

템플릿과 타원정보를 이용한 얼굴검출에 관한 연구

강우석, 김현술, 박남준, 박상희
연세대학교 전기공학과

A Study on Face Detection Using Template Matching and Elliptical Information

Woo-Seok Kang, Hyun-Sool Kim, Nam-Jun Park, Sang-Hui Park
Dept. of Electrical Eng., Yonsei Univ.

Abstract - This paper proposes a new segmentation method of human faces from grey scale images with clutter using a facial template and elliptical structure of the human head. Face detection technique can be applied in many areas of image processing such as face recognition, composition and computer graphics. Until now, many researches about face detection have been conducted, and applications in more complicated conditions are increasing. The general case is more in a complicated background than in a simple one, and a image with not only one face. Research and development of face detection in such a general case are growing rapidly, and the necessity for that is increasing continuously. Sirohey proposed a face detection method using linearized elliptical equation. The method designed in this paper is improved to be applicable even in the more general cases like where the face is much smaller than the image size and with many faces in one image using template matching and elliptic fitting technique.

1. 서 론

얼굴인식기술(face recognition technology : FRT)은 주어진 화상 내의 얼굴을 분석하여 기준의 데이터베이스 안에서 일치하는 얼굴을 찾아내는 기술로서, 보안시스템이나 법 집행목적, 상업적 목적 등으로 사용되어 그 응용범위가 매우 다양하며, 최근 들어 아주 크게 주목받고 있다[1].

얼굴인식은 크게 세 단계로 나뉘어지는데, 첫째 얼굴 검출, 둘째 특징추출, 마지막으로 인식의 단계가 그것이다. 이 중 얼굴검출은 전체 인식시스템에 처리할 얼굴 부분의 데이터만을 제공해 주므로 시스템에서 아주 큰 비중을 차지한다. 얼굴검출은 또한 얼굴인식의 첫 단계로서뿐만 아니라 독립적인 분야로서 많은 응용분야를 가지고 있어서, 연구의 필요성이 늘어가고 있다.

얼굴검출은 입력화상의 배경에 따라 크게 달라진다. 배경이 단순한 경우, 간단한 법칙[2]만을 사용하여 간단하게 얼굴검출이 이루어질 수 있지만, 배경이 복잡한 경우에는 검출이 쉽게 이루어지지 않는다. 동화상의 경우에는 배경이 복잡한 경우에도, 얼굴의 움직임을 단서로 해서 배경과 얼굴을 분리해 낼 수 있다[3]. 정지화상의 경우, 배경이 복잡한 영상, 여러 명의 얼굴이 있는 영상 등에 대한 연구는 현재 활발히 진행중이며 그 필요성이 증대되고 있다. Sirohey는 복잡한 배경에서 선형화된 타원방정식을 이용하여 얼굴의 타원형을 검출하는 방법

을 제시하였다[4].

본 논문에서는 수평 에지를 단서로 인간이 얼굴을 인식하는 가장 간단한 원리를 이용한 얼굴 템플릿을 고안 하여, 배경이 복잡한 영상에서 정면얼굴의 눈과 코·입 영역을 템플릿 정합을 이용하여 찾아내고, 그 부분의 주변에서 얼굴에 해당하는 타원형을 찾아내는 알고리듬을 제안한다. 타원 검출 이전에 얼굴의 눈과 코·입 부분을 찾아내는 전처리과정을 둘으로써 Sirohey의 타원검출법에 비해 좀 더 정확한 타원검출이 가능해지고, 전체 처리시간이 단축되며, 좀 더 일반적인 경우인 여러 명의 얼굴이 존재하는 영상에서도 검출이 가능해졌다.

2. 템플릿 정합 과정

본 논문에서 사용한 템플릿은 Sobel 마스크를 이용하여 얻은 양쪽 눈과 입의 수평 에지를 바탕으로 한 템플릿에 정합하는 부분을 찾아낸다. 이 과정을 그림 1에 나타내었다.

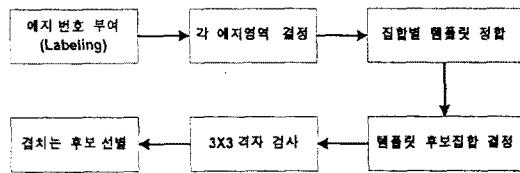


그림 1 템플릿 정합과정

우선 각 수평 에지좌표 (x, y) 들에 번호를 부여한다. 다음 세 개씩의 에지를 묶어서 부여한 번호의 에지 L_i , $i \in \{1, 2, 3\}$ 들마다의 위치 $x_{min}, x_{max}, y_{min}, y_{max}$, 크기 x_{size}, y_{size} , 중심 x_0, y_0 를 구한다. 에지들간의 거리 D_1 과 D_2 는 다음의 식 (1)을 이용해 계산한다.

$$\begin{aligned} D_1 &= |y_2 - y_1|, \\ D_2 &= |x_0 - (x_1 + x_2)/2| \end{aligned} \quad (1)$$

다음 식 (2)와 같이 그 집합과 기본 템플릿과의 비용 함수값 ϵ 을 계산하여 이 값이 문턱값 τ 보다 작고, L_3 가 일부분에 해당하는 에지의 크기일 때, 이를 K 개의 템플릿 후보 T_i 에 포함시킨다. 여기서 문턱값 τ 는 후보 템플릿에 모든 얼굴이 포함될 수 있도록 충분히 크게 하였다.

$$\epsilon = \{D_1 - D_2 \times 1.3\} + \{x_{01} - x_{02}\} + \left\{y_{03} - \frac{(y_{01} + y_{02})}{2}\right\} + \{x_{size_1} - x_{size_2}\} + \{ysize_1 - ysize_2\} \quad (2)$$

이렇게 확보된 후보집합의 영역을 3×3 의 격자로 나누어서 9개의 각 셀 내에 있는 수평 예지의 개수를 세어 셀의 값에 입력시킨다. 이것을 그림 2에 나타내었다.

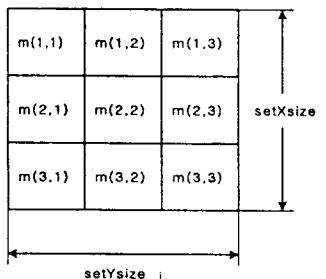


그림 2 정합을 위한 3×3 격자

만약 이 후보영역이 얼굴에 해당한다면 수평 예지는 $m(1,1)$ 과 $m(1,3)$, $m(3,2)$ 에 집중적으로 분포되어 있을 것이다. 이 점을 고려하여 이 후보 템플릿이 얼굴 인지 아닌지를 판단한다. 즉, 한 후보 템플릿 T_i , $i \in \{1, 2, \dots, K\}$ 가 이런 조건을 만족시킨다면 이 템플릿이 얼굴 템플릿에 해당한다고 결정한다. 그러므로 영상 내에 존재하는 얼굴 템플릿의 집합을 F_i , $i \in \{1, 2, \dots, M\}$ 라 했을 때, 식 (3)과 같은 분류가 이루어진다.

$$T_i \subset F_i, i \in \{1, 2, \dots, K\}, j \in \{1, 2, \dots, M\} \quad (3)$$

if

$$\begin{aligned} m(1,1) > T, & m(1,3) > T, & m(2,2) > T, \\ m(1,2) < T, & m(2,1) < T, & m(2,3) < T \end{aligned}$$

여기서 문턱값 T 는 식(4)와 같이 네 셀의 평균값으로 정의한다.

$$T = \frac{(m(1,1) + m(1,3) + m(2,1) + m(2,3))}{4} \quad (4)$$

이러한 과정으로 분류가 끝나면, 최종적인 단계로서 찾아낸 얼굴 템플릿의 영역이 서로 겹치는지 확인하여 그 중 하나만을 최종 얼굴 템플릿으로 결정한다. 이를 가려내기 위해 영역이 겹쳐진 얼굴 템플릿 각각에 대해서 식 (2)에 제시된 값을 구해서 그 중 그 값이 가장 작은 것을 최종 얼굴 템플릿 FT_i , $i \in \{1, 2, \dots, N\}$ 으로 선택한다.

이상의 템플릿 정합과정은 계산량이 매우 적어서 속도가 빠르며, 수평 예지 영상에서 사람이 얼굴을 찾아내는 적관적인 방법 [1][2]을 바탕으로 했다는 특징이 있다.

3. 타원검출 과정

템플릿 정합을 통하여 특정점들을 포함하는 영역을 검출하고 나면 이 영역의 주변영역에 대해서 얼굴 타원을 검출한다. 템플릿과 타원정보를 이용한 얼굴검출 시스템

의 흐름도가 그림 3에 나타나 있다.

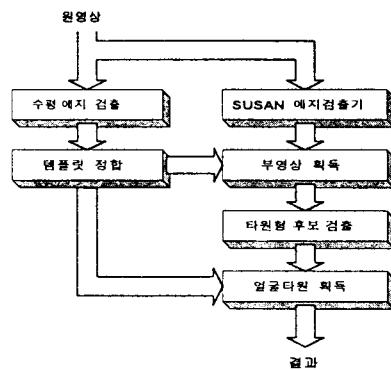


그림 3 얼굴검출 시스템

우선 검출된 예지 영상에서, 템플릿 정합과정에 의해 검출된 얼굴 템플릿 영역의 주변영역만을 포함하는 부영상 을 획득한다. 부영상의 크기는 얼굴 전체가 모두 포함되도록 충분히 크게 선택한다.

우선 부영상의 각 예지에 번호를 부여하고 예지들의 교차점을 없애고, 동일 물체의 경계선끼리 이어준다. 다음 L_1 과 L_2 의 번호가 붙은 한 쌍의 예지들에 대해 선형화된 타원의 방정식 (5)에서 유도된 행렬식 (6)~(7)을 사용하여 타원 파라미터들을 구한다.

$$2x_i a_0 - y_i^2 a_1 + 2y_i a_2 - a_3 = x_i^2 \quad i \in \{1, 2, \dots, N\} \quad (5)$$

$$\begin{pmatrix} 2x_1 & -y_1^2 & 2y_1 & -1 \\ 2x_2 & -y_2^2 & 2y_2 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ 2x_N & -y_N^2 & 2y_N & -1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_N^2 \end{pmatrix} \quad (6)$$

$$X = (A'A)^{-1} A'C \quad (7)$$

이렇게 한 쌍의 예지들에 대해 타원 파라미터 집합 $P_i = \{x_0, y_0, a_i, b_i\}$ 를 구할 때, Sirohey의 논문에서는 장축과 단축의 비율 (a/b)만을 제한조건으로 주었으나, 본 논문에서는 검출된 얼굴 템플릿에 의한 다음 조건이 추가됨으로써 계산량이 줄어들게 되어 많은 속도향상이 이루어지게 된다.

- 1) 이 타원 후보의 영역이 템플릿의 영역을 포함
- 2) 장축, 단축의 크기에 대한 제한
- 3) 중심의 위치에 대한 제한

본 논문에서는 이렇게 선택된 타원 후보들 중, 그것에 속하는 예지들의 총 화소수가 가장 많은 타원 세 개를 순서대로 선택해서, 그 중 이미 검출된 얼굴 템플릿 FT 에 가장 잘 맞는 후보를 얼굴의 타원으로 결정한다. 즉, 다음 식 (8)과 같은 오차가 가장 작은 타원 후보를 선택한다.

$$e = |(a_i - setXsize \times 1.25) + (b_i - setYsize) + (x_0 - (setXmin + setXsize \times 0.25)) + (y_0 - setY0)| \quad (8)$$

4. 실험 및 결과

본 연구에서 대상이 된 실험 영상은 Sirohey의 실험

에서 사용되었던 MIT 얼굴 데이터베이스를 이용한다. 이것은 8비트 흑백 얼굴(128 pixels×120 pixels) 데이터베이스로서, 조명의 위치, 영상에서 차지하는 얼굴의 크기, 얼굴의 회전별로 다양한 영상을 가지고 있으며, 본 실험에서는 12명의 얼굴에 대해 세 가지 조명방향, 세 가지의 얼굴크기로 구성된 총 108개의 얼굴 영상을 대상으로 하였다. 얼굴검출에 있어서의 평가기준은 사실 명확하지 않으므로, 여기서는 전체 얼굴의 주요 윤곽이 거의 다 포함된 경우를 성공한 경우로 판단한다.

4.1 템플릿 정합 결과

수평 예지를 바탕으로 템플릿 정합을 통해서 눈과 입의 특징점을 포함하는 영역을 검출해 낸 결과가 그림 4와 그림 5에 나타나 있다. 표 1에는 108개의 영상에 대해 실험한 결과의 성공률이 나타나 있다.



그림 4 템플릿 정합 결과



그림 5 여러 명의 템플릿 정합 결과

표 1 템플릿 정합 성공률

	성공률
크기 1	81 % (29/36)
크기 2	81 % (29/86)
크기 3	69 % (25/36)
평균	77 %

(크기 1 >> 크기 2 >> 크기 3)

4.2 타원검출 결과

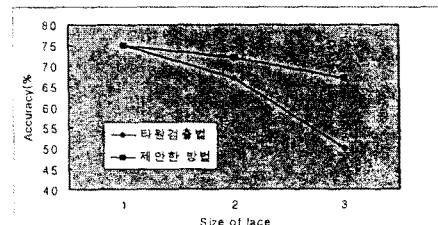
템플릿 주변영역을 포함하는 부영상에서 타원형을 찾아내어 얼굴을 검출한 결과가 그림 6과 그림 7에 나타나 있다. 그리고, Sirohey의 타원검출법과 제안한 방법의 검출성공률과 속도를 비교한 그래프를 그림 8에 나타내었다.



(a) 타원검출법 (b) 제안한 방법
그림 6 얼굴검출의 결과 비교



그림 7 여러 명의 얼굴검출 결과



(a) 검출성공률 비교

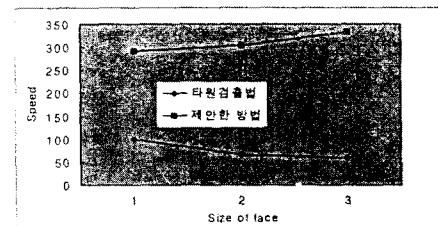


그림 8 검출성공률과 처리속도 그래프

5. 결론

본 논문에서는 인간의 직관적 판단을 근거로 수평예지를 이용한 템플릿 정합을 통하여 특징점을 포함하는 얼굴영역을 찾아낼 수 있었다. 이로써 영상 내에 존재하는 얼굴의 수와 위치, 크기에 대한 정보를 획득하여 타원검출에 이용하였다. 즉, 템플릿 주변영역을 포함하는 예지 부영상으로부터 찾아낸 템플릿에 가장 적합한 타원을 찾아냄으로써, 배경이 복잡하고 영상 내의 얼굴의 크기가 다양하며 여러 명의 얼굴이 존재하는 좀 더 일반적인 영상에 대해 얼굴을 검출할 수 있었다. 특히 영상 내의 얼굴의 크기가 영상크기에 비해 작은 경우에 Sirohey의 타원검출법보다 더욱 높은 성공률을 얻을 수 있었고, 템플릿을 통한 제한조건을 부가함으로써, 처리데이터를 줄임으로써 속도를 개선할 수 있었다.

(참고문헌)

- [1] R. Chellappa, Charles L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces : A Survey," Proc. IEEE, Vol. 83, No. 5, pp. 704-740, May 1995.
- [2] C. Katropoulos and I. Pitas, "Rule-Based Face Detection In Frontal Views," Proc. ICASSP '97, pp. 2537-2540, 1997.
- [3] A. Shio and J. Sklansky, "Segmentation of People in Motion," Proc. IEEE Workshop on Visual Motion, pp. 325-332, 1991.
- [4] S. A. Sirohey, "Human Face Segmentation and Identification," Tech. Rep. CAR-TR-695, Center for Autom. Res., Univ. Maryland, College Park, MD, 1993.