

# 피부색상을 이용한 유해영상 분류기 개발

## Development of an Adult Image Classifier using Skin Color

윤진성\*, 김계영\*, 최형일\*\*  
승실대학교 컴퓨터학과\*, 승실대학교 미디어학과\*\*

Jin-Sung Yoon(kisshajo@chol.com)\*, Gye-Young Kim(gykim11@ssu.ac.kr)\*,  
Hyung-Il Choi(hic@ssu.ac.kr)\*\*

### 요약

최근 인터넷에 유통되는 유해영상이 급증하면서 이들을 자동으로 차단하는 컴퓨터비전 기술의 연구가 활발히 이루어지고 있다. 본 논문에서는 피부색상을 이용한 유해영상 분류도구를 연구 및 개발한다. 제안하는 분류도구는 2단계로 구성되며, 1단계에서는 피부색 분류기를 이용하여 입력영상에서 피부색 영역을 검출하고, 2단계에서는 영역특징 분류기를 이용하여 앞서 검출된 피부색 영역의 비율과 위치 특징을 무해 또는 유해로 분류한다. 피부색 분류기는 히스토그램 모델에 기반하여 무해영상과 유해영상의 RGB 값으로 학습되며, 영역특징 분류기는 SVM(Support Vector Machine)에 기반하여 영상의 29개 지역의 피부색 비율로 학습된다. 실험결과 제안하는 분류기는 92.80%의 검출율(Detection Rate)과 6.73%의 양성오류율(False Positive Rate)을 나타내었다.

■ 중심어 : | 유해영상 | 피부색 검출 | 히스토그램 모델 | SVM |

### Abstract

To classifying and filtering of adult images, in recent the computer vision techniques are actively investigated because rapidly increase for the amount of adult images accessible on the Internet. In this paper, we investigate and develop the tool filtering of adult images using skin color model. The tool is consisting of two steps. In the first step, we use a skin color classifier to extract skin color regions from an image. In the nest step, we use a region feature classifier to determine whether an image is an adult image or not an adult image depending on extracted skin color regions. Using histogram color model, a skin color classifier is trained for RGB color values of adult images and not adult images. Using SVM, a region feature classifier is trained for skin color ratio on 29 regions of adult images. Experimental results show that suggested classifier achieve a detection rate of 92.80% with 6.73% false positives.

■ keyword : | Adult Image | Skin Color Detection | Histogram Model | SVM |

## 1. 서론

최근 인터넷에서 음란 콘텐츠를 통제하는 문제가 사

회적 이슈가 되고 있다. 이에 따라 음란 콘텐츠로부터 청소년들을 보호하기 위한 각종 심의 및 법률들이 강화되고 있으며, 음란 콘텐츠의 자동차단 서비스를 제공하

\* This research is supported by Korea Research Foundation(KRF-2006-005-J03801)

접수번호 : #090225-001

접수일자 : 2009년 02월 25일

심사완료일 : 2009년 03월 23일

교신저자 : 최형일, e-mail : hic@ssu.ac.kr

는 기업체의 수가 증가하고 있다. 동영상은 음란 콘텐츠의 가장 큰 비중을 차지하고 있으며, 대부분의 동영상 자동차단 서비스는 기존 동영상의 크기, 문자 그리고 해시 값 등을 데이터베이스(DB)에 저장하고 게시되는 동영상과 비교하여 차단하는 기술이기 때문에 DB 구축과 갱신에 많은 인력이 필요하고 수정된 동영상과 개인 UCC와 같이 DB 구축이 어려운 동영상에 취약하다는 단점을 가진다. 이에 따라 최근 동영상의 색상이나 패턴정보를 읽어 유해정도를 인식하는 컴퓨터비전 기술의 연구가 활발히 진행되고 있다.

컴퓨터비전 기술에서 가장 유용하게 이용할 수 있는 유해동영상의 특징은 동영상의 각 프레임에 포함된 피부색이다. Shih 등은 RGB 모델에서 Ycbr 모델로 색 공간을 변환한 영상에서 임계치를 설정하여 피부색 영역을 검출하고, 검출된 피부색 영역에서 모양과 위치 특징을 추출하여 DB와 비교하는 유해영상 검증 시스템을 제안하였다[1]. 인터넷에 유통되는 유해영상은 다양한 피부색을 가지기 때문에 검출에 임계치를 이용하는 방법보다 학습모델을 이용하는 방법이 더 효과적일 수 있다. 피부색 학습에 적용할 수 있는 학습모델로써 신경망(Neural network), GMM(Gaussian Mixture Model), 베이저언 망(Bayesian Network) 등이 있으며 [2], 이중 Jones과 Rehg이 제안한 히스토그램 학습 모델은 단순하면서도 매우 효과적이라는 장점이 있는 반면 학습에 상당량의 샘플이 필요하다는 단점이 있다[3]. 그러나 샘플이 더해질수록 신뢰도를 높일 수 있기 때문에 응용에 따라 가장 적합한 모델이 될 수 있다. Zhu 등은 히스토그램이나 GMM으로 검출한 피부색 영역을 GMM를 통해 정제하는 방법을 제안하였다[4]. 이때 형성된 두 개의 가우시안 중에서 피부영역에 해당하는 가우시안과 배경영역에 해당하는 가우시안을 구별하기 위해 가우시안의 파라미터 값과 영역특징을 SVM[5]으로 학습한 분류기를 이용하였다.

본 논문에서는 피부색상을 이용하는 유해영상 분류 도구를 제안한다. 분류 도구는 크게 피부색 분류기와 영역특징 분류기로 구성된다. 피부색 분류기는 유해영상과 무해영상의 색을 학습한 두 개의 히스토그램으로 구성된다. 인터넷에 유통되는 유해영상은 다양한 조명

을 가진다. 본 논문에서는 다양한 조명의 유해영상을 처리하기 위해 정상적인 조명의 유해영상과 비정상적인 조명의 유해영상을 따로 수집하여 각각의 피부색 분류기를 생성한다. 피부색 분류기가 생성되면 베이저언 규칙을 이용하여 빠르게 입력화소를 피부색이나 배경색으로 분류할 수 있다. 유해영상들은 검출된 피부색의 비율과 위치에 관하여 유사성을 가진다. 또한 유해성이 강한 영상은 영상의 중심뿐만 아니라 외곽에서도 피부색이 검출되는 경우가 많다. 본 논문에서는 이러한 유해영상의 특징을 유해영상 분류에 이용하기 위해 영상의 계층적으로 분할하여 구한 21개 지역의 피부색 비율과 대각선으로 분할하여 구한 8개 지역의 피부색 비율을 SVM으로 학습한 영역특징 분류기를 이용한다. 영상이 입력되면 분류 도구는 피부색 분류기로 피부색 영역을 검출하고, 검출된 영역의 특징으로 영역특징 분류기가 입력영상을 무해영상과 유해영상을 분류한다. 다양한 조명의 영상에 대한 실험과 기존 유해영상 분류 기술과의 비교 실험을 통해 제안하는 분류 도구의 성능을 확인한다.

본 논문의 2.1절에서는 히스토그램 모델을 이용한 피부색 학습방법과 검출방법을, 2.2절에서는 유해영상의 피부색이 다양한 범주로 분류될 수 있음을 보이고 이로 인한 문제점 또한 기술한다. 2.3절에서 검출된 피부색 영역의 정제 방법을 기술하며, 2.4절에서 영역특징벡터의 추출방법과 SVM을 통합 학습방법을 기술한다. 3장에서 실험결과를 보이고 4장에서 결론을 맺는다.

## II. 분류기 학습

### 1. 피부색 분류기

#### 1.1 히스토그램 모델

피부색 분류기는 피부색과 배경색을 학습한 두 개의 히스토그램으로 구성된다[3]. 히스토그램 학습방법은 영상처리 분야에서 밝기 값 분석에 사용되는 히스토그램의 생성방법과 유사하다. 차이점은 밝기 값 분석을 위한 히스토그램은 256개의 빈에 영상에서 나타나는 밝기 값의 분포를 기록하지만, 피부색과 배경색 히스토

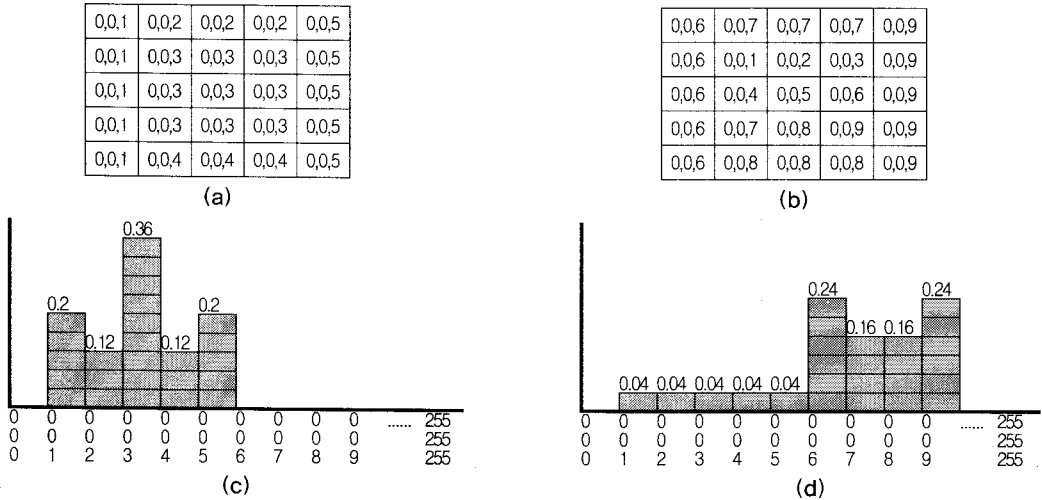


그림 1. 피부색 분류기의 히스토그램 학습 예. (a) 피부색 샘플의 RGB 값, (b) 배경색 샘플의 RGB 값, (c) 피부색 히스토그램, (d) 배경색 히스토그램

그림은 RGB 값의 최대 조합 값인 256×256×256개의 빈에 학습샘플 전체의 RGB 값의 분포를 기록한다는 것이다. [그림 1]에서 예를 들어 히스토그램의 학습과정을 보여준다. (a)와 같은 피부색 샘플이 주어지면 각 화소의 RGB 값의 빈도만큼 해당 빈의 값을 증가시킨다. 예를 들어 (0,0,2)값의 화소는 3개이므로 해당 빈의 값을 3만큼 증가시켜 (c)와 같은 히스토그램이 생성된다. 마찬가지로 (c)와 같은 배경색 샘플이 주어지면 각 색상의 빈도만큼 해당 빈의 값을 증가시켜 (d)와 같은 배경색 히스토그램을 생성된다. 빈 위에 표시된 값은 빈의 값을 전체 값으로 나눈 값이다. 그림에서는 하나의 영상에서 발생하는 색 값의 빈도를 기록하였지만 실제응용에서는 학습샘플 전체에 대한 색 값의 빈도를 기록한다. 본 논문에서는 인터넷을 통해 수집한 유해영상의 피부영역을 포토샵 등의 도구로 지정한 다음 학습과정에서 지정영역의 화소들로 피부색 히스토그램을, 무해영상 전체 화소들로 배경색 히스토그램을 생성하였다. 분류단계에서는 관측 화소를 분류하기 위해 관측 화소의 확률 값을 식 (1)과 같이 계산할 수 있다.

$$P(c) = \frac{\text{count}(c)}{T} \quad (1)$$

여기서  $c$ 는 관측 화소의 RGB 값을,  $\text{count}(c)$ 는 그에 해당하는 히스토그램 빈의 값을 나타내고  $T$ 는 히스토

그램 빈의 전체 누적 값을 나타낸다. 이와 같은 방법으로 관측화소가 피부색과 배경색에 속할 확률을 식 (2)와 같이 계산할 수 있다.

$$P(c/\text{skin}) = \frac{s(c)}{T_s}, \quad P(c/\text{non-skin}) = \frac{n(c)}{T_n} \quad (2)$$

여기서  $s(c)$ 와  $n(c)$ 는 피부색과 배경색 히스토그램 빈의 값이다. 계산된 두 확률 값으로 식 (3)이 성립하면 관측화소를 피부색 화소로 분류된다.

$$\frac{P(c/\text{skin})}{P(c/\text{non-skin})} \geq \theta \quad (3)$$

여기서  $\theta$ 는 검출율과 양성오류를 사이의 상충관계 (trade-off)를 조절하는 임계치로써 식 (4)와 같이 계산할 수 있다.

$$\theta = \frac{c_p P(c/\text{non-skin})}{c_n P(c/\text{skin})} \quad (4)$$

여기서  $c_p$ 와  $c_n$ 는 가중치로써  $c_n$ 이 클수록 피부가 많이 검출되고 반대로  $c_p$ 가 클수록 피부가 적게 검출된다. 예를 들어 분류단계를 구체적으로 설명한다. [그림 1]에서 학습된 두 개의 히스토그램이 주어지고 관측된 화소의 RGB 값이 (0, 0, 2)라고 한다면, 식 (3)에 해당하는 값들을 대입한 결과  $(0.12/0.04=3.0) \geq 0.333$ 이 성립하기 때문에 관측화소는 피부색 화소로 분류된다. 그러나  $c_p$ 와  $c_n$  값을 조절하여 검출되는 피부색의 정도를 조절할

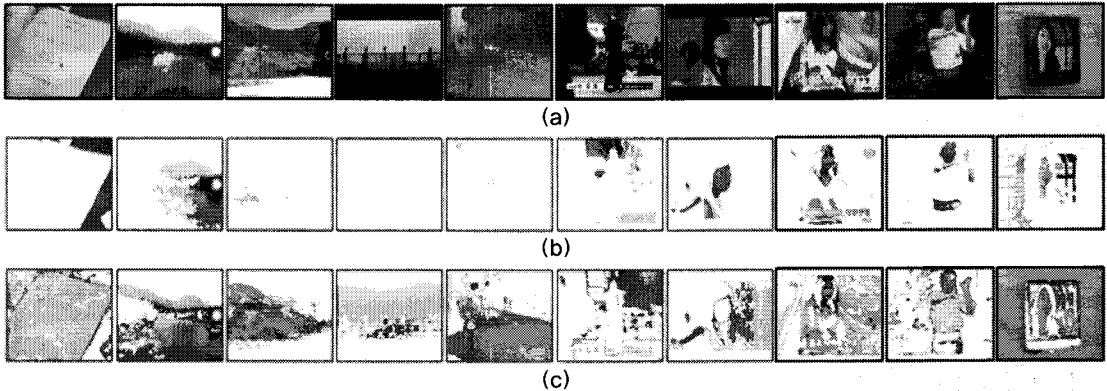


그림 2. 피부색 검출결과. (a) 정상조명 영상, (b) 정상조명의 피부색 검출결과, (c) 비정상조명의 피부색 검출결과

수 있다. 예의 경우  $c_p$ 가  $c_n$ 보다 9배 이상 크면 관측화소는 배경색 화소로 분류될 것이다.

[4]에서 학습화소의 RGB 값을  $32 \times 32 \times 32$ 로 양자화(quantization)하여 학습한 히스토그램이 가장 높은 성능을 나타내었다.  $256^3$ 크기의 히스토그램의 경우 너무 커서 색상 분포를 상세하게 나타내기 어렵기 때문이다. 실험을 통해 학습샘플의 양을 극단적으로 높이면  $256^3$  크기의 히스토그램이 가장 높은 성능을 나타냈지만  $32^3$  크기의 히스토그램에 비해 성능차이가 미세했음을 확인할 수 있었다. 따라서 본 논문에서도 RGB 값을  $32 \times 32 \times 32$ 로 양자하여 학습한  $32^3$  크기의 히스토그램을 사용한다.

### 1.2 유해영상의 조명 특성

인터넷에 유통되는 유해영상들은 다양한 조명환경에서 촬영된다. 특히 정상적이지 않은 조명에서 촬영된 영상의 피부색은 정상조명 영상의 도로, 잔디, 동물 등과 유사한 색을 가지기 때문에 분류기 성능을 저하시키는 요인이 된다. [그림 2]에서 이러한 색의 변화가 야기하는 문제의 예를 보여준다. 첫 번째 행은 정상조명 영상이며 두 번째 행과 세 번째 행은 정상조명 영상에서 검출한 정상조명 피부색과 비정상조명 피부색 영역을 보여준다. 정상조명 피부색은 올바르게 검출되었지만 비정상조명 피부색은 도로, 잔디, 벽 등에서 검출되었다. 영상 촬영 시 빛의 양이 적당하지 않거나, 빛이 반사되어 화소의 색에 많은 양의 녹색과 회색이 포함되기 때문이다. 때문에 여러 조명의 피부색을 하나의 히스토

그램에 학습하면 양성오류율이 높아져 다양한 조명의 유해영상을 올바르게 분류할 수 없다. 색 공간(Color Space)에서의 차이가 너무 크기 때문에 조명에 강건한 YIQ모델이나 Ycbr모델을 사용하여도 큰 효과를 보기 힘들고, 조명성분을 분석하여 색 값을 보정하는 화이트 패치(White Patch) 방법을 적용한 결과 배경영역들의 색에도 영향을 주어 효과적이지 않음을 확인할 수 있었다. 따라서 본 논문에서는 정상조명과 비정상조명 유해영상들을 따로 수집하여 각각의 피부색 분류기를 생성하여 실험하였다.

### 1.3 후처리

피부색 분류기로 검출한 영역에는 사람의 피부영역 뿐만 아니라 도로, 나무, 빌딩 등이 포함될 수 있다. 이들 객체들이 피부색과 유사한 색을 가지기 때문이다. 유해영상 분류 성능 저하에 큰 요인이 되는 문제로써 성능 향상을 위해 객체영역들을 제거하는 과정이 필요하다. 본 논문에서는 객체영역들을 제거하기 위해 질감(Texture)특징을 이용한다. 많은 객체들이 사람의 피부에 비해 복잡한 질감을 가지기 때문이다. 질감분석에 Wavelet Transform, Fourier Transform, Gray-level Co-occurrence Matrix 등의 기술을 이용할 수 있지만 대부분 적절한 크기의 윈도우를 설정하여 질감정도를 분석한다. 이때 문제는 윈도우가 클수록 분석 정확도가 향상되지만 제거되는 영역이 크기가 커져서 원치 않는 영역까지 제거될 수 있다는 것이다.

본 논문에서는 소벨(Sobel) 연산자를 이용하는 간단

한 방법으로 복잡한 질감 영역을 제거한다. 먼저  $((-1, 0, 1), (-2, 0, 2), (-1, 0, 1))$ 의 소벨 연산자와  $((1, 2, 1), (0, 0, 0), (-1, -2, -1))$  소벨 연산자를 입력영상의 그레이(Gray) 영상에 회전(Convolution)적용한다. 그 결과 얻은 에지(Edge)영상은 입력영상에서 복잡한 질감의 영역에서 255에 가까운 값을 나타내며, 단순한 질감 영역에서 0에 가까운 값을 나타낸다. 구해진 에지영상에서  $8 \times 8$  윈도우 단위로 밝기 값의 평균을 계산한다. 평균값을 256으로 나눠 정규화한 값이 임계치보다 크면 윈도우 영역에 포함된 화소들의 값을 255로 설정한다.

복잡한 질감영역을 제거한 영상과 피부색 검출 영상에 침식(erosion) 및 팽창(dilation)연산자를 적용한 영상을 AND 연산한 후 작은 레이블링 영역들을 제거하면 정제된 영상을 얻을 수 있다. [그림 3]에서 후처리 과정을 단계별로 보여준다. 이 같은 후처리 과정은 작은 단위로 복잡한 질감 영역을 제거할 수 있으며 연산비용 또한 낮다.

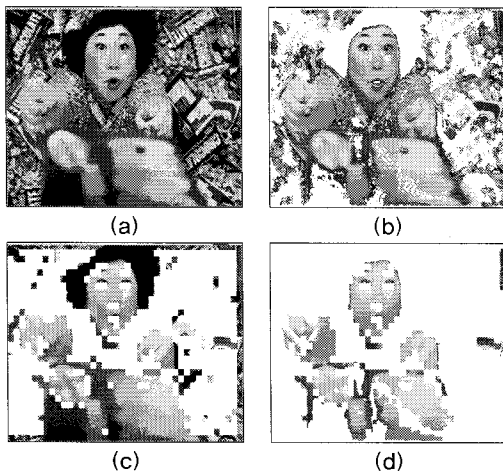


그림 3. 후처리과정. (a) 입력영상(W320×H240), (b) 피부색 검출, (c) 복잡한 질감제거, (d) 영상(b)와 영상(c)의 AND 연산

## 2. 영역특징 분류기

피부색 영역에 따라 입력영상을 무해 또는 유해로 분류하기 위해 피부색 영역의 비율, 모양 그리고 위치 특징 등을 이용할 수 있다. Lee 등은 피부색 영역의 주성분을 분석(PCA)하여 모양특징을 추출하고, 영상 중심

과 가장 큰 피부색 영역 간의 거리차이를 이용하여 위치 특징벡터를 추출하였다. 여기에 피부색 비율까지 특징으로 더한 세 가지 특징을 Adaboost[6]를 통해 학습한 유해영상 분류기를 생성하였다[7].

대부분의 유해영상들이 상당량의 피부색을 포함하기 때문에 피부색 비율은 유해영상 분류에 가장 유용한 특징이다. 모양특징은 매우 다양한 상황들을 표현하는 유해영상의 특성을 나타내기 어렵고, 더욱이 근접 촬영한 유해영상에서는 모양특징을 계산하기 어렵다. 위치특징은 정의가 모호할 수 있다. Lee는 유해영상은 영상중심 지역에서 피부색이 검출될 가능성이 높다고 보았다. 그러나 그렇지 않은 유해영상들도 많으며, 오히려 유해성이 강한 영상일수록 영상 외곽지역에서 피부색이 검출되는 경우가 많았다. [그림 4]에서 영상 외곽지역에서 피부가 검출되는 유해영상의 예들을 보여준다. 보는바와 같이 유해행위와 신체를 나타내는 영상의 경우 피부색이 영상의 상-하단, 좌-우측에서 검출되는 경우, 대각의 형태로 검출되는 경우, 근접하여 촬영한 영상에서 빛의 양이 균등하지 않아 피부색이 한쪽에서만 검출되는 경우가 많다.

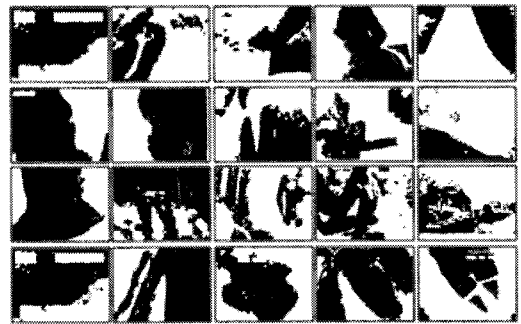


그림 4. 유해영상에서 검출되는 피부색의 위치

본 논문에서는 유해영상에서 검출되는 피부색의 비율과 위치들을 특징으로 이용하기 위해 Shih 등이 제안한 밀집도 기술자(Compact Descriptor)를 확장하여 이용한다. 밀집도 기술자는 [그림 4]의 (a)와 같이 구분되는 25개 지역의 피부색 비율로 구성된다. 피부색 비율은 식 (5) 같이 계산할 수 있다.

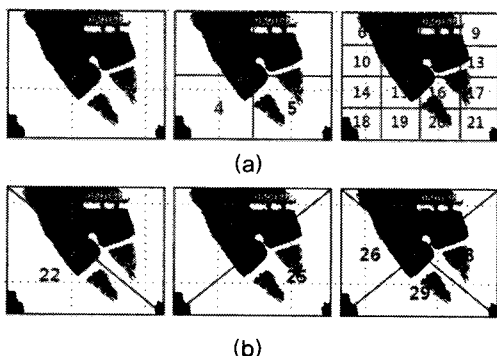


그림 5. 영역특징 벡터.  
 (a) 21개 지역 피부색 비율,  
 (b) 8개 지역 피부색 비율

$$sr_i = \frac{C_i}{W_i \times H_i} \quad (5)$$

여기서  $C_i$ 는  $i$ 번째 지역의 피부색 화소의 수를,  $W_i$ 와  $H_i$ 는  $i$ 번째 영역의 넓이와 높이를 나타낸다. Shin은 피부색 영역의 밀집정도를 나타내기 위해  $sr$ 들을 크기에 따라 정렬하였다. 본 논문에서는 피부색이 검출되는 위치와 비율을 학습하기 밀집도 기술자를 이용하기 때문에  $sr$ 들을 정렬하지 않는다. 따라서 21개 지역의 피부색 비율로 식 (6)과 같은 특징벡터를 얻을 수 있다.

$$v = (sr_1, sr_2, \dots, sr_{25}) \quad (6)$$

상-하단, 좌-우측 그리고 대각의 형태로 검출되는 피부색 영역의 특성을 고려하기 위해 [그림 5] (b)와 같이 두 개의 대각선으로 구분되는 8개 지역의 피부색 비율을 추가한다. [그림 4]의 예를 들면 21개 피부색 비율 특징벡터보다 8개 지역의 피부색 비율 특징벡터가 더 잘 부합할 수 있다. 따라서 본 논문에서는 8개 지역을 추가한 특징벡터를 이용하며 특징벡터는 식 (7)과 같이 나타낼 수 있다.

$$v = (sr_1, sr_2, \dots, sr_{29}) \quad (7)$$

이와 같이 정의한 영역 특징벡터는 검출된 피부색의 비율과 위치 정보를 포함하고 있기 때문에 분류에 효율적인 특징이 될 수 있다.

식 (7)같이 표현되는 고차원의 특징벡터는 피부 영역

에 대한 정보를 잘 나타내지만 각 차원 값의 임계치를 경험적으로 결정하는 일은 매우 어려운 일이다. 이러한 문제에 SVM이 매우 유용한 학습모델로 알려져 있다. SVM은 구조적 위험도를 최소화하는 통계적 학습 모델로써 샘플이 식 (8)과 같이 주어졌을 때 이들을 분리하는 최적 경계면들을 생성한다.

$$\{(x_1, y_1), \dots, (x_n, y_n)\} \quad i = 1..n \quad y_i \in \{-1, 1\} \quad (8)$$

여기서  $x_i$ 는 샘플을 나타내고  $y_i$ 는 범주를 나타낸다. 학습은 비선형의 최적화 문제로 귀결되며 다음 식 (9)과 같이 나타낼 수 있다.

$$f(x) = \text{sgn} \left[ \sum_{i=1}^n y_i \alpha_i K(x \cdot x_i) + b \right] \quad (9)$$

여기서  $\alpha_i$ 는 샘플의 가중치로써, 최적의 분리 경계면들을 찾는 일은 0이 아닌 모든  $\alpha_i$  즉, support vector를 찾는 것과 동일하다. 커널함수  $K(x \cdot x_i)$ 는 비선형 매핑함수의 내적계산을 함수형식으로 치환하여 계산을 간단하게 해주며 본 논문에서는 RBF를 사용하였다.

### III. 실험

성능실험은 Intel-E8200(듀얼코어/2.66GHz) 프로세서, ATI-RadeonHD3450 그래픽처리기, 2Gmb 사양의 PC에서 수행되었으며, 구현에 Visual.NET7.1이 사용되었다.

실험을 위해 인터넷 공유사이트 및 성인사이트 등에서 샘플동영상들을 수집하였다. 무해동영상은 스포츠, UCC, 뉴스, 뮤직비디오, 시사, 동물, 패션쇼, 자연 등의 주제로, 유해동영상은 국내의 제작, 개인제작, 화상채팅 등의 주제로 수집하였다. 수집된 동영상의 장면전환 프레임을 추출하여 320×240크기로 변환하여 샘플영상 DB를 구축하였다.

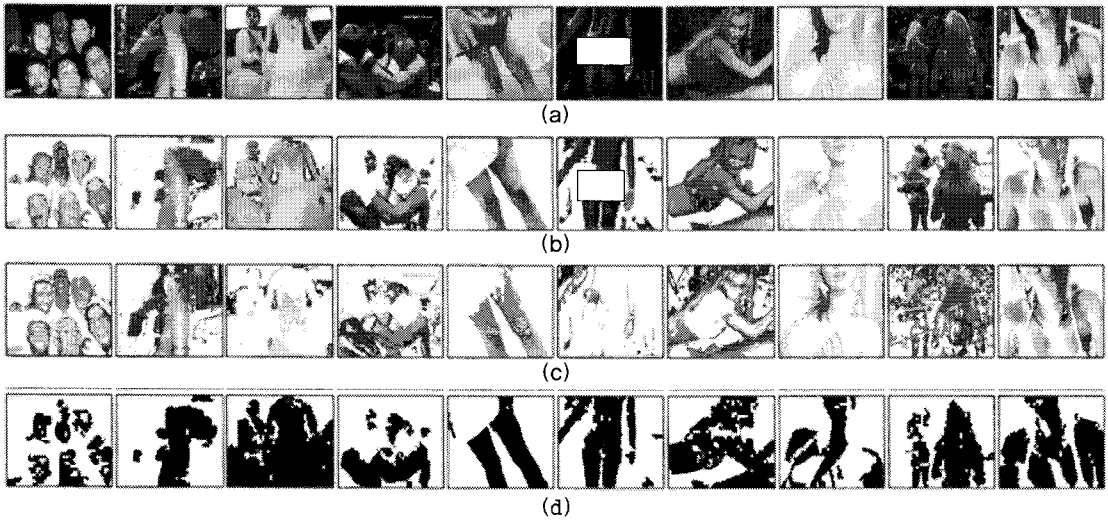


그림 6. 피부색 검출. (a) 입력영상(정상조명), (b) 정상조명의 피부색, (c) 비정상조명의 피부색, (d) 정상조명 피부색 영역 후처리

표 1. 피부색 학습화소 샘플 (천단위)

조명 \ 범주	피부색	배경색
정상	61,888개 픽셀	846,887개 픽셀
비정상	10,549개 픽셀	1,023,237개 픽셀

표 2. 영역특징 학습영상 샘플

조명 \ 범주	유해	무해
정상	1,171개 픽셀	4,904개 픽셀
비정상	735개 픽셀	

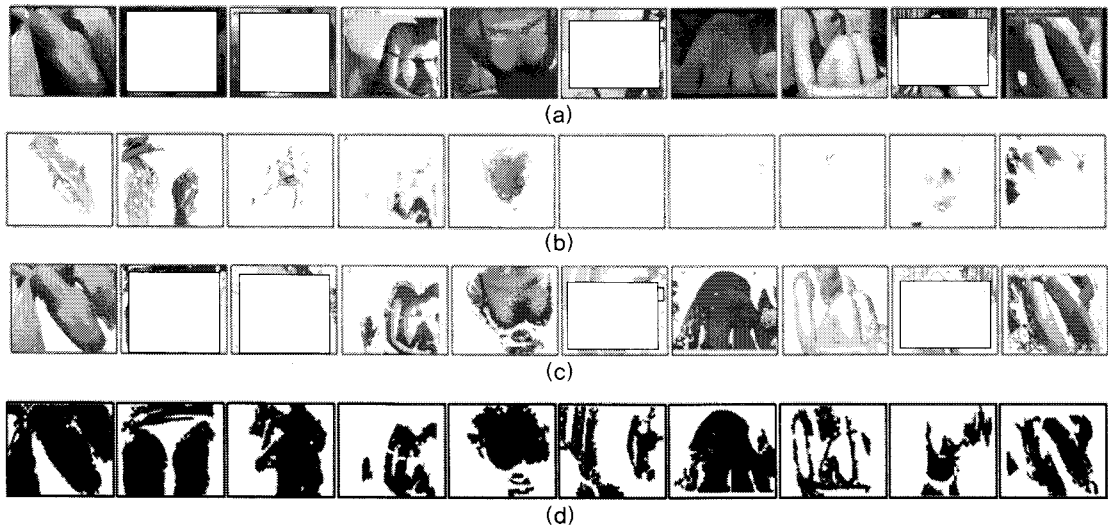


그림 7. 피부색 검출. (a) 입력영상(비정상조명), (b) 정상조명 피부색, (c) 비정상조명 피부색, (d) 비정상조명 피부색 영역 후처리

[표 1]에서 피부색 분류기 생성에 사용된 샘플의 수를 보여준다. 단, 「비정상\배경」으로 「정상\배경」 화소를 사용하였으며 양성오류율을 줄이기 위해 자연 및 거리 영상들을 추가하였다. [표 2]에서 영역특징 분류기 생성에 사용된 학습샘플의 수를 보여준다. 총 6,810개의 영상이 사용되었으며 피부색 분류기로 검출한 지역별 피부색비율을 파일로 저장하여 학습하였다. 이때 21개 특징벡터와 29개 특징벡터를 학습한 두 개의 분류기를 생성하였다. 「비정상\무해」의 경우 「정상\무해」에서 추출한 특징벡터로 대체하였다.

[그림 6]와 [그림 7]에서 피부색 검출결과를 보여준다. 각 그림의 (a)는 입력영상이며, (b)와 (c)는 정상조명 피부색 검출 결과와 비정상조명의 피부색 검출 결과를 보여주며, (d)는 검출된 피부색 영역의 후처리결과를 보여준다. 특이한 점은 비정상조명 피부색 분류기는 정상조명 영상에서도 어느 정도의 피부색을 검출한다.

학습에 사용된 비정상조명 샘플영상들이 정상조명 피부색을 포함하고 있기 때문이다.

다음으로 검출된 피부색 영역을 영역특징 분류기에 입력하여 입력영상을 무해 또는 유해로 분류하는 실험을 한다. 실험영상은 [표 3]와 같이 준비하였다. 표에서 「정상\유해」와 「비정상\유해」는 성행위나 성기가 노출된 유해영상이며, 「정상\무해」 다양한 주제의 모해영상이다. 수영복차림이나 패션쇼 등 상당량의 피부를 노출한 영상들도 포함하고 있다. 「비정상\무해」는 충분한 양의 샘플을 수집하지 못하여 실험에서 제외하였다.

실험에 대한 각 처리과정을 쉽게 확인할 수 있도록 [그림 8]과 같은 썸네일(Thumbnail) 스타일의 테스트 윈도우를 개발하였다. 윈도우는 디렉토리 뷰, 이미지 뷰, 썸네일(Thumbnail) 뷰로 구성되며, 디렉토리 뷰에서 폴더를 선택하면 폴더에 있는 모든 영상들을 썸네일

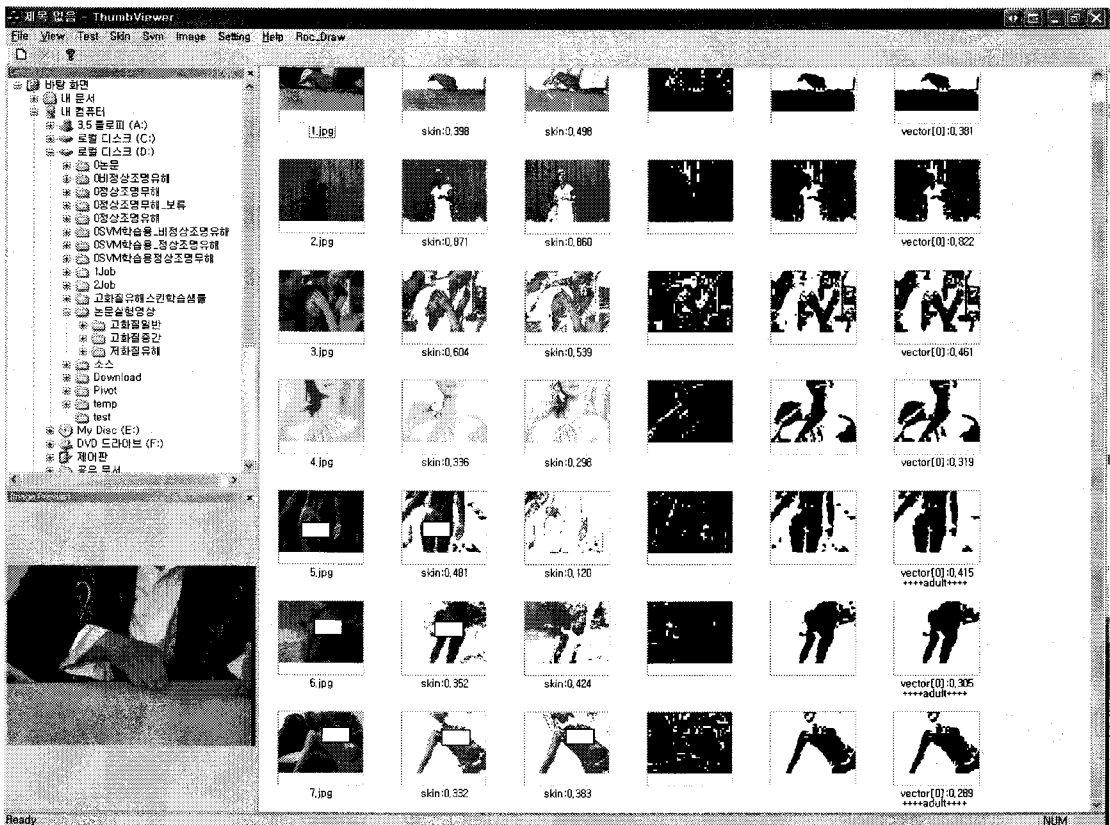


그림 8. 테스트 윈도우



표 3. 성능비교

조명\영상범주	실험영상 개수	Lee의 특징벡터[7]		21개 특징벡터		29개 특징벡터(제안)	
		유해	무해	유해	무해	유해	무해
정상\유해	2,694	2,246 (83.38%)	448 (16.62%)	2,423 (89.94%)	271 (10.06%)	2,500 (92.80%)	194 (7.20%)
정상\무해	3,134	171 (5.47%)	2,963 (94.53%)	265 (8.46%)	2,869 (91.54%)	211 (6.73%)	2,922 (93.27%)
비정상\유해	2,421	1,954 (80.71%)	467 (19.29%)	1,970 (81.37%)	451 (18.63%)	2,058 (85.02%)	363 (14.98%)

류의 첫 번째 행에 출력한다.

테스트 메뉴에서 “Adult Classification”을 선택하면 입력영상에 대한 분류의 전 처리과정과 분류결과가 각 열에 차례대로 출력된다. 자세히 두 번째 열과 세 번째 열은 정상조명의 피부색 검출 영상과, 비정상조명의 피부색 검출 영상을, 네 번째 열은 입력영상에서 복잡한 질감 영역을 제거한 영상(흰색 부분이 제거된 영역)을, 다섯 번째 열은 정상조명의 피부색 검출 영상과 입력영상의 질감 제거 영상에 대한 AND 영상을 출력하며, 작은 레이블영역까지 제거한 영상을 마지막 열에 결과영상으로서 출력한다. 결과영상에 대한 영역특징 벡터가 추출되어 영역특징 분류기에 입력된다. “vector[0]:\*” 문자가 특징벡터의 첫 번째 값 즉, 결과영상의 전체 피부색 비율 값이며, 유해로 분류된 영상의 경우 이어서 “\*\*\*Adult\*\*\*” 문자가 출력된다. 세팅에서 사용할 피부색 분류기, 영역특징 분류기, 그리고 피부색 검출 정도를 조절하는  $c_p$ 와  $c_n$ 의 값을 조절하여 그에 따른 결과를 확인할 수 있다.

1번 영상부터 4번 영상까지 무해영상으로 분류되었으며, 5번 영상부터 7번 영상까지 유해영상으로 분류되었다. 주목할 점은 유해로 분류된 영상 중 일부는 무해로 분류된 영상보다 피부색 비율이 더 낮다는 것이다. 이는 유해영상 분류가 단순히 피부색 비율뿐만 아니라 피부색이 검출된 위치에 따라 이루어진다는 것을 보여준다. 평균수행시간은 0.04초(상수시간)다.

[표 3]에서 기존의 유해영상 분류 기술 중 Lee의 분류기와 본 논문의 분류기를 비교한 결과를 보여준다. Lee는 피부색 검출에 신경망을 이용한 적응적 조명판단 기술을 이용하였지만, 동일한 피부색 검출결과로 비교 실험

하기 위해 영역특징 분류기 생성에 사용된 학습샘플에서 모양 및 위치 특징을 추출하여 Adaboost로 학습하였다. 그 결과, 표에서 보는바와 같이 83.38% 검출율과 16.62%의 양성오류율 나타내었다. Lee는 영상의 중심에서 피부색이 검출될 가능성이 높다고 보았기 때문에 영상의 외곽부분에서 피부색이 검출되는 유해영상들을 무해영상으로 분류하는 경우가 많았다. 이에 비해 본 논문에서 21개 특징벡터(21개 지역 피부색 비율)로 학습한 분류기는 보다 높은 검출율을 보였지만 8.46%의 높은 양성오류율을 나타내었다. 그러나 29개 특징벡터로 학습한 분류기는 검출율과 양성오류율 모두에서 신뢰할만한 수치를 나타내었다. 또한 『비정상\유해』의 경우 21개 특징벡터 대비 29개 특징벡터의 성능 향상 정도가 『정상\유해』에 비해 더 높았다. 이는 제안하는 분류기가 다양한 조명의 유해영상 특히, 유해성이 강한 『비정상\유해』 영상의 분류에 더욱 효과적임을 보여준다.

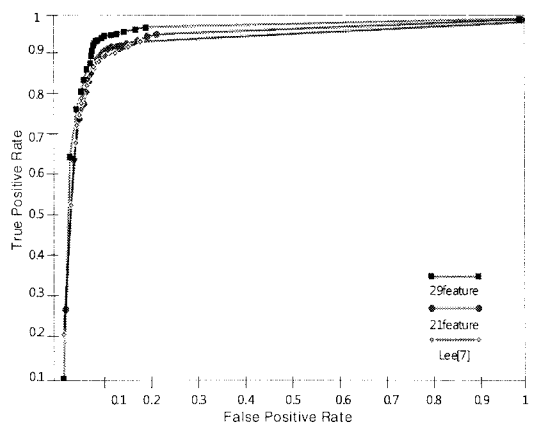


그림 9. ROC 곡선

[그림 9]는 세 가지 분류기에 대한 ROC(receiver operating characteristics)곡선을 비교하여 보여준다. 피부색 검출 단계에서 임계치  $\theta$ 를 20단계로 나눠 구했다.

#### IV. 결론

본 논문에서는 피부색을 이용하여 입력영상의 유해 여부를 판별하는 분류기를 개발하였다. 정상조명과 비정상 조명에 대해 히스토그램 기반 피부색 분류기를 학습하였으며, 피부색 영역을 특징벡터로 하는 SVM기반 영역 분류기를 학습하였다. 실험결과 개발된 분류기가 신뢰할만한 인식률을 나타내었다. 피부와 유사한 색상과 질감을 갖는 가구, 마루, 불빛 등이 양성오류의 주된 원인임을 확인하였으며, 향후 검출된 피부영역을 정제하는 방법과, 입력영상의 조명을 판별할 수 있는 방법에 대해 연구를 지속할 예정이다.

참고로 유해영상 분류에 관심 있는 연구자들에게 실험에 사용된 피부 및 영역특징 학습결과 데이터, 테스트 영상 및 테스트 윈도우 실행과일을 연구목적에 한하여 이메일(kisshajo@chol.com)로 배포할 수 있다.

#### 참고 문헌

[1] J. L. Shih, C. H. Lee, and C. S. Yang, "An adult image identification system employing image retrieval technique," Pattern recognition letters, Vol.28, No.16, pp.2367-2374, 2007.

[2] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," Pattern recognition, Vol.40, No.3, pp.1106-1122, 2007.

[3] M. J. Jones and J. M. Rehg, "Statistical Color Models with Application to Skin Detection," International journal of computer vision, Vol.46, No.1, pp.81-96, 2002.

[4] Q. Zhu, C.-T. Wu, and K.-T. Cheng, "An

Adaptive Skin Model and Its Application to Objectionable Image Filtering," ACM Multimedia, New York, USA, pp.56-63, 2004.

[5] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data mining and knowledge discovery, Vol.2, No.2, pp.121-167, 1998.

[6] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," International journal of computer vision, Vol.57, No.2, pp.137-154, 2004.

[7] J.-S. Lee, Y. M. Kuo and P. C. Chung, "Naked image detection based on adaptive and extensible skin color model," Pattern recognition, Vol.40, No.8, pp.2261-2270, 2007.

#### 저자 소개

윤진성(Jin-Sung Yoon)

정회원



- 1995년 8월 : 군산대학교 컴퓨터 과학과(이학사)
- 1997년 8월 : 송실대학교 전자계산학과(공학석사)
- 2002년 ~ 현재 : 송실대학교 컴퓨터학과 박사과정

<관심분야> : 컴퓨터비전, 3차원 모델링

김계영(Gye-Young Kim)

정회원



- 1990년 2월 : 송실대학교 전자계산학과(공학사)
- 1992년 2월 : 송실대학교 컴퓨터학과(공학석사)
- 1996년 2월 : 송실대학교 컴퓨터학과(공학박사)

• 2001년 ~ 현재 : 송실대학교 컴퓨터학과(교수)

<관심분야> : 컴퓨터비전, 형태인식, 생체인식, 증강 현실, 영상처리

최 형 일(Hyung-Il Choi)

정회원



- 1979년 2월 : 연세대학교 전자공학  
학과(공학사)
- 1982년 2월 : 미시간대학교 전산  
공학과(공학석사)
- 1987년 2월 : 미시간대학교 전산  
공학과(공학박사)

• 1987년 ~ 현재 : 숭실대학교 미디어학과(교수)

<관심분야> : 컴퓨터비전, 퍼지 및 신경망이론, 비디  
오검색, 패턴인식, 인터페이스이전트