 2: Predictive values", *BMJ* **309**(6947): 102, 1994.
- [25] A. Munsell, "A Pigment Color System and Notation", *The American Journal of Psychology* **23**: 236-244, 1912.
- [26] S.S. Beauchemin and J.L. Barron, *The computation of optical flow*, ACM New York, 1995.
- [27] D. DeCarlo and D. Metaxas, "Optical Flow Constraints on Deformable Models with Applications to Face Tracking", *Int'l Journal of Computer Vision* **38**(2): 99-127, 2000.
- [28] S. Basu, I. Essa, and A. Pentland, "Motion Regularization for Model-based Head Tracking", *Proc. of ICPR'96*, 1996.
- [29] J.F. Cohn, A.J. Zlochower, J.J. Lien, and T. Kanade, "Feature-point tracking by optical flow discriminates subtle differences in facial expression", *Proc. of AWFGR'98*, pp.396-401, 1998.
- [30] Y. Yacoob and L.S. Davis, "Recognizing Human Facial Expressions From Long Image Sequences Using Optical Flow", *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18**(6): 636-642, 1996.
- [31] R.E. Kalman, "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering* **82**(1): 35-45, 1960.



- [32] K. Schwerdt and J.L. Crowley, "Robust face tracking using color", *Proc. of AWFGR'00*, pp.90-95, 2000.
- [33] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking", *IEEE Trans. on Pattern Analysis and Machine Intelligence* **25**(5): 564-577, 2003.
- [34] R.J. Qian, M.I. Sezan, and K.E. Matthews, "A robust real-time face tracking algorithm", *Proc. of ICIP'98*, pp.131-135, 1998.
- [35] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-Based Probabilistic Tracking", *Lecture Notes in Computer Science* **2350**: 661-675, 2002.



**김 대 진**

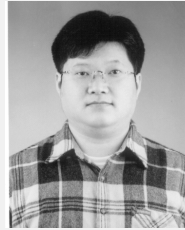
1997 경북대학교 전자공학과 (학사)  
 1999 KAIST 전자전산학과(석사)  
 2004 KAIST 전자전산학과(박사)  
 2004~2006 KAIST 인간친화복  
 지로봇시스템연구센터  
 전문연구원

2006~2007 Multimedia Research Lab, University of  
 Louisville, Louisville, KY 전문연구원

2007~현재 NanoScience Technology Center,  
 University of Central Florida, Orlando, FL 전문연  
 구원

관심분야 : 지능로봇, 인간-로봇-컴퓨터 상호작용, 로  
 봇제어및비전, 얼굴표정인식, 멀티미디어 마  
 이닝

IEEE 정회원



**박 광 현**

1994 KAIST 전자전산학과(학사)  
 1997 KAIST 전자전산학과(석사)  
 2001 KAIST 전자전산학과(박사)  
 2005~2007 KAIST 전자전산학  
 과 BK 초빙교수  
 2008~현재 광운대학교 정보  
 제어공학과 조교수

관심분야 : 인간-로봇 상호작용, 보조공학시스템, 로  
 봇 소프트웨어, 로봇 미디어  
 한국로봇학회 (KRS) 정회원

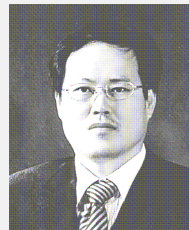


**홍 지 만**

1994 고려대학교 컴퓨터학과  
 (학사)  
 1997 서울대학교 컴퓨터공학  
 과(석사)  
 2003 서울대학교 컴퓨터공학  
 과(박사)

2004~2007 광운대학교 컴퓨터공학과 조교수  
 2007~현재 숭실대학교 컴퓨터학부 조교수

관심분야 : 운영체제, 임베디드 소프트웨어, 실시간  
 시스템, 결합허용 시스템, 센서네트워크



**최 병 욱**

1992 KAIST 전기및전자(박사)  
 1988-2000 LG산전(주), 엘리베  
 이터 연구실장 및 임베디  
 드 시스템 연구팀장  
 2000~2005 선문대학교 제어  
 계측공학과 부교수

2005~현재 서울산업대학교 전기공학과 교수  
 2007~2008 Nayang Technological University, Senior  
 Fellow

관심분야 : 임베디드 시스템, 실시간 시스템, 지능형  
 서비스 로봇 소프트웨어 구조