

얼굴의 geometry 특징을 이용한 다중해상도 템플릿 매칭 얼굴 특징 추출법

윤성욱, 김재민, 조성원, 최경삼, 백성욱
홍익대학교 전기 정보 제어과
전화 : 02-320-1493 / 핸드폰 : 016-379-4265
E-mail : rudra@hanafos.com

Abstract

This paper presents a new template matching method for finding facial feature points. Instead of matching each template to its corresponding feature point separately, the present method matches a set of element templates simultaneously. The set of templates can be placed on the space predefined by the geometrical characteristics of human faces. As a result, the search area for template matching is very small compared with a conventional method. This makes the presented method very robust and accurate. Experiment results show that the presented method results in good performance in various illuminance environments and poses.

I. 서론

생체 인식 분야는 홍채 인식, 지문 인식, 장문 인식, 음성 인식, 얼굴 인식 등 많은 분야로 이루어져 있다. 그 중에서 특히 얼굴 인식은 사용자와 인식기 간의 접촉을 필요로 하지 않는다는 점, 사용자의 편의성이 높다는 점, 특별한 인식 장비를 사용하지 않고 CCTV 등의 널리 보급된 장비를 사용할 수 있다는 점 등, 많은 장점을 가진 생체 인식 분야이다.

얼굴 인식은 크게 PCA를 이용한 방법과 template matching을 이용한 방법으로 나눌 수 있다[1]. 본 논문에서는 얼굴의 geometry 특성과 다중 해상도 방법, illuminance normalize 등을 이용하여 얼굴의 각 부위의 위치를 template matching으로 찾는 방법을 연구, 실험 하였다.

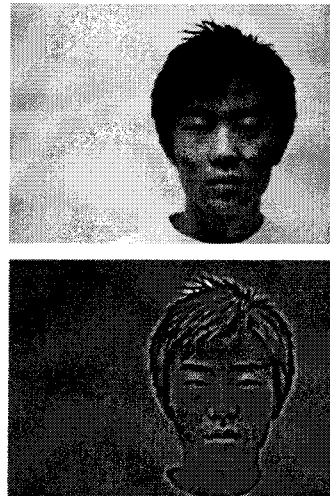
II. Preprocessing

2.1 Illuminance normalization

카메라로부터 획득된 영상은 획득될 당시의 장소, 조명 등 환경에 매우 큰 영향을 받는다. 신뢰도 높은 얼굴 인식을 하기 위해서는 우선 이러한 영향을 최소화해야 한다. 본 논문에서는 조명에 의한 영상의 밝기 왜곡을 최소화하기 위한 방법으로 illuminance normalization을 사용한다.

$$\hat{I}(x, y) = I(x, y) / m$$

여기서 $I(x, y)$ 는 영상의 밝기, m 은 local mean이다. 이미지를 local mean으로 나누면 영상의 밝기 값들이 1 근처에 존재하게 되므로 histogram stretch 방법을 사용하여 0~255로 펼쳐준다.



< 그림 1 > Illumination normalization

2.2 Gaussian down sampling

획득된 영상을 실시간으로 처리할 수 있게 하기

위해서[1], 또 노이즈의 영향을 줄이기 위해서, Gaussian down sampling을 통한 다중해상도 방법을 사용한다.

$$\hat{I}^2(x, y) = \sum_{i=-n}^n \sum_{j=-n}^n \hat{I}(2x+i, 2y+j) f(i, j)$$

여기서 $f(i, j)$ 는 $2n+1$ by $2n+1$ 크기의 normalized gaussian window filter이다.

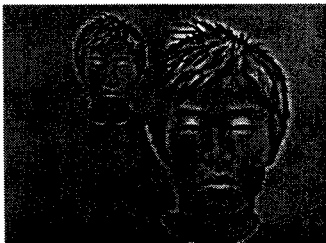
본 연구에서는 아래와 같은 3×3 크기의 filter를 사용하였다.

표 1. 3×3 normalized gaussian window filter

0.0625	0.1250	0.0625
0.1250	0.2500	0.1250
0.0625	0.1250	0.0625

위 filter는 다음 식을 만족한다.

$$\sum_{i=-1}^1 \sum_{j=-1}^1 f(i, j) = 1$$



< 그림 2 > Gaussian down sampling

III. 얼굴의 geometry 특징을 이용한 template matching

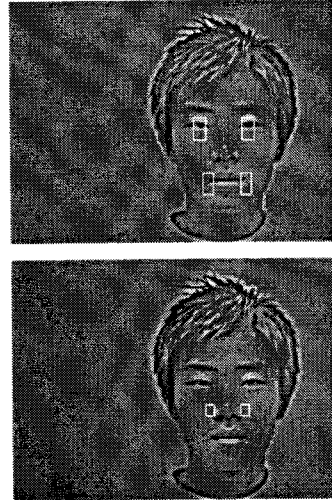
3.1 얼굴의 geometry 특징과 templates

얼굴의 가장 큰 특징이라 할 수 있는 눈, 코, 입을 일반적으로 사람들마다 비슷한 위치에 존재하게 된다. 또한 그 위치들 간의 상호 연관관계, 즉 눈, 코, 입의 기하학적 배열도 대부분 동일하다고 할 수 있다[2].

따라서 이러한 얼굴의 geometry 특징을 이용하면 template matching을 하는데 있어서 searching area를 줄이는 효과를 얻을 수 있을 뿐만 아니라 잘못된 템플

릿 매칭 결과가 나오는 것을 최소화 할 수 있다.

본 연구에서는 얼굴의 geometry 특징을 보다 잘 이용하기 위해 아래 그림과 같은 템플릿을 사용하였다.



< 그림 3 > Templates

Template을 선택할 때 입의 경우에는 웃고 있을 때와 보통 때의 특징 차이가 뚜렷하므로 두 가지 경우 모두에 대해서 각각 템플릿을 선택하였다. 따라서 좌우 눈, 좌우 코, 보통 입의 좌우 꼬리와 웃는 입의 좌우꼬리의 전부 8개의 템플릿을 선택, 저장하였다.

3.2 Matching algorithm

(1) Normalized correlation

카메라로부터 인식을 하기위한 영상이 들어오면 그 영상에 대해서 미리 저장해 놓은 8개의 템플릿에 대하여 다음과 같은 식을 이용하여 normalized correlation 값을 모두 구한다[3].

$$\langle I, T \rangle = \frac{\langle I \cdot T \rangle - \langle I \rangle \langle T \rangle}{\sigma(I) \sigma(T)}$$

where, I : image, T : template,

$\langle \rangle$: average operator,

$\langle I \cdot T \rangle$: inner product,

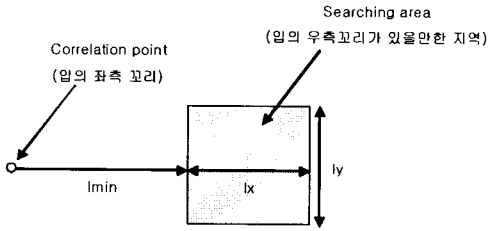
σ : standard deviation

이때 correlation array에 저장되는 값은 -1에서 1사이가 되고, 1에 가까울수록 템플릿과 이미지가 유사함을 뜻한다.

(2) Mouth detection

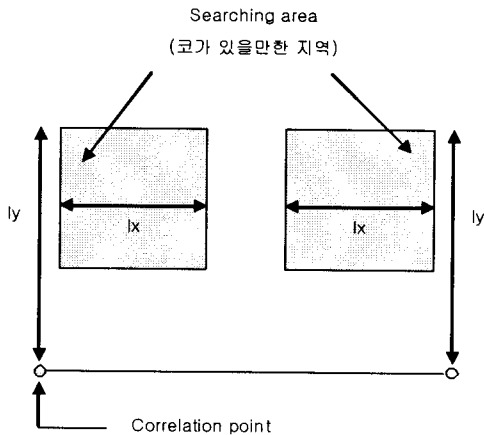
앞에서 구해놓은 8개의 correlation array중에 먼저 좌우 입꼬리의 correlation array 4개를 사용하여 입을 구한다.

입의 경우 좌측과 우측 꼬리 사이의 최소간격(l_{min}), 최대 간격(l_{max}), 좌측 꼬리에 비한 우측 꼬리의 상대적 높이차(l_y)를 다양한 데이터베이스를 통해서 미리 구해놓는다. 이렇게 구해진 값들을 통해 우측 꼬리가 있을만한 searching area를 설정한다. 좌측의 입꼬리로부터 구해진 correlation array의 각 point에 대해 이러한 searching area안에서 우측 꼬리의 correlation 값이 최대가 되는 위치를 찾고 그 위치 값과 두 array의 correlation 값들의 합을 저장한다.



< 그림 4 > Mouth searching

보통의 입에 대해 위의 과정을 마친 후에 웃는 입에 대해서도 같은 과정을 반복한다. 구해진 두 개의 array에 저장된 correlation 값들을 서로 비교한 후 각 point에서 큰 값을 저장하고 그 값에 따른 위치도 역시 저장한다.



< 그림 5 > Nose searching

(3) Mouth-Nose detection

앞의 과정에서 구한 Mouth array에 대해 거리변수

를 코와 입사이의 높이 차이(l_y), 좌우 거리차이(l_x)로 잡고 nose template으로 구한 두 correlation array에 대해 위의 과정을 반복한다.

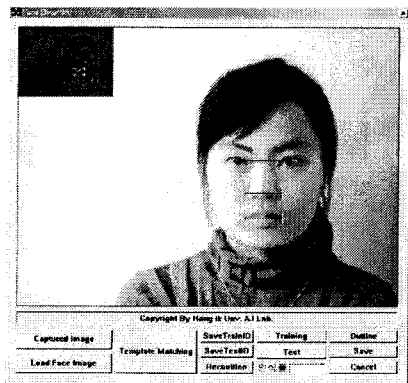
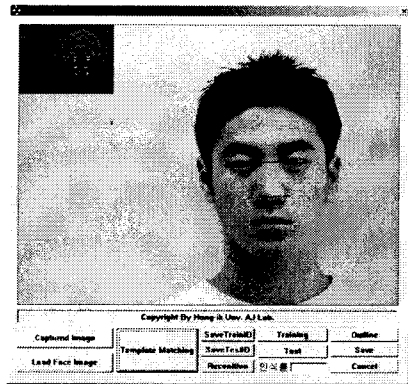
(4) Eye detection

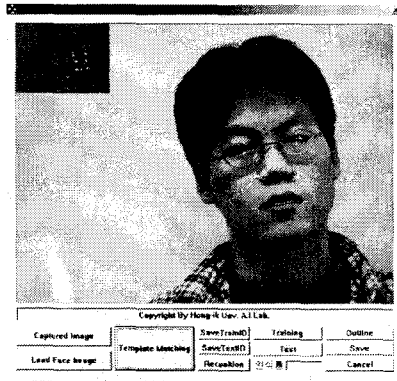
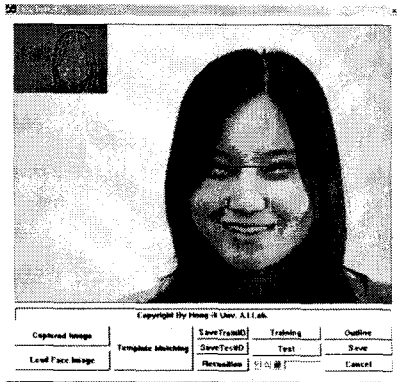
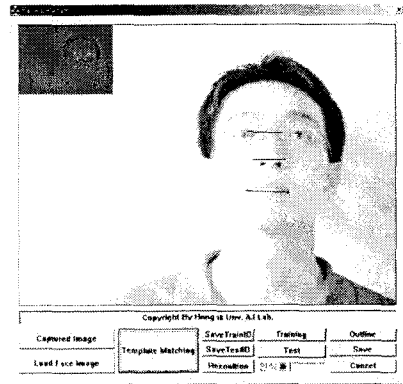
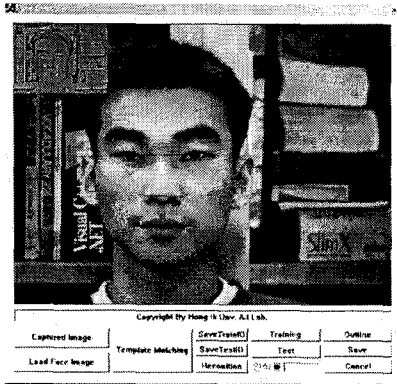
눈 템플릿으로 구한 correlation array에 대해 눈과 코사이의 높이차, 거리차를 이용하여 위와 같은 방식으로 하나의 array에 correlation 값들과 눈의 위치라 예상되는 점을 저장한다.

위의 과정을 모두 거친 후에 저장된 array에서 correlation 값이 가장 값과 연관된 눈, 코, 입의 위치를 선택한다.

IV. 실험 결과

본 연구에서는 75명이 다양한 환경에서 여러 가지 포즈를 취하여 찍은 각각 10장의 영상에 대해서 템플릿 매칭을 하고 그 결과를 직접 확인한 결과 거의 모든 영상에서 눈, 코, 입의 위치를 정확하게 찾아냄을 볼 수 있었다.





V. 결론

실험 결과 본 논문에서 제시한 새로운 템플릿 매칭 방법이 단순히 각각의 템플릿을 따로 적용하는 방식보다 다양한 조명 환경 및 포즈에 대해 매우 안정적으로 동작함을 볼 수 있었다.

본 논문은 얼굴의 geometry 특성을 이용하여 얼굴 주요 부위의 위치를 찾아냄으로써 얼굴인식을 하기 위한 중요한 정보를 추출하는 방법을 제시하였고, 그것은 여러 가지 다양한 얼굴 인식 방법에 응용될 수 있을 것이라 사료된다.

참고문헌

- [1] R. Brunelli and T. Poggio : 'Face Recognition: Features versus Templates', Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 15, Issue 10, Oct. 1993. pp. 1042-1052.
- [2] Ming-Hsuan Yang, D.J. Kriegman and N. Ahuja : 'Detecting Faces in images : a survey', Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume 24, Issue 1, Jan. 2002. pp. 34-58.
- [3] Shinichi Yosimura and Takeo Kanade : 'Fast Template Matching Based on the Normalized Correlation by Using Multiresolution Eigenimages', Proceeding of the IEEE/RSJ/GI International Conference on, vol. 3, 1994, pp. 2086-2093.