

# Affine 변환과 템플릿 정합을 이용한 크기 및 회전 변화 얼굴 인식

정회원 김 문 갑\*

## Recognition of Face with Variable Scale and Rotation Using Affine Transformation and Template Matching

Mun-Gab Kim\* *Regular Member*

요 약

이 논문에서는 affine 변환과 템플릿(template) 정합(matching)을 이용한 얼굴 인식 방법을 제안한다. 얼굴 인식에 사용되는 영상들은 크기와 회전 변화가 존재하는 정면 (frontal view) 얼굴 영상들이다. 회전 및 크기 변화가 존재하는 얼굴 영상을 일정한 크기와 각도로 정규화하기 위해 사용된 affine 변환은 양 눈동자의 중심 좌표를 기준으로 역 방향으로 하였다. 얼굴 인식 시스템 성능을 비교하기 위해 다양한 템플릿 형태와 종류를 사용하였다. 즉, 정규화가 되어진 얼굴 영상에서 눈 부분만을 추출한 템플릿, 눈과 코를 포함한 템플릿, 그리고 타원형 얼굴 템플릿으로 구분하였으며, 템플릿 종류는 개인마다 한 장의 영상으로부터 생성시킨 것과 크기와 회전 변화가 존재하는 다수의 영상으로 생성시킨 평균 템플릿으로 구분하여 실험하였다. 또한, 얼굴의 고유 특징은 그레이 레벨과 그래디언트 영상을 사용하였다. 이 논문에서 사용한 실험 영상에 대해서는 그레이레벨 타원형 평균 템플릿이 좋은 인식률을 나타내었다.

### ABSTRACT

The face recognition method is proposed, which is based affine transformation and template matching. The database is composed of frontal view faces with scale and in-plane rotation variations. Eye points are used to normalize the faces which had variable scale and rotation angle to constant scale and rotation angle by affine transformation. The shape of templates are classified into three models, ellipse of face, rectangle which included the region of eyes, and another rectangle which included the region of eyes and nose. Also, the type of templates are classified into single and average one. A single template is formed by one image which is selected a person face images, and a average template is formed three images with variable scale and rotation angle. Also, the features of face image are used intensity and gradient. To increase matching accuracy, average template used shows the good recognition performance for database which is used in this paper.

### I. 서 론

과학 기술의 발달로 지문, 홍채, 얼굴, 음성 등과 같은 인간의 생물학적 특성을 이용한 생체 측정학 (biometrics)에 대한 연구가 지난 수 십 년간에 활발히 진행되고 있다<sup>[1]</sup>. 특히, 인간의 얼굴을 인식하

는 연구가 컴퓨터비전, 영상처리, 형태인식 등의 분야에서 진행되고 있으며<sup>[2][3]</sup>, 이는 신원 확인, 보안 시스템 등 다양한 분야에서 사용되고 있다.

얼굴 인식은 주어진 정지 또는 동영상에 포함되어 있는 하나 또는 다수의 얼굴들을 등록된 데이터 베이스 내 각각의 얼굴들 사이의 유사성을 측정하

\* 용인송담대학 디지털전자정보과(kimmg@dragon.ysc.ac.kr)  
논문번호: T01005-0326, 접수일자: 2001년 3월 26일

여 그 얼굴들을 식별(identification)하는 것이다. 인간의 얼굴을 식별하는 문제는 인식(recognition)과 확인(verification)의 문제로 분류된다. 얼굴 인식은 주어진 하나의 특징인 얼굴과 등록된 각각의 얼굴들 사이의 계산된 유사도에서 가장 높은 하나의 등록된 얼굴을 최종적인 얼굴 인식의 결과로 출력한다. 그리고 얼굴 확인은 얼굴 인식과정에서 복구된 하나의 등록된 얼굴이 입력된 특징인의 얼굴과 일치하는지를 결정하는 것이다. 이러한 얼굴 인식 시스템<sup>[4][5]</sup>는 단순한 배경과 적절한 조명 조건에서 획득한 영상이라도 입력 장치와 얼굴 인식 대상사이의 거리, 모델의 자세에 따라 얼굴의 크기와 회전 그리고 측정 위치의 이동의 변화가 발생하여 인식하기 어려운 문제점이 있으므로 이를 해결하기 위해 얼굴의 크기와 회전 변화가 발생하는 영상에 대한 얼굴 검출 및 인식 연구가 진행되고 있다. 얼굴의 크기 변화에 대응하기 위해 multitemplate와 multiresolution 기법을 채택하는 템플릿 정합<sup>[4]</sup>, view-based eigenspace를 이용한 방법<sup>[5]</sup> 등 여러 가지가 있는데, 이들은 많은 계산이 필요로 한다. 또한 이 논문에서 사용한 방법으로 사전에 등록된 얼굴의 크기와 회전에 입력 얼굴 영상을 정규화하여 정합시키는 것이다. 얼굴의 크기와 회전 변화를 정규화시키기 위하여 얼굴 구성 요소의 기하학적 위치를 기반으로 affine 변환<sup>[4]</sup> 또는 warping 기법<sup>[5]</sup>이 사용된다. 이들 방법은 기하학적인 왜곡을 수정하는데 사용될 뿐만 아니라, 영상들을 일치시키는데 사용되는 간단한 방법으로 많이 사용되고 있지만 크기와 회전 변화에 불변인 다른 변환 방법에 비해 계산 량의 부담이 있다. Warping 변환은 비선형 기하학적 연산 방법으로 영상에 왜곡을 가져올 수 있지만, affine 변환은 처리된 영상에 임의의 굴곡을 가져오지 않는 변환이므로 이 논문에서는 affine 변환을 이용한다.

이 논문에서는 크기와 회전 변화된 정면 영상들에 대해 다양한 템플릿과 특징들을 이용하여 affine 변환에 의한 얼굴 인식 시스템에 대해 언급한다. 실험은 얼굴 영상에서 양 눈동자의 중심 좌표가 자동적(auto) 또는 인위적(manual)으로 지정되었다는 가정 하에 이루어진다. 이러한 양 눈동자의 위치를 가지고 다양한 템플릿을 형성하였으며, affine 변환을 하기 위한 기하학적 기준 점으로 사용하였다. 영상의 크기 및 각도를 정규화하는 과정에 사용되는 내·외사법은 단순한 최근접 이웃 보간법(nearest neighbor interpolation)을 적용하였다.

이 논문의 구성은 다음과 같다. II장에서는 affine 변환 방법에 대해 설명한다. III장에서는 얼굴의 타원형 모델링<sup>[6][7]</sup> 방법을 설명한다. IV장에서는 얼굴 영상에서 고유 특징 추출과 다양한 템플릿 형상에 대해 설명한다. V장에서는 데이터베이스 구축 방법, 인식 성능 평가, 실험 및 결과를 고찰한다. VI장에서는 결론 및 추후 연구에 대해 기술한다.

## II. Affine 변환

얼굴 영상의 크기와 회전 변화가 다른 두 얼굴 영상을 정규화 시키기 위해서는 식(1)과 같은 affine 변환이 이용된다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} d\cos\theta & d\sin\theta \\ -d\sin\theta & d\cos\theta \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (1)$$

여기서,  $(x_c, y_c)$ 는 얼굴 영상을 회전시키기 위한 영상의 기준좌표로 영상 크기의 중심점이 이용된다.  $(x_n, y_n)$ 은 얼굴 영상의 크기와 회전 변화가 정규화 되는 한 점이다. 그리고  $d$ 는 크기를 보상하기 위한 상수이다. 이 값은 주어진 영상에서 양 눈동자들 사이의 거리들로부터 계산된다. 또한  $\theta$ 는 입력 얼굴 영상과 템플릿 얼굴 영상에서 두 눈동자 사이의 각도의 차이를 나타낸다.

식(1)은 입력 영상으로부터 출력 영상으로의 affine 변환은 순방향(forward mapping)이다.

이 순방향 변환은 영상이 변환될 때, 어떤 위치는 화소가 사상(mapping)되지 않는 홀(hole) 문제가 발생되어, 불필요한 계산이 여러 번 되풀이 될 수 있다. 이 논문에서는 이러한 문제 점을 해결하기 위해 역 방향 변환(reverse mapping)을 사용하였다. 역 방향 affine 변환에 대한 식은 (2)와 같다. 이 변환 방법은 입력 화소의 밝기는 새로운 출력 화소의 위치로 재배치하는 것이다. 즉, 각각의 출력 화소의 위치를 하나씩 처리하여 그것이 변환된 입력 화

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} d\cos\theta & -d\sin\theta \\ d\sin\theta & d\cos\theta \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (2)$$

소의 밝기를 채워나가는 것으로, 순방향 변환에서의 홀 문제가 해결된다. 그림 1은 크기와 회전각이 다른 얼굴 영상들의 순방향 및 역 방향 변환을 이용하여 두 눈동자 사이의 거리가 80화소, 회전각이 0도로 정규화 된 영상들의 예를 보여준다. 그림 (a)에 나타난 얼굴 영상에서  $d$ 와  $\theta$ 는 각각 57화소와

-4도이고, 그림 (d)은 거리가 100화소이며 회전각이 14도인 영상이다. 그림 (b)과 (e)는 (a), (d)영상에 대한 순방향 변환으로 정규화 된 영상이고 (c), (f)는 역 방향 변환된 영상이다.

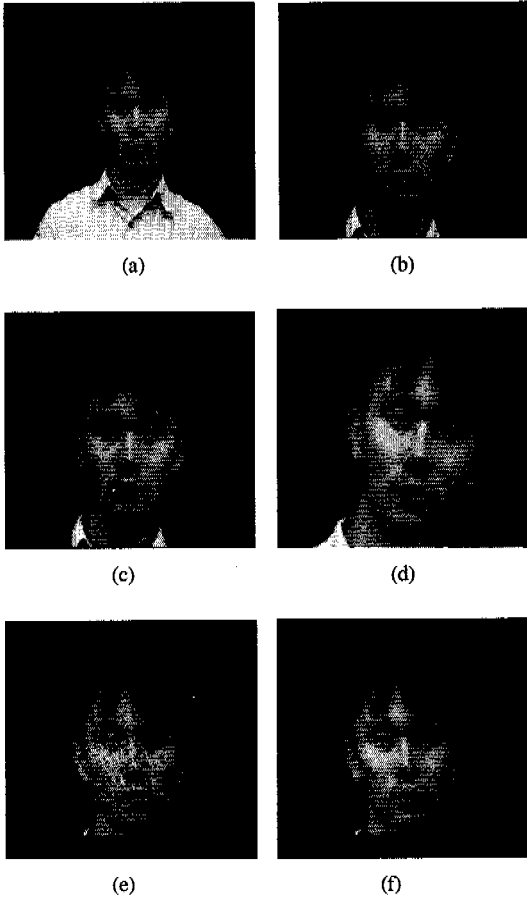


그림 1. 순방향 및 역방향 affine 변환으로 얼굴의 크기와 회전각이 정규화 된 영상의 예.

- (a), (d) 해상도가 512×480인 입력 영상.
- (b), (e) 순방향 affine 변환으로 정규화 된 영상.
- (c), (f) 역방향 affine 변환으로 정규화 된 영상.

### III. 얼굴의 타원형 모델링

입력된 영상에서 양 눈동자의 중심 좌표가 검출되면, 얼굴 영역은 추출된다. 또한 기하학적인 정보를 이용하여 눈을 제외한 나머지 얼굴 구성 요소들의 위치들도 결정된다. 하나의 타원은 그림 2에 나타낸 것과 같이 양 눈동자의 중심 위치 a의 좌표  $(x_l, y_l)$ 와 b의 좌표  $(x_r, y_r)$ 를 근간으로 하는 타원형 얼굴 모델을 적용하여 정면 얼굴에서 가장

두드러진 양 눈과 눈썹, 코 그리고 입 영역이 포함되는 타원형의 얼굴 모델은 식(3)와 같이 결정된다.

$$\frac{((x - x_e)\cos\theta + (y - y_e)\sin\theta)^2}{A^2} + \frac{(-(x - x_e)\sin\theta + (y - y_e)\cos\theta)^2}{B^2} = 1 \quad (3)$$

양 눈동자의 중심 a 및 b와 정삼각형을 형성하는 d의 좌표  $(x_n, y_n)$ 는 코끝의 중심 위치가 된다. 타원의 중심 위치 c는 a, b 그리고 d의 무게 중심의 좌표  $(x_e, y_e)$ 이다. A와 B는

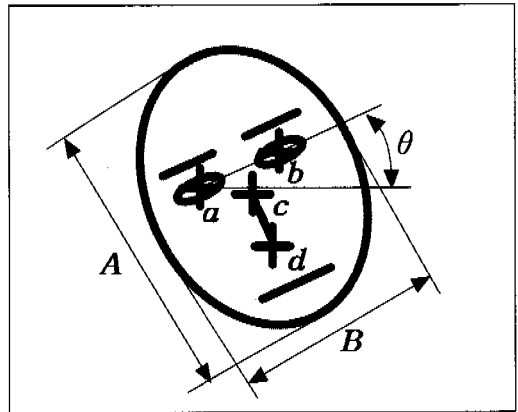


그림 2. 타원형 얼굴 모델

각각 타원의 장축과 단축의 길이로 양 눈동자 사이의 거리를 이용하여 결정한다. 그리고 왼쪽 눈동자를 기준으로 오른쪽 눈동자에 대한 회전각도는  $\theta$ 로 나타낸다. 양 눈동자 사이의 거리 d와 회전각도  $\theta$ 를 이용하여 타원의 얼굴 모델에 필요한 매개변수는 식 (3)에 나타낸 5개의 매개변수들을 결정하기 위하여 왼쪽 눈동자의 좌표와 오른쪽 눈동자의 중심 좌표가 이용된다. 양 눈동자 중심들 사이의 거리는  $d = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2}$ 로 표현된다. 그리고 왼쪽 눈동자의 중심으로부터 수평선을 기준으로 오른쪽 눈동자의 중심이 이루는 각은

$$\theta = \tan^{-1} \left( \frac{y_r - y_l}{x_r - x_l} \right)$$

로 표현된다. 얼굴 영상에서 얼굴의 장축 길이  $A = c_l d$ , 단축 길이  $B = c_s d$ 가 된다. 여기서 양의 상수인  $c_s = \sqrt{\frac{c_l^2}{2} - c_f^2}$ 이다. 이 경우에  $c_l$ 과  $c_f$ 들은 CCD 카메라로부터 영상을 획득하는 경우에 상용으로 판매되는 프레임 그레버의 증횡 비의 특성을 반영하는 양의 상수이다. 얼

굴 영상에서 왼쪽 눈동자의 중심 위치와 오른쪽 눈동자의 중심 위치와 코끝의 중심 위치는 각 변의 길이가  $d$ 인 정삼각형의 형태를 가진다. 따라서 코끝의 중심 좌표는 다음과 같이 결정된다.

$$x_n = \begin{cases} x_c + \frac{\sqrt{3}}{2} d \sin \theta, & \text{if } \theta > 0 \\ x_c - \frac{\sqrt{3}}{2} d \sin \theta, & \text{otherwise} \end{cases} \quad (4)$$

$$y_n = y_c + \frac{\sqrt{3}}{2} d \cos \theta$$

여기서  $(x_c, y_c)$ 는 양 눈동자들의 무게중심의 위치이다. 식 (3)에서 나타난 타원의 중심은 왼쪽과 오른쪽 눈동자의 중심들과 코끝의 중심으로부터 계산되는 무게중심의 위치로 결정된다. 먼저 타원의 중심 위치  $c$ 를 중심으로  $1.7 \times d$  반경이내에 포함되는 영상에 대해서 콘트라스트(contrast)를 개선하기 위하여 국부적인 히스토그램 평활화(histogram equalization)를 수행한다. 그리고 추출된 타원형의 얼굴에서 눈썹 위쪽 영역(양 눈동자의 중심 위치로부터  $d/2$  위쪽 영역)은 모델들의 머리카락 형태에 따라 변화가 다양하게 발생하므로 템플릿 정합의 정확도를 증가시키기 위하여 제거한다. 이와 같은 방법으로 추출된 타원형의 얼굴 영상을 그림 3에 나타내었다.

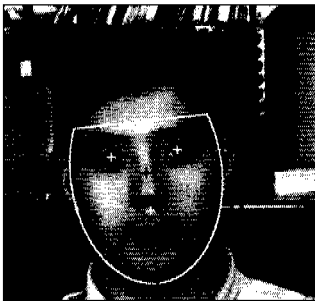


그림 3. 얼굴 영상에 타원 모델을 적용한 예.

#### IV. 얼굴 특징 추출 및 템플릿 형상

##### 1. 얼굴 특징 추출

일반적으로 전체 기반 방법에 의한 얼굴의 특징은 밝기 기반과 그래디언트 기반의 두 가지 방법으로 분류된다. 밝기 기반으로 추출된 얼굴 영상과 동일한 방법으로 다른 조명에서 획득된 입력 영상으로부터 추출되는 얼굴 특징 사이의 유사도가 감소된다고 알려져 있지만<sup>[8]</sup>, 이 논문에서 사용한 얼굴 영상들이 비교적 조명 변화가 적어 두 방법 모두

사용한다. 또한 현재까지 알려져 있는 필터 또는 연산자들로부터 방향성을 가지는 조명 변화에 무관한 특징이 추출되기는 매우 어렵다. 방향성을 가지지 않는 조명의 변화에 대해서는 비교적 강인한 결과를 나타내는 그래디언트의 크기는  $3 \times 3$  Sobel 연산자로부터 추출된다. 그레이 레벨 영상과 그래디언트 영상은 콘트라스트를 개선하기 위하여 얼굴 영상 형상내의 국부적인 히스토그램 평활화를 수행하였다.

##### 2. 다양한 템플릿 형상

이 논문에서 사용한 템플릿 형상은 3가지로 구분된다. 첫 번째 형상은 III장에서 언급한 방법에 의해 구해진 얼굴의 타원형이다. 이는 얼굴 영역에서 고유의 특징을 전 영역에서 추출하는 것이다. 일반적으로 등록된 얼굴 영상이 같은 인종과 성별인 경우는 유사한 얼굴 형상을 가지고 있어 부분 특징을 이용하여 얼굴을 식별하기 어려운 문제점이 있다. 이를 보완하는 방법으로 여러 부분의 부분 특징을 이용하거나 전체 특징을 이용하여 인식률을 높이는 방법이 있다. 두 번째 형상은 얼굴 영상에서 양 눈동자 주위에 대해 일정한 크기를 가지는 것이다. 이 논문에서는 일정한 크기 및 각도로 정규화 된 영상에서 왼쪽 눈동자 좌표를 기점으로 왼쪽 영역으로  $d/1.5$ , 오른쪽 눈동자 좌표를 기점으로 오른쪽영역으로  $d/1.5$ , 양 눈동자의 중심 좌표를 기점으로 위쪽 영역으로  $d/1.5$ , 아래쪽 영역으로  $d/2.0$ 의 크기를 가지는 직사각형 형상을 사용하였다. 여기서 사용한  $d$ 는 정규화 된 양 눈동자 사이의 거리이다. 이 방법은 눈과 눈썹의 모양, 눈과 눈썹 사이의 거리등이 특징으로 나타난다. 그러나 양 눈동자 사이의 거리를 모든 영상에 대해 일정한 거리로 정규화를 하여 양 눈동자사이의 폭이 특징이 될 수 없는 단점이 있다. 세 번째 형상은 양 눈동자 주위와 코를 포함하는 부분 형상이다. 정규화 된 영상에서 눈과 코를 포함하는 일정한 크기의 영상에서는 두 번째 방법에 포함된 특징들과 각 개인마다 눈과 코 사이의 거리, 코의 모양 등에서 차이점이 나타나는 특징을 이용한 것이다. 템플릿 형상은 두 번째 형상의 아래쪽 영역으로 크기가  $d/2.0$ 인 것을  $d$ 로 하였다.

##### 3. 평균 템플릿 얼굴 영상

이 논문에서는 얼굴 인식 시스템의 성능을 향상시키는 방법으로 다수의 영상들로부터 비교적 간단하고 실시간 처리에 적합한 합성 평균 템플릿을 생성하였다. 이것은 개인의 다양한 크기와 회전각도를

가지는 얼굴 영상들을 정규화시킨 후 산술적으로 더한 결과에 대해 평균을 취하는 것이다.

$$I_{avg} = \frac{I_{sng1} + I_{sng2} + I_{sng3}}{3},$$

$$G_{avg} = \frac{G_{sng1} + G_{sng2} + G_{sng3}}{3} \quad (5)$$

여기서  $I_{sng}$ ,  $I_{avg}$ 는 개인 얼굴 영상에서 한 장의 그레이 레벨 영상으로부터 생성된 템플릿, 3장의 그레이 레벨 영상들로부터 생성된 평균 템플릿 영상이다.  $G_{sng}$ ,  $G_{avg}$ 는 그래디언트 영상에 대한 것이다. 그림 4는 그림 1의 영상에 대한 다양한 템플릿 형상과 종류의 예이다. 단일 템플릿에 대해서는 그림 1의 (c)영상을 사용하였다. 여기서 그래디언트 영상은 밝기 레벨을 역으로 하였다.

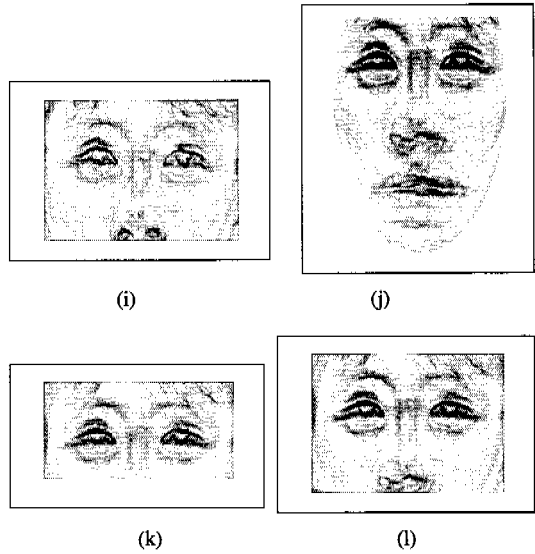
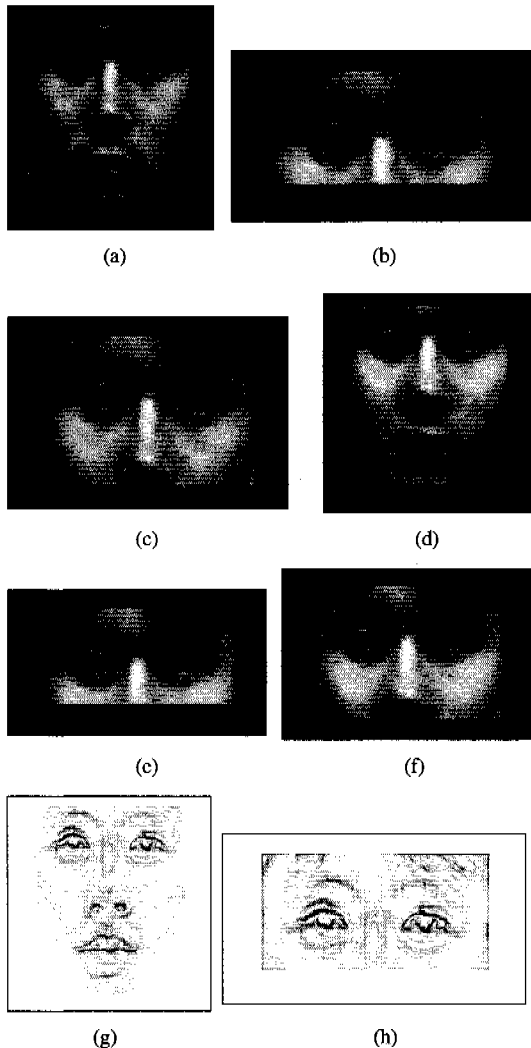


그림 4. 정규화 된 얼굴 영상의 다양한 템플릿 형상과 종류의 예

- (a) 단일 그레이 레벨 타원형 템플릿
- (b) 눈 주위만을 추출한 단일 그레이 레벨 템플릿
- (c) 눈과 코를 포함하는 단일 그레이 레벨 템플릿
- (d) 평균 그레이 레벨 타원형 템플릿
- (e) 눈 주위만을 추출한 평균 그레이 레벨 템플릿
- (f) 눈과 코를 포함하는 평균 그레이 레벨 템플릿
- (g) 단일 그래디언트 타원형 템플릿
- (h) 눈 주위만을 추출한 단일 그래디언트 템플릿
- (i) 눈과 코를 포함하는 단일 그래디언트 템플릿
- (j) 평균 그래디언트 타원형 템플릿
- (k) 눈 주위만을 추출한 평균 그래디언트 템플릿
- (l) 눈과 코를 포함하는 평균 그래디언트 템플릿



#### IV. 실험 및 결과

실험에서는 시스템 처리 시간을 단축하기 위하여 등록된 템플릿들에 대해 양 눈동자의 사이의 거리를 80 화소, 회전 각도를 0도로 정규화를 하였으며, 등록된 다양한 형상의 템플릿과 종류에 해당하는 각각의 얼굴 영상에 평균과 분산을 함께 저장하여 정합과정의 시간을 단축하였다. 입력되는 얼굴 영상에서 양 눈동자 사이의 거리가 80 화소보다 크거나 작을 경우, 거리를 정규화하기 위해 내·외삽 보간법이 필요한데, 정밀한 보간법은 계산량이 많기 때문에 단순한 최근접 이웃 보간법(nearest neighbor interpolation)을 사용하였다. 실험은 Celeron 500MHz 프로세서를 탑재한 개인용 컴퓨터에 Visual C++ 언어로 이루어졌다. 한 장의 얼굴 영상이 입력되어 정규화 된 후 그래디언트 영상을 구해 53명의 템플릿과 정합을 통하여 인식되는 데 약 2.0초가 소요되었다. 그리고 프레임 그래버의 종횡 비의 특성을 반영하는 양의 상수는  $c_f=3.7$ ,  $c_r=1.5$ 를 사용하였다.

1. 얼굴 영상 데이터베이스

CCD 카메라를 이용하여 얼굴의 크기와 회전 변화를 가지는 다양한 얼굴 영상들로 데이터베이스는 구축된다. 실험은 남자 22명과 여자 31명으로 구성되는 53명의 모델로부터 5단계의 크기 변화를 가지며 회전 변화는 크기 변화의 각 단계에서 좌우로 회전된 3장의 영상들로 총 795장의 얼굴 영상을 획득하여 구성하였다. 획득된 영상들은 비교적 일정한 조명 환경에서 크기와 회전 변화가 강조되었다. 실험 대상이 되는 얼굴 영상들의 크기 및 회전각도 변화는  $40 \leq d \leq 120$ ,  $-30^\circ \leq \theta \leq +30^\circ$  이다. 각 모델로부터 획득한 얼굴 영상에서 양 눈동자의 거리가 80화소에 가깝고, 회전 각도가  $\pm 5$ 도 이내의 영상 중에서 하나를 단일 템플릿으로 이용하였고, 단일 템플릿과 양 눈동자 사이의 거리가 기준보다 짧고 회전 각도가  $+10 \sim +20$ 도 이내의 영상, 거리가 기준보다 길고 회전 각도가  $-10 \sim -20$ 도 이내인 영상을 합성 평균 템플릿을 생성하기 위하여 이용하였으며, 나머지 12장은 인식기의 성능을 확인하기 위하여 사용되었다. 그림 5은 실험에 포함되는 얼굴 영상들의 예이다.

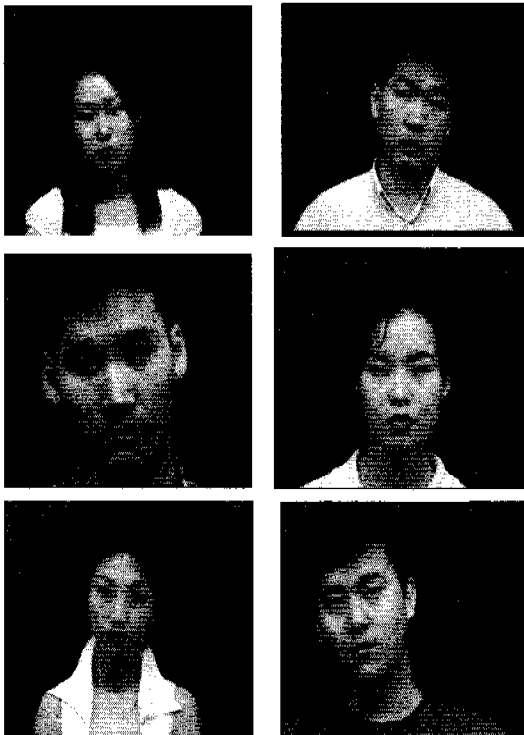


그림 5. 실험에 포함되는 얼굴 영상들의 예

2. 얼굴 인식 성능 평가

실험에서 사용한 유사도 계산은 NCC(normalized cross-correlation)에 의해 이루어진다.

$$NCC = \frac{\langle f(u, v)g(u, v) \rangle - \langle f(u, v) \rangle \langle g(u, v) \rangle}{\sigma_f \sigma_g} \quad (6)$$

여기서  $f(\cdot)$ 는 입력 영상을 나타내며  $g(\cdot)$ 는 템플릿을 나타낸다. 그리고  $\langle \cdot \rangle$ 는 평균을 계산하는 연산자이며  $\sigma_f$ 는  $f(\cdot)$ 의 표준편차를,  $\sigma_g$ 는  $g(\cdot)$ 의 표준편차를 나타낸다. 실험에서는 양 눈동자 위치가 자동적 또는 인위적으로 지정하였다는 가정 하에 이루어진다. 양 눈동자의 위치가 정확하게 지정되지 않았을 경우는 NCC를 이용한 템플릿과 입력 영상과의 유사도 계산에는 오차가 발생할 수 있다. 이러한 문제점을 보완하는 방법으로 NCC 계산과정에서 입력 영상에서 추출된 얼굴 형태는 고정시켜 놓은 상태에서 각 템플릿의 양 눈동자 중심 좌표를  $\pm 2$  화소 이내에서 이동(dithering)하면서 계산된 NCC에서 최고 값을 유사도로 선정하였다.

$$NCC = \max_{\Delta x, \Delta y} \frac{\langle f(x, y)g(x \pm \Delta x, y \pm \Delta y) \rangle - \langle f(x, y) \rangle \langle g(x, y) \rangle}{\sigma_f \sigma_g} \quad (7)$$

다양한 템플릿 형상과 종류 그리고 유사도를 측정하는 방법에 따른 실험 결과는 표 1과 2에 나타내었는데, 표 1은 식 (6)을 이용한 결과이며, 표 2는 식(7)을 이용한 결과이다. 표에서 Top1은 입력 얼굴과 등록된 얼굴 사이의 유사도가 최고 값에 대한 인식률이며, Top3은 입력 얼굴과 유사도가 최상위에 등록되는 3개의 얼굴에 포함되는 경우의 인식률이다. 표 1에서 나타낸 것과 같이 affine 변환과 템플릿 정합을 이용한 얼굴 인식기는 템플릿의 형상과 종류에 따라 다른 결과를 가져왔다. 인식기에서 Top1 인식률은 그래디언트 영상에서는 눈과 코 템플릿 형상이 타원형 템플릿과 비교해 좋은 결과를 나타내었다. 이는 얼굴 전체 영역을 사용하는 타원형 모델에서는 affine 변환이 양 눈동자의 중심 좌표를 기준으로 이루어지므로 눈동자와 거리가 먼 입과 턱 부분에서는 영상이 왜곡되는 현상이 발생되고, 영상 획득 과정에서 모델들에 대해 정확한 정면 얼굴 영상을 획득하지 못한 원인으로 인하여 인식률이 떨어지게 되었다. 그리고 그레이 레벨 영상에서는 타원형 모델에서 좋은 인식률을 보였다. 이

는 그라디언트 영상은 영상의 밝기 변화가 일어나는 부분에서 정보가 존재하므로, 이 논문에서 사용한 조명 변화가 적은 곳에서는 그라디언트 정보가 적다. 이와 반대로 얼굴 영상들의 조명 변화가 적어 그레이 레벨 영상에서는 많은 정보가 존재하므로 위에서 언급한 타원형 모델의 단점을 보완할 수 있다. 표 2는 각 입력 영상에 대해 템플릿을 눈동자 중심 좌표를 기준으로  $\pm 2$ 화소 범위에서 이동시켜 상관계수를 구한 것으로 표 1과 비교하여 부분 템플릿 경우는 대체적으로 약간의 높은 인식을 보였지만, 전체 영역인 타원형 템플릿에서는 인식률의 차이가 거의 없다. 또한, 인식 결과는 한 장의 템플릿 영상을 사용하는 것보다 평균 템플릿을 사용하는 것이 좋은 결과를 나타내었다. 이는 합성 평균 영상이 약간의 affine 변환에 의한 영상 왜곡, 눈동자 위치의 오차 및 모델들의 자세에 대한 문제를 보상할 수 있기 때문이다.

표 1. 다양한 템플릿 형상과 종류에 대한 얼굴 인식 결과 단위(%)

		$G_{eng}$	$G_{avg}$	$I_{eng}$	$I_{avg}$
타원	TOP1	76.88	93.87	86.96	96.70
	TOP3	84.59	95.76	90.88	97.80
눈	TOP1	76.73	92.77	74.06	90.57
	TOP3	83.18	94.97	82.70	92.93
눈과 코	TOP1	77.83	94.34	76.42	93.15
	TOP3	83.18	96.23	85.22	95.60

표 2. 다양한 템플릿 형상과 종류에 디더링을 행한 인식 결과 단위(%)

		$G_{eng}$	$G_{avg}$	$I_{eng}$	$I_{avg}$
타원	TOP1	77.36	93.87	86.95	96.70
	TOP3	84.75	95.60	91.04	97.80
눈	TOP1	84.60	96.38	78.93	94.81
	TOP3	88.84	98.11	86.16	95.44
눈과 코	TOP1	83.81	96.54	79.72	96.23
	TOP3	88.52	98.43	87.58	97.17

## VI. 결론 및 추후 연구

얼굴의 크기와 회전 변화가 있는 정면 영상에 대하여 affine 변환과 템플릿 정합을 이용한 얼굴 인식기에 다양한 템플릿 형상과 종류에 대한 성능 비교를 하였다. Affine 변환을 이용하므로 얼굴의 크기와 회전 변화에 대응하기 위하여 계산량이 증대되는 다수의 템플릿을 이용하는 방법과 multi-resolution 기법을 적용할 필요성이 없게 된다. 얼굴의 크기와 회전변화를 가지는 입력 얼굴 영상과 등록된 얼굴 영상을 비교해보면 그 두 영상 사이에서 국부적으로 밝기와 형태의 왜곡이 종종 발생되어, 상관 정합 방식 인식기의 성능을 저하시키게 된다. 이 논문에서는 이러한 요인들에 기인한 정합 오차를 줄이기 위하여 다수의 영상으로부터 생성되는 평균 템플릿을 이용하였다. 평균 템플릿은 얼굴의 크기와 회전 변화가 서로 다른 영상들로 구성되었다. 따라서 생성된 템플릿은 국부적인 왜곡에 대한 특성들이 암시적으로 반영되어 있으므로 비록 입력 영상에서 국부적인 왜곡이 발생하더라도 그 두 영상 사이에서 계산된 정합 오차는 감소된다.

추후 연구 과제는 affine 변환에 의해 발생하는 왜곡 현상을 제거하는 방법, 입력 영상이 정확한 정면 영상이 아닐 경우 인식률을 높이기 위한 새로운 유사도 측정 방법, 그리고 입력 얼굴 영상에서 자동적으로 눈동자 중심 좌표를 찾는 방법 등이 필요하다.

## 참고 문헌

- [1] S. Pankanti, R.M. Bolle, and A. Jain, *Biometrics: Personal identification in networked society*, Kluwer Academic, Norwell, Mass., 1999.
- [2] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. of The IEEE*, vol. 83, no. 5, 1995.
- [3] A. Samal and P.A. Iyengar, "Automatic recognition and analysis of human faces and facial expression: A survey," *Pattern Recognition*, vol. 25, no. 1, pp.65-77, 1992.
- [4] R. Brunelli and T. Poggio, "Face detection: Features versus templates," *IEEE Trans.*

- Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, 1993.
- [5] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 696-710, 1997.
- [6] H. Nugroho, S. Takahashi, Y. Ooi, and S. Ozawa, "Detecting human face from monocular image sequences by genetic algorithms," *Proceedings of The IEEE*, vol. 4, pp. 2533-2536, 1997.
- [7] 최 일, 진성일, "Log-Polar 사상의 크기와 회전 불변 특성을 이용한 얼굴과 눈 검출," *전자공학 회논문지 제 36권 5편 제 8호*, pp. 968-980, 1999
- [8] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 721-732, 1997.

김 문 갑(Mun-Gab Kim)

정회원



1990년 2월 : 경북대학교

전자공학과 졸업

1992년 2월 : 경북대학교 대학원

전자공학과 졸업

2000년 2월 : 경북대학교 대학원

전자공학과

박사과정수료

1992년 2월~2000년 2월 : 대우중공업(주)

2000년 3월~현재 : 용인송담대학 디지털전자정보과

전임강사

<주관심 분야> 영상신호처리, 컴퓨터비전