
이질적 템플릿 매칭의 융합을 이용한 얼굴 영역 검출

Face Detection Using Fusion of Heterogeneous Template Matching

이경미
덕성여자대학교 컴퓨터공학부

Kyoung-Mi Lee(kmlee@duksung.ac.kr)

요약

본 논문은 영상에 포함된 얼굴을 보다 빠르고 강건하게 검출하기 위해서 이질적 템플릿 매칭의 결과들을 융합하여 얼굴을 검출하는 방법을 제안한다. 먼저 광범위한 조명 환경과 인종을 포괄하는 피부색 모델을 이용해 피부 영역을 검출한다. 그리고 영역 라벨링과 필터링으로 매칭에 필요한 검색 범위를 줄인 후, 피부색과 에지를 이용한 템플릿 매칭을 검출된 영역에 적용한다. 이들 매칭 결과가 융합되어 두 매칭 결과를 동시에 최적으로 만족하는 얼굴이 검출된다. 실험 결과는 제안된 방법이 단일 템플릿을 적용할 때보다 얼굴색과 유사한 배경에서 얼굴을 강건하게 검출하며, 얼굴 후보 영역으로 검색 범위를 줄여 검출 시간을 줄였음을 보여준다. 또한 전역 누산기를 사용하여 템플릿 매칭의 과도한 공간 요구의 문제점을 해결할 수 있었다.

■ **중심어** : | 얼굴 검출 | 템플릿 매칭 | 융합 | 피부색 검출 |

Abstract

For fast and robust face detection, this paper proposes an approach for face detection using fusion of heterogeneous template matching. First, we detect skin regions using a model of skin color which covers various illumination and races. After reducing a search space by region labelling and filtering, we apply template matching with skin color and edge to the detected regions. Finally, we detect a face by finding the best choice of template fusion. Experimental results show the proposed approach is more robust in skin color-like environments than with a single template matching and is fast by reducing a search space to face candidate regions. Also, using a global accumulator can reduce excessive space requirements of template matching.

■ **keyword** : | Face Detection | Template Matching | Fusion | Skin Color Detection |

I. 서론

최근 수십 년 동안 디지털 영상으로부터 사람을 찾아 취하고 있는 자세 및 동작을 인식하는 문제가 활발히 연구되고 있다[1][2]. 일반적으로 사람을 검출하기 위해 인체를 전체적으로 다루는 것 보다는 각 구성요소로 나

누어 다루는 것이 자세나 조명의 변화에 보다 좋은 결과를 줄 수 있다. 인체의 구성요소들 중 얼굴은 인체 검출에 가장 유용한 정보가 될 것이다. 왜냐하면, 얼굴을 구성하고 있는 기관은 비교적 공통된 특성을 찾기 쉬우며, 무엇보다도 다른 어느 부위보다 노출될 가능성이

높기 때문이다. 따라서 얼굴 검출은 사람 검출 및 동작 인식의 첫 단계가 될 것이다[1].

본 논문에서는 광범위한 피부색 모델을 이용해 조명 환경과 인종에 어느 정도 변화를 허용한 얼굴 검출을 가능하게 한다. 그리고 검출된 피부색 영역을 연결 요소로 그룹화를 한 후, 불필요한 배경이나 얼굴 이외의 부분을 미리 제거하여 후보 영역을 결정한다. 이는 템플릿 매칭에 필요한 검색 범위를 크게 줄임으로써 전체적인 얼굴 검출 시간을 감소시켜 시간 효율성을 높일 수 있다. 마지막으로 이질적인 두 (피부색과 예지) 템플릿의 매칭을 융합함으로써, 단일 템플릿을 사용했을 때 보다 얼굴 영역 검출의 정확성을 높일 수 있다. 제안하는 방법은 Garcia와 Tziritas의 방법[5]과 CMU 얼굴 검출기[21]와 얼굴 검출 결과를 비교하였다.

II. 관련 연구

영상에서 얼굴 검출을 위한 다양한 방법이 연구되었다. Yang *et al.*의 논문[3]에 따라, 영상에서의 얼굴 검출 방법은 크게 4 종류로 나눌 수 있다: 지식 기반 방법, 특징 불변 방법, 템플릿 매칭 방법, 외형 기반 방법. 우선, 지식 기반 방법은 얼굴이 무엇으로 구성되었는지에 대한 지식으로부터 규칙을 만들고 이를 부호화한다. Yang과 Huang는 세 개의 규칙 계층으로 이루어진 계층적 지식기반 방법을 이용하여 얼굴을 검출하였다. 높은 계층의 틀은 얼굴에 대한 일반적인 모습을 물로 만든 반면, 낮은 계층은 얼굴에 대한 자세한 특징을 물로 만들었다[8]. 그러나 지식을 규칙으로 표현하는 것은 쉽지 않을 뿐더러, 가능한 모든 경우를 고려하여야 하므로 서로 다른 자세의 얼굴 검출이 어렵다.

지식 기반 방법이 규칙으로 특징을 표현하는 하향식 방법이라면, 특징 불변 방법은 얼굴의 특징을 먼저 찾아 얼굴을 검출하는 상향식 방법이다. 특징 불변 알고리즘들은 자세, 관점, 조명이 다르더라도 불변하는 얼굴의 구조적 특징을 이용하여 얼굴의 위치를 찾는 것을 목표로 한다. Leung *et al.*은 5개의 얼굴 특징(두 눈, 두 콧구멍, 인중)과 그들 사이의 매칭으로 묘사된 얼굴

을 찾는 확률적 방법을 개발하였다[9]. Dai와 Nakano는 16×16의 부분 영상에 대해 이차 확률 특징으로 계산된 텍스처를 이용하여 얼굴을 검출하였다[10]. 피부색은 얼굴검출에 가장 효과적으로 이용되는 특징이다. 피부색 분포에 대한 Gaussian 조합을 이용하여 가장 피부색 같은 후보 영역을 예측할 수 있었다[11]. 그러나 피부색은 피부색 하나만으로 충분히 얼굴을 검출해 내기는 어렵기 때문에 형태, 움직임과 같은 정보를 결합하는 연구가 이루어지고 있다. Sobottka와 Pitas는 화소의 색상을 이용하여 피부색 같은 화소를 검출한 후, 검출된 화소를 연결한 영역의 모양을 이용하여 얼굴을 검출하는 방법을 제안하였다[12]. 이러한 특징 불변 방법은 영상의 특징들은 조명, 잡음 등으로 손상되기 쉬운 단점을 가지고 있다.

템플릿 매칭 방법은 얼굴에 대한 여러 (주로 정면 얼굴) 패턴들을 미리 정의한다. Craw *et al.*은 정면 얼굴에 대한 형태 템플릿을 정의하였다. Sobel 필터를 이용하여 얼굴 형태, 눈, 눈썹, 입술 등의 특징을 포함시켰다[13]. Lantis *et al.*은 형태와 색상 정보를 모두 이용하여 얼굴을 표현하였다. 눈, 코, 볼 등의 외곽선을 훈련시킨 active shape model을 이용하여 얼굴의 위치를 찾았다[14]. 그러나 다양한 크기와 자세, 형태를 효과적으로 처리하기 위해서는 다양한 패턴을 미리 정의하여야 하므로 시간 비용이 매우 높다.

템플릿 매칭이 템플릿을 미리 정의하는 것과는 반대로, 외형 기반 방법은 영상의 다양한 얼굴 예제로부터 학습을 통해 템플릿을 만든다. 통계적 방법이나 인공지능 기법을 이용하여 얼굴의 특성을 찾아 학습시켜서 얼굴을 검출하게 된다. Turk와 Pentland는 주성분 분석을 이용하여 얼굴을 검출하였다[15]. 얼굴 영상에서 주성분 분석을 통해 얻어진 eigenface를 영상에 투영하여 가장 얼굴같은 부분이 있는 위치를 찾아낸다. Osuna *et al.*은 처음으로 얼굴 검출에 SVM을 적용하였다[16]. 그들은 얼굴 영상과 얼굴이 아닌 영상을 훈련시켜 구별할 때 예측 오차를 최소화시키는 최적화 문제를 해결하였다.

Yang *et al.*이 분류한 얼굴 검출 이외에 최근에는 얼굴 검출을 위해 데이터를 융합을 시도하는 연구가 많이

이루어지고 있다. 대부분은 얼굴의 구성요소들 또는 구성요소를 포함하고 있는 부분 영상을 융합하는 방법에 대한 연구이거나[17][18], 동영상에서 색상과 움직임 정보를 융합해 얼굴을 검출하는 연구도 있었다[19][20].

본 논문에서는 정지 영상에서 얼굴을 검출하기 위해 피부색을 이용하여 얼굴의 후보 위치를 찾은 후, 템플릿 매칭 방법을 사용한다. 기존의 템플릿 매칭 방법은 구현하기 쉽지만, 시간적 효율성의 문제와 단일 정보만을 이용한 템플릿 매칭은 얼굴의 특징과 유사한 배경을 포함하는 경우에 검출율이 현저히 저하되는 문제점을 갖고 있다[4]. 따라서 본 논문에서는 얼굴의 후보 위치를 포함하는 피부색 정보와 얼굴의 구성요소 특징을 포함하는 에지정보를 이용한 이질적 템플릿 매칭을 융합하여 주어진 영상에서 빠르고 강건하게 얼굴을 검출하는 방법을 소개한다. 즉, 서로 다른 특징을 포함하는 이질적 템플릿의 융합을 통해 매칭 결과를 모두 만족하는 최적의 얼굴을 검출하는 것이다.

III. 얼굴 영역 검출

본 장에서는 이질적 템플릿 매칭을 이용하여 얼굴을 검출하는 방법을 소개한다. [그림 1]은 본 논문에서 제안된 얼굴 검출 알고리즘을 보여주고 있다.

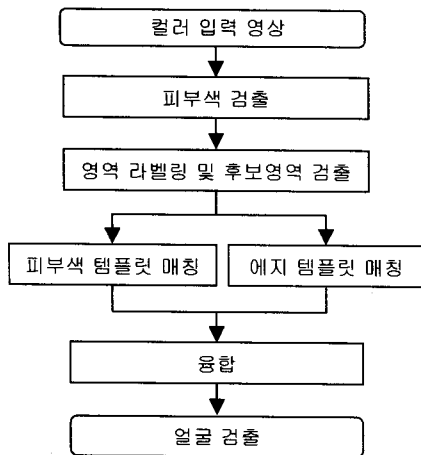


그림 1. 제안된 알고리즘의 처리 단계

1. 피부색 영역 검출

우선, 입력 영상 [그림 2]이 주어지면 미디언 필터를 이용하여 RGB 색공간에서 평활화를 시킨다. 다음은, Garcia와 Tziritas의 논문[5]에서 적용된 피부색 모델을 이용하여 피부색 영역을 결정한다. Garcia와 Tziritas는 YCbCr 색공간에서의 피부색 모델이 HSV 색공간에서의 피부색 모델보다 조밀하다고 보고했다. 따라서 우리는 YCbCr 색공간에서 식 (1)을 이용해서 피부색 영역을 결정하였다. 식 (1)의 YCbCr 모델은 다양한 인종, 조명 환경 등 광범위한 피부색 범위를 가지는 950개의 이미지에서 얻어진 피부색 샘플을 분석하여 얻어진 것이다[5].



그림 2. 실험 영상

$$\text{if } (Y > 128) \begin{cases} \theta_1 = -2 + \frac{256 - Y}{16}, \\ \theta_2 = 20 + \frac{256 - Y}{16}, \\ \theta_3 = 6, \theta_4 = -8 \end{cases}$$

$$\text{if } (Y \leq 128) \begin{cases} \theta_1 = 6, \theta_2 = 12, \\ \theta_3 = 2 + \frac{Y}{32}, \theta_4 = -16 + \frac{Y}{16} \end{cases} \quad (1)$$

$$\begin{aligned} Cr &\geq -2(Cb + 24), Cr \geq -(Cb + 17), \\ Cr &\geq -4(Cb + 32), Cr \geq 2.5(Cb + \theta_1), \\ Cr &\geq \theta_3, Cr \geq 0.5(\theta_4 - Cb), \\ Cr &\leq \frac{220 - Cb}{6}, Cr \leq \frac{4}{3}(\theta_2 - Cb) \end{aligned}$$

$$I'(x,y) = \begin{cases} I(x,y) & I(x,y) \text{가 식(1)을 만족하면,} \\ 0 & \text{그렇지 않으면} \end{cases}$$

여기서 Y, Cb, Cr 은 각각 YCbCr 모델에서의 값이며, $I(x,y)$ 는 영상 I 의 (x,y) 번째 화소의 값을 말한다. [그림 3]은 [그림 2]에서 피부영역을 검출한 결과 영상이다. [그림 3]에서 보듯 색상 정보만으로는 불필요한 배경 부분이 완전하게 제거되지 않아 얼굴 영역 검출에 불충분하며 추가적인 처리 단계를 필요로 한다.

2. 얼굴 후보영역 검출

다음 단계는 영상의 전 영역을 대상으로 하는 템플릿 매칭보다 빠른 검출 시간을 얻기 위해, 영상에서 얼굴 후보영역을 검출하는 과정이다. [그림 3]과 같이 피부색이 검출된 영상에 대해, 피부색 픽셀의 4-인접 픽셀 중 피부색 픽셀을 연결 요소로 정한다. 영역의 그룹화를 마치면, 주어진 영상에 피부 연결 영역들의 리스트가 만들어지고, 다음과 같은 조건에 따라 얼굴일 확률이 낮은 영역을 제거함으로써 얼굴 후보 영역을 결정한다.



그림 3. 피부색 검출

- 가로 또는 세로의 길이가 매우 작은 영역
- 면적(= 인접 영역 내 픽셀 수)이 매우 작은 영역
- 비율(= 가로/세로)이 매우 적거나 큰 영역
- 밀집도(= 면적/인접 사각형의 크기)가 매우 작은 영역

이런 필터링과정은 매칭의 검색 범위를 줄여 주며, 피부색과 비슷한 배경으로부터 얼굴을 분리해 내어 얼굴

의 위치를 찾기 쉽게 한다[5]. [그림 4]는 [그림 3]의 피부색 영역에서 검출된 얼굴 후보 영역을 보여주고 있다.

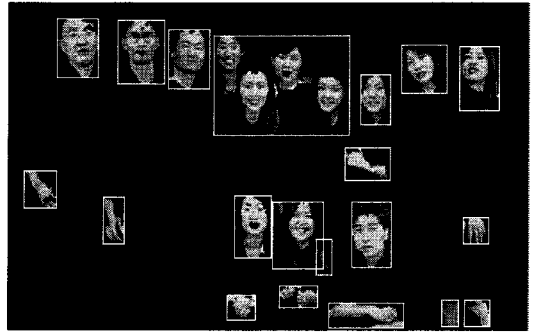


그림 4. 얼굴 후보영역

3. 피부색 템플릿을 이용한 매칭

피부색 템플릿을 만들기 위해서 다양한 인종과 피부색의 얼굴과 다양한 조명 환경을 포함하는 40개 (남자 20명과 여자 20명)의 영상으로부터 얼굴 샘플영상을 취했다. 40개의 얼굴 샘플영상에 대해 미디언 필터링과 피부색 검출(III-2장 참조)을 실행한 후, 20×20으로 크기 정규화를 시키고 피부색 평균값을 구하여 피부색 템플릿을 만든다[그림 5].

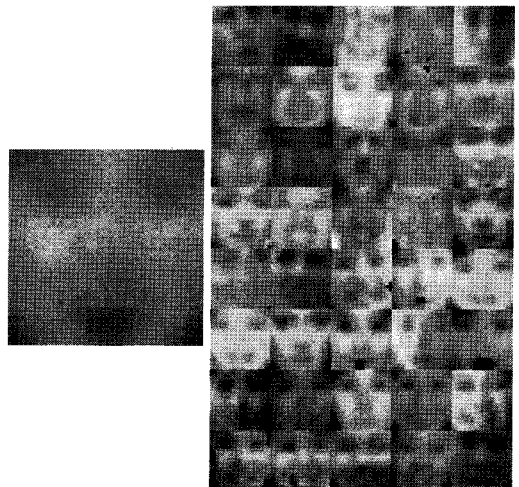


그림 5. 피부색 템플릿과 사용된 피부색 영상들



그림 6. 피부색 템플릿 매칭 후 영상

$$SkinTemplate(x,y) = \frac{1}{n} \sum_{i=0}^{n-1} T_i'(x,y)$$

$$T_i'(x,y) = \begin{cases} T_i(x,y) & T_i(x,y) \text{가 식(1)을 만족하면,} \\ 0 & \text{그렇지 않으면} \end{cases}$$

여기서 $T_i(x,y)$ 는 i 번째 템플릿 T_i 의 (x,y) 번째 화소를 의미한다. [그림 5]의 왼쪽 영상은 피부색 템플릿으로 40개의 얼굴 샘플에 대한 두 눈과 입술, 인종의 공통적인 색상 특성을 모두 포함하고 있다. 피부색 템플릿은 평균화 과정을 통해 서로 다른 조명, 다양한 피부색 등의 정보는 줄이는 반면, 두 눈과 입술, 인종과 같은 공통적인 정보는 강조하는 효과를 보이고 있다.

III-2장에서 구한 각 인접 영역에 대해, 피부색 영역과 템플릿의 상관관계를 계산하였다. 일반적으로 얼굴의 λ 보다 적게 나타난 경우는 얼굴로 판단하기 어려우므로, 템플릿이 검출된 피부색 영역에 적어도 λ 이상 겹치는 경우만 얼굴로 인정한다.

$$TM_1(I'(x,y), SkinTemplate(x,y)),$$

if $I'(x,y) \neq 0$ (2)

여기서 $TM_1(I, T)$ 는 I 와 T 의 피부색 템플릿 매칭을 이용한 유사도 관계로, 피부색인 경우에만 ($I'(x,y) \neq 0$) 계산에 포함시킨다.

4. 에지 템플릿을 이용한 매칭

다음으로, 에지 템플릿[그림 7]을 만들기 위하여 40개의 샘플 영상을 평활화한 후, Sobel 연산자를 이용하여 에지를 검출한다. 얼굴에서 만들어지는 에지는 주로 눈, 눈썹, 입술 등에 위치하는 가로 에지이다. 피부색 템플릿과 같이, 20×20 크기로 정규화한 후, 가로 에지를 가진 영상의 평균값으로 에지 템플릿 $EdgeTemplate(x,y)$ 을 만든다[그림 7].

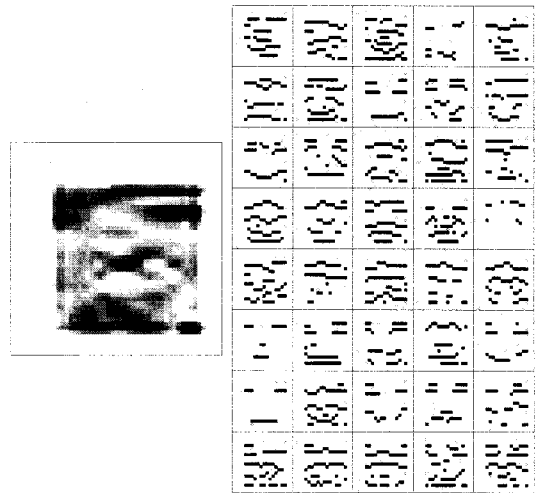


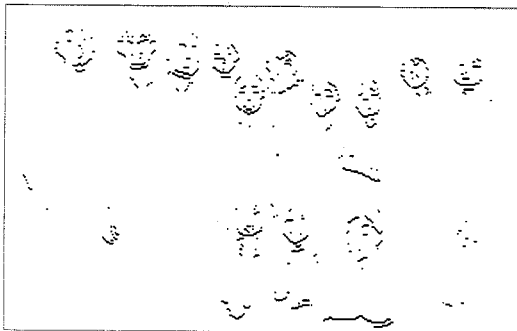
그림 7. 에지 템플릿과 사용된 에지 영상들

영상이 주어지면, 마찬가지로 평활화한 후, Sobel 연산자를 이용하여 가로 에지만을 구하고, 피부색 템플릿 매칭과 마찬가지로 템플릿 매칭을 이용하여 유사도(식(2))를 계산한다.

$$TM_2(I'(x,y), EdgeTemplate(x,y)),$$

if $I'(x,y) \neq 0$ (3)

[그림 8(a)]은 [그림 2]의 Sobel 가로 에지 영상들, [그림 8(b)]은 [그림 8(a)]에 에지 템플릿 [그림 7]을 적용한 결과 영상이다.



(a)



(b)

그림 8. (a) 피부색 영역에 대한 가로 에지 영상과 (b) 에지 템플릿 매칭 후 검출된 얼굴 영역

IV. 이질적 템플릿 매칭의 융합

1. 템플릿 매칭의 융합

템플릿 매칭은 특정 공간에서 주어진 템플릿과 영상 사이의 유사도를 계산하는 방법으로, 주로 영상의 단일 특징에 대해 수행된다. 그러나 단일 특징을 이용한 템플릿 매칭은 좋은 검출 결과를 기대하기 어렵다. 예를 들어, 얼굴 검출을 위해 피부색 템플릿 매칭을 이용하면 평균 피부색과 가장 유사한 곳을 찾는데 (Ⅲ-3장), 손목이나 목 등도 함께 검출될 뿐만 아니라 피부색과 유사한 배경을 구분하지 못한다. 또, 가로 에지만을 이용한 템플릿인 경우엔 평균 가로 에지와 가장 유사한 곳을 찾는데 (Ⅲ-4장), 벽들과 같이 가로선을 가진 배경이나 물체를 얼굴로 오인식하게 된다. 이처럼 어느 하나의 특징으로 이루어진 템플릿을 이용하여 정확하게

얼굴 영역을 검출하기는 쉽지 않다.

이 문제를 해결하기 위해서, 본 논문에서는 이질적 템플릿을 이용한 매칭을 융합하는 방법으로 상호 보완될 수 있는 방법을 제안한다. 주어진 영상에 대해서, 식 (2)와 (3)인 피부색 템플릿 매칭(TM_1)과 에지 템플릿 매칭(TM_2)을 각각 실행한 뒤, 융합할 수 있다. 그러나 어떤 특징이 검출에서 더 중요한 역할을 하는지 알 수 없으므로, 이질적 특징에 대한 템플릿 매칭의 주의 깊은 통합이 필요하다. 주어진 영상과 각 템플릿 사이의 매칭 결과가 검출을 결정하는데 가지는 중요도가 다르므로, 그 중요도에 따라 매칭 결과를 융합하여 최종 결정을 내린다. 두 템플릿 매칭의 결과를 융합하여 전체적으로 얼굴일 확률은 아래와 같이 계산될 수 있다.

$$D(x,y) = \prod_{i=1}^2 \gamma TM_i(x,y) \quad (4)$$

위에서 사용된 계수 γ 는 가중치를 말한다. 이 논문에서는 두 매칭 결과를 융합하기 위해서, 선형적 결합을 이용하였다.

여기서 강조할 점은 우리의 목적은 단일 템플릿을 이용하여 최적의 얼굴 위치를 검출하는 것이 필요한 것이 아니라, 템플릿 융합을 통해 $D(x,y)$ 의 값을 최대화시키는 최적의 얼굴 위치를 구하는 것이다.

2. 공간 효율적 템플릿 매칭

영상은 다양한 크기의 얼굴을 포함하므로 주어진 템플릿을 이용하여 이들을 모두 검출할 수 있어야 한다. 본 논문에서는 다양한 크기의 얼굴을 검출하기 위하여, [그림 9]에서 보는 것과 같이 입력 영상을 일정한 비율 (1.1배)로 축소하여 피라미드 영상을 만들었다[6]. 템플릿보다 큰 얼굴을 검출하기 위해서 영상을 축소한 후, 반복적으로 두 개의 템플릿 매칭과 융합을 적용하여 얼굴을 검출한다.

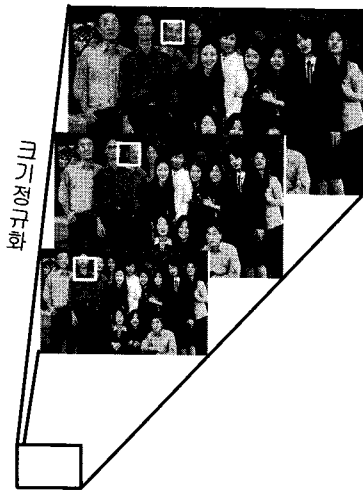


그림 9. 입력 영상에 대한 피라미드 영상

그런데 템플릿 매칭은 적용되는 각 템플릿마다 매칭 결과를 저장하기 위해 영상의 크기와 동일한 누산기를 사용한다. 영상이 주어진 템플릿과 거의 매칭되는 특징을 포함하고 있을 때, 누산기는 최고치를 갖는다. 모든 누산기가 각각의 최고점을 갖지만, 값의 크기는 누산기마다 다르다. 본 논문은 다양한 크기의 얼굴 템플릿 중에서 최적의 크기와의 매칭만을 찾는 것이 목적이므로, 모든 크기에 대한 융합 결과를 저장하는 대신에, 전역 공간을 두어 최고의 융합 결과만을 저장하면 많은 저장 공간의 요구를 피할 수 있다[7].

$$\begin{aligned} & \text{if } D_j(x,y) > \text{Global}[0](x,y) \\ & \text{Global}[0](x,y) = D_j(x,y) \\ & \text{Global}[1](x,y) = j \end{aligned}$$

여기서 $D_j(x,y)$ 은 식 (4)를 이용하여 j 번째 크기 템플릿에 의해 융합된 결과이고, 전역누산기 $\text{Global}(x,y)$ 은 영상 $I(x,y)$ 에서 모든 템플릿 융합 중 최대값을 갖는 템플릿에 관한 정보를 포함한다. $\text{Global}[0]$ 은 최대값을, $\text{Global}[1]$ 은 최대값일 때의 템플릿 크기 비율을 저장한다. 또한 중복된 템플릿 매칭을 막기 위해서는 템플릿의 높은 값들 중 최고치를 찾으면서 다른

값들은 지워주는 최고치 선명화(sharpening)을 수행해야 한다. [그림 10]은 이질적 템플릿 융합에 의해 검출된 얼굴을 표시하고 있다.

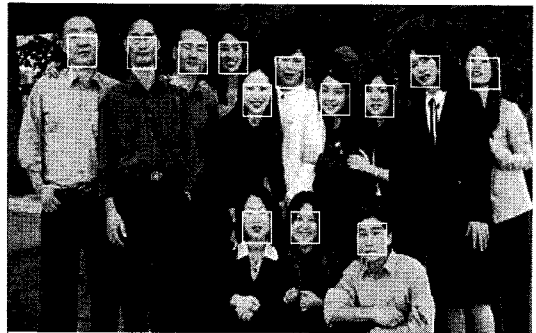


그림 10. 이질적 템플릿 융합 후 검출된 얼굴 영역

V. 실험 결과

제안된 알고리즘은 운영체제 Windows 2000 XP 상에서 JAVA를 이용하여 구현되었으며, 실험은 Pentium-IV 1.8 GHz 인 CPU와 512 MB 메모리 사양의 Windows 2000 XP 환경에서 수행하였다. 실험에 사용된 테스트 영상은 인터넷에서 얻은 200개의 영상을 사용하였다. 영상은 다양한 인종, 얼굴 크기, 사람 수, 조명 환경을 포함한다. 본 논문은 20×20부터 40×40 크기까지의 얼굴을 검출하도록 하였기 때문에, 큰 얼굴은 최대한으로 얼굴의 부분을 검출하게 되었다. 또한, 본 논문에서는 정면 얼굴을 대상으로 실험 대상으로 하였지만, 가로 에지 성분을 포함한 경우 측면얼굴도 검출될 수 있으므로, 상하좌우 22.5° 범위 안에서의 회전을 허용한 얼굴 검출을 목표로 한다. [그림 11]은 제안된 방법으로 얼굴을 검출한 결과 영상이다.

표 1. 템플릿 매칭 알고리즘의 수행 시간

단위 sec (초)	매칭 융합	단일 템플릿
전체 영상	101.8	223.0
피부색 영역만	22.4	49.0

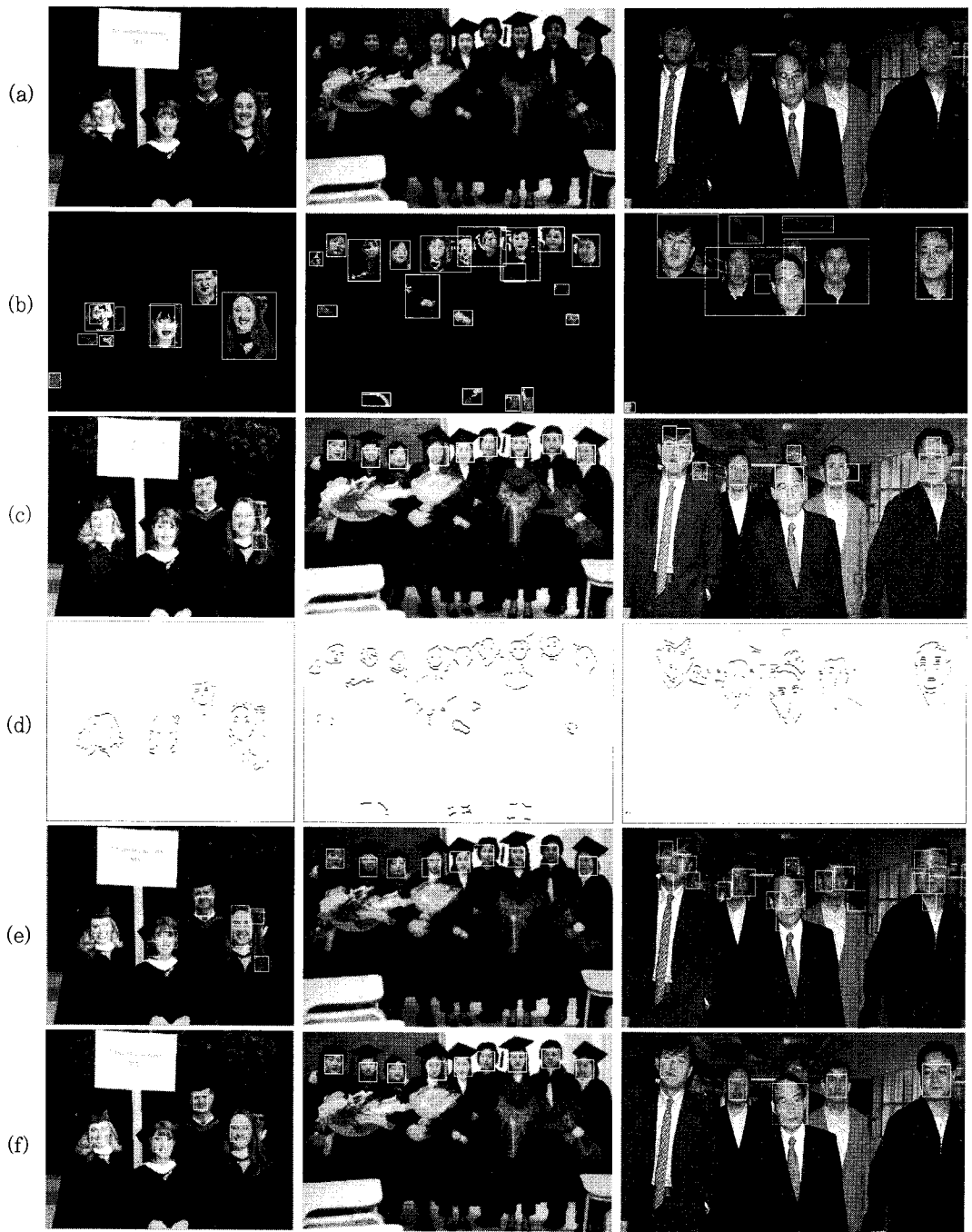


그림 11. 제안된 방법을 이용한 얼굴 영역 검출 결과 : (a) 실험 영상, (b) 병합된 얼굴 후보 영역, (c) 피부색 템플릿 매칭 후 영상, (d) 가로 에지 영상, (e) 에지 템플릿 매칭 후 검출된 얼굴 영역, (f) 이질적 템플릿 융합 후 검출된 얼굴 영역

[표 1]은 200개 테스트 영상에 대해 전체 영상과 피부색 부분 영역에 템플릿 매칭을 수행한 시간 결과를 보여주고 있다. [표 1]은 피부색 영역에만 적용하면 평균 78%의 시간 효율을 얻을 수 있음을 보여주고 있다. 또한 매칭 융합을 적용하더라도 단일 템플릿을 전체 영상에 대입한 경우보다 평균 54%의 시간 효율을 얻을 수 있었다.

표 2. 얼굴 검출에 대한 혼돈 행렬(confusion matrix) †

	CLASS		매칭 융합	피부색 템플릿	에지 템플릿
	TRUE	PREDICTED			
TP	FACE	FACE	756	757	692
FP	FACE	NONE	81	80	145
FN	NONE	FACE	113	175	146

† FACE는 얼굴을, NONE은 얼굴이 아닌 경우를 나타냄. TP(True Positive)는 실제 얼굴이 얼굴로 검출된 경우, FP(False Positive)는 얼굴이 검출되지 않은 경우, FN(False Negative)는 얼굴이 아닌 것이 얼굴로 검출된 경우를 나타냄.

[표 2]는 테스트 영상에 대한 템플릿 매칭의 개수를 보여주고 있다. 피부색을 이용하는 경우는 단일 매칭과 매칭 융합에 대해 모두 평균 90% 이상의 얼굴(FACE)을 얼굴로 검출하였지만, 에지만을 이용하는 경우에는 얼굴 중 평균 17.3%를 검출하지 못했다 (NONE).

알고리즘의 분할 성능을 평가하기 위해서 민감도(sensitivity)와 정확도(Accuracy)를 사용하였다. 민감도는 검출된 얼굴 중에 실제 얼굴이 검출될 확률로 다음과 같이 정의된다:

$$\text{민감도} = \frac{TP}{TP + FN}$$

정확도는 실제 얼굴 중에서 알고리즘에 의해 검출될 확률이다:

$$\text{정확도} = \frac{TP}{TP + FP}$$

표 3. 템플릿 매칭 알고리즘의 민감도와 정확도

	매칭 융합	피부색 템플릿	에지 템플릿
민감도	87.0	81.2	82.6
정확도	90.3	90.4	82.7

[표 3]은 제안된 알고리즘의 민감도와 정확도를 나타낸다. 피부색 템플릿만 이용한 경우는 평균 81.2%, 에지 템플릿만 이용한 경우는 평균 82.7%, 매칭 융합을 사용한 경우는 평균 87.0%의 민감도를 보였다. 즉, 피부색만을 이용하는 경우는 얼굴을 잘 찾는 반면, 얼굴이 아닌 노출된 신체 부분인 손, 목, 다리 등도 함께 검출되었다. 에지만을 이용하는 경우는 얼굴이 아닌 신체 부분이 상대적으로 덜 검출되는 반면, 그만큼 얼굴도 잘 찾지 못하였다. 반면, 매칭 융합은 피부색 템플릿 정도로 얼굴을 검출하지만, 얼굴이 아닌 신체 부위를 제외시키는 능력이 뛰어났다.

표 4. 실험 영상에서 얼굴 검출에 대한 비교

	제안된 방법	Garcia [5]	CMU 얼굴검출기[21]
민감도	87.0	85.9	89.1
정확도	90.3	87.5	84.1

제안된 방법의 얼굴 검출 결과를 비교하기 위해 컬러 영상을 대상으로 연구한 Garcia와 Tziritis의 방법[5]과 CMU 얼굴 검출기[21]를 이용하였다. [표 4]는 200개의 테스트 영상에 대한 얼굴 검출의 민감도와 정확도를 보여주고 있다. CMU 얼굴 검출기는 상대적으로 민감도는 높은 반면, 정확도는 떨어졌고, Garcia와 Tziritis의 방법은 CMU 얼굴검출기 보다 정확도는 높으나 제안하는 방법보다 정확도가 떨어졌다.



(a) (b)

그림 12. 오검출 예

그러나 여전히 남아있는 문제는 얼굴이 아닌지만 얼굴과 유사한 특징을 가진 물체(손)에 대한 검출이다 [그림 12(a)]. 또는 얼굴이지만 빛이나 음영에 의해 얼굴 특징이 제대로 검출되지 못한 경우도 있었다[그림 12(b)].

VI. 결론

본 논문에서 제안한 방법은 이질적 템플릿의 매칭을 융합하여 최적의 얼굴 영역을 검출하는 것이다. 단일 템플릿을 적용할 때보다 얼굴색과 유사한 배경에서 얼굴을 강건하게 검출하며 얼굴 후보 영역을 미리 결정하여 얼굴 영역 검출 시간을 줄여줄 수 있음을 실험을 통해 보였다. 또한, 전역누산기의 사용으로 템플릿 매칭의 문제점 중 하나인 과도한 공간 요구에 대한 문제를 해결할 수 있었다.

본 논문에서는 템플릿 매칭의 융합이 가중치의 선형 결합 방식을 이용하기 때문에, 가중치의 값에 따라 피부색이나 가로 예지에 민감하게 반응한다. 따라서 이 문제는 신경망 등을 이용한 비선형 결합 방식에 의한 보다 효율적인 융합으로 해결할 수 있으며, 더불어 가중치를 학습에 의해 자동적으로 구할 수 있다. 또한, 정면 얼굴뿐만 아니라, 측면 얼굴도 함께 검출하기 위해서, 이질적 템플릿 매칭의 융합을 확장하는 노력이 필요하다.

참고 문헌

- [1] P. F. Delzenszwalb and D. P. Huttenlocher, "Efficient matching of pictorial structures," In proc. CVPR, pp.66-73, 2000.
- [2] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," IEEE Tran. on PAMI, Vol.23, No.4, pp.49-361, 2001.
- [3] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," IEEE Tran. on PAMI, Vol.24, No.1, pp.34-58, 2002.
- [4] D. Chai and K.N. Ngan, "Locating facial region of a head-and-shoulders color image," In proc. AFGR, pp.586-591, 1991.
- [5] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," IEEE Tran. on Multimedia, Vol.1, No.3, pp.264-276, 1999.
- [6] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," IEEE Tran. on PAMI, Vol.20, No.1, pp.23-38, 1998.
- [7] 이경미, W. N. Street, "다중해상도 탐색을 이용한 반복 일반화 허프 변환," 정보과학회 논문지:소프트웨어 및 응용, 제30권, 제9-10호, pp.973-982, 2003.
- [8] G. Yang and T. S. Huang, "Human face detection in complex background," Pattern Recognition, Vol.27, No.1, pp.53-63, 1994.
- [9] T. Leung, M. Burl, and P. Perona, "Finding faces in cluttered scenes using labelled random graph matching," In Proc. CV, pp.637-644, 1995.
- [10] Y. Dai and Y. Nakano, "Face-texture model based SGLD and its application," Pattern Recognition, Vol.29, No.6, pp.1007-1017, 1996.
- [11] T. Poggio and K. K. Sung, "Finding human faces with a gaussian mixture distribution-based face model," In proc. ACCV, pp.437-446, 1995.
- [12] K. Sobottka and I. Pitas. "Looking for faces and facial features in color images," PRIA: Advances in Mathematical Theory and Applications, Vol.7, No.1, pp.124-137, 1997.
- [13] I. Craw, D. Tock, and A. Bennett, "Finding face features," In proc. ECCV, pp.92-96, 1992.
- [14] A. Lanitis, C. J. Taylor, and T. F. Cootes, "An automatic face identification system using flexible appearance models," Image and Vision Computing, Vol.13, No.5, pp.393-401, 1995.
- [15] M. Turk and A. Pentland, "Face recognition using eigenfaces," In proc. CVPR, pp.586-591, 1991.
- [16] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: An application to face detection," In proc. CVPR, pp.130-136, 1997.

- [17] B. Xie, D. Comaniciu, V. Ramesh, M. Simon, and T. Boulton, "Component fusion for face detection in the presence of heteroscedastic noise," In proc. conf. of the German Society for Pattern Recognition, pp.434-441, 2003.
- [18] A. E. Savakis and D. Higgs, "Parts-based face detection at multiple views," In proc. VISAPP, pp.298-301, 2007.
- [19] D. Butler, C. McCool, M. McKay, S. Lowther, V. Chandran, and S. Sridharan, "Robust face localisation using motion, colour and fusion," In proc. DICTA, pp.899-908, 2003.
- [20] J. Zapata and R. Ruiz, "Biometric and color features fusion for face detection and tracking in natural video sequences," In proc IWINAC (2), pp.72-80, 2007.
- [21] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," IEEE Trans. on PAMI, Vol.20, No.1, pp.23-28, 1998.

저자 소개

이 경 미(Kyoung-Mi Lee)

정회원



- 1993년 2월 : 덕성여자대학교 전산학과 (이학사)
- 1996년 2월 : 연세대학교 전산학과 (이학석사)
- 2001년 12월 : 미 The Univ. of Iowa, 전산학과 (전산학박사)
- 2003년 2월 ~ 현재 : 덕성여자대학교 컴퓨터공학부 교수

<관심분야> : 내용기반 영상검색, 멀티미디어 처리, 교육용 소프트웨어, 감시시스템