

論文99-36S-8-10

# Log-Polar 사상의 크기와 회전 불변 특성을 이용한 얼굴과 눈 검출

## (Automatic Face and Eyes Detection: A Scale and Rotation Invariant Approach based on Log-Polar Mapping)

崔 一 \* , 秦 成 一 \*

(Il Choi and Sung-Il Chien)

## 요 약

얼굴과 얼굴 구성 요소를 자동으로 검출하는 것은 얼굴 인식 시스템에서 반드시 필요한 과정이다. 본 논문에서는 얼굴의 크기와 회전 변화가 동시에 존재하는 영상에서 log-polar 사상(mapping)에 기반한 단일 log-polar 얼굴 템플릿(template)을 사용하는 밝기 기반 템플릿 정합(matching)으로 얼굴과 눈을 자동으로 검출하는 새로운 방법을 제안한다. 템플릿 정합에서는 입력 영상의 크기와 회전 변화를 템플릿의 크기와 회전 각도에 대하여 정규화시키는 과정이 필요하다. 그런데 인간 시각 시스템의 space-variant 특성과 유사한 구조를 가지는 log-polar 사상은 입력 영상의 크기와 회전 변화를 log-polar 평면상에서 수평과 수직 순환 이동으로 변화시키는 특성을 갖는다. 이러한 특성을 이용하여 입력 상반신 영상의 fixation point들에서 사상된 후보 log-polar 얼굴들을 log-polar 평면상에서 수평과 수직 이동시켜 템플릿과 정합한다. 따라서 제안한 방법은 얼굴의 크기와 회전 변화에 대응하기 위하여 계산량이 증대되는 multitemplate와 multiresolution 기법들을 적용할 필요성이 없게된다. 이러한 크기와 회전 불변 정합을 이용하여 제안된 방법은 얼굴과 눈을 동시에 검출할 수 있다. 얼굴의 크기와 회전 변화가 존재하는 상반신 영상 795장에 제안한 방법을 적용하여 98.7%의 얼굴과 눈의 검출율을 달성하였다.

## Abstract

Detecting human face and facial landmarks automatically in an image is an essential step to a fully automatic face recognition system. In this paper, we present a new approach to detect automatically face and its eyes of input image with scale and rotation variations of faces by using an intensity based template matching with a single log-polar face template. In a template-based matching it is necessary to normalize the scale changes and rotations of an input image to a template ones. The log-polar mapping which simulates space-variant human visual system converts scale changes and rotations of input image into constant horizontal and cyclic vertical shifts in the output plane. Intelligent use of this property allows us to shift of the candidate log-polar faces mapped at various fixation points of an input image to be matched to a template over the log-polar plane. Thus, the proposed method eliminates the need of adapting multitemplate and multiresolution schemes, which inevitably give rise to intensive computation involved to cope with scale and rotation variations of faces. Through this scale and rotation invariant matching, the developed method can lead to detecting face and its eyes simultaneously. Experimental results on a database of 795 images show over 98% detection rate.

\* 正會員, 慶北大學校 電子電氣 工學部  
(School of Electronic and Electrical Engineering,  
Kyungpook National University)

※ 본 연구는 정보 통신부 초고속망 과제(계약 번호  
98-42)의 부분적인 지원으로 수행되었습니다.  
接受日字:1999年6月9日, 수정완료일:1999年7月26日

## I. 서론

최근 인공지능에 관한 연구는 인간이 가지고 있는 인식 능력을 컴퓨터에 이식하는데 많은 노력을 기울이고 있다. 특히 얼굴 인식<sup>[1-3]</sup> 기술은 영상 처리, 형태 인식 그리고 컴퓨터 비전 전반에 걸쳐 지난 20년간 활발히 연구되고 있는 분야이다. 얼굴 인식 시스템은 신원 확인, 보안 시스템 그리고 man-machine 인터페이스 등에 광범위하게 사용되고 있다. 입력 영상으로부터 얼굴을 검출하는 단계와 검출된 얼굴을 사진에 등록된 템플릿과 비교하여 인식하는 과정으로 얼굴 인식 시스템은 구성된다. 얼굴 검출은 배경으로부터 발생하는 다양한 종류의 형태와 사람의 얼굴을 구분하는 과정을 포함하므로 근본적으로 형태 인식의 문제이다. 단순한 배경과 적절한 조명 조건에서 획득한 영상 일지라도 카메라와 대상 모델 사이의 거리, 모델의 키와 자세에 따라서 얼굴은 크기와 회전 그리고 위치 이동의 변화가 동시에 발생하게 된다. 주어진 영상에서 이러한 변화에 관한 사전 지식을 획득하기가 용이하지 않으므로 컴퓨터를 이용하여 자동으로 얼굴을 검출하는 문제는 매우 어렵다. 그리고 검출된 얼굴을 사진에 등록된 템플릿과 정합하여 얼굴을 인식하는 시스템<sup>[4,5]</sup>은 검출된 얼굴을 템플릿에 대하여 크기와 회전 변화를 정규화시키는 부가적인 과정이 필요하다. 이러한 경우에 필요한 정보는 얼굴 구성 요소의 위치-예를 들어 양 눈의 위치, 양 눈 사이의 거리 그리고 회전 각도-로부터 획득한다. 따라서 자동으로 얼굴을 인식하는 시스템에서는 얼굴과 얼굴 구성 요소를 검출하는 과정이 반드시 필요하게 된다.

얼굴의 크기와 회전 변화가 발생하는 영상에서 얼굴과 얼굴 구성 요소를 검출하기 위하여 최근 활발한 연구가 진행되어 왔다. 얼굴의 크기 변화에 대응하기 위하여 multitemplate와 multiresolution 기법을 채택하는 템플릿 정합<sup>[4]</sup>, view-based eigenspace를 이용한 방법<sup>[5]</sup>, 그리고 space-variant 영상의 한 형태인 retinal 샘플링(sampling)을 통하여 구축한 SOFM(self organizing feature maps) 필터뱅크(filter banks)들을 적용하는 방법<sup>[6]</sup>들은 많은 계산량을 필요로 한다. 그리고 multiresolution 기법을 근간으로 신경망을 적용한 방법<sup>[7]</sup>과 통계적 특성과 신경망을 결합한 방법<sup>[8]</sup>, log-polar 영상에 higher order local autocorrelation 계수를 추출하여 검출

하는 방법<sup>[9]</sup>들은 얼굴 구성 요소를 검출하는 별도의 과정이 필요하다. 또한 이러한 방법들<sup>[4-9]</sup>은 얼굴의 크기와 회전이 동시에 발생하는 영상에서 얼굴과 얼굴 구성 요소를 검출하기는 어렵다. multitemplate와 multiresolution 기법을 이용하지 않는 수평과 수직 투영 기법<sup>[10]</sup>은 모델의 의상, 자세 그리고 배경의 영향에 따라서 검출 성능이 저하되며, deformable 템플릿을 사용하는 방법<sup>[11]</sup>은 입력 영상에서 검출 대상이 되는 물체의 대략적인 초기 영역을 설정하는 과정이 필요하다. 그리고 물체의 대칭성(symmetry)을 이용하는 방법<sup>[12]</sup>은 얼굴의 크기 변화에 대응하여 가우시안(Gaussian) 거리 가중(distance weighting) 함수의 분산을 적절히 결정하기가 어렵다.

본 논문에서는 얼굴 인식 시스템을 근간으로 하는 보안 시스템에 적용하기 위하여, 얼굴의 크기와 회전 변화가 동시에 발생하는 상반신 영상에서 fixation point들로부터 추출한 후보 log-polar 얼굴 영상들을 단일 템플릿과 정합하여 얼굴과 눈을 동시에 검출한다. 먼저 한 사람(남자 또는 여자)의 눈, 코 그리고 입이 포함된 얼굴 영상에서 양 눈의 위치를 기반으로 하는 타원형의 얼굴 모델을 적용하여 템플릿에 사용될 얼굴을 추출한다. 추출된 타원형 얼굴 영상의 다양한 위치에서 fixation point를 선정하여 실험한 결과로부터 최적으로 결정된 양 눈 사이의 중심점에서 log-polar 사상<sup>[13-15]</sup>한 영상을 최종적인 log-polar 얼굴 템플릿으로 사용한다. 다음 단계는 입력 영상에서 템플릿과 정합 대상이 되는 후보 log-polar 얼굴 영상들을 추출한다. 그런데 눈은 얼굴 영상에서 대칭 위치에 쌍으로 존재하며 모델의 표정과 자세 등의 변화가 발생할 경우에도 상대적으로 어두운 밝기를 가지는 두드러진 영역으로 나타난다. 또한 조명 조건과 배경을 적절히 통제할 수 있는 보안 시스템에 적용하는 얼굴 검출 시스템이므로 눈은 영역 분할 방법(segmentation)으로 비교적 용이하게 추출할 수 있다. 영역 분할된 영상에서 눈동자로 간주되는 고립 영역들을 추출하여 눈동자 후보 쌍들을 선택한다. 선택된 후보 쌍들로부터 타원형의 얼굴 모델과 log-polar 사상을 적용하여 후보 log-polar 얼굴 영상들을 추출한다.

마지막으로, 추출된 후보 log-polar 얼굴들과 템플릿을 정합하여 최종적인 얼굴을 검출한다. 그런데 각각 표피상에서 후보 얼굴들과 템플릿에 사용된 얼굴 사이에는 크기와 회전 변화의 차이가 발생한다. 인간

시각 시스템과 유사한 구조를 가지는 log-polar 사상은 직각 좌표상의 물체의 크기와 회전 변화를 log-polar 좌표상에서 각각 수평 이동과 수직 순환 이동으로 변화시키는 특성을 가지고 있다. 이러한 log-polar 사상의 특성과 얼굴의 크기와 회전 변화를 측정할 수 있는 양 눈 사이의 거리와 회전 각도를 이용하여 입력 영상에서 추출한 후보 log-polar 얼굴 영상들을 한 장의 크기와 회전 각도가 고정된 템플릿에 대하여 log-polar 좌표상에서 수평과 수직 이동 차이를 정렬한다. 정렬된 후보 log-polar 영상들은 템플릿에 대하여 크기와 회전 불변 특성을 가지게 되므로 계산량이 증대되는 multitemplate와 multiresolution 기법들을 적용할 필요성이 없게 된다. 정렬된 후보 log-polar 영상들을 단일 템플릿과 정합하여 정합 오차가 최소가 되는 후보 얼굴을 최종적인 얼굴로 결정한다. 양 눈동자 쌍의 위치를 근간으로 타원형의 얼굴 모델을 적용하여 후보 얼굴을 추출하므로 제안한 방법은 얼굴과 눈을 동시에 검출할 수 있게 된다.

본 논문의 구성은 다음과 같다. 얼굴의 크기와 회전 변화에 무관한 특징을 추출할 수 있는 log-polar 사상에 대하여 II장에서 설명한다. III장에서는 얼굴의 크기와 회전 변화에 무관하게 얼굴과 눈을 검출하는 방법에 대하여 상세하게 설명한다. IV장에서는 제안한 방법의 실험 결과를 고찰한다. 마지막으로 V장에서는 결론을 기술한다.

## II. Log-Polar 사상

Log-Polar 사상은 fixation point로부터 거리가 멀어질수록 샘플링 밀도가 감소되는 space-variant 방법의 가장 간단한 형태이며 인간 시각 시스템과 유사한 구조를 가진다. 그림 1에 나타낸 바와 같이 직각 좌표상의 한 점  $(x, y)$ 의 밝기값을 log-polar 좌표상의 한 점  $(u, v)$ 의 밝기값으로 사상한다<sup>[13-15]</sup>.

$$(u, v) = \left( \log \frac{r}{r_0}, \theta \right) \quad (1)$$

여기서  $r$ 은 fixation point로부터  $(x, y)$ 까지의 유클리드 거리,  $r_0$ 는 fixation point에 가장 근접한 반경, 그리고  $\theta$ 는 fixation point와  $(x, y)$ 간의 회전 각도이다. 직각 좌표상에서 물체의 크기가  $a$ 배 변화하면 log-polar 좌표상에서  $u = \log r + \log a - \log r_0$ 가 된

다.  $a > 1$ 이면  $u$ 축의 오른쪽으로 이동하고 그 반대이면 왼쪽으로 이동한다. 그리고 회전 변화는  $v$ 축의 방향으로 수직 순환 이동으로 나타난다. 즉 직각 좌표상에서 물체의 크기와 회전 변화가 발생하면 log-polar 좌표상에서는 각각 수평과 수직 이동의 차이로 변환되어 나타난다. 따라서 직각 좌표상에서 크기와 회전 변화가 다른 두 물체를 정규화시키는 것은 log-polar 좌표상에서 두 영상 사이의 수평과 수직 이동 차이를 정렬시키는 것과 동일하다.

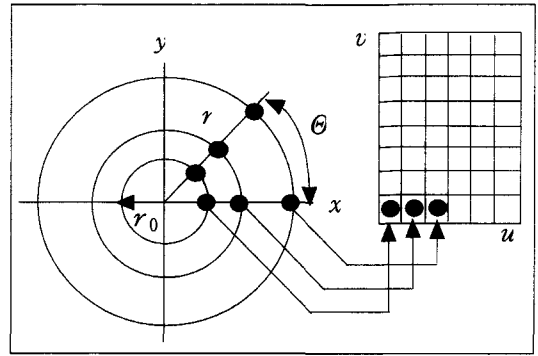


그림 1. Log-Polar 사상

Fig. 1. Log-Polar Mapping.

식 (1)에서  $(x, y)$ 가 주어지고  $(u, v)$ 를 계산하는 forward 사상을 사용하면 log-polar 영상에서는 hole들과 overlap들이 발생한다<sup>[16]</sup>. 본 논문에서는 inverse 사상 방법을 사용하여 hole들과 overlap들이 발생되지 않는 log-polar 영상을 추출한다. 먼저 직각 좌표상에서 log-polar 샘플링의 fixation point를  $(x_c, y_c)$ 로 두자. 식 (1)에서  $r_0$ 의 영향을 고려하기 위하여 log-polar 좌표상에서 수평축 좌표값  $u = u_0, \dots, U-1$ 로 두자. 여기서  $u_0 \geq 0$  이다. Log-Polar 좌표상의 수직축 좌표값  $v = 0, \dots, V-1$ 로 두자. 그리고 직각 좌표계에서 log-polar 사상되는 최대 반경은  $X$ 로 두자. 직각 좌표상에서 log-polar 사상되는 위치를 결정하는 거리 basis를  $b = 10^{\lfloor \log X/U \rfloor}$ 로, 각도 basis를  $\theta = 2\pi/V$ 로 두자. Log-Polar 좌표상의 한 점  $(u, v)$ 로부터 다음과 같은 inverse 사상으로 직각 좌표상의 한 점  $(x, y)$ 를 구한다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \lfloor b^u \begin{bmatrix} \cos \theta v \\ \sin \theta v \end{bmatrix} + 0.5 \rfloor \quad (2)$$

직각 좌표상의 한 점  $(x, y)$ 에서 밝기값을 획득하여 log-polar 좌표상의 한 점  $(u, v)$ 의 밝기값으로 결정

한다. 이러한 과정을 통하여 최종적으로 hole들과 overlap들이 발생되지 않는 log-polar 영상을 추출할 수 있게 된다. 그리고 고정된 X값에서 U와 V를 가변시켜 원 영상으로부터 다양한 해상도를 가지는 log-polar 영상을 추출할 수 있다.

### III. 크기와 회전 불변 접근 방법

인간 시각 시스템과 유사한 구조를 가지는 log-polar 사상은 직각 좌표상의 크기와 회전 변화를 log-polar 평면상에서 각각 수평과 수직 순환 이동으로 변화시키는 특성을 가진다. 이러한 특성과 입력 영상으로부터 추출된 후보 얼굴들과 단일 템플릿에 사용된 얼굴 영상의 양 눈동자 사이의 거리와 회전 각도를 이용하여 얼굴의 크기와 회전 변화에 무관한 특징을 가지는 후보 log-polar 얼굴 영상들을 log-polar 평면상에서 추출한다. 추출된 후보 log-polar 영상들은 템플릿에 대하여 크기와 회전 변화에 무관한 특징을 가지므로 제안한 방법은 multitemplate와 multiresolution 기법을 사용할 필요성이 없게된다.

#### 1. log-polar 얼굴 템플릿

실험 모델에 포함되지 않는 남자 한 사람의 얼굴 영상을 템플릿으로 사용한다. 먼저 그림 2에 나타낸바와 같이 양 눈동자의 중심 위치 a와 b를 근간으로 하는 타원형 얼굴 모델<sup>[17]</sup>을 적용하여 정면 얼굴에서 가장 두드러진 양 눈과 눈썹, 코 그리고 입 영역이 포함되는 타원형의 얼굴을 추출한다.

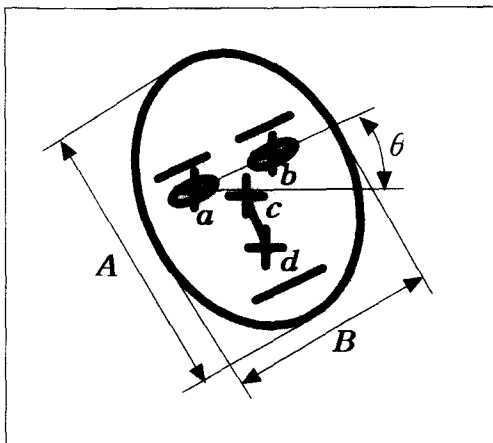


그림 2. 타원형 얼굴 모델  
Fig. 2. Ellipse model of face.

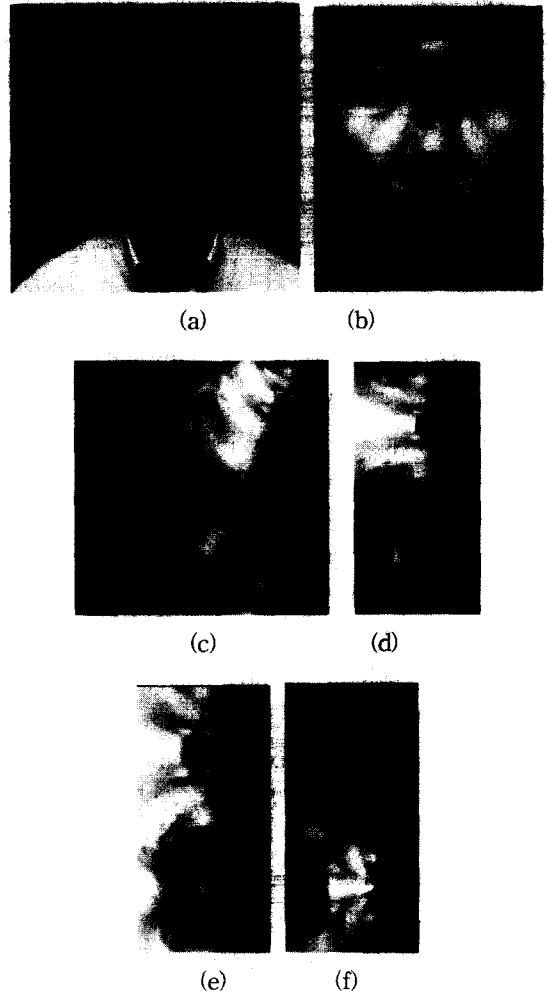


그림 3. Log-Polar 템플릿 얼굴 영상: (a) 양 눈동자 사이의 거리 83화소와 회전 각도 1.4°를 가지는 얼굴 영상, (b) (a)의 영상에서 타원형 얼굴 모델을 적용하여 추출된 얼굴 영상, (c) (b)의 영상에서 fixation point를 왼쪽 눈동자의 중심 점으로 선택하여 해상도 256×256으로 log-polar 사상된 영상, (d)-(f) (b) 영상에서 a, b, 그리고 c의 fixation point들에서 해상도 128×256으로 log-polar 사상된 템플릿 영상.

Fig. 3. The log-polar face template. (a) Face image with interocular distance 83 pixels and rotation angle 1.4°. (b) Face image extracted using the elliptic face model. (c) The log-polar image with a resolution 256×256, in which extracted image is mapped into the log-polar plane at the center of the left pupil as a fixation point. (d)-(f) The log-polar template images with a resolution 128×256, in which the image (b) is log-polar mapped at three fixation points a, b, and c respectively.

그림 2에서  $a$ 와  $b$ 는 각각 왼쪽과 오른쪽 눈동자의 중심 위치이다. 양 눈동자의 중심  $a$  및  $b$ 와 정삼각형을 형성하는  $d$ 는 코끝의 중심 위치가 된다. 타원의 중심 위치  $c$ 는  $a$ ,  $b$  그리고  $d$ 의 무게중심 좌표이다.  $A$ 와  $B$ 는 각각 타원의 장축과 단축의 길이로 양 눈동자 사이의 거리를 이용하여 결정한다. 그리고 왼쪽 눈동자를 기준으로 오른쪽 눈동자에 대한 회전각도는  $\theta$ 로 나타낸다.

먼저 타원의 중심 위치  $c$ 를 중심으로 적절한 반경 이내에 포함되는 영상에 대하여 콘트라스트(contrast)를 개선하기 위하여 국부적인 히스토그램 평활화(histogram equalization)를 수행한다. 그리고 추출된 타원형의 얼굴에서 눈썹 위쪽 영역은 모델들의 머리카락 형태에 따라 변화가 다양하게 발생되므로 템플릿 정합의 정확도를 증가시키기 위하여 제거한다.

이와 같은 방법으로 그림 3의 (a) 영상으로부터 추출된 타원형의 얼굴 영상을 그림 3의 (b)에 나타낸다. 그런데 얼굴 영역에서 fixation point는 입의 위치를 선택할 수 있다. 예를 들어 그림 3의 (b)에서 왼쪽 눈동자 영역의 중심점을 fixation point로 식 (2)에서  $u_0$ 는 0으로, 그리고  $X$  및  $U$ 와  $V$ 를 256으로 선정하여 사상된 log-polar 영상은 그림 3의 (c)에 나타낸 바와 같이 왼쪽 눈동자 영역의 밀도가 높게 샘플링되어 눈동자의 형태가 잘 표현되지 않는다. 그리고 오른쪽 눈동자 영역은 관심 영역의 중심으로부터 멀어지므로 샘플링 밀도가 감소되어 눈동자의 형태가 축소되어 표현된다. 따라서 템플릿 정합을 할 경우에 정합 오차를 증가시키는 요인이 된다.

본 논문에서는 얼굴을 구성하는 눈, 코, 그리고 입 영역의 형태와 위치가 log-polar 사상을 수행한 후에도 직각 좌표상에서와 같이 비교적 두드러지게 표현될 수 있는 위치를 fixation point로 선택한다. 그런데 양 눈동자 영역 사이의 중심점과 입 영역의 중심점을 수직으로 연결하는 선을 기준으로 얼굴 영역은 대칭 관계를 나타낸다. 이러한 특성을 이용하여 본 논문에서는 fixation point를 그림 3의 (b)에서 검은색 십자선으로 나타낸 바와 같이 얼굴을 수직 이등분하는 선에 위치한 세점 - 양 눈동자 영역 사이의 중심점  $a$ , 양 눈동자 영역의 중심점들과 코 끝 중심점 사이의 무게 중심점  $b$ , 그리고 코 끝의 중심점  $c$  - 을 사용한다. 그런데 직각 좌표상에서 선택한 fixation point들의 근접한 영역( $r < r_0$ )에서는 중요한 얼굴 구성 요소가

존재하지 않으며 비교적 동일한 밝기값을 가진다. 따라서 fixation point들의 근접한 영역을 제거하여도 눈, 코, 그리고 입 영역은 원래의 해상도를 유지하므로 정합 결과에는 영향을 미치지 않는다. 직각 좌표상에서 이 영역을 제거한 영상을 log-polar 사상하면 log-polar 영상의 해상도를 감소시킬 수 있게 된다. 또한 정합에 필요한 수행 시간을 단축시킬 수 있게 된다. 본 논문에서는 fixation point에서 반경이 양 눈동자 사이의 거리 1/4이내에 포함되는 영역을 제거한다. 이 영역은 log-polar 좌표상의  $0 \leq u \leq 127$ 에 해당된다. 즉 식 (2)에서  $u_0 = 128$ 이 된다. 선택한 fixation point들에서 식 (2)를 이용하여 각각 해상도  $128 \times 256$ 으로 log-polar 사상된 최종적인 템플릿 영상들을 그림 3의 (d), (e) 그리고 (f)에 나타내었다. 이 경우에  $U$ 와  $V$  및  $X$ 는 256이다. 그림 3의 (d)를 살펴보면 fixation point  $a$ 에 가까운 양 눈동자는 강조되면 거리에 있는 눈과 입은 상대적으로 축소되어 표현되었다. 그리고 그림 3의 (e)에 나타낸 바와 같이 얼굴의 중심 위치인  $b$ 에서 사상된 영상은 눈, 코 그리고 입 영역이 비슷한 크기로 표현된다. 그리고 코끝의 중심인  $c$ 에서 사상된 영상은 그림 3의 (f)에 나타낸 바와 같이 코와 입 영역은 강조되고, 눈 영역은 상대적으로 축소되어 표현되었다. 그런데 적절한 조명 조건에서 획득한 얼굴 영상에서는 눈 영역이 가장 두드러진 특징을 나타낸다. 따라서 log-polar 좌표상에서 눈 영역의 특징이 현저하게 추출되는  $a$ 에서 log-polar 사상한 영상을 본 논문에서 사용하는 최종적인 템플릿으로 결정한다.

## 2. 후보 log-polar 얼굴 추출

입력 영상에서 후보 log-polar 얼굴들을 추출하기 위하여 타원형 얼굴 모델의 근간이 되는 눈동자 후보 쌍의 위치를 찾는다. 먼저 눈동자 영역이 될 수 있는 고립 영역들을 영역 분할과 영역 레이블링(labeling)으로 추출한다. 영역 분할된 영상에서 양 눈동자 영역의 형태는 유사하다. 그리고 눈동자는 얼굴 영역에서 대칭 위치에 쌍으로 존재한다. 이러한 사전 지식들을 활용하여 추출된 고립 영역들로부터 눈동자 쌍이 될 수 있는 후보 쌍들을 찾는다. 후보 쌍들의 순위를 결정하기 위하여 후보 쌍을 구성하는 두 고립 영역 사이의 비유사도를 영역 분할된 영상에서 두 고립 영역의 특징을 이용하여 계산한다. 그리고 적절한 순위 이내에 포함되는 후보 쌍들에서 템플릿과 동일한 타원형

얼굴 모델로부터 후보 얼굴들을 추출한다. 추출된 후보 얼굴에서 후보 쌍을 구성하는 두 고립 영역 사이의 중심 위치를 fixation point로 log-polar 사상하여 후보 log-polar 얼굴들을 추출한다.

(1) 영역 분할과 레이블링

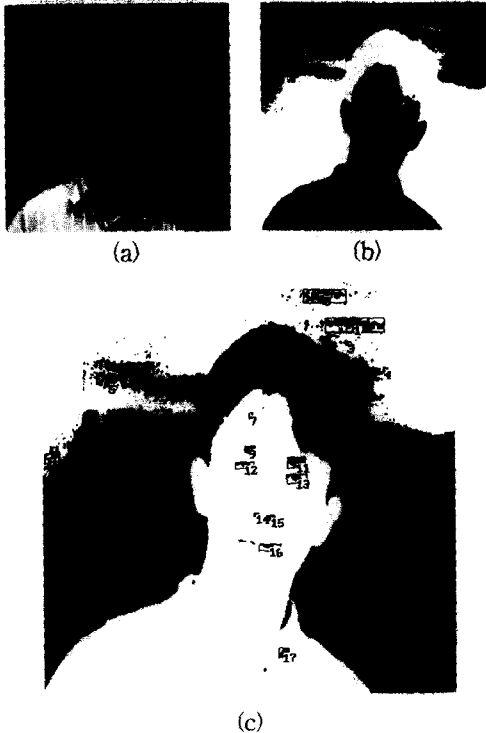


그림 4. (a) 입력 상반신 얼굴 영상, (b) 고정된 문턱 값으로 영역 분할한 영상, (c) (b) 영상의 고립 영역들을 Pavlidis 알고리즘으로 레이블링한 영상

Fig. 4. (a) The input upright frontal face image. (b) The segmented image obtained by thresholding with a fixed value. (c) The labelled image, in which the isolate regions are found by Pavlidis algorithm.

영상의 콘트라스트를 개선시키고 카메라 입력 이득을 보상하기 위하여 입력 영상에서 히스토그램 평활화를 수행한다. 그리고 잡음의 영향을 줄이기 위하여 가우시안 스무딩(smoothing)을 적용한다. 히스토그램 평활화된 영상에서 후보 얼굴 영역을 추출하는 근간이 되는 양 눈동자 영역을 찾기 위하여 그레이 레벨(gray level) 히스토그램을 이용하는 Otsu 방법<sup>[18]</sup>으로 영역을 분할한다. 그런데 히스토그램 평활화를 수행하면 입력 영상의 히스토그램은 비교적 평탄하게 된다. 이 경우에 Otsu 방법으로 결정되는 문턱 값은

최대 그레이 레벨의 1/2이 된다. 본 논문에서 사용되는 입력 영상의 최고 그레이 레벨은 '255'이므로 영역을 분할하는 문턱 값은 고정된 '128'로 된다. 따라서 Otsu 방법을 이식하지 않으므로 영역을 분할하는 수행하는 시간을 단축시킬 수 있게 된다.

그림 4의 (b)는 그림 4의 (a) 영상을 히스토그램 평활화와 가우시안 스무딩하여 고정된 문턱 값 '128'로 영역 분할된 영상을 보여준다. 고립 영역의 위치와 최소 외접 사각형(minimum bounding rectangle; MBR)은 영역 분할된 영상에서 영역 레이블링으로 획득한다. 영역 레이블링 방법은 raster-scan 방식의 Pavlidis 알고리즘<sup>[19]</sup>을 사용한다. 따라서 고립 영역의 번호는 그림 4의 (c)에 나타난 바와 같이 영상의 왼쪽 위 부분으로부터 오른쪽 아래 부분까지 순서에 따라 결정된다.

(2) 사전 지식에 기반한 규칙

눈동자는 적절한 조명 조건에서 획득한 얼굴 영상에서 비교적 어두운 밝기를 가지는 두드러진 영역이다. 따라서 눈동자 영역의 모양과 면적은 영역 분할된 영상에서 얼굴의 크기와 회전 변화에 따라서 비교적 일정하게 나타난다. 그리고 눈동자는 얼굴 영역에서 좌우 대칭 위치에 쌍으로 존재한다. 이러한 사전 지식들을 이용하여 추출한 고립 영역들에서 눈동자 쌍이 될 수 있는 후보 쌍들을 선택한다.

첫번째 과정에서는 눈동자 영역으로 간주되는 고립 영역들을 선택한다. 고립 영역의 모양과 면적은 각각 MBR의 가로 길이와 세로 길이의 비인 종횡비(aspect ratio)와 MBR의 면적으로 나타낼 수 있다. 눈동자 영역은 영역 분할된 영상에서 비교적 일정한 모양과 면적을 가진다. 따라서 눈동자가 될 수 있는 고립 영역들에서 MBR의 종횡비와 MBR의 면적은 제한된 범위의 값을 가지게 된다.  $i$  번째 고립 영역에서 MBR의 종횡비는  $R_i$ 로 MBR의 면적은  $S_i$ 로 두자. 눈동자로 간주될 수 있는 고립 영역들은 다음의 규칙으로 선택한다.

$$(R_{\min} \leq R_i \leq R_{\max}) \cap (S_{\min} \leq S_i \leq S_{\max}) \quad (3)$$

눈동자 영역들이 포함되도록  $R_{\min}$ ,  $R_{\max}$ ,  $S_{\min}$ , 그리고  $S_{\max}$ 를 고정된 값으로 적절하게 설정한다. 눈동자로 간주될 수 없는 고립 영역들은 제외되므로 눈동자가 될 가능성이 있는 후보 쌍의 수를 감소시킬 수

있게 된다.

두 번째 과정은 식 (3)을 만족하는 고립 영역들로부터 눈동자 쌍이 될 수 있는 후보 쌍들을 선택한다. 눈동자는 얼굴 영역에서 좌우 대칭 위치에 쌍으로 존재한다. 따라서 눈동자 쌍이 될 수 있는 후보 쌍은 2개의 고립 영역으로 구성할 수 있다. 고립 영역들의 수를  $M$ 으로 두자. 만약  $M < 2$ 이면 눈동자 검출이 실패한 경우이다. 그리고  $M \geq 2$ 이면 후보 쌍의 수는 최대  $M \cdot (M-1)/2$ 가 된다. 후보 쌍  $(j, k)$ 를 구성하는  $j$ 번째와  $k$ 번째 고립 영역간의 유클리드 거리와 회전 각도를 각각  $d_{jk}$ 와  $\theta_{jk}$ 로 두자. 그런데 얼굴의 크기는 양 눈동자 중심 사이의 거리로, 회전 각도는 수평선에 대하여 왼쪽 눈동자의 중심을 기준으로 오른쪽 눈동자의 중심에 대한 각도로 표현할 수 있다. 똑바로 선 상반신 영상에서 양 눈동자 사이의 거리와 회전 각도의 범위는 제한적이다. 따라서 눈동자 쌍이 될 수 있는 후보 쌍들은 다음의 규칙을 이용하여 선택할 수 있다.

$$(d_{\min} \leq d_{jk} \leq d_{\max}) \cap (\theta_{\min} \leq \theta_{jk} \leq \theta_{\max}) \quad (4)$$

유클리드 거리의 최소값  $d_{\min}$ 과 최대값  $d_{\max}$ , 그리고 회전 각도의 최소값  $\theta_{\min}$ 과 최대값  $\theta_{\max}$ 는 눈동자 쌍이 포함되도록 고정된 적절한 값을 사용한다.

### (3) 후보 쌍의 순위

후보 쌍의 중심 위치에서 타원형 얼굴 모델로부터 후보 log-polar 얼굴 영상을 추출하여 템플릿 정합으로 제한한 방법은 얼굴과 눈을 검출한다. 따라서 얼굴과 눈을 검출하는 수행 시간은 템플릿 정합에 참여하는 후보 log-polar 얼굴들의 수에 비례하여 증가된다. 따라서 식 (4)를 만족하는 후보 쌍들에서 눈동자 쌍이 포함되도록 적절한 수만큼의 후보 쌍들을 선택하는 것이 바람직하다.

먼저 후보 쌍을 구성하는 두 고립 영역의 비유사도를 측정한다. 영역 분할된 영상에서 눈동자 쌍을 구성하는 두 고립 영역의 모양과 면적은 유사하다. 따라서 이러한 특징을 이용하여 후보 쌍을 구성하는 두 고립 영역의 비유사도를 측정할 수 있게 된다. 영역 분할된 영상에서 후보 쌍  $(j, k)$ 를 구성하는  $j$ 번째와  $k$ 번째 고립 영역에서 MBR의 종횡비  $R_j$ 와  $R_k$ , 두 고립 영역의 면적  $A_j$ 와  $A_k$ , 그리고 두 고립 영역의 외곽선 길이  $L_j$ 와  $L_k$ 로부터 비유사도  $c_{jk}$ 는 다음과 같이 계산된다.

$$c_{jk} = w_1 \times |R_j - R_k| + w_2 \times |A_j - A_k| + w_3 \times |L_j - L_k| \quad (5)$$

여기서 외곽선 길이는 8방향 contour 추적 알고리즘<sup>[20]</sup>을 사용하여 계산한다. 후보 쌍을 구성하는 두 고립 영역들에서 추출한 특징들의 중요성에 부가되는 가중치는 각각  $w_1$ 과  $w_2$  그리고  $w_3$ 을 사용한다. 가중치  $w_1$ 과  $w_2$  그리고  $w_3$ 은 실험적으로 적절한 선택한 고정된 값을 사용한다. 비유사도의 값이 작을수록 후보 쌍을 구성하는 두 고립 영역의 유사도는 커진다.

모든 후보 쌍들에서  $c(j, k)$ 의 값이 적은 순서대로 후보 쌍을 sorting하는 방법을 "sorting method I"으로 두자. 그러나  $M$ 개의 고립 영역중에서 임의로 선택한 한 개의 고립 영역이 포함되는 후보 쌍들의 일부 또는 모든 순위가 상위에 등록되는 경우가 종종 발생하게 된다. 예를 들어 눈동자와 배경 영역에서 발생하는 고립 영역들로 구성되는 후보 쌍들과 눈동자와 눈썹으로 구성되는 후보 쌍의  $c(j, k)$ 가 눈동자 쌍의  $c(j, k)$ 보다 적은 값을 가지면 눈동자 쌍의 순위는 상대적으로 하위에 등록된다. 따라서 얼굴과 얼굴 구성 요소의 검출율을 증가시키기 위해서는 템플릿 정합에 참여시키는 후보 log-polar 얼굴들의 수를 증가시켜야 한다. 그러므로 전체 시스템의 수행 시간은 상대적으로 증가하게 된다.

이러한 문제를 해결하기 위하여, 본 논문에서는  $M$ 개의 고립 영역 각각에서 반드시 한 개의 후보 쌍이 상위의 순위에 등록되도록 sorting하는 다음과 같은 "sorting method II"를 제안한다.

**단계 1.**  $M$ 개의 고립 영역중 한 개의 고립 영역을 선택한다. 선택된 고립 영역이 포함되는 후보 쌍들의  $c(j, k)$ 를 비교하여 최소값을 가지는 후보 쌍을 선택한다. 이러한 과정을 나머지  $M-1$ 개의 고립 영역에 대해 반복하여 수행하면  $M$ 개의 고립 영역은  $M$ 개의 후보 쌍에 반드시 포함된다. 선택된  $M$ 개의 후보 쌍들의  $c(j, k)$ 를 sorting하여  $p(q)$ 에 저장한다. 여기서  $q=1, \dots, M, \dots, M \cdot (M-1)/2$ 로 후보 쌍의 순위를 나타내며  $p(q)$ 는 해당되는 순위의 후보 쌍  $(j, k)$ 를 나타낸다.

**단계 2.** 단계 1에서 선택된 후보 쌍을 제외한 나머지 후보 쌍들에서 단계 1을 반복한다. 만약  $M$ 이 홀수이면 단계 1을  $(M-1)/2$ 번, 짝수이면  $M/2$ 번 반복한다.

이러한 과정을 반복하여 모든 후보 쌍들의 순위를 할당하게 된다.

상위에 등록된 후보 쌍부터 순서대로 적절한 수만큼의 후보 쌍들을 선택한다. 선택된 후보 쌍을 구성하는 두 고립 영역의 위치를 근간으로 템플릿과 동일한 타원형 얼굴 모델을 적용하여 그림 5의 (a), (b), 그리고 (c)에 나타난 바와 같이 후보 얼굴들을 추출한다. 그림 5의 (a), (b), 그리고 (c)는 그림 4의 (c)에서 각각 고립 영역 '12'와 '13', '11'과 '12', 그리고 '9'와 '11'로 구성되는 후보 쌍의 중심 위치에서 그림 4의 (a) 영상으로부터 국부적인 히스토그램 평활화를 수행하여 타원형의 얼굴 모델로부터 추출된 영상을 보여준다. 여기서 작은 흰색 상자와 검은색 십자선은 각각 후보 쌍의 위치와 fixation point를 나타낸다. 그리고 추출된 후보 얼굴들의 fixation point에서 식 (2)를 이용하여 템플릿 정합에 참여시키는 후보 log-polar 얼굴들을 생성한다.

### 3. 크기와 회전 불변 정합

눈동자 쌍을 구성하는 두 고립 영역 사이의 거리와 회전 각도는 입력 영상에서 얼굴의 크기와 회전 변화에 대응하여 나타난다. 직각 좌표상에서 템플릿과 눈동자 후보 쌍의 중심점에서 추출한 후보 얼굴 사이에는 크기와 회전 각도의 차이가 발생한다. 따라서 log-polar 좌표상에서 두 영상 사이에는 수평과 수직 이동의 차이가 존재하게 된다. Log-Polar 좌표상에서 두 물체간의 수평과 수직 이동 차이를 보상하면 직각 좌표상에서 두 물체간에 존재하는 크기와 회전 변화에 무관한 특징을 추출할 수 있게 된다.

이러한 수평과 수직 이동 차이에 무관하게 물체를 검출할 수 있는 방법은 추출된 log-polar 영상을 푸리에(Fourier) 변환하여 정합하는 방법과 템플릿을 수평과 수직으로 이동시키면서 정합하는 방법이 있다. 푸리에 변환은 물체의 이동 변화를 위상의 변화로 나타낸다. 따라서 물체의 이동 변화가 발생한 영상에 푸

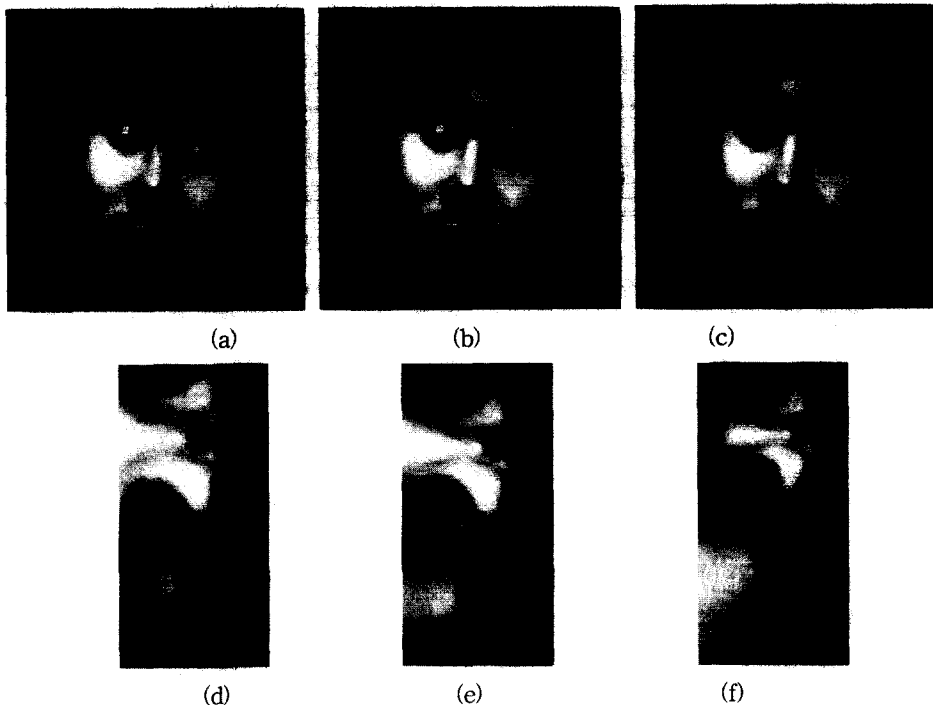


그림 5. 템플릿에 대하여 크기와 회전 변화에 무관한 후보 log-polar 얼굴 영상 생성 과정: (a)-(c) 타원형의 얼굴 모델을 적용하여 추출한 후보 얼굴 영상, (d)-(f) (a)-(c) 영상의 fixation point에서 추출된 log-polar 영상을 템플릿에 대하여 크기와 회전 변화를 정렬한 후보 log-polar 얼굴 영상

Fig. 5. Illustration of scale and rotation invariant matching of the candidate log-polar face images to a template: (a)-(c) The candidate face images extracted using the elliptic face model. (d)-(f) The candidate log-polar face images with a resolution  $128 \times 256$ , in which (a)-(c) images are mapped into the log-polar plane at each fixation point and the resultant log-polar images are realigned using the correction equations to obtain scale and rotation invariance.



리에 변환을 적용하여 크기만 고려하면 물체의 이동 변화에 무관한 특징을 추출할 수 있게 된다. 그러나 푸리에 변환은 시스템의 수행 시간을 증가시키게 된다<sup>[13]</sup>. 마찬가지로 크기와 회전 각도가 고정된 템플릿을 추출한 log-polar 영상에 가능한 모든 수평과 수직 이동 변화량으로 이동시키면서 정합을 수행하므로 계산 시간이 증대된다.

눈동자 후보 쌍을 구성하는 두 고립 영역 사이의 거리와 회전 각도, 그리고 템플릿에 사용하는 얼굴 영상에서 양 눈동자 사이의 거리와 회전 각도를 이용하여 제안한 방법은 두 log-polar 영상을 정렬시킨다. 그리고 정렬된 두 영상을 한 번의 정합으로만 정합 결과를 계산한다. 따라서 시스템의 수행 시간을 단축시킬 수 있게 된다. 먼저 log-polar 사상에서 사용되는 거리와 각도의 basis를 이용하여 두 log-polar 영상 사이에 존재하는 수평 이동 차이  $\Delta u$ 와 수직 이동 차이  $\Delta v$ 를 각각 다음과 같이 계산한다.

$$\Delta u = \lfloor \frac{\log d_{jk} - \log d_t}{\log b} + 0.5 \rfloor \quad (6)$$

$$\Delta v = \lfloor \frac{(\theta_{jk} - \theta_t)V}{2\pi} + 0.5 \rfloor \quad (7)$$

여기서  $d_t$ 와  $\theta_t$ 는 템플릿의 얼굴 영상에서 양 눈동자 사이의 거리와 회전 각도이다. 수평 이동 차이  $\Delta u$ 의 부호가 음이면 추출된 영상이 템플릿 보다 크기가 작은 경우이고, 양이면 그 반대의 경우이다. 눈동자 후보 쌍의 중심에서 추출한 log-polar 영상을  $\Delta u$ 의 부호가 음이면 오른쪽으로, 양이면 왼쪽으로  $|\Delta u|$ 만큼 이동시켜 템플릿 영상에 대하여 크기를 정렬시킨다. 그리고 수직 이동 차이  $\Delta v$ 의 부호가 음이면 추출된 영상의 회전 각도는 템플릿보다 작고, 양의 부호를 가지면 템플릿보다 큰 회전 각도를 가지는 경우이다. 수직 이동 차이  $\Delta v$ 의 부호가 음이면 위쪽으로, 양이면 아래쪽으로 추출한 log-polar 영상을  $|\Delta v|$ 만큼 이동시켜 템플릿과 회전 각도의 차이를 보상한다. 정렬된 영상에서 템플릿과 마찬가지로  $0 \leq u < u_0$ 에 해당되는 영역을 제거한다.

그림 5의 (a), (b), 그리고 (c)에 나타난 영상에서 검은색 십자선으로 표시된 위치를 fixation point로 식 (2)를 이용하여 사상된 후보 log-polar 얼굴 영상을 추출한다. 이 경우에 식 (2)에서 사용한  $u_0$ 는 0이며  $X$ 와  $U$  그리고  $V$ 는 256이다. 그리고 식 (6)과

(7)를 이용하여 추출된 log-polar 얼굴 영상들을 템플릿 영상에 대하여 정렬한다. 정렬된 후보 log-polar 얼굴 영상들의  $0 \leq u < u_0 (=128)$ 와  $0 \leq v < V$ 에 해당되는 영역을 제거한 최종적인 후보 log-polar 얼굴들을 그림 5의 (d), (e), 그리고 (f)에 나타내었다. 그림 5의 (d)에 나타낸 바와 같이 눈동자 쌍의 중심 위치에서 log-polar 사상된 영상을 템플릿에 대하여 정렬된 영상과 그림 3의 (d)에 나타낸 템플릿 영상을 비교해보면 얼굴을 구성하는 각 요소의 위치가 두 영상에서 정렬되었음을 알 수 있다. 마찬가지로 그림 5의 (e)와 (f)에 나타낸 바와 같이 후보 log-polar 얼굴 영상들을 생성할 수 있다. 이러한 방법으로 정렬된 후보 log-polar 얼굴 영상은 템플릿에 대하여 크기와 회전 변화에 무관한 특징을 가지게 된다.

마지막으로, 정렬된 두 영상을 템플릿 정합한 결과를 비교하여 정합 오차가 최소가 되는 후보 쌍을 검출하여 타원형 얼굴 모델로부터 얼굴과 눈을 동시에 검출한다. 템플릿 log-polar 영상을  $I_t(u, v)$ 로 두자. 비 유사도의 순위 결정 방법에서 "sorting method II"로 선택한 눈동자 후보 쌍의 집합을  $\Omega = \{(j, k) \in \mathcal{Q} \mid 1 \leq j \leq Q\}$ 로 두자. 여기서  $\mathcal{Q}$ 의 최대값은  $M \cdot (M-1)/2$ 가 된다. 선택된 후보 쌍에서 추출한 후보 log-polar 얼굴 영상을 템플릿에 대하여 크기와 회전 변화를 정렬시킨 영상을  $I_{jk}(u, v)$ 로 두자. 정렬된 두 log-polar 영상을 다음과 같은 MAD (mean absolute difference)를 이용하여 정합 정도를 계산한다.

$$(j, k)^* = \arg \min_{(j, k) \in \Omega} \{ \sum_u \sum_v |I_{jk}(u, v) - I_t(u, v)| \} \quad (8)$$

정합 오차가 최소가 되는 눈동자 쌍  $(j, k)^*$ 로부터 타원형의 얼굴 모델을 이용하여 얼굴과 눈을 동시에 검출하게 된다. 여기서 사용된 템플릿은 얼굴을 표현하는 주요 특징이 되는 눈, 코 그리고 입을 포함한다. 따라서 제안한 방법은 배경에서 발생하는 후보 쌍들과 눈동자와 눈썹으로 구성되는 후보 쌍들이 존재하는 영상에서도 견실하게 얼굴과 눈을 검출할 수 있다.

#### IV. 실험 결과 및 고찰

실험에 사용하는 얼굴 영상들은 후백 CCD 카메라 (NTSC 방식)로부터 약 1.5m 떨어진 거리에서 획득



그림 6. 실험에 사용한 얼굴 영상의 예  
Fig. 6. Sample images of our face database.

한  $512 \times 480$  해상도를 가지는 256 그레이 레벨 영상이다. 얼굴 영상은 총 53명(남자 21명, 여자 32명)의 각 모델로부터 상반신 영상 15장을 획득하여 총 795장의 실험 영상들을 구축하였다. 각 모델의 상반신 영상 15장은 5종류의 크기에 대하여 각각 정면과 좌우 회전된 3종류로 구성하였다. 획득한 얼굴 영상의 크기와 회전 변화는 각각 양 눈동자 중심 사이 거리가 약 40에서 140화소이며 회전 각도는  $\pm 25$ 도의 범위에 포함된다. 이 경우에 모델의 자세는 구두로 지시하였으며 카메라의 초점과 모델의 시선을 일치시키는 특별한 방법은 사용하지 않았다. 단 안경을 착용한 모델일 경우 안경을 착용하지 않도록 요청하였다. 조명과 배경은 적절하게 조절하였다. 이와 같은 과정으로 획득한 얼굴 영상의 예를 그림 6에 나타내었다.

제한한 방법의 성능을 확인하기 위하여 Pentium II 333MHz 프로세서를 탑재한 개인용 컴퓨터에 C 언어로 이식하여 795장의 상반신 영상에 대하여 실험하였다. 먼저 얼굴 검출이 성공한 경우와 실패한 경우를 결정하기 위하여 수동으로 실험에 사용된 795장의 영상에서 눈동자의 중심을 찾았다. 검출된 눈동자 쌍을 구성하는 두 고립 영역이 수동으로 결정한 양 눈동자의 중심을 포함하면 얼굴 검출이 성공한 경우로 그렇지 않으면 실패한 경우로 결정하였다. 본 논문에서 제한한 눈동자 후보 쌍의 순위를 결정하는 "sorting

method II"의 타당성을 입증하기 위하여 템플릿 정합에 참여시키는 후보 얼굴의 수에 따른 얼굴의 검출율을 그림 7에 나타내었다. 여기서 템플릿과 눈동자 후보 쌍에서 추출한 log-polar 영상의 해상도는  $128 \times 256$ 이다. 그리고 log-polar 사상에 사용한 fixation point는 입력 영상에서 추출한 눈동자 후보 쌍에서는 두 고립 영역 사이의 중심점이며 템플릿에서는 양 눈동자 사이의 중심점이다.

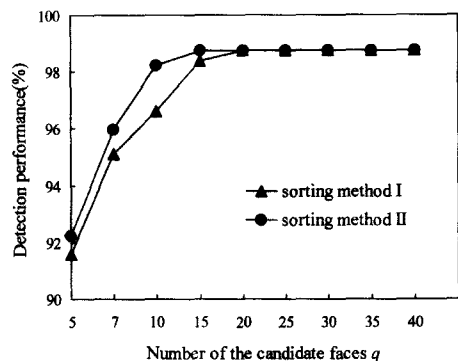


그림 7. 후보 얼굴들의 수에 대한 두 종류의 sorting 방법에 따르는 검출 성능

Fig. 7. Detection performance of two sorting methods against the number of candidate faces.

그림 7에서 "sorting method II"는 후보 쌍의 순

위를  $Q \leq 15$ 로 제한하면 98.7%,  $Q \leq 25$ 로 제한하면 98.7%의 검출율을 나타낸다. 그리고 "sorting method I"은 후보 쌍의 순위를  $Q \leq 15$ 로 제한하면 98.3%,  $Q \leq 25$ 로 제한하면 98.7%의 검출율을 나타낸다. 두 경우를 비교해보면 제안한 "sorting method II" 방법이 template 정합에 참여하는 후보 얼굴의 수를 적게 선택하여도 얼굴과 눈의 검출 성능이 우수함을 알 수 있다.

표 1. Fixation point **a**, **b**, 그리고 **c**에 따르는 검출율(%)  
Table 1. Detection performance(%) with three fixation points **a**, **b**, and **c**.

| Fixation points | Template sizes |        |         |
|-----------------|----------------|--------|---------|
|                 | 32×64          | 64×128 | 128×256 |
| <b>a</b>        | 98.1           | 98.7   | 98.7    |
| <b>b</b>        | 96.7           | 98.4   | 98.6    |
| <b>c</b>        | 91.0           | 91.0   | 91.0    |

Fixation point를 양 눈동자 사이의 중심점 **a**, 얼굴의 중심점 **b**, 그리고 코끝의 중심 위치 **c**로 선택할 경우에 얼굴 검출율을 표 1에 나타내었다. 표 1을 살펴보면 양 눈동자 사이의 중심점을 fixation point로 선정할 경우에 가장 높은 얼굴 검출율을 나타낸다. 그리고 양 눈동자 사이의 중심 위치로부터 거리가 멀어질수록 얼굴 검출율이 감소됨을 알 수 있다. 이러한 결과로부터 본 논문에서 결정한 양 눈동자 사이의 중심점이 fixation point로 타당함을 확인하였다.

표 2. 얼굴의 크기 변화에 따르는 검출된 눈 중심 위치의 정확도

Table 2. Accuracy in estimating the center position of detected eyes according to face sizes.

| Face Sizes (interocular distance) | Average error (pixels) | standard deviation (pixels) |
|-----------------------------------|------------------------|-----------------------------|
| 40 - 60                           | 1.3                    | 0.4                         |
| 60 - 80                           | 1.6                    | 0.5                         |
| 80 - 100                          | 2.1                    | 0.6                         |
| 100 - 120                         | 2.6                    | 0.7                         |
| 120 - 140                         | 3.2                    | 0.8                         |

Log-Polar 영상의 해상도를  $32 \times 64$ ,  $64 \times 128$ , 그리고  $128 \times 256$ 일 경우에 얼굴과 눈의 검출율은 각각 98.0%, 98.7%, 그리고 98.7%를 나타내었다. 그리고 검출된 눈동자 위치의 정확도를 확인하기 위하여, 수동으로 결정한 눈동자의 무게중심과 제안한 방법으로 검출된 눈동자 무게중심 사이의 유클리드 거리로 계산한 오차를 5단계의 얼굴 크기 변화에 대응하여 표 2에 나타내었다.

표 2에서 사용한 log-polar 영상의 해상도는  $128 \times 256$ 이다. 표 2를 살펴보면 얼굴의 크기가 커지면 오차가 증가되는 현상을 나타낸다. 그런데 눈동자의 가로 길이는 양 눈동자 사이 거리의 약 1/2배에 해당된다. 따라서 각 단계의 얼굴 크기에서 발생하는 오차는 비교적 타당함을 알 수 있다.

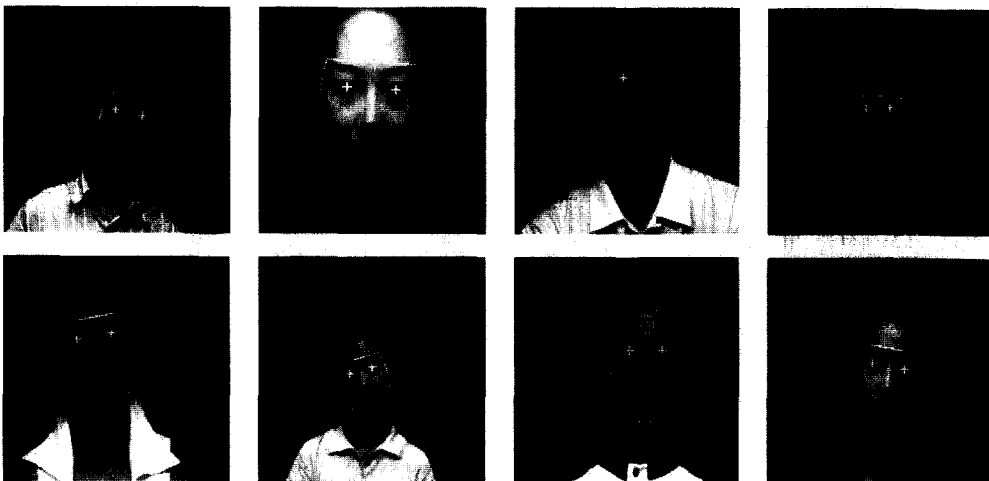


그림 8. 검출된 얼굴과 눈의 예  
Fig. 8. Examples of their detected faces and eyes.

그리고 평균 수행 시간은 실험에 사용한 3종류의 해상도에 대해서 최소 2초에서 최대 12초이며 평균 7초이었다. 본 논문에서 사용한 얼굴 영상에서 양 눈동자 사이의 최소 거리와 최대 거리의 비는 약 3.5이다. 이러한 큰 변화에 대응하기 위해 식 (3)과 (4)에서 사용한 최소값과 최대값의 범위는 큰 폭을 가지도록 하였다. 따라서 배경에서 발생하는 눈동자 후보 쌍의 수가 증가하게 된다. 만약 얼굴의 크기와 회전 변화를 적절히 제한하면 상대적으로 배경에서 발생하는 눈동자 후보 쌍의 수를 감소시킬 수 있으므로 검출 시스템의 수행 시간을 단축시킬 수 있게 된다.

본 논문에서 사용한 얼굴 영상에 제안한 방법을 적용하여 검출한 얼굴과 눈의 예를 그림 8에 나타내었다. 제안한 방법은 입력 영상에서 눈동자 쌍을 근간으로 타원형의 얼굴 모델을 적용하여 후보 얼굴들을 추출하므로 정합 과정에서 얼굴과 눈을 동시에 검출한다. 따라서 눈을 검출하기 위한 부가적인 과정이 필요하지 않게 된다.

## V. 결 론

본 논문에서는 log-polar 사상의 특성을 기반으로 단일 템플릿을 사용하는 템플릿 정합으로 얼굴과 눈을 자동으로 검출하는 새로운 방법을 제안하였다. 입력 영상에서 선택된 fixation point들로부터 추출한 다양한 크기와 회전 변화를 가지는 후보 얼굴들을 제안한 크기와 회전 불변 접근 방법으로 단일 템플릿에 대하여 log-polar 평면상에서 크기와 회전 변화를 정규화할 수 있게된다. 따라서 얼굴의 크기와 회전 변화에 대응하기 위하여 계산량이 증대되는 다수의 템플릿을 이용하는 방법과 multiresolution 기법을 적용할 필요가 없게된다. 제안한 방법을 얼굴의 크기와 회전 변화가 동시에 존재하는 상반신 영상 795장에 적용하여 98.7%의 얼굴과 눈의 검출율을 달성하였다.

추후 연구 과제는 본 논문에서 제안된 크기와 회전 불변 정합을 이용하여 얼굴 인식 시스템을 구축하는 것이다.

## 참 고 문 헌

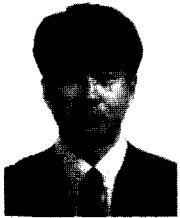
[ 1 ] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expression: A survey," *Pattern*

*Recognition*, vol. 25, no. 1, pp. 65-77, 1992.

- [ 2 ] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of The IEEE*, vol. 83, no. 5, 1995.
- [ 3 ] D. Valentin and G. W. Cottrell, "Connectionist models of face processing: A survey," *Pattern Recognition*, vol. 27, no. 9, pp. 1209-1230, 1994.
- [ 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 ] B. Takacs and H. Wechsler, "Detection of faces and facial landmarks using iconic filter banks", *Pattern Recognition*, vol. 30, no. 10, pp. 1623-1636, 1997.
- [ 7 ] H. A. Rowley, S. Baluja, and T. Kande, "Neural network-based face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38, 1998.
- [ 8 ] Kah-Kay Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, 1998.
- [ 9 ] K. Hotta, T. Kurita, and T. Mishima, "Scale invariant face detection method using higher-order local autocorrelation features extracted from Log-Polar image," *Proceedings of The 2nd International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, pp. 70-75, Apr. 14-16, 1998.
- [ 10 ] X. Jia and M. Nixon, "Extending the feature vector for automatic face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.

- 17, no. 12, pp. 1167-1176, 1995.
- [ 11 ] Kin-Man Lam and H. Yan, "Locating and extracting the eye in human face images," *Pattern Recognition*, vol. 29, no. 5, pp. 771-779, 1996.
- [ 12 ] D. Reisfeld, H. Wolfson, and Y. Yeshurun, "Context-free attentional operators: The generalized symmetry Transform," *International Journal of Computer Vision*, 14, pp. 119-130, 1995.
- [ 13 ] R. A. Messner and H. H. Szu, "An image processing architecture for real time generation of scale and rotation invariant patterns," *Computer Vision, Graphics and Image Processing*, 31, pp. 50-66, 1985.
- [ 14 ] C. Capurro, F. Panera, and G. Sandini, "Dynamic vergence using Log-Polar images," *International Journal of Computer Vision*, 24(1), pp. 79-94, 1997.
- [ 15 ] B. B. Bederson, R. S. Wallace, and E. L. Schwartz, "A miniaturized space-variant active vision system: Cortex-I," *Artificial Neural Networks for Speech and Vision*, 24, pp. 429-456, 1993.
- [ 16 ] G. Wolberg, *Digital Image Warping*, pp. 42-45, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [ 17 ] 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.
- [ 18 ] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 62-66, Jan. 1979.
- [ 19 ] T. Pavlidis, *Algorithm for graphics and image processing*, Rockville, Md.: Computer Science Press, 1982.
- [ 20 ] D. H. Ballard, *Computer vision*, pp. 143-146, New Jersey, Prentice-Hall, Inc., 1982.

## 저 자 소 개



## 崔 一(正會員)

1986년 경북대학교 전자공학과(공학사). 1988년 경북대학교 대학원 전자공학과(공학석사). 1988년 ~ 1994년 국방과학연구소 연구원. 1996년 ~ 1997년 한국표준과학연구원 부설 천문대 선임 연구원. 1994

년 ~ 현재 경북대학교 대학원 전자공학과 박사과정 재학중. 주관심분야는 형태인식, 컴퓨터 비전

## 秦 成 一(正會員) 第 32卷 S編 第 9號 參照

경북대학교 전자전기 공학부 교수