

Face Detection using Template Matching and Ellipse Fitting

鄭泰允* · 金賢述** · 姜羽錫*** · 朴相嗜[§]
(Tae-Yun Jung · Hyun-Sool Kim · Woo-Seok Kang · Sang-Hui Park)

Abstract - This paper proposes a new detection method of human faces in grey scale images with cluttered background 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, etc. Until now, many researches about face detection have been done, and applications in more complicated conditions are increasing. The existing technique proposed by Sirohey shows relatively good performance in image with cluttered background, but can apply only to image with one face and needs much computation time. The proposed method is designed to reduce complexity and be applied even in the image with several faces by introducing template matching as preprocess. The results show that the proposed method produces more correct detection rate and needs less computation time than the existing one.

Key Words : 얼굴 검출, 템플릿 정합, ellipse fitting

1. 서 론

얼굴 인식 기술(face recognition technology ; FRT)은 주어진 화상 내의 얼굴을 분석하여 기존의 데이터베이스 안에서 일치하는 얼굴을 찾아내는 기술로서, 보안 시스템이나 법집행목적, 상업적 목적 등으로 사용되어 그 응용 범위가 매우 다양하며, 최근 들어 아주 크게 주목받고 있다[1].

얼굴 인식은 크게 세 단계로 나뉘어지는데, 첫째 얼굴 검출, 둘째 특징추출, 마지막으로 인식의 단계가 그것이다. 이중 얼굴 검출은 전체 인식 시스템에 처리할 얼굴부분의 데이터만을 제공해 주므로 시스템에서 아주 큰 비중을 차지한다. 또한 얼굴 검출은 얼굴인식의 첫 단계로서 뿐만 아니라 독립적인 분야로서 많은 응용분야를 가지고 있어서, 연구의 필요성이 늘어가고 있다.

얼굴 검출은 입력 화상의 배경에 따라 크게 달라진다. 배경이 단순한 경우, 간단한 기술[2]만을 사용하여 간단하게 얼굴 검출이 이루어질 수 있지만, 배경이 복잡한 경우에는 검출이 쉽게 이루어지지 않는다. 동화상의 경우에는 배경이 복잡한 경우에도, 얼굴의 움직임만을 단서로 해서 배경과 얼굴을 분리해 낼 수 있다[3]. 정지 화상의 경우, 배경이 복잡한 영상, 여러 명의 얼굴이 있는 영상 등에 대한 연구는 현재

활발히 진행중이며 그 필요성이 증대되고 있다. 특히 Sirohey는 복잡한 배경에서 선형화된 타원방정식을 이용하여 얼굴의 타원형을 검출하는 방법을 제시하였다[4].

본 논문에서는 수평 에지를 단서로 인간이 얼굴을 인식하는 가장 간단한 원리를 이용한 얼굴 템플릿을 고안하여, 배경이 복잡한 영상에서 정면 얼굴의 눈과 코·입 영역을 템플릿 정합을 이용하여 찾아내고, 그 부분의 주변에서 얼굴에 해당하는 타원형을 찾아내는 알고리즘을 제안한다. 타원 검출 이전에 얼굴의 눈과 코·입 부분을 찾아내는 전처리 과정을 둬으로써 Sirohey의 타원 검출법에 비해 좀 더 정확한 타원 검출이 가능해지고, 전체 처리 시간이 단축되며, 기존의 방법으로는 불가능한 좀 더 일반적인 경우인 여러 명의 얼굴이 존재하는 영상에서도 검출이 가능해졌다.

2. 템플릿 정합과정

본 논문에서 사용한 템플릿 정합은 Sobel 마스크를 이용하여 얻은 양쪽 눈과 입의 수평 에지를 바탕으로 그림 2의 기본 얼굴 템플릿에 정합하는 부분을 찾아낸다. 템플릿 정합과정은 그림 1과 같다.

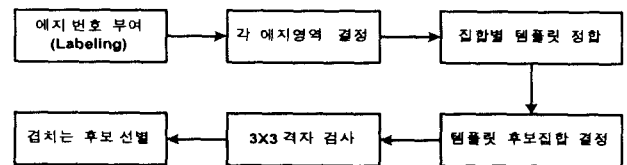


그림 1 템플릿 정합과정
Fig. 1 Template matching process

* 正 會 員 : 三星電子 中央研究所 博士課程

** 正 會 員 : 延世大 電氣·컴퓨터工學科 博士課程

*** 正 會 員 : 신도리코 R&D 센터

§ 正 會 員 : 延世大 電氣·컴퓨터工學科 教授·工博

接受日字 : 1999年 9月 20日

最終完了 : 1999年 10月 21日

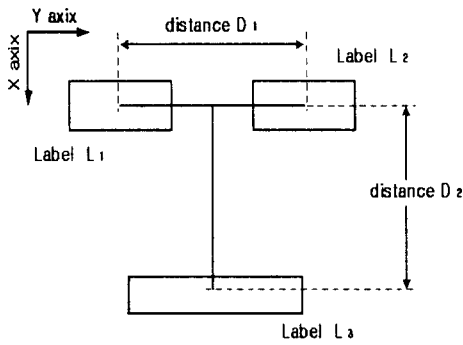


그림 2 기본 얼굴 템플릿
Fig. 2 Basic facial template

먼저 각 수평 에지화소 (x, y) 들에 번호를 부여한다. 이 과정은 $\sqrt{2}$ 보다 멀리 떨어진 에지 화소들은 분리된 에지로 간주하여 분리되지 않은 에지 화소들끼리 하나의 번호를 붙여가는 방식으로 이루어진다. 다음 모든 경우를 고려해서 세 개씩의 에지를 묶은 후 부여한 번호의 에지 $L_i, i \in \{1, 2, 3\}$ 들에 대해 위치 $x_{min_i}, x_{max_i}, y_{min_i}, y_{max_i}$ 와 크기 x_{size_i}, y_{size_i} , 중심 x_{C_i}, y_{C_i} 를 다음과 같이 구한다.

$$\begin{aligned} \text{center} : x_{C_i} &= (x_{max_i} + x_{min_i}) / 2 \quad \forall i \in \{1, 2, 3\} \\ y_{C_i} &= (y_{max_i} + y_{min_i}) / 2 \quad \forall i \in \{1, 2, 3\} \\ \text{size} : x_{size_i} &= (x_{max_i} - x_{min_i}) \quad \forall i \in \{1, 2, 3\} \\ y_{size_i} &= (y_{max_i} - y_{min_i}) \quad \forall i \in \{1, 2, 3\} \end{aligned} \quad (1)$$

에지들 간의 거리 D_1 과 D_2 는 다음의 식 (2)을 이용해 계산한다.

$$\begin{aligned} D_1 &= y_{C_2} - y_{C_1} \\ D_2 &= x_{C_3} - (x_{C_1} + x_{C_2}) / 2 \end{aligned} \quad (2)$$

위에서 얻은 결과를 다음 식 (3)에 적용하여 그 집합과 기본 템플릿과의 비용 함수값 ϵ 을 계산하고 이 값이 문턱값 τ 보다 작은 경우, L_3 의 크기가 임부분에 해당하는 에지의 크기라고 판단될 때, 이를 K 개의 템플릿 후보 T_i 에 포함시킨다. 여기서 문턱값 τ 는 후보 템플릿에 모든 얼굴이 포함될 수 있도록 충분히 크게 하였다.

$$\begin{aligned} \epsilon = & \{ D_1 - D_2 \times 1.3 \} + \{ x_{C_1} - x_{C_2} \} \\ & + \left\{ y_{C_3} - \frac{(y_{C_1} + y_{C_2})}{2} \right\} \\ & + \{ x_{size_1} - x_{size_2} \} + \{ y_{size_1} - y_{size_2} \} \end{aligned} \quad (3)$$

이렇게 확보된 후보 집합의 영역을 3×3 의 격자로 나누어서 9개의 각 셀 내에 있는 수평 에지의 개수를 세어 그림 3 과 같이 셀의 값에 입력시킨다.

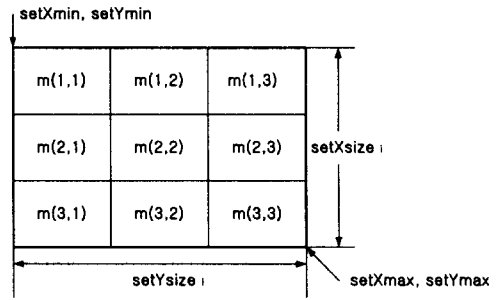


그림 3 정합을 위한 3×3 격자
Fig. 3 3×3 grid for matching

만약 이 후보 영역이 얼굴에 해당한다면 수평 에지는 $m(1,1)$ 과 $m(1,3), m(3,2)$ 에 집중적으로 분포되어 있을 것이다. 이러한 내용을 고려하여 이 후보 템플릿이 얼굴인지 아닌지를 판단한다. 즉, 하나의 후보 템플릿 $T_i, i \in \{1, 2, \dots, K\}$ 가 이런 조건을 만족시킨다면 이 템플릿이 얼굴 템플릿에 해당한다고 판단한다. 그러므로 영상 내에 존재하는 얼굴 템플릿의 집합을 $F_i, i \in \{1, 2, \dots, M\}$ 라 했을 때, 식 (4)와 같은 분류가 이루어진다.

$$T_i \subset F_j, \quad i \in \{1, 2, \dots, K\}, \quad j \in \{1, 2, \dots, M\} \quad (4)$$

if

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

여기서 문턱값 T 는 식(5)와 같이 4개의 셀의 평균값으로 정의한다.

$$T = \frac{(m(1,1) + m(1,3) + m(2,1) + m(2,3))}{4} \quad (5)$$

이러한 과정으로 분류가 끝나면, 최종적인 단계로서 찾아낸 얼굴 템플릿의 영역이 서로 겹치는지를 확인하여 그 중 하나만을 최종 얼굴 템플릿으로 결정한다. 이들을 가려내기 위해 영역이 겹쳐진 얼굴 템플릿 각각에 대해서 식 (3)에 제시된 값을 구하고 그 중 $FT_i, i \in \{1, 2, \dots, M\}$ 값이 가장 작은 것을 최종 얼굴 템플릿으로 선택한다.

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

3. 타원검출 과정

템플릿 정합을 통하여 특징점들을 포함하는 영역을 검출하고 나면 이 영역의 주변영역에 대해서 얼굴 타원을 검출한다. 템플릿과 타원정보를 이용한 얼굴검출 시스템의 흐름도가 그림 4에 나타나 있다.

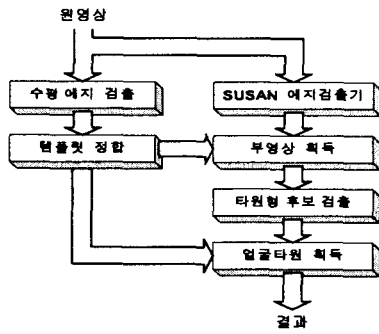


그림 4 얼굴검출 시스템

Fig. 4 Face detection system

우선 SUSAN 에지 검출기를 사용하여 검출된 에지 영상에서[5], 템플릿 정합과정에 의해 검출된 얼굴 템플릿 영역의 주변 영역만을 포함하는 부영상을 획득한다. 부영상의 크기는 얼굴 전체가 모두 포함되도록 충분히 크게 선택한다.

우선 부영상의 각 에지에 번호를 부여하고 에지들의 교차점을 없애고, 동일 물체의 경계선끼리 이어준다. 다음 L_i 와 L_j 의 번호가 붙은 한 쌍의 에지들에 대해 선형화된 타원의 방정식 (6)에서 유도된 행렬식 (7)~(8)을 사용하여 타원 파라미터들을 구한다.

$$2x_i a_0 - y_i^2 a_1 + 2y_i a_2 - a_3 = x_i^2 \quad i \in \{1, 2, \dots, M\} \quad (6)$$

$$\begin{pmatrix} 2x_1 & -y_1^2 & 2y_1 & -1 \\ 2x_2 & -y_2^2 & 2y_2 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ 2x_N & -y_N^2 & 2y_N & -1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_N^2 \end{pmatrix} \quad (7)$$

$$X = (A^T A)^{-1} A^T C \quad (8)$$

이렇게 한 쌍의 에지들에 대해 타원 파라미터 집합 $P_i = \{x_{0i}, y_{0i}, a_i, b_i\}$ 를 구할 때, 기존의 방법에서는 장축과 단축의 비율 (a/b)만을 제한조건으로 주었으나, 본 논문에서는 검출된 얼굴 템플릿을 이용한 다음 조건이 추가됨으로써 계산량이 줄어들게 되어 많은 속도향상이 이루어지게 된다. 또한 기존의 타원 검출법으로는 불가능한 2개 이상의 얼굴의 검출이 가능하게 되었다.

- 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 에지들의 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 (9)와 같은 오차가 가장 작은 타원 후보를 선택한다.

$$e = | (a_i - setXsize \times 1.25) + (b_i - setYsize) + (x_{Ci} - (setXmin + setXsize \times 0.25)) + (y_{Ci} - setYc) | \quad (9)$$

여기서,

$$setYc = \frac{setYmin + setYmax}{2}$$

4. 실험 및 결과 고찰

본 연구에서 대상이 된 실험 영상은 MIT 얼굴 데이터베이스를 이용한다. 이것은 8비트 흑백 얼굴(128×120 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 에지를 바탕으로 템플릿 정합을 통해서 눈과 입의 특징점들을 포함하는 영역을 검출해 낸 결과가 그림 5와 그림 6에 나타나 있다. 표 1에는 108개의 영상에 대해 실험한 결과의 성공률이 나타나 있다.

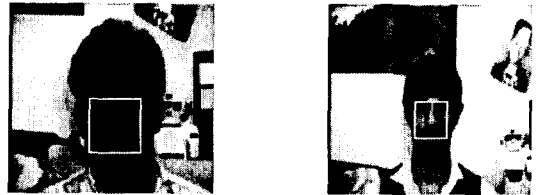


그림 5 템플릿 정합결과

Fig. 5 Results of template matching

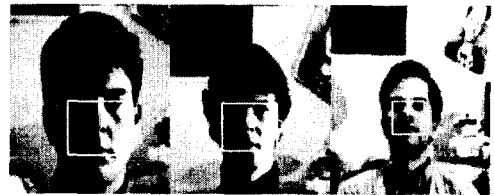


그림 6 3명의 얼굴을 포함한 영상에 대한 템플릿 정합결과

Fig. 6 Result of template matching for image with 3 faces

표 1 템플릿 정합 성공률

Table 1 Success rate of template matching

	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 7과 그림 8에 나타나 있다. 그리고, 기존의 Sirohey의 타원검출법과 제안한 방법의 검출 성공률과 속도를 비교한 그래프를 그림 9에 나타내었다.

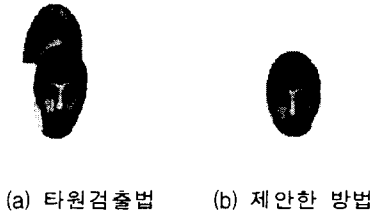


그림 7 얼굴검출의 결과 비교
Fig. 7 Result comparison of face detection

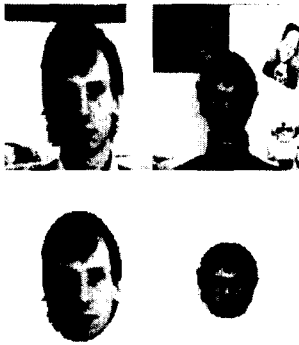
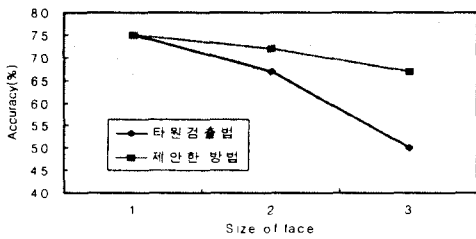
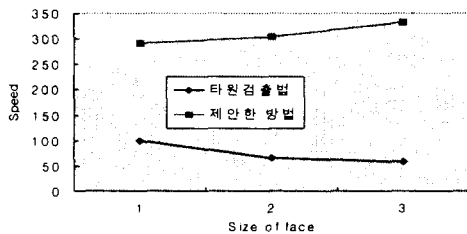


그림 8 2명의 얼굴을 포함한 영상에 대한 얼굴검출 결과
Fig. 8 Result of face detection for image with 2 faces



(a) 검출성공률



(b) 처리속도

그림 9 검출성공률과 처리속도 비교
Fig. 9 Comparisons of detection rate and processing speed

5. 결 론

본 연구에서는 배경이 복잡하고 여러 얼굴이 존재하는 정면 얼굴 영상에서도 얼굴을 검출할 수 있도록 템플릿과 타원 정보를 이용한 새로운 기법을 제안하였다. 템플릿 정합에서는 수평 에지를 이용하여 특징점들을 포함하는 얼굴 영역을 찾아낼 수 있었다. 이로써 영상 내에 존재하는 얼굴의 수와 위치, 크기에 대한 정보를 획득할 수 있었으며 이를 타원 검출에 이용하였다. 즉, 템플릿 주변영역을 포함하는 에지 부영상에 대해서 타원 검출법을 적용하고 찾아낸 타원중 템플릿에 가장 적합한 타원을 선택함으로써, 배경이 복잡하고 영상 내의 얼굴의 크기가 다양하며 여러 명의 얼굴이 존재하는 좀 더 일반적인 영상에 대해서도 얼굴을 검출할 수 있었다. 특히 영상 내의 얼굴의 크기가 영상크기에 비해 작은 경우에 기존의 Sirohey의 타원 검출법보다 더욱 높은 성공률을 얻을 수 있었고, 템플릿을 통한 제한조건을 부가함으로써, 처리 데이터를 줄임으로써 속도를 크게 개선할 수 있었다.

참 고 문 헌

- [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.
- [5] S. M. Smith, and J. M. Brady, "SUSAN - A New Approach to Low Level Image Processing," *Int. Journal of Computer Vision*, Vol. 23, No. 1, pp. 45-78, May 1997.
- [6] A. Yuille, D. Cohen, and P. Hallinan, "Feature Extraction from Faces using Deformable Templates," *Proc. IEEE Computer Soc. Conf. on Computer Vision and Patt. Recog.*, pp. 104-109, 1989.