

# 적응적 피부영역 검출을 이용한 얼굴탐지

## (Face Detection using Adaptive Skin Region Extraction)

황 대 동<sup>\*</sup>      박 영 재<sup>\*</sup>      김 계 영<sup>\*\*</sup>  
 (Dae-Dong Hwang)    (Young-Jae Park)    (Gye-Young Kim)

**요약** 본 논문에서는 입력영상에서 적응적으로 피부색상 모델을 생성하여 얼굴을 탐지하는 방법을 제안한다. 제안하는 방법의 기본적인 절차는 먼저 눈의 특징을 인공신경망에 적용하여 눈 후보를 찾은 후, 그 주변의 색상을 이용하여 피부영역의 색상값 분포를 찾는다. 그 다음은 피부영역으로 검출된 색상값 분포를 이용하여 얼굴영역을 산출하고, 해당 얼굴영역 내에서 입 후보를 찾아 눈 후보와 입 후보의 구조적인 관계가 얼굴 구조와의 일치여부를 판단하여 얼굴영역을 검증하는 과정을 거친다. 이 방법은 눈을 찾아서 피부영역을 적응적으로 검출하기 때문에 기존의 얼굴탐지 방법들의 문제인 피부색상의 왜곡으로 인한 오검출을 해결하였다. 실험은 눈 탐지와, 피부 탐지, 입 탐지, 얼굴탐지에 대해 각각 수행하였다. 실험을 통하여 기존의 주요 방법들 보다 우수한 결과를 보였다.

**키워드** : 영상이해, 영상처리, 얼굴탐지, 눈탐지, 피부영역탐지

**Abstract** In this paper, we propose a method about producing skin color model adaptively in input image and face detection. The principle process which we proposed is finding eyes candidates by applying the eye features to neural network, and then using the around color to find the distribution of color value. There will be a verification process that producing face region by using color value distribution which is detected as skin region and find mouth candidate in corresponding face region; if eye candidate and mouth candidate's connection structure is similar with face structure, then it can be judged as a face. Because this method can detect skin region adaptively by finding eyes, we solve the rate of false positive about the distorted skin color which is used by existing face detection methods. The experiment was performed about detecting the eye, the skin, the mouth and the face individually. The results revealed that the proposed technique is better than the traditional techniques.

**Key words** : Computer Vision, Image Processing, Face Detection, Eye Detection, Skin Detection

### 1. 서론

영상에서 인체를 인식하는 기술은 보안, 통계, 군사, 의료 등 수많은 분야에서 널리 사용된다. 잡음과 복잡한

배경으로 구성된 영상에서 인체를 인식하기 위해서는 탐지가 선행되어야 하는데, 이러한 탐지과정에 사용되는 가장 강력한 특징 중 하나가 바로 얼굴이다. 손과 발의 개수를 보고 단지 몇 명의 사람이 있다고 추측만 할 뿐, 정확한 인원을 알 수 없으며, 몸통 부분은 의복을 입기 때문에 특징이 친차만별이고 팔, 다리의 움직임은 형태가 일정하지 않아 쉽게 추출할 수 없다. 하지만 얼굴은 마스크, 안경 등 몇몇 경우를 제한하면, 탐지에 큰 영향을 미치는 물체가 없기 때문에 다른 신체부위와 달리 특징이 쉽게 변하지 않고, 신체에 유일하게 하나만 존재하며, 크기 및 형태도 비교적 일정하다. 또한 나이, 성별, 감정과 같은 정보도 얻을 수 있어 생체인식 분야에서 가장 큰 비중을 차지하는 영역 중 하나이다.

현재까지 연구된 얼굴탐지기술은 여러 가지 방법이 있다. 그중에서 Kawaguchi와 Rizon[1]은 영상의 명암도와 에지를 이용한 방법을 제안하였고, Song과 Liu[2]도 에지를 이용하여 얼굴탐지를 시도하였으며, Brunelli

\* 이 논문은 2006년 정부(교육인적자원부)의 지원으로 한국 학술진흥재단의 지원을 받아 수행된 연구(KRF-2006-005-J03801)

<sup>\*</sup> 학생회원 : 숭실대학교 컴퓨터학과  
 chabert1@nate.com  
 webmaster@kuly.org

<sup>\*\*</sup> 종신회원 : 숭실대학교 컴퓨터학과 교수  
 gykim11@ssu.ac.kr

논문접수 : 2009년 7월 22일  
 심사완료 : 2009년 10월 21일

Copyright©2010 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨터의 실제 및 테더 제16권 제1호(2010.1)

와 Poggio[3]은 형판 정합을 통하여 얼굴을 탐지하는 방법을 제안하였다. 하지만 위에서 제안한 방법들은 영상의 해상도, 대비, 밝기에 영향을 많이 받는다는 단점을 가지고 있다. Pentland[4]는 고유영역을 이용하여 탐지하는 방법을 제안하였으며, Viola와 Jones[5], Fröba와 Ernst[6]는 AdaBoost를 이용한 얼굴탐지방법을 제안하였는데, 이 방법은 회전된 얼굴영상의 탐지에 적합하지 않고, 명암값을 Haar-like 특징으로 사용하기 때문에 얼굴과 유사한 그림자, 무늬, 주름 등에서도 얼굴이라고 탐지되는 경우가 많이 발생하는 단점이 있다.

Hsu와 Abdel-Mottaleb[7]은 얼굴구성요소인 눈과 입을 이용한 얼굴탐지방법을 제안하였다. 이 방법은 눈과 입의 색차성분과 휘도성분을 이용하여 눈 후보와 입 후보를 검출한 후 후보들의 조합을 통해 얼굴을 탐지하는 방법으로 회전된 얼굴에도 강건하게 탐지된다. 하지만 배경, 모발의 일부 등 얼굴구성요소가 아닌 특징들도 구성요소 후보로 검출되는 경우가 많아서 오검출된 구성요소가 많은 경우 연산속도가 상당히 느리다는 단점이 있다. 따라서 Hsu는 피부색을 검출하여 얼굴영역을 지정한 후 영역 내부의 구성요소만을 조합하는 방법을 이용하여 필요없는 구성요소 후보들을 제거하였다. 피부색상이 정확히 표현 및 분류될 수 있다면 피부색상 정보를 사용한 얼굴영역 추출은 매우 효과적인 방법이다. 하지만 현재까지 제안된 피부검출 방법으로는 인종별, 개인별 피부색의 차이, 조명의 색, 피부색과 유사한 색상의 배경, 색조 화장 및 분장, 광학기기에 의한 색 왜곡 등 다양한 인자들에 민감하므로 좋은 피부검출 결과를 얻기가 어렵다. 따라서 피부가 잘 검출되지 않는 이러한 영상에서 구성요소 기반의 얼굴탐지 방법은 당연히 좋은 결과를 얻을 수가 없다.

본 논문에서는 눈을 탐지하고 해당 눈 주변영역에서 피부색 샘플을 추출하여 추출된 샘플과 유사한 색상특징을 갖는 영역을 피부영역으로 사용하고 해당 영역에서 입을 찾아 얼굴을 검증하는 방식으로 기존의 구성요소 기반 얼굴탐지방법의 문제점인 피부색의 왜곡에 의한 오탐지를 증가 문제를 해결한 강건한 탐지 방법을 제안한다.

그림 1은 본 논문에서 제안한 방법의 개요도이다. 눈 후보영역 탐지 단계에서는 눈의 휘도와 색상특징을 이용한 EyeMap을 통해 1차 눈 후보영역을 탐지한 후, 인공신경망을 이용하여 잡음영역을 제거하는 과정을 수행한다. 피부영역 검출 단계에서는 선정된 1차 눈 후보의 주변영역에서 피부영역의 색상값 분포를 추출하고, 그와 유사한 색상값 분포를 갖는 피부영역들을 레이블링한다. 이렇게 레이블링 된 피부영역에 눈 후보가 하나 이상 존재한다면 얼굴 후보영역으로 선정한다. 입 후보영역 탐지 단계에서는 선정된 얼굴영역 후보의 피부색상 정

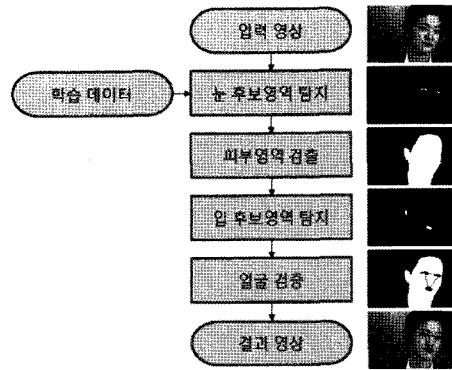


그림 1 시스템 개요도

보와 입의 특징을 이용하여 입 후보영역을 탐지하며, 얼굴 검증 단계에서는 선정된 얼굴 후보영역에서 눈 후보와 입 후보를 직선으로 각각 연결하고 해당 연결 구조가 얼굴 구조와 유사한지 판단하여 최종적으로 얼굴임을 검증한다.

본 논문은 전체 7장으로 구성되어 있다. 1장에서는 연구 배경 및 목적, 관련된 연구와 제안하는 방법에 대해서 설명하고, 2장에서는 눈의 특징과 인공신경망을 이용하여 눈 후보영역을 찾는 방법에 대해 설명한다. 3장에서는 탐지된 눈 후보영역을 이용하여 피부영역을 찾는 방법에 대해 설명하며, 4장에서는 피부영역의 색상정보와 입의 특징을 이용하여 입을 찾는 방법을 설명한다. 5장에서는 검출된 얼굴 영역 후보 내에서 눈 영역 후보, 입 영역 후보의 구조적 특징을 이용하여 검증하는 방법에 대해 설명한다. 6장에서는 제안한 방법과 기존의 방법을 동일한 조건에서 각각 실험하여 우수성과 문제점에 대해 살펴보고, 마지막으로 7장에서는 실험을 토대로 연구에 대한 결론을 내리고 향후 연구 방법에 대한 제시를 한다.

## 2. EyeMap과 인공신경망을 이용한 눈 후보영역 탐지

본 논문에서는 눈을 이용하여 얼굴의 피부영역을 찾고 최종 검증까지 하기 때문에 눈이 가장 중요한 요소라고 할 수 있다. 따라서 눈의 특징을 이용하여 1차적으로 후보를 선별한 후, 오탐지율을 낮추기 위하여 인공신경망을 이용한 검증 단계를 추가하였다.

우선 1차 눈 후보를 선별하기 위하여 눈과 그 주변의 명암 대비를 이용한 방법과 눈의 색상 특징을 이용한 방법을 사용한다. 이러한 특징을 이용하여 눈을 찾는 방법으로는 Hsu[7]가 제안한 EyeMap이 있다. EyeMap은 YCbCr 색상 모델에서 연산이 이루어지기 때문에 RGB 색상 모델을 식 (1)을 이용하여 YCbCr 색상 모델로 변환한다.

$$\begin{aligned} Y &= 16 + (65.738)R + (129.057)G + (25.064)B \\ Cb &= 128 + (-37.945)R + (-74.494)G + (112.439)B \\ Cr &= 128 + (112.439)R + (-94.154)G + (-18.285)B \end{aligned} \quad (1)$$

정면에서 바라본 눈은 검은자와 흰자로 이루어져 있어 명암 대비가 다른 영역에 비해 크게 나타난다. 이러한 특징은 EyeMapL을 이용하여 얻을 수 있다. 하지만 본 논문에서는 EyeMapL의 느린 속도를 보완하면서도 유사한 결과를 얻기 위해 EyeMapL을 수정한 식 (2)를 이용하여 눈을 찾는다.

$$EyeMapL = \frac{Y(x,y) \oplus g_{\sigma}(x,y)}{Y(x,y) + 1} \quad (2)$$

$\oplus$ 는 팽창연산을 의미한다. 그림 2(a)와 같이 입력영상  $Y$ 에 팽창연산을 적용한 결과를 입력영상으로 나누면 명암값이 높은 흰자가 명암값이 낮은 검은자와 속눈썹으로 팽창되어 눈 영역의 명암값이 전체적으로 높아진 영역을 원본 입력영상의 눈동자와 속눈썹같이 명암값이 낮은 영역으로 나누게 되므로 다른 영역에 비해 상대적으로 높은 값을 가지게 된다. Hsu[7]가 제안한 EyeMapL은 팽창연산을 적용한 영상에 침식연산을 적용한 연산을 나누었기 때문에 눈 영역으로 산출된 영역이 눈의 범위를 벗어나 모발이나 눈썹같은 영역에 붙어버리는 문제점이 발생하였다. 반면에, 개선한 EyeMapL은 눈의 범위를 벗어나지 않기 때문에 이러한 문제를 해결하였으며, 침식연산을 수행하지 않기 때문에 연산에 걸리는 시간도 감소되었다.

입력영상  $Y$ 에 팽창연산을 적용한 결과 그림 2(b)에 입력영상 그림 2(a)를 나누면 명암값이 높은 흰자가 명암값이 낮은 검은자와 속눈썹으로 팽창되어 눈 영역의 명암값이 전체적으로 높아진 영역에서 원본 입력영상의 눈동자와 속눈썹같이 명암값이 낮은 영역이 나뉘지게 되므로 그림 2(c)와 같이 다른 영역에 비해 상대적으로 높은 값을 가지게 된다. Hsu[7]가 제안한 EyeMapL은 팽창연산을 적용한 영상에 침식연산을 적용한 연산을 나누었기 때문에 눈 영역으로 산출된 영역이 눈의 범위를 벗어나 모발이나 눈썹같은 영역에 붙어버리는 문제점이 발생하였다. 반면에, 개선한 방법은 눈의 범위를 벗어나지 않기 때문에 이러한 문제를 해결하였으며, 침식연산을 수행하지 않기 때문에 연산에 걸리는 시간도 감소되었다.

색상 특징을 이용하여 눈을 찾는 방법은 Hsu[7]가 제안한 EyeMapC를 사용한다. EyeMapC는 식 (3)을 이용하여 구할 수 있다.

$$EyeMapC = \frac{1}{3} (C_b^2) + (\widetilde{C}_r)^2 + (C_b/C_r) \quad (3)$$

$C_b$ 는 YCbCr 색상모델에서의 파란색의 분포를 나타내는  $C_b$ 값을 의미하며,  $C_r$ 은 붉은색의 분포를 나타내는

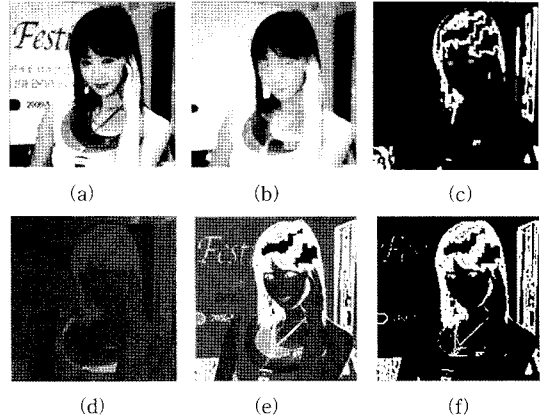


그림 2 EyeMap의 과정. (a) 원본 영상, (b) 팽창연산을 적용시킨 영상, (c) 영상(b)에서 영상(a)를 나눈 결과, (d) EyeMapL, (e) 영상(c)와 영상(d)의 AND연산 결과, (f) 영상(e)의 이진화 적용 영상

$Cr$ 값을 의미한다.  $C_b^2$ 과  $(\widetilde{C}_r)^2$ ,  $(C_b/C_r)$ 은 0에서 255사이의 값으로 정규화하여 사용한다. 이 방법은 일반적으로 눈의 색이  $C_b$ 값이 높고,  $Cr$ 값이 낮다는 관찰에 근거한 방식이다. 이렇게 산출한 EyeMapL과 EyeMapC를 식 (4)를 이용하여 그림 2(e)와 같이 EyeMap으로 만든다.

$$EyeMap = (EyeMapL) AND (EyeMapC) \quad (4)$$

EyeMap을 이진화하여 눈의 후보영역을 추출한 뒤, 레이블링 하여 1차 눈 후보 영역을 생성한다. 이렇게 생성한 눈 후보영역은 그림 2(f)와 같이 눈 뿐만 아니라 모발의 일부, 옷의 무늬, 배경 영역을 포함하고 있으며, 이러한 잡음을 인공신경망을 이용하여 제거한다.

인공신경망은 뉴런으로 불리는 단위처리 객체로 이루어져 있으며, 각각의 뉴런은 입력신호에 대한 연산을 수행한 후, 연결고리를 통하여 다음 뉴런에 정보를 전달하는 형태로 구성되어 있다. 각 연결고리에는 가중치가 할당되어 있으며, 다음 뉴런으로 전해지는 데이터에 가중치가 고려된 값이 전달되는 형태를 띈다. 신경망을 특정한 목적으로 사용하기 위해서는 가중치를 목적에 맞게 변경해 주어야 하며, 신경망의 입력값에 실제 출력값을 유도하는 목표값을 적용하여, 실제 출력값과 목표값의 오차를 최소화하기 위해 가중치를 수정하는 일련의 과정을 학습이라 한다. 학습 및 인식을 위해 인공신경망에 입력되는 눈 후보영역의 특징은 총 4개이며 그림 3과 같은 눈 후보영역의 최소경계사각형과 최소경계사각형의 3배에 해당하는 주변영역에서의 명암값과 에지 개수를 이용하여 산출한다. 특징은 식 (5)를 이용하여 눈 후보영역 내부의 에지의 개수를 눈 후보영역 외부, 즉 피부영역의 에지의 개수와 나누어 정규화한 값 ESM과

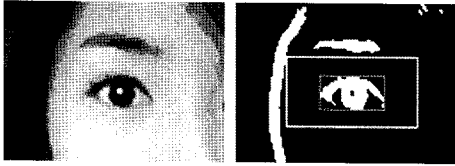


그림 3 눈 후보영역의 최소경계사각형과 눈 주변영역의 예

식 (6)을 이용하여 눈 후보영역 최소경계사각형 내부의 최대 명암값을 최소 명암값으로 나누어 정규화한 값 BSM, 식 (7)을 이용하여 눈 후보영역의 장축과 단축의 비율 CLR, 식 (8)을 이용하여 눈 후보영역 외부의 에지를 팽창연산시킨 후, 에지영역이 아닌 영역의 Cr값의 평균과 Cb값의 평균을 차연산한 값 SSR을 각각 산출하여 사용한다.

$$ESM = \sum_{i \in ER} e(i) / \sum_{j \in ER} e(j) \quad (5)$$

$$BSM = \sum_{i \in ER} \max(b(i)) / \sum_{j \in ER} \min(b(i)) \quad (6)$$

$$\begin{aligned} & \text{if}(width(ER) > height(ER)) \\ & \quad CLR = height(ER) / width(ER) \\ & \text{otherwise} \\ & \quad CLR = width(ER) / height(ER) \end{aligned} \quad (7)$$

$$\begin{aligned} SSR &= \sum_{i \in DER} C_r / \sum_{i \in DER} C_b \\ DER &= (N - ER) - \sum_{i \in ER} e(i) \oplus k(i) \end{aligned} \quad (8)$$

식 (5)~(8)에서 ER은 눈 후보영역의 최소경계사각형을, (N-ER)과  $\notin ER$ 은 눈 주변영역을 의미한다. e와 b는 각각 에지와 명암값을 나타낸다. 식 (5)는 눈 후보영역이 피부영역에 비해 에지가 조밀하고 많이 존재한다는 점에서 착안하였고, 식 (6)은 흰자의 명암값과 검은자의 명암값의 차이가 확연하게 나타난다는 점에서 착안한 방법이다. 식 (7)은 눈 후보영역 최소경계사각형의 장축과 단축의 비율은 대체로 일정하다는 점, 식 (8)은 피부영역 Cr값과 Cb값의 차는 대체로 일정하다는 관찰에 근거한 방식이다. 하지만 눈 주변영역에는 피부영역 뿐만 아니라, 눈썹, 모발과 같은 개체가 존재하며 올바르게 추출하지 않는 결과를 유도할 수 있다. 피부영역은 에지가 잘 산출되지 않는 반면에, 눈썹이나 모발은 에지가 비교적 뚜렷하게 산출된다는 특징을 바탕으로, 에지를 구한 후, 팽창연산을 적용하여 모발과 눈썹 그리고 그 영향이 미치는 주변영역을 찾고, 팽창된 에지가 아닌 영역을 라벨링하여 가장 크기가 큰 영역에서만 식 (8)을 적용한다(그림 4).

이렇게 추출된 ESM, BSM, CLR, SSR을 인공신경망에 대입하여 학습과 최종 눈 후보 탐지를 수행한다. 본

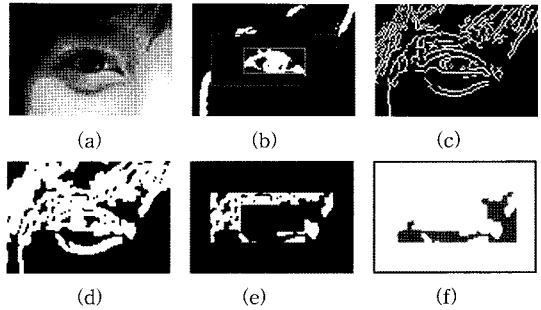


그림 4 Cr-Cb 값을 구하기 위한 전처리. (a) 원본 영상, (b) 눈 후보영역과 주변영역, (c) 캐니 에지 영상, (d) 영상(c)에 팽창연산을 적용한 결과, (e) 주변영역에서 본 팽창연산, (f) 레이블링 후 가장 큰 개체 추출 결과

논문에서는 잡음의 효과적인 제거를 위하여 잡음영역에 대해서도 동일한 방법으로 학습데이터를 산출하였다. 인공신경망의 구조는 입력노드 4개, 은닉노드 12개, 출력노드 1개로 그림 5와 같이 구성되어 있다. 눈 영역의 특징 데이터는 목표값을 1, 잡음영역의 특징 데이터는 목표값을 0으로 설정하여 탐지과정 결과 1개의 출력층에서 0~1 사이의 값을 출력함으로써 정확도를 백분율로 산출할 수 있도록 하였다. 따라서 학습 과정을 거쳐 가중치가 눈 영역 탐지에 적합하게 변경된 인공신경망을 생성한 후, 식 (5)~(8)을 이용하여 산출한 값을 인공신경망에 통과시키면 해당 후보가 눈에 가까운 확률이 산출되며, 이 산출값이 임계값 이하이면 눈 후보에서 제외하고 임계값 이상인 영역만 2차 눈 후보영역으로 사용한다(그림 6).

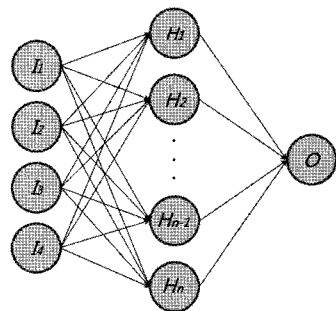


그림 5 인공신경망의 구조

### 3. 눈 주변 색상을 이용한 적응적 피부영역 검출

본 논문에서는 얼굴영역을 탐지하기 위해 피부영역을 검출하여 사용한다. 이러한 피부영역을 검출하는 기존의 방법은 대부분이 색상을 이용하는 방법이다. 이는 일반적으로 알려진 사람 피부색상을 이용하여 전체 영상에

서 찾는 방법으로 간단하지만 많은 문제를 가지고 있다. 우선 피부색은 인종, 개인, 환경 등에 의해 차이가 존재한다는 것이다. 그리고 색조화장, 조명, 카메라의 색 왜곡, 소프트웨어에 의한 후처리 등으로 인해서 피부색이라고 할 수 있는 범위가 넓어진다. 이러한 범위를 모두 수렴하기 위해 피부색의 탐지 분포를 크게 하면 나무, 돌, 동물의 가죽 등 배경영역의 잡음이 피부영역으로 검출되는 경우가 발생하며, 탐지 범위를 좁히면 피부영역이 검출되지 않는 경우가 발생한다. 따라서 본 논문에서는 눈 주변영역의 피부색 분포를 이용하여 얼굴의 피부색을 탐지하는 적응적 피부영역 검출방법을 사용하여 얼굴영역 후보로 사용될 피부영역을 검출한다.

적응적 피부영역 검출방법은 우선 눈 주변의 피부로부터 색상 모델을 만든다. 색상모델은 눈 주변영역에서 피부색을 검출한 다음, YCbCr공간의 Cb와 Cr을 사용하여 식 (9)와 같이 만들어진다. 식 (9)에서 Cb'는 영상 전체의 Cb값을 누적하여 평균을 취한 값이며 Cr'은 Cr 값을 누적하여 평균을 취한값이다. 이것들은 피부영역을 검출할 때 기준값으로 사용하게 된다.

$$C_b' = \frac{1}{N} \sum_{i=0}^N C_b(i), \quad C_r' = \frac{1}{N} \sum_{i=0}^N C_r(i) \quad (9)$$

$C_r/C_b$  공간에서 피부 색상 모델을 만들면 색의 분포가 어떠한 기준점으로부터 타원의 형태로 분포하는 것을 알 수 있다. 하지만 영상에 따라 타원이 위치하는 곳이 다르기 때문에 영상에서 직접 피부 샘플을 만들어 적응적인 피부 색상 모델을 제작한 것이다. Hsu[7]에 따르면 타원을 이용하여 피부 영역을 검출할 때 식 (10), (11)을 사용한다. 먼저 현재의 Cb, Cr값을 피부 색상 모델을 원점으로 하는 좌표로 이동시키고 경험적인 값을 이용하여 회전을 시킨다. 타원을 회전시킬 때 약간의 오차가 발생하며 오차를 보정하기 위해  $ec_x$ 와  $ec_y$ 를 사용하여 오차를 보정한다.

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (10)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} C_b - C_b' \\ C_r - C_r' \end{bmatrix} \quad (11)$$

$C_b'$ ,  $C_r'$ 은 앞서 구한 색상 모델의 결과 값이며  $\theta = 2.53(\text{radian})$ ,  $ec_x = 1.0$ ,  $ec_y = 2.0$ ,  $a = 25.39$ ,  $b = 14.03$ 의 값을 이용하였다. 이렇게 검출한 피부영역을 레이블링하여 얼굴영역 후보로 사용한다(그림 7).

#### 4. 입 후보영역 탐지

본 논문에서는 눈과 함께 입을 이용하여 얼굴 후보영역이 얼굴임을 검증한다. 입을 검출하기 위한 방법으로



그림 6 인공신경망 적용 결과의 예. (a) EyeMap을 이용하여 산출한 1차 후보영역, (b) 인공신경망을 이용한 잡음 제거 결과(임계값 0.85 적용)

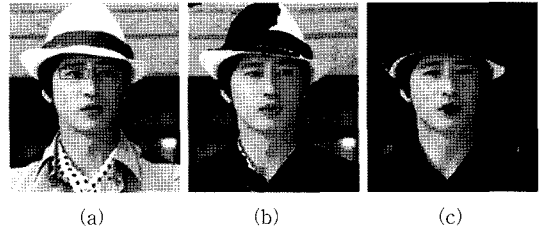


그림 7 피부 추출 결과. (a) 원본 영상, (b) Hsu가 제안한 피부영역 탐지 결과, (c) 영상(a)의 눈 주변영역(파란색 사각형)의 피부색을 이용한 피부탐지 결과

는 Hsu[7]가 제안한 LipMap을 사용한다. LipMap은 입술의 색상 특징을 이용하여 검출하는 방법으로, 식 (12)와 같이 산출한다.

$$LipMap = C_r^2 (C_r^2 - \eta \cdot C_r / C_b)^2 \quad (12)$$

EyeMap과 같이  $C_r^2$ 는 0에서 255사이의 값으로 정규화 한다. 입술은 일반적으로 붉은색을 띠며, 그 붉은색의 강도는 개인, 인종마다 다르지만 피부색보다 높게 나타난다는 특징이 있다. 따라서  $C_r^2$ 은  $C_r/C_b$ 보다 높은 값을 가지게 되며, 여기에 피부색을 의미하는  $\eta$ 를 적용함으로써 피부색보다  $C_r^2$ 이 높은 영역을 찾는다.  $\eta$ 은 식 (13)과 같다.

$$\eta = 0.95 \frac{\frac{1}{N} \sum_{(x,y) \in FG} C_r(x,y)^2}{\frac{1}{N} \sum_{(x,y) \in FG} C_r(x,y) / C_b(x,y)} \quad (13)$$

FG는 피부영역을 의미한다. 즉 피부영역의  $C_r$ 값과  $C_b$ 값으로  $\eta$ 를 산출하며, 눈 후보 주변영역에서 추출한 피부영역을 사용하여 산출한다.

이렇게 탐지된 입 후보영역(그림 8)이 피부색으로 산출된 얼굴 후보영역 내부에 존재하지 않는 경우 잡음으로 판단하여 제거하고, 얼굴 후보영역 내부에 존재하는 후보들을 이용하여 얼굴을 검증하게 된다.

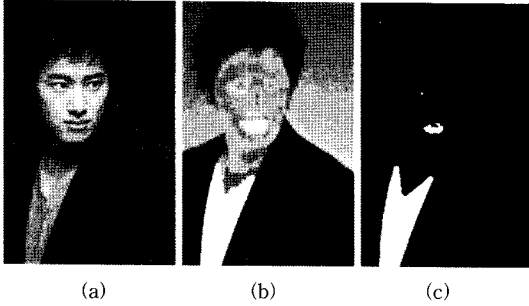


그림 8 LipMap의 과정. (a) 원본 영상, (b) LipMap, (c) 영상(b)를 이진화한 영상

5. 얼굴 검증

Lin[9]은 사람의 얼굴에서 눈과 입을 직선으로 연결해 보면 정삼각형에 가까운 삼각형이 나타난다는 특징(그림 9)을 바탕으로 얼굴탐지방법을 제안하였다. 본 논문에서는 이러한 삼각형 구조 기반 특징을 개선한 방법을 이용하여 얼굴을 검증하는 방법을 사용한다.

먼저 각 얼굴후보영역에서 두 눈 후보  $E_1$ 과  $E_2$ 를 연결하는 직선  $\overline{E_1E_2}$ 를 생성한다. 얼굴후보영역 내에 눈 후보가 3개 이상 존재한다면 모든 눈 후보에 대해 1:1 대응하는 직선들을 생성하고, 두 눈 후보가 존재하지 않는다면 해당 얼굴후보영역은 얼굴이 아닌 것으로 판단하고 제거한다. 그 후, 얼굴후보영역 안에 존재하는 입 후보  $M$ 과 각 눈 후보  $E_i$ 를 연결하는 직선을 생성한다. 정삼각형은 모든 변의 길이가 동일하다는 특징을 가지고 있기 때문에, 직선  $\overline{E_1E_2}$ 와  $\overline{E_1M}$ ,  $\overline{E_2M}$ 의 길이를 각각 비교하는 과정을 통해 두 눈과 입이 이루는 구조적 특징과 유사한 후보들의 조합을 찾는다. 이것은 식 (14)와 같다.

$$\alpha \cdot l(\overline{E_1E_2}) \leq l(\overline{E_1M}) + l(\overline{E_2M}) \leq \beta \cdot l(\overline{E_1E_2}) \quad (14)$$

$$\theta_{\min} \leq \left| \theta_{\overline{E_1E_2}} - \theta_{\overline{E_iM}} \right| \leq \theta_{\max} \quad (15)$$

$$\theta_{\overline{AB}} = (y_A - y_B) / (x_A - x_B)$$

$$FACE = \min(\sqrt{(x_{tc} - x_{sc})^2} + (y_{tc} - y_{sc})) \quad (16)$$

함수  $l$ 은 해당 직선의 길이를 나타낸다. 눈과 입을 각각 직선으로 연결한 구조가 정삼각형과 동일하다면, 직선  $\overline{E_1M}$ 과  $\overline{E_2M}$ 의 길이는  $\overline{E_1E_2}$ 의 길이의 두배와 동일하다. 하지만 실제로 측정해본 결과 눈과 입을 연결한 구조는 정확한 정삼각형으로 볼 수 없으며, 개인차, 표정, 얼굴의 회전 등으로 인해 조금씩 다른 모습을 보인다. 때문에 본 논문에서는 직선  $\overline{E_1E_2}$ 의 길이에 상수  $\alpha$ 와  $\beta$ 를 곱하여 삼각형 구조의 범위를 설정하였다. 본 논

문에서는  $\alpha = 1.4$ ,  $\beta = 2.5$ 를 사용하였다. 이는 얼굴영상 100장에서 직선  $\overline{E_1E_2}$ 과  $\overline{E_1M}$ ,  $\overline{E_2M}$ 의 길이를 측정하여 최대 비율과 최소 비율을 산출하고 간략화한 것이다.

정삼각형의 다른 특징으로는 3개의 각이 모두  $60^\circ$ 로 일정하다는 것이다. 따라서 식 (15)와 같이 눈과 눈을 잇는 직선  $\overline{E_1E_2}$ 와 각 눈과 입을 잇는 직선  $\overline{E_iM}$ 의 각을 산출한 후 차연산하여 나온 값이 임계각도  $\theta_{\min}$ 과  $\theta_{\max}$  안에 존재하지 않으면 제거하였다. 본 논문에서는 임계각도  $\theta_{\min} = 55^\circ$ ,  $\theta_{\max} = 83.5^\circ$ 를 사용하였다.  $\theta_{\min}$ 은 인종과 성별, 개인차에 의해 직선  $\overline{E_1E_2}$ 와 직선  $\overline{E_iM}$ 의 각도가 정확히  $60^\circ$ 가 되지 않고  $\pm 5^\circ$  정도 차이가 존재하는 사실을 감안하여 산출하였다.  $\theta_{\max}$ 는 촬영장비가 정지해 있다는 가정하에 얼굴이 목을 기준으로 회전될 때마다 상대적으로 직선  $\overline{E_1E_2}$ 와 직선  $\overline{E_iM}$ 의 각도가 변하며 최대  $83.5^\circ$ 까지 커진다는 관찰에 근거하였다.

한 얼굴영역에 여러 눈 후보와 입 후보가 존재한다면 위의 두 과정을 거친 후에도 두 개 이상의 삼각형 구조가 존재할 수 있다. 이러한 경우 얼굴로 검증되었다고 볼 수 있지만, 정확하게 눈과 입을 가리키는 삼각형을 제외한 다른 삼각형들은 눈 후보 대신 눈썹이나, 모발의 일부를 가리고 있으며, 눈이나 입의 정보가 추가적으로 필요한 프로그램에 응용 시 문제점이 발생한다. 따라서 본 논문에서는 식 (16)을 이용하여 하나만 선택하는 방법을 제안한다.  $tc$ 는 삼각형의 무게중심,  $sc$ 는 얼굴 후보영역의 무게중심을 의미한다. 즉, 두 무게중심의 거리가 가장 가까운 하나의 후보만을 선택하여 최종적으로 하나의 얼굴 후보영역에 하나의 삼각형 구조만이 존재하게 된다.

눈, 입 후보가 존재하지 않거나 삼각형 구조 기반 특징을 갖고 있지 않는 얼굴 후보영역은 제거되었기 때문에 최종적으로 삼각형 구조 기반 특징이 존재하는 얼굴 후보영역만 남아있으며, 이 후보영역들을 얼굴로 판단한다.

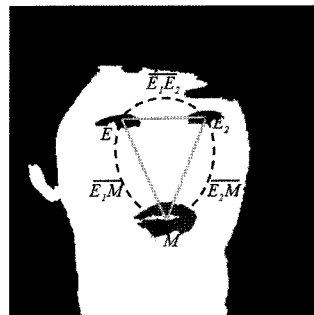


그림 9 얼굴의 삼각형 구조 기반 특징

6. 실험 및 결과

본 실험에는 Intel Pentium-4 3.0Ghz CPU, 1GB 메모리, nVIDIA Geforce 9500 GT 그래픽카드를 탑재한 IBM 호환기종 컴퓨터가 사용되었으며, 운영체제는 Microsoft Windows XP SP3, 사용언어는 Mathworks Matlab 7.6을 사용하였다. 실험영상으로는 인터넷을 이용하여 수집한 인물영상 337장과 인물이 없는 자연영상 81장, 자체적으로 촬영한 인물영상 124장으로 구성하였다. 인공신경망의 학습에 사용된 데이터는 눈 영역에서 추출한 데이터 200개와 잡음영역에서 추출한 데이터 50개를 5000회 반복 학습시켜 사용하였다.

본 논문에서는 얼굴탐지에 대한 실험 뿐만 아니라 기존의 방법과 제안한 방법의 차이를 자세히 알 수 있도록 눈 후보 탐지 성능평가와 피부영역 탐지 성능평가, 입 후보 탐지 성능평가를 각각 수행하였다.

표 1은 눈 후보 탐지에 대한 실험 결과를 보여준다. 탐지율은 영상에서 눈만 탐지되었을 경우의 확률을 100%로 보고, 오탐지는 잡음 개수의 평균을 산출하였다. 기존의 방법은 잡음을 피부영역만을 이용해서 제거하기 때문에, EyeMap만을 수행하였을 경우 상당히 많은 잡음이 나타나는 것을 볼 수 있다. 또한 그림 10(f)와 같이 팽창연산에 의해 모발의 일부에 붙어버리는 경우 눈이라고 볼 수 없기에 탐지율이 제안한 방법에 비해 저조하였다. 반면에 우리가 제안한 방법은 EyeMapL

표 1 눈 후보 탐지에 대한 실험 결과

인물영상	HSU 방법		제안한 방법	
	탐지율(%)	오탐지(개)	탐지율(%)	오탐지(개)
인물영상	73.45	28.42	94.25	10.58

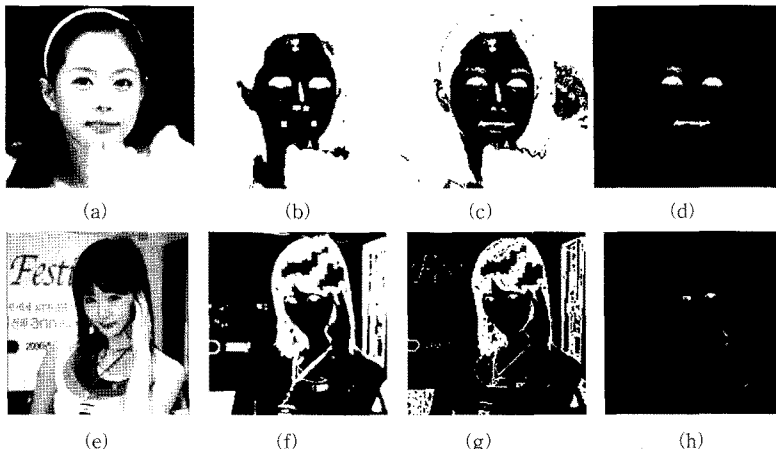


그림 10 눈 후보 탐지 결과 영상. (a),(e)원본 영상, (b),(f)기존의 방법으로 탐지한 눈 후보, (c),(g)제안한 방법으로 탐지한 눈 후보, (d),(h)인공신경망을 적용하여 잡음을 제거한 영상

표 2 피부영역 탐지에 대한 실험 결과

인물영상	HSU 방법	제안한 방법
	오차율(%)	오차율(%)
인물영상	18.22	4.35

을 개선한 방법으로 인해 눈 후보가 잡음에 붙어버리는 경우가 드물었으며, 인공신경망을 통해 잡음을 한번 더 제거하므로 오탐지율이 기존의 방법에 비해 감소하였다. 인공신경망으로 인해 제대로 추출된 눈이 잡음으로 판단되는 경우가 일부 발생하였는데, 이러한 문제는 학습의 양을 늘리면 해결될 수 있을 것으로 판단된다.

표 2는 피부영역 탐지에 대한 실험 결과를 보여준다. 수작업으로 추출한 피부영역과 탐지된 피부영역이 95% 이상 일치할 경우의 오차율을 0%로 하였다. 제안한 방법은 18.22%로 나온 반면에 제안한 방법은 오차율이 상당히 감소한 4.35%로 나온것을 볼 수 있다. 특히 조명 또는 영상처리 프로그램의 후보정도로 인해 피부가 탐지되지 않는 경우나 배경색이 피부색과 유사한 영상 등 기존 방법의 검출율이 0%에 가까운 영상에서도 100%에 가까운 우수한 탐지율을 보인다. 이는 기존의 방법이 이미 실험과 관찰로 산출된 고정된 임계값으로 피부를 탐지하기 때문에 그림 11(b)와 같이 해당 임계값을 벗어난 피부색을 탐지하지 못하거나, 그림 11(e)와 같이 임계값에 포함되는 배경색을 걸러낼 수 없는 반면에, 제안한 방법은 눈 주변의 피부색을 이용하여 임계값을 수정하기 때문에 피부영역을 강건히 찾으며, 배경색도 상당부분 제거하기 때문이다.

표 3은 입 후보 탐지에 대한 실험 결과를 보여준다. 기존의 방법과 제안한 방법 모두 Hsu가 제안한 LipMap을 이용하여 입을 탐지하기 때문에 유사한 탐지율과 오탐지

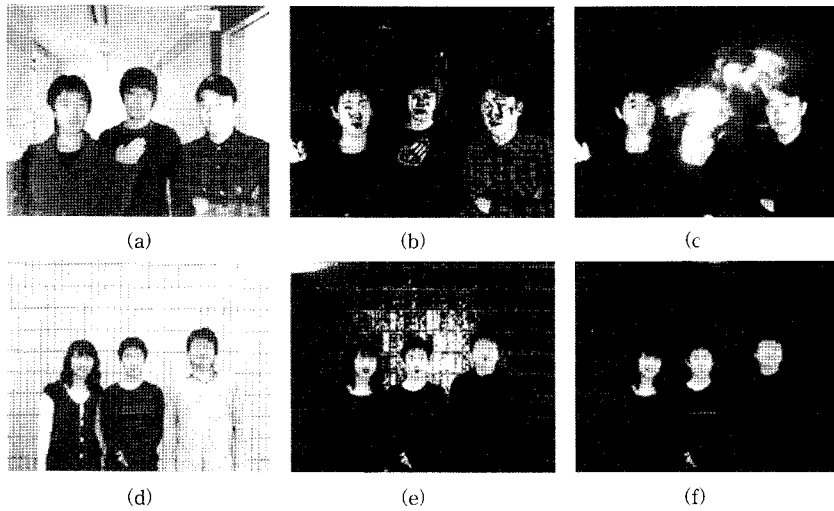


그림 11 피부영역 탐지 결과 영상. (a),(d)원본 영상, (b),(e)기존의 방법으로 탐지한 피부영역, (c),(f)제안한 방법으로 탐지한 피부영역

표 3 입 후보 탐지에 대한 실험 결과

	HSU 방법		제안한 방법	
	탐지(%)	오탐지(개)	탐지(%)	오탐지(개)
인물영상	93.45	3.15	94.60	2.68

표 4 얼굴탐지에 대한 실험 결과

	HSU 방법		제안한 방법	
	탐지	오탐지	탐지	오탐지
인물영상	88.17	12.53	92.25	7.27
자연영상	-	9.94	-	2.17

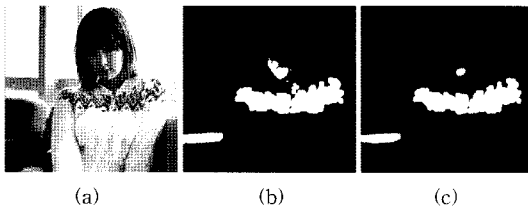


그림 12 입 후보 탐지 결과. (a) 원본 영상, (b) 기존의 방법으로 탐지한 입 후보, (c) 제안한 방법으로 탐지한 입 후보

울을 보이지만, LipMap은 피부영역의 탐지율에 영향을 받는  $\eta$ 값을 포함하기 때문에 그림 12(b)와 같이 입의 범위를 벗어나거나 미탐지한 경우가 발생하기도 하였으며 제안한 방법이 대부분의 영상에서 우수한 성능을 보인다.

표 4는 기존의 방법과 제안한 방법을 실험영상으로 실험한 결과이다. 실험영상은 얼굴이 존재하는 영상과 오탐지율을 측정하기 위해 얼굴이 존재하지 않는 자연 영상 두 개로 나누어 실험하였다. 기존의 방법, 제안한 방법 모두 높은 탐지율을 보였다. 하지만 기존의 방법은 얼굴이 아닌 배경영역에서 오탐지 하는 경우, 예를 들면 염색한 모발영역에서 에지가 강한 부분을 눈으로 인식하거나, 배경의 광고판에서 검은 글씨를 눈, 빨간 글씨를 입으로 탐지하는 경우가 발생하였다. 반면에 제안한

방법은 인공신경망에서 눈의 특징을 이용하여 후보를 상당수 단축하고, 적응적 피부영역 검출방법을 사용하여 피부영역이 비교적 정확하게 검출되었기 때문에 이러한 문제점들을 상당수 해결하였다.

그림 13은 기존의 방법과 제안한 방법으로 실험한 결과영상이다. 얼굴영역으로 탐지된 부분은 눈과 입을 꼭 지점으로 하는 삼각형이 표기되도록 하였다. 기존의 방법은 피부영역의 범위가 모발까지 포함하는 현상이 발생하여, 눈을 잘 못찾는 문제점이 발생하였으며, EyeMap 적용 시, 산출한 영역이 원래의 눈 영역을 벗어나 모발에 붙어버려 눈 영역으로 탐지하지 못하거나 모발영역을 눈으로 인식하는 예상치 못한 결과가 나타났다. 또한 사람의 얼굴임에도 불구하고 피부영역이 탐지되지 않아 찾지 못하거나 배경영역에서 얼굴이 탐지된 오탐지 현상도 일부 발생하였다. 반면에 제안한 방법은 피부영역이 비교적 정확하게 검출되어 강하게 눈과 입을 볼 수 있다.

자연영상에는 얼굴이 존재하지 않는 모발, 피부색과 유사한 나무, 돌 등이 존재하며, 심지어는 해질녘의 노을도 피부색으로 추출하는 경우가 발생한다. 이러한 영역, 특히 나무와 같이 피부색과 유사하고 독특한 질감과 무늬가 있는 개체에서는 EyeMap과 LipMap을 적용하면 유사한 특징의 잡음이 생성되는데 기존의 방법은 이러한





그림 13 실험 결과 영상

영역도 눈과 입이라고 인식하고 얼굴을 검출한 것을 볼 수 있다. 하지만 제안한 방법은 동일한 영상에서 인공신경망을 통해 눈 후보 검출을 시도하였으나 학습시 눈과 유사한 특징을 가지고 있는 영역을 제외하도록 하여 얼굴을 오탐지하는 경우가 거의 발생하지 않았다.

### 7. 결론 및 향후연구

구성요소 기반 얼굴탐지는 다른 얼굴탐지 방법에 비해 회전된 얼굴을 강건하게 탐지하며, 얼굴탐지와 동시에 눈과 입 등 구성요소의 위치정보를 함께 추출하기 때문에, 눈과 입 정보가 요구되는 콘텐츠에 유용하게 쓰일 수 있는 방법이다. 하지만 기존의 방법은 피부영역이 잘 추출되지 않는 저해상도 영상 및 다양한 색의 조명에서 촬영한 영상에서는 좋은 결과를 얻기 힘들어 구성요소 기반 얼굴탐지 방법의 여러 장점에도 불구하고 다른 얼굴탐지 방법에 비해 응용되는 경우가 적었다. 본 연구는 적응적 피부영역 검출 방법을 이용하여 피부색이 잘 추출되지 않는 영상에서도 강건하게 얼굴을 탐지하는 방법을 목표로 하였으며, 따라서 기존의 피부검출 방법인 단순히 색상에 임계값을 적용하여 피부영역 탐지하는 방법을 탈피하고, 눈을 이용하여 해당 인물의 피부색을 추출, 추출된 피부색을 바탕으로 얼굴 후보영역

을 생성하고 얼굴을 탐지하는 방법을 제안하였다. 실험 결과 기존의 방법과는 다르게 조명, 개인차로 인해 피부영역이 제대로 검출되지 않는 영상과 피부색과 유사한 배경을 가지고 있는 영상에서도 우수한 성능을 보여주었다.

하지만 인공신경망을 이용한 눈 검출 시, 일부 잡음영역이 여전히 눈으로 검출되는 경우가 발생하였으며, 이는 오작동의 원인이 되었다. 학습 데이터의 양을 증가시키는 것으로 어느정도 해결할 수 있지만, 보다 우수한 특징을 찾아내어 사용하는 것이 근본적인 문제 해결 방안으로 보인다. 따라서 본 논문에서 제시한 방법 뿐만 아니라 다른 방법으로 특징을 산출하여 탐지하는 연구가 지속되어야 하며, EyeMap, LipMap과 같은 얼굴구성요소 후보 영역의 산출 방법을 개선하여 보다 빠르고 잡음이 없이 정확하게 산출하는 방법에 대한 연구도 병행한다면, 좋은 결과가 있을 것이라 판단한다.

### 참고 문헌

- [1] T. Kawaguchi, and M. Rizon, "Iris detection using intensity and edge information," *Pattern Recognition*, vol.36, no.22, pp.549-562, 2003.
- [2] J. Song, Z. Chi, and J. Liu, "A robust eye detection method using combined binary edge and intensity information," *Pattern Recognition*, vol.39, no.6, pp.1110-1125, 2006.
- [3] R. Brunelli, and T. Poggio, "Face recognition: features versus templates," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.15, no.10, pp.1042-1052, 1993.
- [4] A. Pentland, B. Moghaddam, and Thad Starner, "View-based and modular eigenspaces for face recognition," In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.84-91, 1994.
- [5] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.511-518, 2001.
- [6] B. Fröba, and A. Ernst, "Face- Detection with the Modified Census Transform," In *Proceedings of the IEEE Conference on Automatic Face and Gesture Recognition*, pp.91-96, 2004.
- [7] R.L. Hsu, M. Abdel-Mottaleb, "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, no.5, pp.696-706, 2002.
- [8] P. T. Jackway, "Scale-Space Properties of the Multiscale Morphological Dilation-Erosion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.18, no.1, pp.38-51, 1996.
- [9] C. Lin, K.C. Fan, "Triangle-based approach to the detection of human face," *Pattern Recognition*

- Society*, vol.34, pp.1271-1284, 2001.
- [10] B. Heisele, T. Serre, M. Pontil, T. Poggio, "Component-based face detection," *IEEE Conf. on Computer Vision and Pattern Recognition*, vol.1, pp.657-662, 2001.
- [11] M. Abdel-Mottaleb, A. Elgammal, "Face Detection in Complex Environments from Color Images," *IEEE Conf. Image Processing*, pp.622-626, 1999.
- [12] J. Shih, C. Lee, and C. Yang, "An Adult Image Identification System Employing Image Retrieval Technique," *Pattern Recognition Letters*, vol.28, pp.2367-2374, 2007.
- [13] K. M. Lee, "Component-based detection and verification," *Pattern Recognition Letters*, vol.29, pp.200-214, 2008.
- [14] C. Lin, "Face Detection in complicated backgrounds and different illumination conditions by using YCbCr color space and neural network," *Pattern Recognition Letters*, vol.28, pp.2190-2200, 2007.
- [15] S. Liapis, G. Tziritas, "Color and texture image retrieval using chromaticity histogram and wavelet frames," *IEEE Trans. Multimedia* 6, vol.5, pp.676-686, 2004.
- [16] J. S. Lee, Y. M. Kuo, P. C. Chung, E. L. Chen, "Naked image detection based on adaptive and extensible skin color model," *Pattern Recognition Society*, vol.40, pp.2261-2270, 2007.
- [17] Kyung-Min Cho, Jeong-Hun Jang, Ki-Sang Hong, "Adaptive skin-color filter," *Pattern Recognition Society*, vol.34, pp.1067-1073, 2001.
- [18] J. Song, Z. Chi, J. Liu, "A robust eye detection method using combined binary edge and intensity information," *Pattern Recognition*, vol.39, no.6, pp.1110-1125, 2006.
- [19] N. Otus, "A threshold selection method from gray-level histogram," *IEEE Trans. Systems Man Cybernet. SMC-8*, pp.62-66, 1978.
- [20] Jau-Ling Shih, Chang-Hsing Lee, Chang-Shen Yang, "An adult images identification system employing image retrieval technique," *Pattern Recognition Letters*, vol.28, pp.2367-2374, 2007.
- [21] Y. J. Park, S. W. Jang, G. Y. Kim, "A Study on Extraction of Skin Region and Lip Using Skin Color of Eye Zone," *Journal of KSCI*, vol.14, no.4, pp.19-30, 2009.



박 영 재

2005년 청운대학교 컴퓨터학과(공학사)  
2008년 숭실대학교 컴퓨터학과(공학석사)  
2008년~현재 숭실대학교 컴퓨터학과 박사과정 재학중. 관심분야는 컴퓨터비전, 영상처리, 패턴인식



김 계 영

1990년 숭실대학교 전자계산학과(공학사)  
1992년 숭실대학교 컴퓨터학과(공학석사)  
1996년 숭실대학교 컴퓨터학과(공학박사)  
2001년~현재 숭실대학교 컴퓨터학과 부교수. 관심분야는 컴퓨터비전, 형태인식, 생체인식, 증강현실, 영상처리



황 대 동

2007년 한경대학교 컴퓨터공학과(공학사)  
2007년~현재 숭실대학교 컴퓨터학과 석사과정 재학중. 관심분야는 컴퓨터비전, 패턴인식, 생체인식