

論文 2002-39SP-2-1

로그폴라 사상과 어파인 변환을 이용한 새로운 템플릿 기반 얼굴 인식 (New Template Based Face Recognition Using Log-polar Mapping and Affine Transformation)

金文甲*, 崔一**, 秦成一***

(Mun-Gab Kim, Il Choi, and Sung-Il Chien)

요약

이 논문에서는 크기와 영상 평면상에서 회전 (in-plane rotation) 변화를 가지는 정면 얼굴 영상의 인식 성능을 향상시키기 위하여, 새로운 템플릿 (template) 기반 접근 방법들을 제안한다. 인식 성능을 향상시키기 위한 템플릿들은 크기와 회전 변화가 다른 다수의 영상들을 선형 또는 비선형 연산에 의하여 생성된다. 얼굴의 크기와 영상 평면에서 회전 변화에 무관한 얼굴의 특징을 추출하기 위하여 어파인 (affine) 변환, 로그폴라 (log-polar) 사상, 그리고 로그폴라 영상에 기반한 FFT들이 이용된다. 제안된 방법들은 인식률과 수행 시간 측면에서 비교된다. 실험 결과로부터 제안된 템플릿을 이용한 방법들의 인식률이 한 장의 영상으로 생성된 템플릿을 이용한 방법들의 인식률보다 우수함을 나타낸다. 어파인 변환을 이용한 방법의 인식률이 로그폴라 사상을 이용한 방법과 로그폴라 영상에 기반한 FFT 방법의 인식률보다 우수하며, 수행 시간 측면에서는 로그폴라 사상을 이용한 방법이 가장 빠르다.

Abstract

This paper presents the new template based human face recognition methods to improve the recognition performance against scale and in-plane rotation variations of face images. To enhance the recognition performance, the templates are generated by linear or nonlinear operation on multiple images including different scales and rotations of faces. As the invariant features to allow for scale and rotation variations of face images, we adopt the affine transformation, the log-polar mapping, and the log-polar image based FFT. The proposed recognition methods are evaluated in terms of the recognition rate and the processing time. Experimental results show that the proposed template based methods lead to higher recognition rate than the single image based one. The affine transformation based face recognition method shows marginally higher recognition rate than those of the log-polar mapping based method and the log-polar image based FFT, while, in the aspect of processing time, the log-polar mapping based method is the fastest one.

Key Word : 로그폴라 사상, 어파인 변환, FFT, 템플릿 기반 얼굴 인식 log-polar mapping, affine transformation, FFT, template-based face recognition.

* 正會員, 龍仁松潭大學 디지털電子情報科
(Yong-In Songdam College, Dept. of Digital Electronics & Information)

** 正會員, 龜尾1大學 情報通信科
(Kumi College, Dept. of Information and Communica-

tions)

*** 正會員, 慶北大學校 電子電氣컴퓨터學部
(Kyungpook National University, School of Electrical Engineering and Computer Science)

接受日字: 2001年10月23日, 수정완료일: 2002年2月18日

I. 서 론

생체 측정학 분야에서 얼굴 인식은 최근에 활발히 연구되고 있다. 얼굴 인식은 범죄자 식별, 은행, 개인의 신원 확인 등 다양한 분야에서 응용되고 있다^[1-2]. 얼굴 인식은 특징 기반 정합(feature-based matching)과 템플릿 기반 정합(template-based matching)으로 분류된다^[3]. 특징 기반 정합은 얼굴 구성 요소들의 주요 기준점 사이의 기하학적 정보^[3, 5] 또는 가비 웨이브렛(Gabor wavelet) 계수를 이용하는 방법^[8] 등이 있다. 템플릿 기반 정합은 얼굴 전체 영역의 그레이 레벨을 직접 이용하는 방법^[3], K-L 변환으로부터 추출되는 고유벡터(eigenvector)를 이용하는 방법^[4, 6] 등이 있다.

이 논문에서는 비교적 조명 변화가 적은 환경에서 획득된 크기와 회전 변화를 가지는 정면 얼굴 영상을 인식하기 위하여 새로운 템플릿 기반 정합 방법을 제안한다. 고전적인 템플릿 기반 정합에서는 등록되는 모델을 대표하는 한 장의 영상을 이용하여 템플릿을 생성한다^[3]. 그런데 이 논문에서는 얼굴 영상의 크기와 회전 각도가 다른 다수의 영상들을 이용하여 템플릿을 생성하는 방법을 제안한다. 크기와 회전각이 서로 다른 영상들은 어파인 변환(affine transformation)^[7]에 의해 정규화가 되거나 로그폴라 사상(log-polar mapping)^[10-12]에 의해 로그폴라 영상으로 변환된다. 그리고 템플릿과 입력 영상 사이에서 조명의 변화가 발생하는 경우에 정합 오차를 줄이기 위하여 그래디언트 영상^[9]을 이용하여 템플릿을 생성한다.

이 논문에서는 어파인 변환과 로그폴라 사상 방법에 따라서 한 장의 영상으로부터 생성된 템플릿과 제안하는 방법으로 생성된 템플릿들 사이의 인식 성능을 비교 분석한다. 이 경우에서 얼굴 데이터베이스는 51명의 얼굴 모델로부터 얼굴의 크기와 회전각이 서로 다른 765장의 영상들로 구성된다. 입력 영상과 등록된 템플릿 사이의 정합은 정규화된 교차 상관 계수(normalized cross-correlation; NCC)를 이용한다. 실험 결과로부터 제안된 템플릿들을 이용하는 방법의 인식률이 한 장의 영상을 이용한 템플릿보다 매우 우수함을 나타낸다. 그리고 인식률은 어파인 변환에 기반한 방법이 가장 우수하며, 가장 빠른 수행 시간을 나타내는 방법은 로그폴라 사상에 기반한 것이다.

이 논문의 나머지 구성은 다음과 같다. 제안하는 그

래디언트 기반 템플릿들의 생성 방법은 2장에서 상세하게 설명한다. 국부적으로 발생하는 왜곡에 대하여 정합 오차를 감소시킬 수 있는 디더링(dithering)과 FFT 방법들은 3장에서 기술한다. 제안하는 얼굴 인식 방법의 실험 및 고찰은 4장에서 기술한다. 마지막으로 5장에서는 결론을 기술한다.

II. 템플릿 생성

제안하는 템플릿이 생성되는 방법은 그림 1에 나타난 바와 같다. 그림 1을 살펴보면 전처리 과정에서는 얼굴의 크기와 회전 각도가 서로 다른 영상들로부터 타원형 얼굴 영역^[11, 12]이 추출된다. 추출된 얼굴 영역에 대하여 콘트라스트 스트레칭(contrast stretching) 과정을 수행한다. 이 영상으로부터 그래디언트 영상이 추출된다. 얼굴의 크기와 회전 변화에 불변인 특징을 추출하는 과정에서는 전처리 과정에서 추출된 그래디언트 얼굴 영상들에 대하여 어파인 변환 또는 로그폴라 사상을 적용한다. 국부적인 왜곡에 무관한 특징을 추출하기 위하여, 어파인 변환 또는 로그폴라 사상으로 정규화 및 변환된 영상들은 각 방법에 따라서 단순하게 산술적으로 더해진다. 더해진 한 장의 영상에 대하여 매우 간단한 선형 또는 비선형 연산으로부터 템플릿이 최종적으로 생성된다.

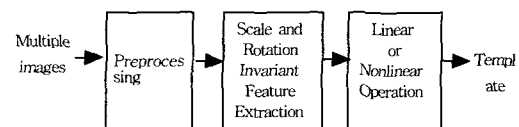


그림 1. 다수의 영상들을 이용하여 제안하는 템플릿을 생성하는 방법

Fig. 1. Proposed template generation method using multiple images.

1. 크기 및 회전 불변 특징 추출

제안하는 템플릿 생성 방법은 동일한 모델에 대하여 템플릿을 생성하기 위하여, 서로 다른 크기와 회전 각도를 가지는 3장의 영상들을 이용한다. 이 경우에 얼굴의 크기는 두 눈동자 사이의 거리(d)로 두고, 회전각은 왼쪽 눈동자의 중심으로부터의 수평선과 왼쪽 눈동자와 오른쪽 눈동자의 중심들을 연결하는 선사이의 각(θ)으로 둔다. 템플릿을 생성하기 위하여 사용하는 영상들을 선택하는 방법은 다음과 같다. 첫 번째 영상은

d 가 템플릿들을 생성하기 위한 영상들의 평균에 근접하며, θ 는 비교적 0°에 가까운 영상이 선택된다. 이러한 특성을 가지는 영상은 한 장의 영상으로부터 템플릿을 생성시키는 영상으로도 이용된다. 두 번째 영상은 d 가 평균보다 작으며, θ 는 양의 큰 값을 가지는 영상이 선택된다. 세 번째 영상은 d 가 평균보다 크며, θ 는 음의 큰 값을 가지는 영상이 선택된다. 위에서 언급한 기준에 의거하여 실험에 이용된 임의의 한 모델에 대한 영상들은 그림 2에 나타낸다. 이 영상들의 해상도는 512(가로) × 480(세로) 화소이다. 그림 2의 영상들에서 d 와 θ 는 왼쪽부터 각각 74.1과 3.1°, 50.5와 13.8°, 90.5와 -13.4°이다.



그림 2. 한 모델에 대하여 크기와 회전 각도가 서로 다른 영상들
Fig. 2. Example of images with different scales and in-plane rotations for one model.

템플릿을 생성하기 위하여 선택된 영상들에 대하여 시간에 따라서 변화가 발생할 수 있는 머리카락 영역을 제거하고, 얼굴의 주요 구성 요소들인 눈썹, 눈, 코, 입만 포함되도록 하는 타원형 얼굴 모델^[11, 12]이 적용된다. 추출된 영상의 예는 그림 3의 (b)에 나타낸다. 이 경우에 필요한 두 눈동자의 중심 좌표 값들은 수동으로 결정한다. 이는 눈동자의 중심 좌표를 검출하는 과정^[13, 14]에서 발생할 수 있는 오차를 제거하여 제안하는 방법들의 공정한 평가를 하기 위한 것이다. 방향성이 없는 조명 변화에 정합 오차를 감소시킬 수 있다고 알려져 있는 그래디언트 영상들^[9]은 추출된 타원형의 얼굴로부터 3×3 Sobel 연산자에 의해 생성되며, 이것의 예는 그림 3의 (c)에 나타낸다. 그림 3의 (c) 영상의 명암은 전시 목적을 위해 역으로 한다. 이들 영상들에 대하여 크기와 회전 각도가 불변인 특징들을 추출하기 위하여, 어파인 변환과 로그폴라 사상이 적용된다. 어파인 변환은 특정한 크기와 회전각을 가지는 직각 좌표상의 영상을 다른 크기와 회전각을 가지는 영상으로 변환한다. 로그폴라 사상은 특정한 크기와 회전각을 가지는 직각 좌표상의 영상을 로그폴라 좌표상의 영상

으로 변환한다. 이것은 시선 고착점(fixation point)의 주위에서는 샘플링 밀도가 조밀하고, 시선 고착점에서 거리가 멀어질수록 샘플링 밀도가 감소되는 비선형 사상이다.

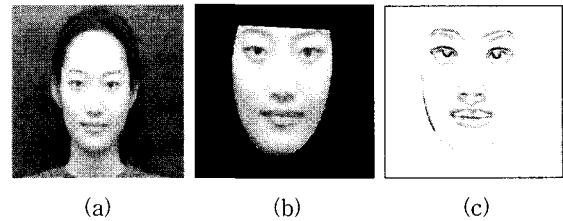


그림 3. (a) 원 영상 (b) 타원형 모델이 적용되어 추출된 영상, 전시 목적을 위해 확대 (c) (b)의 그래디언트 영상
Fig. 3. (a) Original image. (b) Face image extracted by applying the ellipse face model and enlarged for displaying purpose. (c) Gradient image of (b).

원 영상에서 임의의 한 점은 (x, y) 로 두고, 변환된 영상에서의 한 점은 (x', y') 로 두자. 그리고 (x_c, y_c) 는 변환되는 영상의 기준 좌표이다. 이 경우에 어파인 변환은 다음과 같다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (1)$$

여기서 s 는 정규화된 영상에서의 두 눈동자 사이의 거리(d_s)와 기준 영상에서의 두 눈동자 사이의 거리(d_r) 사이의 비율이다. 그리고 ϕ 는 기준 영상에서의 두 눈동자 사이의 회전각(θ_r)과 정규화된 영상에서의 두 눈동자 사이의 회전각(θ_s) 사이의 차이를 나타낸다. 실험에 이용된 d_r 과 θ_r 은 각각 70과 0°이다.

직각 좌표 상에서 임의의 두 물체 사이에 존재하는 크기와 회전각의 차이는 로그폴라 좌표 상에서 각각 수평 이동과 수직 순환 이동의 차이로 표현된다. 직각 좌표 상에서 크기와 회전각이 서로 다른 두 영상을 정규화시키는 과정은 로그폴라 좌표 상에서 그 두 영상을 수평과 수직 순환 이동시켜 정렬하는 과정으로 대응된다. 직각 좌표 상에서 시선 고착 점을 (x_c, y_c) 으로 두자. 로그폴라 영상에서 한 화소의 수평과 수직의 위치를 나타내는 인덱스를 각각 $u=1 \dots U$, $v=1 \dots V$ 로 두면, 로그폴라 사상^[12]은 다음과 같다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + l b^{\rho_u} \begin{bmatrix} \cos(v\psi) \\ \sin(v\psi) \end{bmatrix} \quad (2)$$

여기서 $b = 10^{\frac{\log_{10}(r_m)}{u}}$ 이며, $\rho_u = K \ln r_u$ 이다. 이 경우에 r_m 은 직각 좌표 상에서 시선 고착점으로부터 최대 거리를 가지는 방사선의 크기이며, 이 논문에서는 256을 이용한다. $u_m (= U)$ 은 로그폴라 평면상에서 수평 해상도의 최대 크기를 나타내는데, 이 논문에서는 256을 이용한다. 크기 요소(scale factor) $K = \ln b^{-1}$ 이며, r_u 는 직각 좌표상의 시선 고착점으로부터 방사선 상의 임의의 한 점까지의 거리이다. 그리고 $\psi = \frac{2\pi}{V}$ 이며, 이 논문에서는 $V = 256$ 을 사용한다.

식 (1)과 식 (2)에서 사용한 기준 점과 시선 고착점은 두 눈동자 중심 좌표의 중간 좌표를 이용한다. 그런데 선정된 시선 고착점을 기준으로 일정한 반경 이내의 영역에서는 얼굴을 구성하는 주요 구성 요소가 존재하지 않는다. 이 논문에서는 이 영역을 제거한 나머지 영역만을 이용한다. 이와 같은 과정을 통하여 생성된 로그폴라 영상에서 정합 시간을 단축하기 위하여, 로그폴라 사상된 모든 영상들은 수평축과 수직축으로 일정한 거리와 각도가 되도록 정렬된다. 이 논문에서 사용한 로그폴라 영상의 해상도는 시선 고착점 주위가 제거된 $128(U) \times 256(V)$ 가 사용된다. 이 해상도에 의해 어파인 변환 기반 방법과 로그폴라 사상 기반 방법의 연산 과정에 참여하는 화소의 수의 차를 최소화하기 위하여 실험에서 d_s 을 70으로 한다. 그림 2의 왼쪽 영상에 대하여 식 (1)과 (2)를 적용하여 변환된 영상들의 예는 그림 4에 나타낸다.

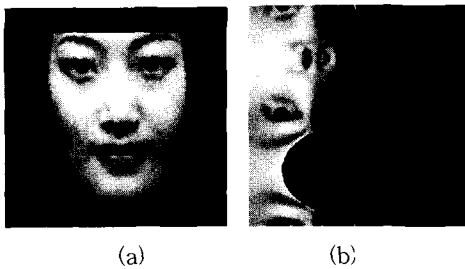


그림 4. 정규화된 영상의 예 (a) 어파인 변환된 영상 (b) 로그폴라 사상된 영상

Fig. 4. Example of normalized face image. (a) Affine transformed image. (b) Log-polar mapped image.

2. 템플릿 생성

크기, 회전, 조명 등의 변화로부터 발생하는 국부적인 왜곡의 영향으로 증가되는 정합 오차를 감소시키기 위하여, 나머지 두 단계의 과정을 새롭게 도입한다. 그 과정은 다음과 같다. 어파인 변환 또는 로그폴라 사상된 영상들은 단순히 모두 더해진다. 그리고 더해진 값들을 선형 또는 비선형 연산을 통해 템플릿이 생성된다. 이 과정에서 선형 연산으로부터 생성되는 템플릿은 평균 템플릿이라 하며, 비선형 연산에 의하여 생성되는 템플릿은 포화 템플릿(saturation template)이라 한다. 각 템플릿을 생성시키는 구체적인 방법은 다음과 같다. 평균 템플릿은 어파인 변환 또는 로그폴라 사상된 3장의 그래디언트 영상들로부터 같은 위치에 있는 화소 값들을 산술적으로 더하여 평균을 취하는 것이다.

$$G_{avg} = \frac{1}{3} \sum_{i=1}^3 G^i_{sng} \quad (3)$$

여기서 G_{avg} 는 평균 그래디언트 템플릿을 나타내며, G_{sng} 는 한 장의 영상으로부터 생성되어 단일 템플릿으로 이용되는 그래디언트 영상을 나타낸다.

포화 템플릿은 어파인 변환 또는 로그폴라 사상된 3장의 그래디언트 영상들을 단순히 더한다. 그리고 더해진 값이 미리 정해진 최대 값보다 크면 그 값으로 화소 값을 결정하고, 그렇지 않으면 더한 값으로 결정하는 것이다. 이 논문에서는 최고 값을 255로 한다.

$$G_{sat} = \begin{cases} 255, & \text{if } \sum_{i=1}^3 G^i_{sng} \geq 255 \\ \sum_{i=1}^3 G^i_{sng}, & \text{otherwise} \end{cases} \quad (4)$$

여기서 G_{sat} 는 포화 그래디언트 템플릿을 나타낸다.

위에서 언급한 방법들을 그림 2의 영상들에 적용하여 생성된 템플릿들은 그림 5에 나타낸다. 그림 5의 영상들의 명암은 전시 목적에 의해 역으로 한다. 그림 5의 (a)에 나타난 영상들은 어파인 변환에 의해 생성된 템플릿들이며, 그림 5의 (b)에 나타난 영상들은 로그폴라 사상에 의해 생성된 템플릿들이다. 그림 5의 왼쪽부터 한 장의 영상으로부터 생성된 템플릿이며, 그 다음은 평균 템플릿이며, 가장 오른쪽은 포화 템플릿이다. 그림 5를 살펴보면 가장 왼쪽에 나타난 단일 템플릿보다 평균 템플릿은 다소 평활화(smoothing) 되었음을 알 수 있다. 그리고 가장 오른쪽의 포화 템플릿은 나머지 두 종류의 템플릿보다 얼굴 구성 요소들의 윤곽선이

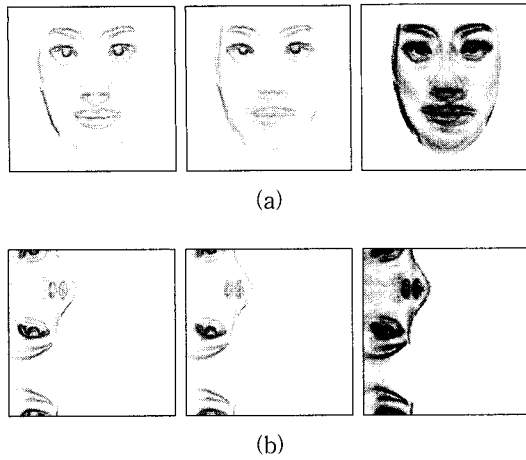


그림 5. 제안하는 방법에 의해 생성된 템플릿들의 예. 첫 번째는 단일 템플릿 (G_{sng}), 두 번째는 평균 템플릿 (G_{avg}), 세 번째는 포화 템플릿 (G_{sat}). (a) 어파인 변환이 적용 (b) 로그폴라 사상이 적용

Fig. 5. Example of the proposed templates. First image is single template (G_{sng}), second one is average template (G_{avg}) and third one is saturation template (G_{sat}). (a) Generated by affine transformation. (b) Generated by log-polar mapping.

뚜껑게 나타남을 알 수 있다. 이는 정합과정에서 왜곡 현상을 억제시킬 수 있다.

III. 디더링과 FFT

이 논문에서는 얼굴의 크기와 회전각이 서로 다른 영상들에 어파인 변환 또는 로그폴라 사상을 적용하기 위하여, 두 눈동자들의 중심 좌표를 수동으로 검출하여 이들 좌표들의 중간 좌표를 이용한다. 그런데 눈동자의 중심을 검출하는 자동화된 방법은 얼굴의 크기, 회전, 조명 변화 등의 요인들로 인하여 눈동자의 위치가 정확하게 검출되지 않는 경우가 종종 발생하게 된다. 이러한 경우에 템플릿과 입력 영상에 어파인 변환이나 로그폴라 사상을 적용하여도 얼굴 영역에서의 국부적인 왜곡으로 인하여, 그 두 영상 사이의 정합 오차가 증가될 수 있게된다. 이 문제를 해결하기 위하여, 디더링과 FFT 방법을 이용한다.

이 논문에서는 입력 영상과 템플릿 사이의 정합 오차를 감소시키기 위하여, 템플릿을 입력 영상의 평면에

서 이동시키는 것을 디더링으로 정의한다. 로그폴라 사상을 기반으로 하는 정합에서의 디더링은 템플릿을 수평으로 좌우 이동하며, 수직으로는 상하 순환 이동시킨다. 어파인 변환과 로그폴라 사상 기반 방법 사이의 공정한 평가를 위해 로그폴라 평면에서의 디더링이 직각 좌표상에서는 크기와 회전 변화로 나타나므로, 어파인 변환 기반 방법에서의 디더링은 템플릿의 크기와 회전 각도를 변화시킨다. 디더링에 의해 계산된 유사도의 값 가운데에서 최고 값이 최종적으로 두 영상 사이의 유사도로 결정된다. 서로 다른 크기와 회전 변화를 가지는 동일한 형태의 두 영상의 FFT에 의한 스펙트럼의 크기는 동일하지 않게 된다. 다만 크기와 회전각이 동일하고 수평과 수직의 이동량이 서로 다른 두 영상의 FFT에 의한 계산된 스펙트럼의 크기는 동일하게 된다. 그런데 로그폴라 사상을 기반으로 하는 정합에서 디더링은 계산량의 증가를 가져오게 된다. 이러한 FFT의 성질과 디더링에서 계산량이 증가되는 문제를 해결하기 위하여 이 논문에서는 정렬된 로그폴라 영상에 FFT를 적용한다. FFT를 적용하여 계산된 스펙트럼의 크기 가운데에서 얼굴 영상의 특징으로 인하여 DC 성분에 해당되는 스펙트럼 크기는 다른 주파수 영역에서의 스펙트럼 크기보다 상대적으로 매우 커지게 된다. 따라서 서로 다른 모델 사이의 식별력이 매우 저하되는 현상이 발생하게 된다. 이 논문에서는 스펙트럼의 크기 값 가운데에서 DC 성분의 영역을 중심으로 3×3 화소 영역을 제거한 나머지 스펙트럼 성분을 이용하여 입력 영상과 템플릿 영상 사이의 유사도를 측정한다.

IV. 실험 결과 및 고찰

제안하는 얼굴 인식 방법은 서로 다른 크기와 회전 각을 가지는 영상들에 어파인 변환이 적용되어 생성된 템플릿을 이용하는 방법, 로그폴라 사상이 적용되어 생성된 템플릿을 이용하는 방법, 그리고 로그폴라 영상에 기반한 FFT를 이용하는 방법으로 분류된다. 그리고 각 방법에서 다시 한 장의 영상을 이용하여 생성된 단일 템플릿, 3장의 영상들로부터 생성된 평균 및 포화 템플릿을 이용하는 방법들로 분류된다. 제안하는 방법들 사이의 성능을 비교하기 위하여 51명의 모델로부터 얼굴 영상을 획득하여 인식 실험에 이용한다. 제안된 방법들은 Visual C++ 언어로 구현하며, Celeron 500MHz를 가지는 PC를 이용하여 인식 실험을 수행한다. 템플릿

영상과 입력되는 얼굴 영상의 유사도 계산에 이용되는 NCC는 다음 식 (5)와 같다. 식 (5.1)은 어파인 변환 기반 방법에 사용되며, 식 (5.2)는 로그폴라 사상 기반 방법에 이용된다.

$$NCC = \begin{cases} \max_{l_z, \Delta\phi} \frac{\langle f(x, y)g(x, y, l_z, \Delta\phi) \rangle - \langle f(x, y) \rangle \langle g(x, y, l_z, \Delta\phi) \rangle}{\sigma_f \sigma_g} & (5.1) \\ \max_{\Delta u, \Delta v} \frac{\langle f(u, v)g(u \pm \Delta u, v \pm \Delta v) \rangle - \langle f(u, v) \rangle \langle g(u \pm \Delta u, v \pm \Delta v) \rangle}{\sigma_f \sigma_g} & (5.2) \end{cases}$$

여기서 $f(\cdot)$ 는 어파인 변환 또는 로그폴라 사상에 의하여 생성된 입력 영상을 나타내며, $g(\cdot)$ 는 템플릿을 나타낸다. 그리고 $\langle \cdot \rangle$ 는 평균을 계산하는 연산자이며, σ_f 는 $f(\cdot)$ 의 표준편차를, σ_g 는 $g(\cdot)$ 의 표준편차를 나타낸다. (5.1)의 l_z 와 $\Delta\phi$, 식 (5.2)의 Δu 와 Δv 는 각각 직각 좌표상과 로그폴라 평면상에서 디더링의 계수를 나타낸다. 이 논문에서는 사전 실험을 통하여 $\Delta u = \pm 1$, $\Delta v = \pm 2$ 로 결정하였다. 직각 좌표계와 로그폴라 좌표계 사이의 디더링 관계는 다음과 같다.

$$\Delta\phi = \phi \times \Delta v \quad (6)$$

$$l_z = b^{\Delta u} \quad (7)$$

여기서, $\Delta\phi$ 는 직각 좌표계에서 회전각이며, l_z 는 크기 요소이다. 만일 $l_z > 1$ 이면 템플릿이 커지고, 반대는 템플릿이 줄어든다.

1. 얼굴 영상 데이터베이스

이 논문에서는 얼굴의 크기와 회전 변화를 가지는 다양한 정면 얼굴 영상들로 데이터베이스를 구축한다. 참여한 모델들은 한국인이며, 남자 22명과 여자 29명으로 구성된다. 얼굴 영상을 획득하기 위하여 Pulnix사의 TM7CN 모델인 흑백 CCD 카메라와 이스라엘의 Bar-Gold Electronics 사의 IVP-150 모델의 프레임 그래버를 이용한다. 각 모델로부터 크기와 회전 변화를 가지는 얼굴 영상을 획득하기 위하여, 크기 변화는 두 눈동자 사이의 거리가 약 40화소에서 120화소까지의 분포를 가지도록 5단계로 분류한다. 그리고 크기 변화의 각 단계에서 약 -25° 에서 25° 까지 3단계의 회전 변화를 가지도록 한다. 이러한 방법으로 51명의 모델로부터 765장의 영상을 획득한다. 각 모델로부터 획득된 15장의 얼굴 영상에서 3장의 영상은 템플릿을 생성하기 위하여 이용되며, 12장은 인식 방법의 성능을 확인하기 위하여 이용된다. 데이터베이스에 포함되는 얼굴 영상의 예는 그림 6에 나타내며, 그림 7에서는 실험에 사



그림 6. 실험에 이용된 다양한 크기와 평면상의 회전 각도를 가지는 얼굴 영상들의 예

Fig. 6. Example of sample images with various scales and rotations of faces in test database.

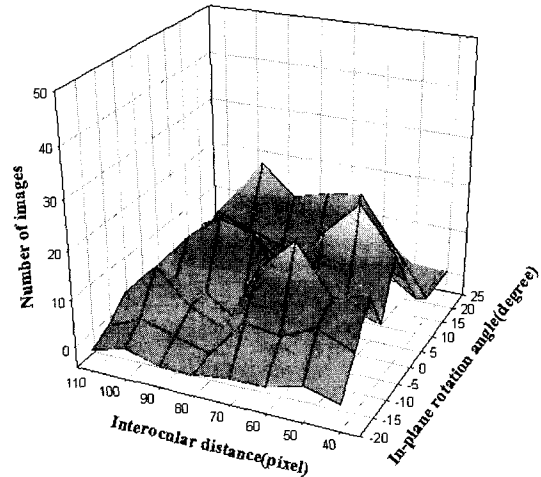


그림 7. 실험에 사용된 얼굴 영상의 거리 및 회전 각도에 대한 히스토그램

Fig. 7. Histogram of test images in terms of distances and rotation angles of face images in test database.

용된 얼굴 영상들의 두 눈동자 사이의 거리 및 회전 각도에 따르는 영상의 수에 대한 히스토그램을 나타낸다.

2. 결과 및 고찰

등록된 모델들 가운데에서 임의의 얼굴 영상이 입력되면 사전에 등록된 템플릿과의 유사도가 식 (5)로부터 계산되어진다. 계산된 유사도를 가운데에서 가장 높은 값을 가지는 템플릿이 입력된 얼굴 영상과 동일한 모델인 경우의 인식률은 TOP1으로 계산된다. 그리고 계산된 유사도들로부터 상위 3번째까지 등록된 템플릿

가운데에서 입력된 얼굴 영상과 동일한 모델이 포함되는 경우의 인식률은 TOP3로 계산된다. 이와 같은 과정으로부터 제안된 방법들에 대한 실험 결과는 표 1에 나타낸다.

표 1. 제안된 방법들의 인식률
Table 1. Recognition results of the proposed methods.

Template	Recognition rate(%)	Log-polar mapping based method		Log-polar image based FFT method	Affine transformation based method	
		No Dithering	Dithering		No Dithering	Dithering
G_{avg}	Top1	73.20	79.41	84.48	78.27	83.50
	Top3	80.07	86.93	89.05	84.64	87.91
G_{avg}	Top1	91.18	94.44	96.24	95.43	96.40
	Top3	94.12	95.75	97.06	97.55	97.71
G_{cat}	Top1	95.43	96.24	93.14	95.26	96.57
	Top3	96.90	97.88	96.08	97.71	98.20

표 1에 나타낸 바와 같이 입력 얼굴의 크기와 회전 변화에 대한 실험 결과에서는 단일 영상으로부터 생성된 템플릿을 이용하는 방법보다 이 논문에서 제안된 평균 템플릿 또는 포화 템플릿을 이용하는 방법이 모든 방법들에서 높은 인식률을 나타낸다. 이는 얼굴의 크기와 회전 변화가 서로 다른 영상들을 결합하여 생성된 평균 및 포화 템플릿은 국부적인 왜곡에 대한 특성들이 암시적으로 반영되어 비록 입력 영상에서 국부적인 왜곡이 발생하더라도 그 두 영상 사이에서 계산된 정합 오차는 감소되기 때문이다.

디더링을 수행하는 인식 방법들에서는 어파인 변환 기반 방법은 포화 템플릿을 이용하는 것이 평균 템플릿을 이용하는 것보다 약간의 높은 인식률을 보이지만, 로그폴라 사상 기반 방법에서는 포화 템플릿을 이용하는 것이 평균 템플릿을 이용한 것보다 약 2%의 높은 인식률을 나타낸다. 이것은 포화 템플릿이 약간 평활된 평균 템플릿보다 얼굴의 특징을 조금 더 잘 나타내기 때문이다.

또한, 인식 과정에서 디더링의 수행 유무에 따르는 인식 결과에서는 디더링을 수행한 것이 수행하지 않은 것보다 모든 방법들에서 높은 인식 결과를 보인다. 세 장의 영상들로 생성된 템플릿이 입력 얼굴 영상의 크기와 회전 변화, 조명의 변화, 잡음 영향, 로그폴라 또는

어파인 변환 과정에서 발생하는 국부적인 왜곡을 모두 흡수하기는 어렵다. 이 경우에 디더링을 수행하면 템플릿과 서로 다른 국부 왜곡을 가지는 입력 얼굴 영상 사이의 정합 오차를 다소 감소시킬 수 있기 때문이다.

제안된 방법들 가운데에서는 어파인 변환을 이용한 것이 다른 두 방법들보다 조금 높은 인식 결과를 보인다. 이는 직각 좌표 상에서 비선형 표본화를 통하여 데이터 감축이 발생하는 로그폴라 사상보다 어파인 변환에서는 데이터의 감소가 일어나지 않기 때문이다.

제안된 얼굴 인식 방법들은 입력 영상들의 크기와 회전 각도에 따른 인식률을 보기 위하여, 그림 8에서 크기 변화를 8 단계, 회전 변화는 10 단계로 나누어 인식 결과를 나타낸다. 그림 8의 (a), (b)를 살펴보면 인식 방법과는 무관하게 제안된 평균과 포화 템플릿이 이용되는 경우의 인식률은 크기와 회전 변화를 가지는 입력 영상들에 대해서도 상대적으로 안정된 결과를 보였다. 반면, 한 장의 영상으로 생성된 템플릿을 이용하는 방법들의 성능을 크기와 회전 변화에 따라 많은 변동이 있음을 알 수 있다. 로그폴라 영상 기반 FFT를 이용한 방법에서도 다른 두 방법과 유사한 결과를 나타내었다.

각 인식 방법들의 수행 시간은 표 2에 나타낸다. 표 2에 나타낸 수행 시간은 한 장의 얼굴 영상과 두 눈동자들의 좌표들이 입력된 후, 하나의 템플릿과의 정합에 소요되는 시간으로 입력 612장 영상들에 대한 평균이다. 이 경우에 수행 시간을 단축시키기 위하여, 식 (5)에 나타낸 NCC에서 템플릿과 디더링된 템플릿들의 평균과 분산은 오프라인(off-line)으로 사전에 계산된다. 표 2에서 디더링을 수행하지 않은 방법은 입력 영상과 하나의 템플릿 사이의 정합이며, 디더링을 수행하는 방법은 입력 영상과 15장의 디더링된 템플릿들과의 정합하여 최고의 유사도를 가지는 템플릿을 찾는 것까지의 소요 시간이다. 로그폴라 사상은 로그폴라 평면의 해상도와 최대 반경이 사전에 결정되면, 로그폴라 사상된 좌표들을 오프라인에서 미리 계산되어, 룩업 테이블(look-up table)에 저장될 수 있다. 이것은 입력 영상의 크기와 회전 변화에 무관하다. 그러나, 어파인 변환은 입력 영상의 크기와/또는 회전 각도가 기준 값과 다르면 매번 변환이 일어난다. 표 2에서 나타낸 것과 같이 디더링을 수행하지 않은 방법에서는 로그폴라 사상 기반 방법이 어파인 변환 기반 방법보다 약 21 msec 빠르고, 디더링을 수행하는 방법에서는 약 25 msec 빠르

다. 인식의 전 과정에서 소요되는 시간의 70~90%가 전처리 과정(타원 추출, 그래디언트 추출)에서 소요된다. 제안하는 세 가지 방법 중에서 로그폴라 영상 기반 FFT 방법이 가장 많은 시간이 소요되는 데, 이는 FFT 연산 때문이다. 그러나, FFT을 수행하는 전용 하드웨어를 사용한다면 시간이 단축될 것이다.

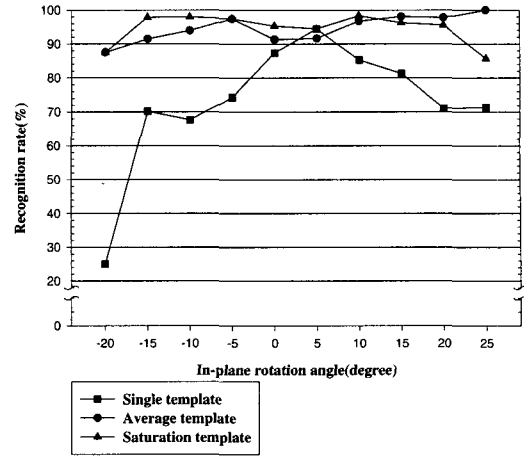
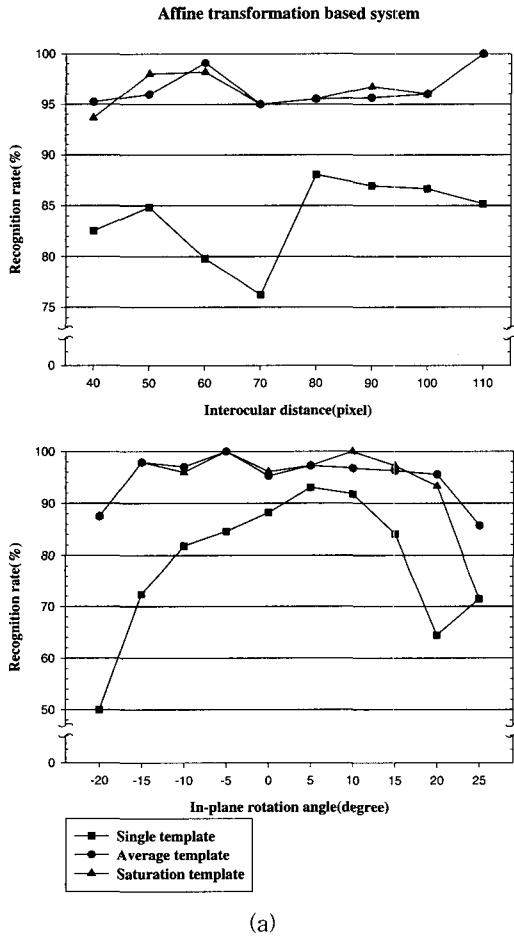
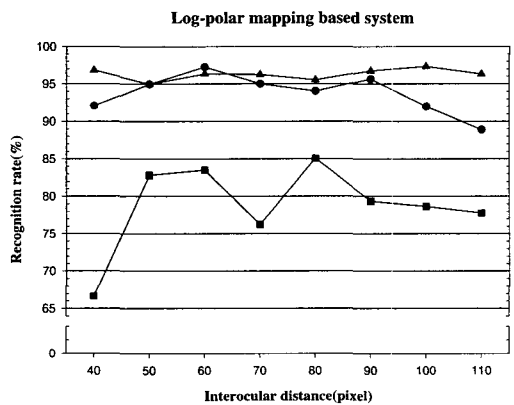


그림 8. 얼굴의 크기와 평면상의 회전 변화에 따르는 방법들의 인식률 (a) 어파인 변환에 기반한 방법 (b) 로그폴라 사상에 기반한 방법
Fig. 8. Recognition rate according to scale and in-plane rotation variations of face images. (a) Affine transformation based method. (b) Log-polar mapping based method.

표 2. 한 모델에 대한 제안된 방법들의 평균 수행 시간

Table 2. Average processing time of the proposed methods per model.

Approaches	Processing time	
	No Dithering	Dithering
Log-polar mapping based method	228.14 msec	264.82 msec
Log-polar image based FFT method	481.20 msec	
Affine transformation based method	248.73 msec	303.17 msec



V. 결 론

이 논문에서는 다양한 크기와 평면상에서 회전 변화가 존재하는 정면 얼굴 영상들에 대하여 새로운 템플릿 기반 정합을 제안하였다. 템플릿 기반 정합에서 가장 중요한 요소인 템플릿을 생성시키기 위하여, 얼굴 영상의 크기와 회전 각도가 다른 다수의 영상들로부터 평균과 포화 템플릿을 생성하는 방법을 제안하였다. 크기와 회전 각도가 다른 영상들에 불변인 특징을 가지는 어파인 변환, 로그폴라 사상, 그리고 로그폴라 영상

기반 FFT 방법을 적용하였다. 각 방법들은 제안된 템플릿들을 이용한 것이 한 장의 영상으로부터 생성된 단일 템플릿을 이용한 것보다 인식 성능이 우수함을 실험 결과로부터 확인하였다. 또한 디더링을 수행하는 것이 수행하지 않는 것보다 모든 방법들에서 좋은 결과를 보였다. 어파인 변환을 이용한 방법의 인식률이 로그폴라 사상을 이용한 방법과 로그폴라 영상에 기반한 FFT 방법보다 우수하며, 수행 시간 측면에서는 로그폴라 사상을 이용한 방법이 가장 빨랐다.

참 고 문 헌

- [1] S. Pankanti, R.M. Bolle, and A. Jain, *Biometrics: Personal identification in networked society*, Kluwer Academic, Norwell, Mass., 1999.
- [2] P.J. Philips, A. Martin, C.L. Wilson, and M. Przybocki, "An introduction to evaluating biometric systems," *IEEE Computer: Innovative Technology for Computer Personals*, pp. 56~64, Feb. 2000.
- [3] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042~1052, 1993.
- [4] 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.
- [5] X. Jia and M.S. Nixon, "Extending the feature vector for automatic face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 12, pp. 1162~1176, 1995.
- [6] J. Zhang, Y. Yan, and M. Lades, "Face recognition: Eigenface, elastic matching, and neural nets," *Proc. of The IEEE*, vol. 85, no. 9, pp. 1423~1435, 1997.
- [7] G. Wolberg, *Digital Image Warping*, pp. 42~45, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [8] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775~779, 1997.
- [9] 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.
- [10] M. Bolduc and M.D. Levine, "A review of biologically motivated space-variant data reduction models for robotic vision," *Computer Vision and Image Understanding*, vol. 69, no. 2, pp. 170~184, 1998.
- [11] S.-I. Chien and I. Choi, "Face and facial landmarks location based on log-polar mapping," S.-W. Lee, H.H. Buelthoff, and T. Poggio Eds., *BMCV 2000, Lecture Notes in Computer Science 1811*, pp. 379~386, 2000.
- [12] 최 일, "로그폴라 영상 변환을 이용한 얼굴 검출 및 인식", 경북대학교 박사학위 논문, 2000년 12월
- [13] K. -M. Lam, H. Yan, "Location and Extracting the Eye in Human Images," *Pattern Recognition*, vol. 29, no. 5, pp. 771~779, 1996.
- [14] J. Y. Deng, F. Lai, "Region-Based Template Deformation and Masking for Eye-Feature Extraction and Description," *Pattern Recognition*, vol. 30, no. 5, pp. 403~419, 1997.

저 자 소 개



金 文 甲(正會員)

1990年 2月 : 경북대학교 전자공학과 졸업(공학사). 1992年 2月 : 경북대학교 대학원 전자공학과 졸업(공학석사). 2000年 2月 : 경북대학교 대학원 전자공학과 박사과정수료. 1992年 2月~2000年 2月 : 대우

중공업(주). 2000年 3月~현재 : 용인송담대학 디지털전자정보과 전임강사. <관심분야 : Computer vision, Pattern recognition>

崔 一(正會員) 第39卷 SP編 第1號 參照

현재 : 구미1대학 정보통신과

秦 成 一(正會員) 第39卷 SP編 第1號 參照

현재 : 경북대학교 전자전기컴퓨터학부