

논문 2004-41SP-6-18

# 얼굴 인식을 개선을 위한 선형이동 능동카메라 시스템기반 얼굴포즈 보정 기술

(Sliding Active Camera-based Face Pose Compensation for Enhanced  
Face Recognition)

장 승 호\*, 김 영 옥\*\*, 박 창 우\*\*, 박 장 한\*\*\*, 남궁재찬\*\*\*, 백 준 기\*\*\*\*

(Sung-ho Jang, Young-Ouk Kim, Chang-Woo Park, Chang-Han Park,  
Jae-Chan Namkung, and Joonki Paik)

## 요 약

최근 지능형 로봇에 대한 관심이 모아지고 있다. 지능형 로봇의 가장 큰 특징은 사용자를 추적, 인식하고 그 결과를 기반으로 상호활동적인 대응을 할 수 있다는 점이다. 얼굴인식이 다른 생체인식과의 비교에서 장점을 가질 수 있는 점은 비 강제성과 비 접촉성을 들 수 있다. 그러나 얼굴인식은 얼굴 취득단계부터 차원의 감소가 발생하고 인식하고자 하는 얼굴 및 주변 환경 변화가 매우 심하기 때문에 다른 생체인식에 비하여 인식이 낮다. 얼굴인식의 성능을 저하시키는 요인들로는 조명변화, 포즈변화, 표정변화, 카메라와의 거리 등을 들 수 있다. 본 논문에서는 실제 환경에서 얼굴 인식 성능에 가장 많은 영향을 미치는 포즈변화에 대응하기 위하여 새로운 선형이동 능동형 카메라를 개발하여, 정면 얼굴에 근접한 영상을 취득하고 주성분 분석 및 Hidden Markov Model 알고리즘을 이용하여 인식률을 개선하고자 한다. 제한된 방법은 지능형 보안시스템 및 모바일 로봇에 적용하는 것을 목표로 개발 되었지만, 높은 정확도의 얼굴인식을 요구하는 응용분야에 널리 적용할 수가 있다.

## Abstract

Recently, we have remarkable developments in intelligent robot systems. The remarkable features of intelligent robot are that it can track user and is able to do face recognition, which is vital for many surveillance-based systems. The advantage of face recognition compared with other biometrics recognition is that coerciveness and contact that usually exist when we acquire characteristics do not exist in face recognition. However, the accuracy of face recognition is lower than other biometric recognition due to the decreasing in dimension from image acquisition step and various changes associated with face pose and background. There are many factors that deteriorate performance of face recognition such as the distance from camera to the face, changes in lighting, pose change, and change of facial expression. In this paper, we implement a new sliding active camera system to prevent various pose variation that influence face recognition performance and acquired frontal face images using PCA and HMM method to improve the face recognition. This proposed face recognition algorithm can be used for intelligent surveillance system and mobile robot system.

**Keywords** : sliding active cameras, convex-hull, fuzzy inference, principal component analysis, face recognition

\* 학생회원, \*\*\*\* 정회원, 중앙대학교  
(Chung-Ang University)

\*\* 정회원, 전자부품연구원  
(Korea Electronics Technology Institute)

\*\*\* 정회원, 광운대학교  
(kwangwoon University)

※ 이 연구는 과학기술부 국가지정연구실 사업과 교육  
부 BK21 사업 지원으로 수행되었음

접수일자: 2004년7월21일, 수정완료일: 2004년8월9일

## I. 서 론

생체인식이란 사람 고유의 음성, 홍채, 망막, 혈관, 서명, 지문 등을 이용하여 개인에 따른 특징을 추출하고 이 데이터를 이용하여 타인의 특징과 비교하는 방식으로 인증하는 기술이다. 생체적인 특징을 복제하거나 위조하기가 어렵고 휴대할 필요가 없다는 점에서 열쇠나

카드식 출입통제 시스템등과 같은 보안시스템을 대신할 차세대 보안 기술로 연구 되고 있는 분야이다. 최근에는 활용성 및 편리성이 뛰어난 얼굴인식 기술이 새로운 생체인식 기술로 자리를 잡아가고 있는 추세이다.

얼굴인식의 단계는 얼굴영역 검출, 특징 추출 및 데이터베이스(database; DB) 구축, 그리고 정합 알고리즘의 3단계로 나눌 수 있으며, 얼굴인식 방법은 크게 특징점 기반 방법과 영상 기반 방법으로 이루어져 있다. 특징점 기반의 방법은 얼굴을 구별할 수 있게 하는 특징점을 추출 한 후 특징점의 기하학적인 관계를 분석하거나 눈, 코 입 등의 텍스처를 비교하는 방법으로 패턴 인식 분야에서 오랜 시간동안 연구되어 오고 있으며, 영상 기반 방법은 얼굴 영상 전체를 이용하여 특징량을 계산하는 방법이다.<sup>[1]</sup>

효과적인 특징추출은 얼굴인식의 문제를 해결하는 중요한 요소가 된다. 얼굴영상의 특징은 경계선, 등고선, 영상 영역의 시각적 특징,<sup>[2]</sup> 히스토그램, 모멘트 등의 통계적 특징,<sup>[3]</sup> 푸리에변환 등의 변환 계수적 특징,<sup>[4]</sup> 주성분 분석법(principal component analysis; PCA)<sup>[5-11]</sup> 등의 대수적 특징으로 나누어진다.

이 중에서도 통계적인 방법을 이용하는 PCA, 신경회로망(neural networks; NN),<sup>[12]</sup> support vector machines (SVM),<sup>[13]</sup> linear discriminant analysis(LDA)<sup>[14]</sup> 방법이 널리 사용되어 왔다.

또한 이외에도 모델기반 매칭방법을 사용한 elastic graph matching(EGM),<sup>[15]</sup> 확률을 이용하여 특징을 알아내는 hidden markov models(HMM)<sup>[16]</sup> 등이 얼굴인식 분야에서 많이 쓰이고 있다. 얼굴영상이 정면영상으로 구성된 DB에 정면 얼굴을 입력하는 경우 인식률은 상당히 높다고 알려져 있지만, 이 방법은 전체적인 얼굴영상 정보에 기반하여 얼굴을 모델링 하기 때문에 얼굴의 포즈변화나 조명변화, 표정변화 등 얼굴의 전체적인 영상정보가 취득되지 않으면 인식률이 급격하게 저하되는 단점을 가지고 있다.

최근에는 이러한 경향을 보완하기 위하여 얼굴의 전체적인 특징이 아닌 지역적 특징을 통계적으로 모델링하는 local feature analysis(LFA)<sup>[17]</sup>, independent component analysis(ICA)<sup>[18]</sup>방법을 이용하여 얼굴의 국소적인 변화에 보다 강건하게 대처하기 위한 방법들이 소개 되고 있다.

얼굴인식의 성능에 가장 큰 영향을 미치는 얼굴 포즈변화를 보정하기 위해서는 정확한 포즈 추정이 필수적이다.<sup>[19-20]</sup> 포즈변화는 카메라의 축 시스템에 따라서 세 개

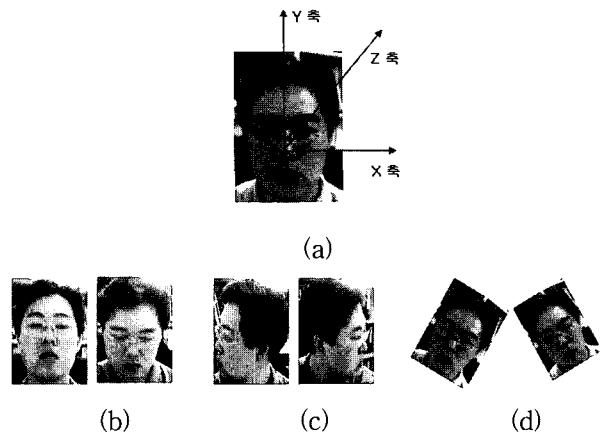


그림 1. 얼굴 포즈 변환 : (a) 정면 영상, (b) X 축 변환, (c) Y 축 변환, (d) Z 축 변환

Fig. 1. Pose variation of face : (a) Frontal Image (b) X axis (c) Y axis (d) Z axis.

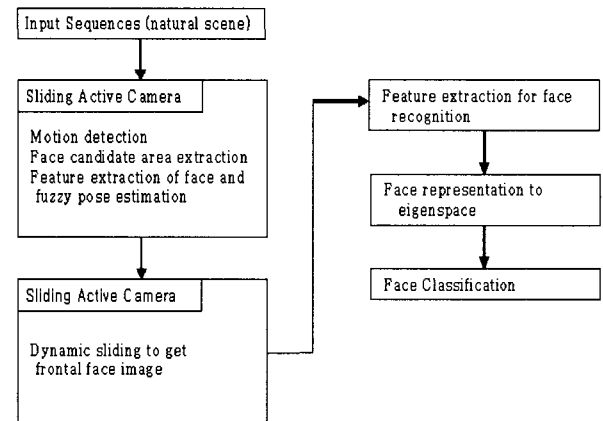


그림 2. 제안 알고리즘의 블록도

Fig. 2. Block diagram of the proposed algorithm.

의 포즈변화로 분류할 수 있으며, 첫째는 X축으로, 두 번째는 Y축으로, 마지막으로 Z축으로 회전되는 변환이다. 이 변화 중 Y 축으로의 회전이 가장 얼굴인식률에 미치는 영향이 크다. 그림 1은 정면 영상으로부터 어파인 변환(affine transform) 되어진 영상들을 보여주고 있다.

본 논문에서는 선형이동이 가능한 팬틸드(pan/tilt; PT) 카메라 모듈을 이용하여 포즈 별로 구성된 DB에서 전통적인 PCA 알고리즘을 이용하여 얼굴의 Y 축 회전이 발생하는지를 추정할 수 있다. 추정된 회전을 기준으로 카메라가 Y 축 방향으로 선형이동하고 PT 모듈에서 패닝(panning)이 되면 얼굴포즈 각도를  $\pm 5^\circ$  정도로 보정된 영상을 취득 할 수 있다. 그림 2에서는 포즈변화에 따른 얼굴인식에 대한 제안된 알고리즘 순서도를 나타내었다.

본 논문의 II장에서는 선형이동 능동카메라(sliding active camera; SAC) 시스템을 기반으로 실시간 얼굴

영역을 추적할 통하여 SAC 시스템에 내장된 PT 모듈을 효과적으로 제어하고,<sup>[21]</sup> convex-hull을 이용한 객체 가려짐 및 얼굴색 그리고 옷을 입은 사람에 대해서도 강인하게 얼굴영역을 검출 할 수 있는 방법을 제안하였다. III장에서는 II장의 얼굴영역 추적을 기반으로 퍼지 추론(fuzzy inference) 및 SAC 시스템을 이용한 얼굴포즈 보정과 보정된 정면영상을 다층에지맵을 이용하여 얼굴인식을 위한 DB 구축에 대하여 소개한다. IV장에서는 III장에서 취득된 정면 얼굴을 얼굴인식의 보편적인 알고리즘인 PCA와 HMM 알고리즘을 이용하여 얼굴인식을 수행하며 기존의 얼굴보정을 하지 않고 포즈별로 구성된 DB와의 인식률을 비교하였다. V장에서 결론 및 향후 연구방향에 대해서 기술한다.

## II. 선형이동 능동카메라 시스템

### 1. 선형이동 능동카메라

본 논문에서 제안하는 SAC는 실시간으로 얼굴영역을 추적할 수 있는 PT모듈과 얼굴 포즈를 추정하고 추정된 각도를 변환하여 정면영상을 취득하기 위한 선형이동 모듈로 구성되어 있다. 이 시스템은 기존 고정 카메라의 경우 한정된 시야(fields of view; FOV)로 인하여 감시 영역이 협소해 짐으로서 한정된 감시 영역만을 설정 할 수 있으나, PT 및 Y 축 이동을 이용하여 카메라의 FOV를 Y 축으로  $\pm 90^\circ$ (pan), X 축으로  $\pm 45^\circ$ (tilt)의 동작 범위를 가짐으로서 광역의 공간에서 객체의 움직임을 추적 및 감시 할 수 있다는 장점을 가지고 있다.

얼굴인식의 경우 실제 환경에서 인식률에 영향을 받는 여러가지 요소가 존재한다. 주변 조명의 변화, 얼굴 포즈 및 표정변화, 시간흐름, 얼굴영상의 해상도에 따라서 인식률은 저하된다. 포즈변화의 경우 X, Y, Z 축 방

향 변화 중에서 Y 축 방향의 포즈변화가 얼굴인식성에 가장 많은 영향을 미친다. Y 축 방향의 경우 얼굴포즈 각에 따라서 얼굴의 전역적인 특징이 나타나지 않기 때문에 인식률이 저하된다.

표 1과 그림 3<sup>[21]</sup>은 Y 축 방향의 얼굴포즈 변화에 따른 얼굴 인식률의 통계를 보여주고 있다.

그림 3에서 보는 것과 같이  $\pm 25^\circ$  사이에서는 인식률이 비교적 높는데 그 이상으로 각이 커지게 되면 얼굴 인식률이 떨어지는 것을 볼 수가 있다. 따라서 이런 문제점을 극복하기 위하여 SAC를 이용하여 Y 축 방향으로의 포즈를 추정하고 포즈별 각도에 해당하는 만큼 선형이동 하면 정면 얼굴 영상에 가까운 영상을 취득 할 수 있어 인식률을 높일 수 있다. 그림 4는 SAC를 보여주고 있다. 이론적으로 본 논문에서 제안하는 시스템은 거의  $\pm 80^\circ$  까지 얼굴포즈에 대하여 보정을 할 수 있다. 하지만 카메라와 얼굴사이의 거리가 멀어질수록 sliding 축의 길이 길어지기 때문에 제한점을 갖는다.

그림 5에서는 실제 구현된 SAC 시스템의 구조도이며 선형이동 길이가 1.2m이고, 추적 객체가 SAC에 0.6m 이내에 떨어져 있을 경우에 한하여 얼굴의 Y 축으로 회전 되어진 각에 대하여  $\pm 45^\circ$ 까지 보정이 가능하다.

이러한 하드웨어 방식의 포즈보정 방법은 모바일 로

표 1. 포즈별 인식률

Table 1. Summary of the pose test.

Pose (R, L)	1st Match (%)	1* 10 Match (%)	Manual Aligning Required (%)
90°L	N/A	N/A	100.0
60°L	34.5	71.0	13.5
40°L	65.0	91.0	4.5
25°L	95.0	99.5	2.5
15°L	97.5	100.0	0.5
0	100.0	100.0	0.0
15°R	99.0	99.5	0.0
25°R	90.5	99.5	2.0
40°R	61.5	87.5	4.5
60°R	27.5	65.0	11.0
90°R	N/A	N/A	100.0

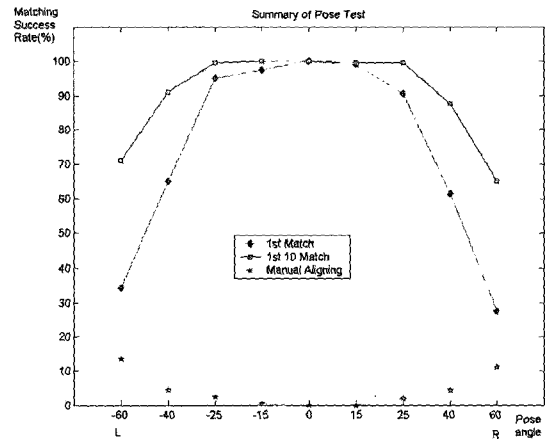


그림 3. 포즈별 인식률 그래프

Fig. 3. Summary of the pose test graph.

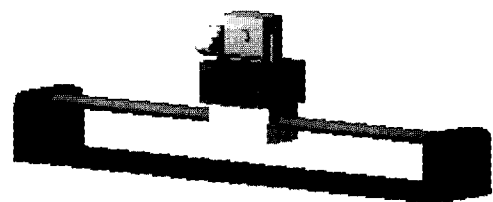


그림 4. 선형이동 능동카메라

Fig. 4. sliding active camera.

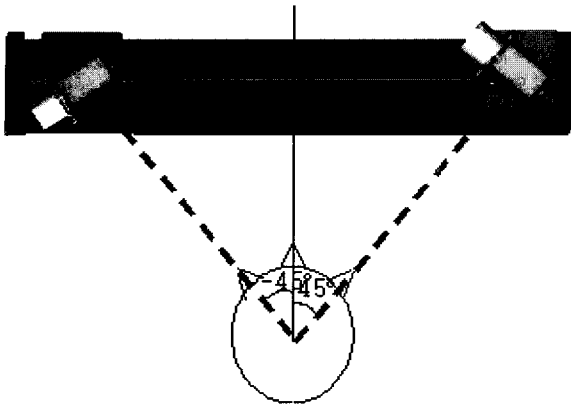


그림 5. 포즈 정규화를 위한 선형이동 능동카메라의 이동  
 Fig. 5. Moving the sliding active camera for pose localization.

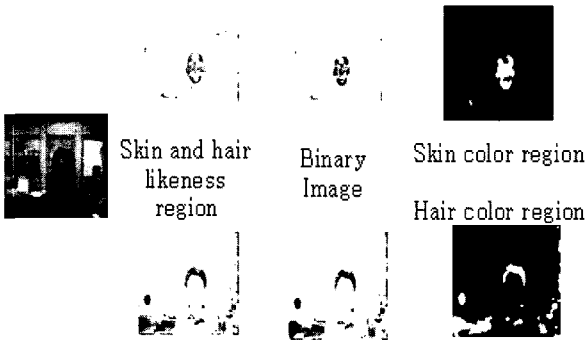


그림 6. 피부와 머리카락의 색상 범위  
 Fig. 6. Skin and hair color region.

봇 시스템에 적용했을 경우 제한점 없이 모든 Y 축 회전에 대하여 정면 얼굴 영상을 취득 할 수 있다는 장점이 있고, 3차원 얼굴영상 합성 및 기하학적인 정보를 이용한 얼굴합성 방식에 비하여 정확하고 빠른 정면영상을 취득 할 수 있는 장점이 있다.

2. Convex-Hull을 이용한 얼굴영역 검출<sup>[22]</sup>

추적 시스템에서 많이 사용하는 정보로는 (i) 칼라 분포, (ii) 움직임 정보, (iii) 객체의 윤곽 등이 있다. 본 논문에서는 탐지 하려는 객체의 영역에서의 피부 색상 정보와 머리카락의 색상 정보를 알아내는 방법을 적용한다. 이를 통하여 얼굴영역에 대해 convex-hull을 적용하여 얼굴 영역을 검출한다. YCbCr의 칼라 모델에서 convex-hull의 Cb, Cr을 동양인에 맞게 얼굴 피부 색상 범위를  $R_{cb} = [77 \ 127]$ ,  $R_{cr} = [133 \ 173]$ 로 정의한다. 얼굴 영역의 피부 색 정보(chrominance component)를 얻어내는 방법은 식 (1)과 같다.  $S(x, y)$  는 얼굴영역에서의 피부 색상 정보를 이용하여 convex-hull을 씌운

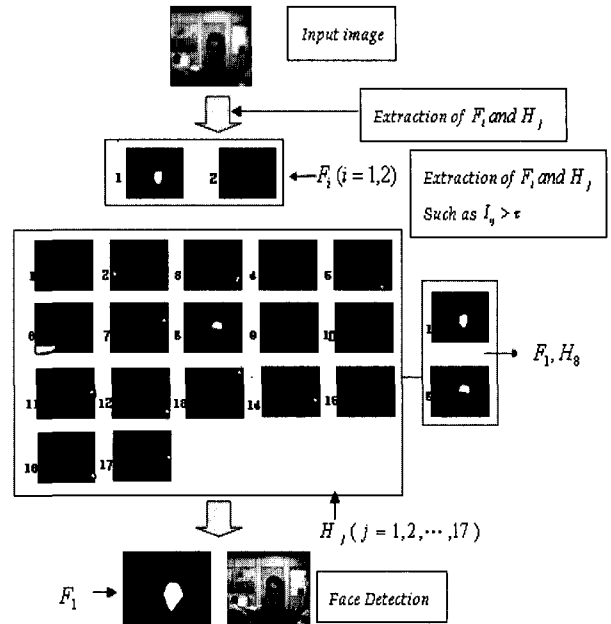


그림 7. 얼굴 검출 과정  
 Fig. 7. Face detection procedure.

부분을 다음 식과 같이 나타낸다.

$$S(x, y) = \begin{cases} 1, & \text{if } [Ch(x, y) \in R_{Cb}] \cap [Cr(x, y) \in R_{Cr}] \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

얼굴영역에서의 머리카락 부분의 색상 정보를 RGB로 나눌 수 있으며, 이 또한 RGB 각각의 범위를 동양인의 머리카락 색상 정보에 맞게  $RR = [0 \ 30]$ ,  $RG = [0 \ 30]$ ,  $RB = [0 \ 30]$ 로 정의된다. 얼굴영역의 머리카락의 밝기 정보(luminance component)를 얻어내는 방법은 식 (2)와 같다.

$$H(x, y) = \begin{cases} 1, & \text{if } [R(x, y) \in R_R] \cap [G(x, y) \in R_G] \cap [B(x, y) \in R_B] \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$H(x, y)$  는 얼굴영역에서의 머리카락의 색상정보를 이용하여 convex-hull을 씌운 부분을 나타낸다. 그림 6은 식 (1)과 (2)를 통하여 얻은 얼굴영역의 피부 색상 정보와 머리카락의 밝기 정보를 통하여 두 영역에 대하여 각각 분할한 것을 보여주고 있다.

본 논문에서는 각 블롭(blob)들을 포함하는 convex-hull을 구해서 이들 사이의 관계를 통하여 얼굴 영역을 분할하는 방식을 제안한다. 식 (3)은 convex-hull을 이용하여 피부 색상 정보를 통한 convex-hull 영역과 머리카락의 convex-hull 영역에서의 공통된 원소 집합을 보여주고 있다.



그림 8. 얼굴영역 분할  
Fig. 8. Face segmentation.

$$I_{ij} = F_i \cap H_j, \quad (i = 1 \sim n, j = 1 \sim m)$$

$$\text{The values of pixels in } F_i = \begin{cases} 1, & n[I_{ij}] > \tau \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

여기서  $n[*]$ 는 집합에 대한 화소 수,  $H_j$ 는 머리카락 영역을 포함하는 convex-hull 내부의 화소들의 집합,  $F_i$ 는 얼굴영역을 포함하는 convex-hull 내부 화소들의 집합,  $\tau$ 는 임계치를 나타낸다. 그림 7에서는 얼굴영역을 검출하는 과정을 보여주고 있다.

그림 8에서는 얼굴영역의 피부 convex-hull 영역과 머리카락의 convex-hull 영역을 합한 convex-hull을 통하여 얼굴영역을 분할한 영상을 보여주고 있다.

### III. 포즈 추정 및 보정을 위한 퍼지 추론

#### 1. 퍼지 추론 엔진

앞 절의 convex hull을 이용한 사람의 얼굴 후보 영역을 분할하는 과정에서 SAC시스템의 PT 모듈은 실시간으로 얼굴영역을 검출하는 동작이 되고 있다. 그림 1의 (b), (c)는 각각 X 축, Y 축을 기준으로 얼굴이 회전 될 수 있음을 보여준다. 본 논문에서는 얼굴인식물에 가장 영향을 미치는 Y 축의 회전을 보정할 수 있는 방법에 대해서 설명한다.

먼저 정면 얼굴 영상을 중심으로 15° 씩 회전 된 7종류의 각도에 대해서 DB를 구축하고, 앞에서 기술된 convex-hull에 의하여 검출된 얼굴영역을 입력으로 사용한다. Y 축으로 회전된 각도를 추정하기 위하여 Mamdani의 퍼지추론(Mamdani's fuzzy inference)을 사용하며, 그 입력은 PCA나 HMM의 출력을 사용한다.

포즈추정을 위한 얼굴영상의 각도는 각각 -45°, -30°, -15°, 0°, 15°, 35°, 45° 의 7개의 DB를 정의하고, 그 DB가 그림 9에 보인다.

SAC가 퍼지추론 방법을 통하여 입력된 얼굴이 Y 축으로의 회전 값을 연속적인 값으로 추정하기 때문에 각각의 각도별 영상이 많이 필요하지는 않다. 본 논문에서는 7개의 각도에 따라 각각 5명의 영상을 통하여 퍼지추



(L45) (L30) (L15) (0) (R15) (R30) (R45)

그림 9. 포즈별 데이터베이스  
Fig. 9. Face database of pose estimation.

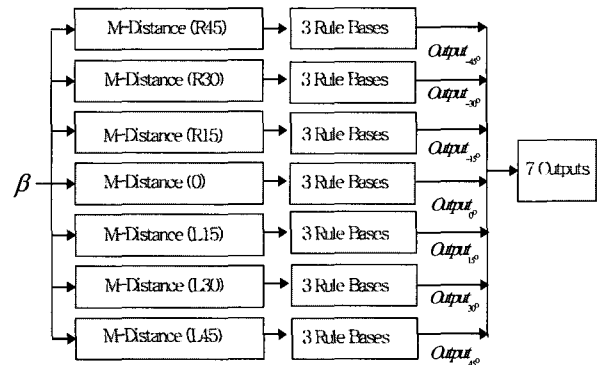


그림 10. 포즈별 퍼지 추론 방법  
Fig. 10. Pose estimation using fuzzy inference.

론을 이용하여 포즈추정 및 선형이동 축의 이동 거리를 결정한다. 그림 10은 SAC의 이동 거리를 알아보기 위한 퍼지추론 방법이다

$\beta$ 는 입력영상을 각도별로 구성된 DB를 바탕으로 PCA를 통한 계수 벡터이며, 식 (4)는 각각의 퍼지추론들의 규칙기반(rule bases)에 따른 각도별 클래스 사이의 Mahalanobis 거리를 계산하여 7개의 퍼지추론의 입력 값으로 할당한다.

$$\text{Output}_\theta = \begin{cases} 1, & D_\theta(\beta) = 0 \\ 0.5, & D_\theta(\beta) \text{ is small,} \\ 0, & D_\theta(\beta) \text{ is large} \end{cases} \quad (4)$$

여기서  $D_\theta(*)$ 는 입력에 대한 각 각도별 클래스 사이의 Mahalanobis 거리의 값을 나타내고,  $\theta$ 는 -45°, -30°, -15°, 0°, 15°, 35°, 45° 이다. 출력 값 '1'은 최대의 클래스 소속도를 의미하며, '0'은 최소의 클래스 소속도를 의미한다.

그림 11은 퍼지추론에 대한 입력 소속 함수(member-ship function)를 보여주고 있다. 여기서 MaxD는 실험 값에 의해 취득된 각각의 7개 입력 값에 대한 Mahalano

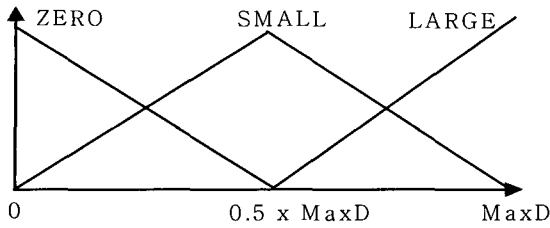


그림 11. 소속 함수  
Fig. 11. Membership function.

-bis 거리의 값이 최대가 된다.

포즈별 SAC의 이동 거리를 알기 위해서 Singleton fuzzifier와 평균 defuzzifier를 이용하여 입력영상에 대해서 소속 함수를 이용한 퍼지추론 방법의 규칙에 따라 일정 범위의 소속도를 구한다. 소속도에 따라서 식 (5)를 통하여 각각의 포즈별 소속 값을 곱하여, 얼굴포즈의 각도를 얻어낸다. 이를 통하여 SAC의 이동된 거리를 추론할 수 있다.

$$\begin{aligned} \text{Face Pose Angle} = & -45^\circ(\text{Output}_{-45^\circ}) + -30^\circ(\text{Output}_{-30^\circ}) + -15^\circ(\text{Output}_{-15^\circ}) \\ & + 0^\circ(\text{Output}_0) + 15^\circ(\text{Output}_{15^\circ}) + 30^\circ(\text{Output}_{30^\circ}) + 45^\circ(\text{Output}_{45^\circ}) \end{aligned} \quad (5)$$

2. 정면 얼굴 및 다층에지맵을 이용한 얼굴 DB 구축<sup>[23]</sup>

본 논문에서는 컬러영상에 무관한 얼굴요소의 명도특성을 이용한 다층에지맵(multilayered edge map)을 통해 얼굴검출 및 얼굴 요소를 추정하는 접근방법을 제시한다. 두 눈은 주변 얼굴영역보다 어두운 특성을 갖는다. 보통의 눈 부근의 경우 어두운 작은 영역이 수평방향으로 보다 큰 길이를 가지게 되는 흡사 블롭 형태의 수평에지와 같은 모양으로 얻어진다. 입 영역도 마찬가지로 눈 영역과 같은 주변보다 어두운 영역이 상당히 존재하며, 눈 영역에 비해 그 영역넓이가 다양하나, 그 크기는 대체로 눈 영역 크기와 크게 차이 나지 않는다.

이러한 얼굴요소의 명도특성을 고려하여 방향성 블롭 템플릿(directional blob template)을 구성하여 얼굴요소 가능위치를 추정하기로 한다.

템플릿의 크기는 검출하려는 얼굴크기를 고려하여 결정하되, 그림 9와 같이 눈 크기와 대강 비슷한 정도로 가로가 더 긴 비율로 선택한다. 영상 크기가  $W \times H$ 인 영상 내부의 화소  $P(x, y)$ 에서 템플릿 중심에 놓이는 화소를  $P_{cent} = (x_c, y_c)$ 로 정의하고,  $P_{cent}$ 를 기준으로  $W_{FF} \times H_{FF}$  크기의 얼굴요소 방향성 템플릿의 8 방향의 화소들의 평균밝기  $\bar{I}_{Dir}$  를 계산하여  $I_{cent}(P(x_c, y_c))$ 의 밝기 값)와

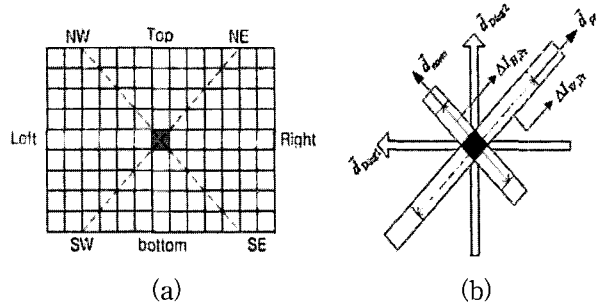


그림 12. (a) 방향성 템플릿, (b) 에지 맵에 대한 새로운 방향성  
Fig. 12. (a) Directional intensity template (b) Principal direction for creation of multilayered edge maps.

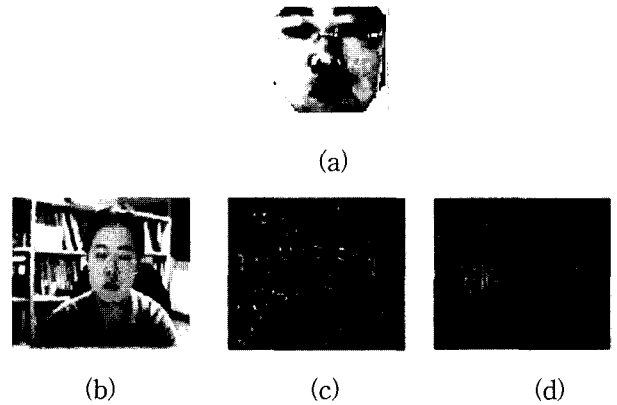


그림 13. (a) 데이터베이스 (60\*60) (b) 원 영상 (c) 에지 맵 (d) 소벨 에지  
Fig. 13. (a) Database (60\*60) (b) Original image (c) multilayered edge (d) Sobel edge.

의 밝기 편차 값을 얻는다. 이러한 8 방향 템플릿의 예를 그림 12(a)에서 보여준다. 가장 밝기편차가 큰 쪽을 주성분 방향(principal direction)  $\bar{d}_{pr}$ 로 결정하고, 이때의 편차 크기  $\bar{d}_{pr}$ 을 얻는다.

각 화소의 편차 주성분 방향  $\bar{d}_{pr}$ 과 이때의 편차크기  $\bar{d}_{pr}$ 을 이용하여 그림 12(b)와 같이 방향성 템플릿을 전체 영상에 다시 적용한다.  $\bar{d}_{pr}$  방향의 양쪽으로 평균영상 밝기를 계산한  $\Delta I_{W,pr}$ 와 주성분 방향 편차크기에 적당한 가중치를 할당한 값을 문턱치로 하여 이보다 높을 때 +1레벨의 에지 강도를 할당한다. 다음으로 주 방향의 수직방향  $\bar{d}_{Norm}$ 으로 다시 가중치 할당된 값보다 높은 픽셀일 경우 다시 +1(총+2)레벨의 에지 강도를 부여한다. 두 방향의 대각방향  $\bar{d}_{Diag1,2}$ 에 대해서도 마찬가지로 +1의 에지강도를 부여한다. 이러한 과정을 통해서 전체 영상은 4단계의 에지강도 조건을 만족하는 경

우에 따라 다른 강도의 밝기를 가지는 에지맵 이 작성되며, 최대 +4 레벨까지 부여되는 에지강도 위치에 가장 밝게 그레이 에지가 생성되게 된다. 그림 13은 다층 에지맵 영상을 이용한 얼굴영역 분할을 보여주고 있으며, 이를 통하여 얼굴 DB를 구축하였다.

**IV. 주성분 분석과 HMM을 이용한 얼굴인식 성능 평가**

본 논문에서는 SAC에 내장된 PT 모듈을 이용하여 실시간으로 얼굴영역을 추적하면서 Y 축으로 회전된 각도를 피쳐추론 엔진으로 추정하고, 선형이동 축을 제어하여 거의 정면에 가까운 얼굴 영상 취득이 가능한 시스템을 제안했다.

제안한 방법으로 취득된 얼굴영상으로 DB를 구축하고, 일반적으로 얼굴인식 시스템에 많이 사용되고 있는 PCA방법과 HMM방법을 이용하여 성능평가를 하였으며, 입력 영상으로는 실내 환경에서 Y축으로 회전된 다양한 얼굴영상을 사용하였다.

PCA는 다차원 신호를 분석하여, 데이터의 차원을 낮추는 데 많이 이용되는 다변량 통계 분석 방법이다. 주어진 데이터를 분산이 최대가 되는 축으로 변환하는 것으로, 이 새로운 차원에서의 특징벡터를 얻어내고, 입력 영상과 각각의 영상별로의 특징벡터에 관하여 Mahala-nobis 거리를 계산하여 가장 작은 결과 값을 가지게 되는 영상이 인식되게 되는 것이다. HMM은 개인별로의 포즈에 따라 얼굴영역을 눈, 코, 입의 3 단계 영역으로 오버래핑 되게 나눈 후, 상태 천이 확률을 구하고, DB를 이용하여 관측 벡터(observation vector)를 구해낸다. 이를 통하여 HMM을 트레이닝 시킴으로써 대표 얼굴영상과 각각 개인의 얼굴 영상의 특징을 모델링 할 수 있다.<sup>[16]</sup>

**1. 포즈 변환에 실험 결과**

그림 14는 직접 제작된 SAC 시스템을 보여주고 있다.

본 시스템은 200W급 AC 서보모터와 전용 드라이버 및 선형 가이드(linear guide)를 사용했으며, 선형이동 축의 제어는 Delta Tau사의 전용 제어기(PMAC)를 이용하여 PT 모듈의 위치를 피드백하여 제어하는 방식을 사용했다. 선형이동은 0.2~0.8 m/sec의 속도로 구동된다.

실험에 사용된 데이터베이스는 2, 3장에 기술된 방법으로 취득했으며, 영상의 크기는 60×60을 사용하였다. -45°, -30°, -15°, 0°, 15°, 35°, 45° 의 각각의 포즈별로 100명의 DB를 구축하였다. 총 700개의 영상 DB를 사용하여, 실험 장치는 각각 포즈별로 156개의 입력 영상으로

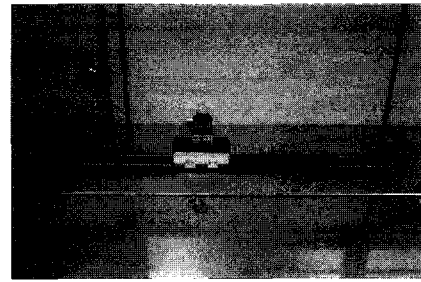


그림 14. 선형이동 능동카메라 시스템  
Fig. 14. sliding active camera system.

표 2. 포즈변환에 따른 실험 결과

Table 2. Results showing the performance of pose variation.

포즈 변환 (R, L)	테스트 영상갯수	정확도 (%)
40~50°R	156	92.3
25~35°R	156	95.6
10~20°R	156	98.1
5°R~5°L	156	99.2
10~20°L	156	98.3
25~35°L	156	93.4
40~50°L	156	92.5

표 3. 포즈변환에 따른 적용 알고리즘 비

Table 3. Comparative evaluation of different pose variation methods.

포즈 변환 (R, L)	*Our method (%)	*PCA (%)	**HMM (%)	**PCA (%)
40~50°R	90.2	34.5	70.6	60.2
25~35°R	93.4	65.1	88.2	78.6
10~20°R	96.4	92.0	93.8	93.0
5°R~5°L	98.2	97.4	98.0	97.8
10~20°L	96.2	92.2	93.6	93.1
25~35°L	93.5	65.0	87.9	79.5
40~50°L	90.8	35.0	71.0	62.7

\* : 70 명이 정면영상만 가지고 있을 경우.

\*\* : 70 명이 각각 포즈별 영상을 가지고 있을 경우.

실험하였다. 본 실험에서의 실험 장치는 고정된 각이 아니고 일정 범위(-50°~50°)내에서 Y 축으로 회전되어 있는 얼굴영상을 입력으로 사용하였다.

표 2에서 나타난 결과와 같이 얼굴후보영역에 대한 검출의 확률이 높을수록 포즈 추정이 정확한 것을 볼 수 있으며, 포즈추정 정확도가 90%를 상회할 경우 포즈추정 오차는 ±3° 이하로 추정됨을 알 수 있었다.

**2. 얼굴 인식 결과**

본 실험은 PCA와 HMM 알고리즘 방법을 SAC 시스

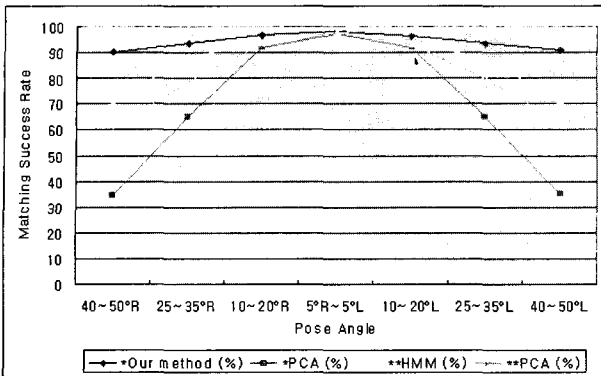


그림 15. 포즈변화에 따른 적용 알고리즘 비교 그래프  
Fig. 15. Comparative evaluation of different pose variation methods graph.



그림 16. 실험에 의한 결과 영상들  
Fig. 16. Experimental result images.

템에 적용하여 포즈보정에 따른 방법과 포즈를 보정하지 않은 경우를 비교하였다. 본 논문에서 제안하고 있는 방법을 적용하였을 경우에 다른 방법에 비해서 좀 더 인식률이 높다는 것을 볼 수 있었다. 또한 PCA 알고리즘을 적용한 결과를 보게 되면 70 명이 각각 포즈별 영상을 가지고 있을 경우에 좀 더 좋은 인식률을 보였다.

HMM을 적용한 실험 역시 좋은 결과를 얻었으나, 본 논문에서 제안된 방법을 사용하여 비교하였을 경우 포즈변화에 따른 인식률에 있어서는 다른 알고리즘에 비하여 좋은 결과를 얻을 수 있음을 알게 되었다.

그림 16은 포즈변환 되어진 영상으로부터 선형이동 능동카메라를 통해 얻어진 정면 영상들을 보여주고 있다.

### V. 결 론

본 논문에서는 얼굴포즈 변화에 따른 인식률저하, 그중에서도 특히 얼굴인식에 가장 많은 영향을 미치는 Y 축으로 회전된 얼굴영상을 보정하기 위하여 포즈추정과 새로운 SAC 시스템을 제작하여 정면 얼굴 영상을 취득한 결과를 제시하였다.

또한 이 시스템을 기반으로 실제 환경에서 객체추적과 얼굴영역분할, 퍼지추론을 이용한 카메라 이동 및 얼굴인식이 가능한 통합된 시스템을 구현하였고, SAC 시스템을 이용하여 PCA와 HMM 알고리즘을 적용하여 인식 성능을 실험하였으며, 이를 통하여 영상합성이나 스테레오 영상처리에 의한 포즈보정방법보다는 향상된 인식결과를

얻을 수 있었다.

본 논문에서 제안된 포즈보정방법을 이동형 로봇에 적용할 경우 제한없이 사용자의 정면 얼굴을 취득할 수 있는 장점이 있다. 향후에는 모바일 로봇에 적용을 염두에 둔 처리 속도개선 및 주변 환경 변화에 강인한 포즈추정 및 보정에 관한 연구가 진행되어야 한다.

### 참 고 문 헌

- [1] J. Zhang, Y. Yan and M. Lades, "Pace recognition: eigenface, elastic matching, and neural nets," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1423-1435, 1997.
- [2] O. Nakamura, S. mathur and T. Minami, "Identification of human faces based on isodensity maps," *Pattern Recognition*, vol. 24, pp. 236-272, 1991.
- [3] C. H. The and R. T. Chin, "On image analysis by the method of moments," *IEEE Trans. Patt. Anal. Machine Intell.* vol. 10, pp. 496-513, 1988.
- [4] S. R. Dubois and F. H. Grantz, "A autoregressive model approach to two-dimensional shape classification," *IEEE Trans. Patt. Anal. Machine Intell.* vol. 8, pp. 56-65, 1986.
- [5] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal, Cognitive Neuroscience*, vol. 3, pp 72-86, 1991.
- [6] A. J. O'Toole, H. Abdi, K. A. Deffenbacher and D. Valentin, "Low dimensional representation of faces in high dimensions of the face space," *J. Opt. Soc. Am. A*, vol. 10, pp. 405-411, 1993.
- [7] L. Sirovich and M. Kirby, "Low dimensional procedure for the characterization of human faces," *J. Opt. Soc. Am.* vol. 4, pp. 519-524, 1987.
- [8] Z. Q. Hong, "Algebraic feature extraction of images for recognition," *Pattern Recognition*, vol. 24, pp. 211-219, 1991.
- [9] L. Sirivich and M.kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Amer.*, vol. 4, pp. 519-524, 1987.
- [10] S. Gong, S. J. McKenna, A. Psarrou, "Dynamic Vision," Imperial College press, 2000.
- [11] E. Oja, "Principal components, minor components, and linear neural networks," *Neural Networks*,



- vol. 5, pp. 927-935, 1992.
- [12] H. Rowley, S. Baluja and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 20, no. 1, pp. 203-208, 1998.
- [13] E. Osuna, R. Freund and F. Girosi, "Training support vector machines: an application to face detection," *Proc. IEEE Conf. Computer Vision, Pattern Recognition*, pp. 130-136, 1997.
- [14] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. PAMI*, vol. 19, no. 7, pp. 711-720, 1997.
- [15] L. Wiskott, J. Fellous, N. Krger and C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Trans. PAMI*, vol. 19, pp. 775-779, 1997.
- [16] F. Samaria and S. Young, "HMM based architecture for face identification," *Image and Vision Computing*, vol. 12, pp. 537-543, 1994.
- [17] P. Penev and J. Attick, "Local Feature Analysis: a general statistical theory for object representation," *Network: Computation in Neural Systems*, vol. 7, no. 3, pp. 447-500, 1996.
- [18] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287-314, 1994.
- [19] C. W. Park, Y. O. Kim, H. K. Sung, "Multiple Face Segmentation and Tracking based on Robust Hausdorff Distance Matching," *Int. Journal of Fuzzy Logic and Int. System*, vol. 3, no. 1, pp. 87-92, 2003.
- [20] L. M. Brown, Y. L. Tian, "Comparative study of coarse head pose estimation," *IEEE Workshop on Motion and Video Computing*, pp. 125-130, 2002.
- [21] Y. O. Kim, J. k. Paik, J. G. Heo, A. Koschan, B. Abidi and M. Abidi, "Automatic Face Region Tracking for Highly Accurate Face Recognition in Unconstrained Environments," *IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 29-36, 2003.
- [22] C. W. Park, E. Kim and M. Park, "Human Face Detection via Characterized Convex Regional Relationship in Color Images," *IEICE Trans. INF. & SYST*, vol. E86-D, no. 4, pp. 759-762, 2003.
- [23] T. G. Lee, S. K. Park and M. S. Kim, "Face Detection and Recognition with Multiple Appearance Models for Mobile Robot Application," *Proceedings ICCAS*, pp. 215-218, 2002.

— 저 자 소 개 —



**장 승 호**(학생회원)  
 2003년 2월 중앙대학교 전자전기  
 공학부 학사졸업  
 2003년 3월~2003년 6월  
 삼성종합기술원  
 2003년 9월~현재 중앙대학교  
 영상공학과 석사과정  
 2004년 5월~현재 전자부품연구원 정밀기기연구  
 센터 지능형 로봇사업단 위촉연구원  
 <주관심분야: 얼굴인식, 비디오 감시 시스템, 객  
 체 기반 압축, 컴퓨터 비전>



**박 창 우**(정회원)  
 1997년 고려대학교 전자공학과  
 학사졸업  
 1999년 연세대학교 전자공학과  
 석사졸업  
 2003년 연세대학교 전자공학과  
 박사졸업  
 2003년~현재 전자부품연구원 정밀기기연구센터  
 선임연구원  
 <주관심분야: 지능제어, 로보틱스, 컴퓨터비전>



**남공 재 찬**(정회원)  
 1970년 2월 인하대학교  
 전기공학과 학사졸업  
 1976년 8월 인하대학교 대학원  
 전자공학과 석사졸업  
 1982년 2월 인하대학교 대학원  
 전자공학과 박사졸업  
 1979년 3월~현재 광운대학교 컴퓨터공학과 교수  
 <주관심분야: 신경망, 영상처리, 문자인식>



**김 영 옥**(정회원)  
 1995년 2월 홍익대학교  
 전기공학과 학사졸업  
 1997년 2월 홍익대학교 대학원  
 전기 공학과 석사졸업  
 1997년 7월~현재 전자부품연구원  
 정밀기기연구센터지능형  
 로봇사업단 선임연구원  
 2003년 3월~현재 중앙대학교 영상공학과 박사과정  
 <주관심분야: 2D/3D 영상해석 및 영상인식, HCI,  
 지능형로봇비전>



**박 장 한**(정회원)  
 1997년 2월 방송통신대학교 전산  
 계산학과(독학학위) 학사  
 졸업  
 1999년 8월 광운대학교 전산대학원  
 소프트웨어공학과 석사졸  
 업  
 2001년 3월~현재 광운대학교  
 대학원 컴퓨터공학과  
 박사과정  
 <주관심분야: 스테레오비전, 3D영상, 영상처리, 영  
 상인식>



**백 준 기**(정회원)  
 1984년 서울대학교 제어계측  
 공학과 학사졸업  
 1987년 노스웨스턴대학교 전기 및  
 컴퓨터 공학과 석사졸업  
 1990년 노스웨스턴대학교 전기 및  
 컴퓨터 공학과 박사졸업  
 1993년~현재 중앙대학교 첨단영상대학원  
 영상공학과 교수  
 <주관심분야: 영상복원, 신호처리, 반도체>