

---

# 강건한 다인종 얼굴 검출을 위한 통합 3D 피부색 모델

## Integrated 3D Skin Color Model for Robust Skin Color Detection of Various Races

---

박경미, 김영봉  
부경대학교 정보공학과

Gyeong-Mi Park(miya11@pknu.ac.kr), Young-Bong Kim(ybkim@pknu.ac.kr)

---

### 요약

올바른 피부색 검출은 사람의 얼굴 검출 및 동작 분석에서 매우 중요한 전처리과정에 속한다. 피부 검출은 일반적으로 화소의 칼라 공간을 Non-RGB로 변형하고, 피부색의 조명 요소를 제거한 다음 피부색 분포 모델에 의해 Skin과 Non-Skin으로 분류하는 3단계로 진행된다. 이는 피부색 검출이 칼라 공간, 조명 요소의 존재 여부, 피부 모델링 방법에 따라 수행 성능에 많은 영향을 받기 때문이다. 본 연구에서는 조명 조건에 따라 피부색 모델의 범위에 차이가 있다는 사실에 기초하여 다양한 조명 조건과 복잡한 배경을 가진 영상에서 효과적으로 다인종의 피부색을 분류해내기 위한 3차원 피부색 모델을 제시하고자 한다. 제안된 피부색 모델은 화소의 칼라 공간을 YCbCr 공간으로 변형하고, 각 요소(Y, Cb, Cr)의 값에 의한 3차원 피부색 모델을 형성한다. 다인종의 피부색을 함께 분할하기 위해 인종(백인, 흑인, 황인)별 피부색 모델을 먼저 생성한 후 각각의 모델에서 피부색 확률에 따라 결합한 다인종을 위한 통합 모델을 생성하였다. 또한 우리는 적은 양의 훈련 데이터로 피부색 영역을 올바르게 검출할 수 있도록 여러 단계의 피부색 영역을 설정하였다.

■ 중심어 : | 피부색 모델 | 피부색 검출 | 피부 화소 분할 |

### Abstract

The correct detection of skin color is an important preliminary process in fields of face detection and human motion analysis. It is generally performed by three steps: transforming the pixel color to a non-RGB color space, dropping the illuminance component of skin color, and classifying the pixels by the skin color distribution model. Skin detection depends on by various factors such as color space, presence of the illumination, skin modeling method. In this paper we propose a 3d skin color model that can segment pixels with several ethnic skin color from images with various illumination condition and complicated backgrounds. This proposed skin color model are formed with each components(Y, Cb, Cr) which transform pixel color to YCbCr color space. In order to segment the skin color of several ethnic groups together, we first create the skin color model of each ethnic group, and then merge the skin color model using its skin color probability. Further, proposed model makes several steps of skin color areas that can help to classify proper skin color areas using small training data.

■ keyword : | Skin Color Modeling | Skin Color Detection | Skin Pixel Segmentation |

## 1. 서론

최근 활발하게 연구되고 있는 휴먼 컴퓨터 인터페이스 분야 연구에서 사람의 손이나 얼굴을 인식하는 것은 매우 중요하다. 사람의 얼굴이나 손 등을 인식하기 위해서는 입력영상에서 피부 영역을 세그멘테이션(Skin-region segmentation)하는 전처리 단계가 먼저 선행되어야 한다. 피부 영역 세그멘테이션은 얼굴 인식시스템, 제스처 인식시스템, 손 추적시스템, 영상 필터링 등 여러 응용분야에서 전체적인 실행 시간을 줄이기 위한 전처리과정으로 활용되고 있다.

기존의 피부 영역 검출 기술의 대부분은 칼라 영상에서 각각 화소의 색상정보를 조사하여 Skin 화소 또는 Non-Skin 화소로 분류 하는 방법을 사용하고 있다. 이 방법은 피부색이 다른 물체와 구분이 되는 일관된 색상을 가지고 있다는 이론적 배경에 근거를 두고 있다. 그러나 피부 화소의 색상 정보는 조명의 세기 및 방향에 매우 민감하다. 또한 인터넷에서 수집한 분장사진과 같이 피부색에 변형이 가해진 경우나 피부색과 유사한 색의 객체가 포함된 영상에서 피부색을 제대로 검출하지 못하는 한계를 가지고 있다. 이런 문제를 해결하기 위해 최근 Skin 화소와 Non-Skin 화소 분류 알고리즘에 대한 많은 논문이 발표 되어왔다. 이들 알고리즘에는 고정된 칼라 입계 영역을 사용하는 방법과 히스토그램 기술을 사용하는 베이지언 방법, 가우시안 분류 방법, Neural Net을 이용한 학습을 통하여 피부 화소를 분류하는 MLP(Multilayer Perceptrons) 등이 있다.

고정된 피부색 입계영역을 사용하는 방법으로 Sobottka와 Pitas[1]는 HSI 칼라공간에서 조명요소 I를 제거한 HS 평면을 사용하였고, D.Chai[2]는 YCbCr 칼라공간에서 역시 밝기 값 Y를 제거한 CbCr에 대한 고정된 입계값을 사용하였다.

Chang[3][4]과 Jones and Rehg[5] 등은 RGB 또는 YCbCr 칼라 공간에서 히스토그램 기술을 사용한 베이시언(Bayesian) 분류 방법을 제안하였다. 이 방법은 Skin 화소와 Non-Skin 화소에 대한 분류를 위해 통계적 패턴 분류에서 확립된 최소 비용을 위한 베이시언 결정 룰을 기반으로 하고 있다.

Terrillon[6]은 단일 가우시안(Unimodel Gaussian) 모델[7][8]과 혼합 가우시안(Gaussian Mixture) 모델[9][10]을 비교하였다. 가우시안 모델은 피부 화소를 분류할 때 확률분포에 근거한 확률밀도 추정 값에 따라 구분을 하고, 입계치 구간 정의가 확률 분포를 이용하여 보다 명확하게 표현될 수 있다. 가우시안 혼합 모델은 복수 개의 가우시안 확률밀도함수에 의해 피부 화소를 모델링하는 방법이다.

MLP(Multilayer Perceptrons) 방법은 뉴럴 네트워크(NN: Neural Networks) 기술을 사용한 학습을 통해 Skin과 Non-Skin 화소를 분류하는 방법이다. 이 방법은 복잡한 비선형 입/출력 관계를 학습하는 능력과 주어진 데이터를 일반화하는 능력 때문에 패턴 인식에서 주로 사용되고 있다. Karlekar와 Desai [11], Phung [12]은 피부 검출을 위해 CbCr 칼라공간에서 MLP를 사용하였고, Sahbi와 Boujema[13]은 RGB칼라 공간에서 3계층화된 NN을 사용하여 피부 분류를 수행하였다.

위에서 살펴본 피부 영역 검출 방법들에서 고정된 입계 영역을 사용하는 기법들은 피부색과 유사한 영역이 영상에 포함된 경우 피부 화소 검출이 효율적이지 못하며, 조명 요소가 제거되었기 때문에 다양한 조명조건에 강건하지 못하다. 베이지언 방법은 여러 연구자들에 의해 피부 화소 검출 효율이 높은 것으로 증명되었으나 많은 훈련 데이터와 메모리 자원을 요구하는 단점이 있다. 가우시안 분류 방법은 적은 양의 훈련 데이터를 필요로 하지만 지수와 루트 계산을 요구하며, MLP는 피부 화소 검출 효율성이 높으며, 적은 메모리 자원을 요구하지만 학습을 위한 Neural Network 기술을 필요로 하며, 피부색 검출보다는 패턴 인식에 주로 사용되고 있다.

본 논문에서는 적은 양의 훈련 데이터를 사용하여 다양한 조명 조건과 복잡한 배경을 가진 영상에서 다인종에 대해 강건한 피부 영역을 검출할 수 있는 3차원 피부색 검출 모델을 제안한다. 피부색은 조명 조건에 따라 조금씩 변하기 때문에 제안된 알고리즘에서는 조명 변화에 강건한 피부 검출을 위해 YCbCr 칼라공간에서 조명요소를 제거하지 않고 각 요소(Y, Cb, Cr) 값을 고려한 3차원 피부색 모델을 만든다. 또한 다인종의 피부

색을 고려하기 위해서 먼저 각 인종별(백인, 흑인, 황인) 모델을 생성하고, 각 인종별 모델에서 피부색 확률값을 고려하여 통합한 다인종 피부색 모델을 만든다. 제안한 방법은 적은 양의 훈련 데이터를 사용하여 3차원 모델을 만들기 용이하고, 조명 요소의 포함으로 다양한 조명의 변화를 가진 영상과 피부색과 유사한 복잡한 배경을 가진 정지 영상에서 기존의 피부색 모델보다 효과적으로 다인종의 피부색 영역을 분할시킬 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 3차원 Skin Color 모델 생성하고 3장에서는 생성된 3차원 피부색 모델을 이용하여 일반 정지 영상에서 피부색을 검출하는 과정을 다루며, 4장에서 제안된 방법과 기존의 방법 비교 실험과 실험 결과를 고찰하고, 5장에서는 결론 및 향후 과제에 대해 기술한다.

### II. 3차원 피부색 모델 생성

본 논문에서 제안하는 다인종의 피부색 검출을 위한 알고리즘은 [그림 1]에서 보이는 것과 같이 크게 두 부분으로 나누어져 있다. 첫 번째는 [그림 1]의 왼쪽에서 주어졌듯이 입력 영상으로부터 3차원 피부색 모델을 생성하는 과정이고 두 번째는 구현된 3차원 피부색 모델을 이용하여 칼라 영상으로부터 피부 영역을 분할하는 과정으로 구성된다.

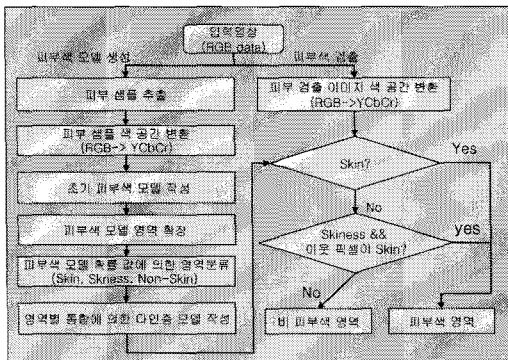


그림 1. 전체 시스템 흐름도

### 2.1 초기 3차원 피부색 모델 작성

3차원 피부색 모델 생성 과정은 영상으로부터 피부색 샘플 화소를 선택하고, 선택된 샘플 화소를 피부 검출에 유리한 칼라 공간으로 변환하고, 변환된 칼라 공간에서 샘플 화소에 의해 모델을 작성하는 3단계로 이루어진다. 첫 번째는 다양한 조명 조건을 가진 각각의 인종별 피부색이 포함된 550장의 RGB 칼라 영상에서 훈련 데이터로 사용될 피부색 샘플을 추출하는 과정이다. 피부색 샘플은 얼굴의 일부(눈, 입술, 눈썹, 머리카락 제외) 영역과 손, 팔 등을 포함한 화소를 [그림 2]와 같이 선택하였다.

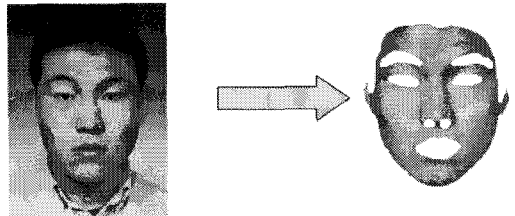


그림 2. Skin Pixel Sample 선택

두 번째 단계는 선택된 피부 샘플 영역의 칼라 공간을 YCbCr 칼라공간으로 변환하는 작업이다. YCbCr 칼라 공간은 색차 성분인 Cb와 Cr이 다른 칼라 공간에 비해 피부 영역의 범위가 좁은 영역에 밀집되어 분포되기 때문에 비피부 영역이 피부영역으로 검출되는 양을 줄일 수 있는 특징이 있다. 본 연구에서도 이 특징을 사용하여 피부색 영역을 검출하기 위한 칼라공간으로 선택하였다. 피부 샘플 데이터를 RGB 칼라 공간에서 YCbCr 칼라 공간[14]으로 변환은 식(1)에 의해 수행된다.

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.29900R + 0.58700G + 0.11400B \\ -0.16874R - 0.33126G + 0.50000B \\ 0.50000R - 0.41869G - 0.08131B \end{pmatrix}, \dots (1)$$

세 번째 단계는 YCbCr 칼라 공간으로 변환된 피부 샘플 데이터를 이용하여 3차원 피부 모델을 작성하는 과정이다. 3차원 피부색 모델을 작성하기 위해서 피부 샘플 데이터의 각 화소의 Y, Cb, Cr 값을 인덱스로 사용하여 3차원 히스토그램  $h[Y][Cb][Cr]$ 을 식(2)를 사용

하여 구한다.

$$h[Y][Cb][Cr] = \begin{cases} +1, I(Y, Cb, Cr) = SP \\ +0, I(Y, Cb, Cr) = NSP \end{cases}$$

$$\text{단 } (0 \leq Y, Cb, Cr \leq 255), \dots\dots\dots(2)$$

위 식에서  $I(Y, Cb, Cr)$ 는 피부 샘플 영상의 각 화소를 YCbCr 값으로 나타낸 것이다. SP(Skin Pixel)는 피부 샘플 영상에서 피부색으로 선택된 화소를 나타내고, NSP(Non-Skin Pixel)는 피부색에서 제외된 화소를 나타낸 것이다. 식(2)에 의해 학습된 결과에서 각각 히스토그램 누적 값이 1 이상인 것은 피부 화소가 한번 이상 나타난 것이다. 이들 누적 값들을 이용하여 초기 3차원 피부색 모델을 작성하였다.

초기 3차원 피부색 모델의 좌표는 각각 밝기 값 Y를 가로축인 x좌표로 지정하고, Cr을 세로축인 y좌표로, Cb를 깊이인 z축 좌표로 지정하였다. [그림 3]은 YCbCr 칼라 공간에서 식(2)의 학습 결과에 의해 작성된 백인, 흑인, 황인 초기 모델과 세 인종별 모델을 통합한 초기 3차원 피부색 모델을 나타낸 것이다. [그림 3]의 (a)는 가로 축 x(Y), 세로 축 y(Cr), 깊이 z(Cb)로 3차원 공간 표시하여 각 Y값에 대한 Cr값의 분포를 나타낸 것이다. [그림 3]의 (b)는 (a)를 x축을 기준으로 90도 회전시켜 Y(가로축) 값에 대한 Cb(세로축)값 분포를 나타낸 것이다. [그림 4]는 [그림 3]의 3차원 피부색 모델에서 밝기 값(Y)이 각각 90, 128, 200일 때, Cb(x축) 값과 Cr(y축) 값에 대한 분포를 백인, 흑인, 황인 모델과 다인종 통합 모델의 단면도를 나타낸 것이다.

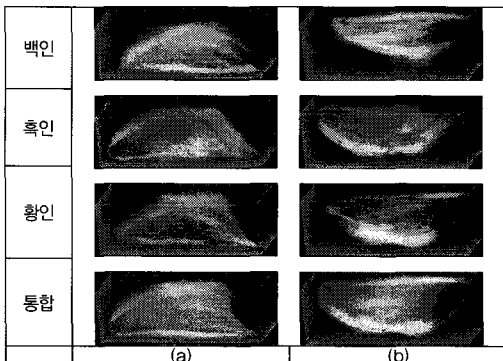


그림 3. 3D Skin Color 모델

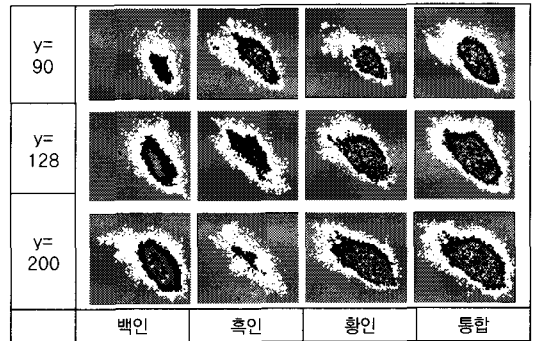


그림 4. 인종별, 밝기 값 Y에 의한 Cb, Cr 단면도

[그림 3]의 3차원 피부색 모델에서 Y값이 90, 128, 200일 때의 단면도가 [그림 4]에 나타나 있다. 단면도에서 검정색으로 표시된 부분은 히스토그램에 높은 값을 갖는 부분으로 피부색일 확률이 높은 곳, 흰색으로 표시된 부분은 낮은 확률을 갖는 곳이다. 본 연구에서는 높은 피부색 확률을 갖는 검정색 부분은 피부색(Skin) 영역, 흰색으로 표현된 중간 정도의 확률을 갖는 곳은 피부색 후보(Skinness) 영역으로, 나머지 회색으로 표시된 부분은 비 피부색(Non-Skin) 영역으로 지정하였다.

위의 모델에서 백인의 모델은 다른 모델에 비해 밝기 값 Y가 높은 곳에 피부 화소의 Cb, Cr의 범위가 넓게 분포되어 있다. 그러나 흑인 모델은 밝기 값이 낮은 곳에 피부색 분포가 다른 모델들 보다 넓은 것을 알 수 있다. 또한 황인의 경우 백인 모델과 흑인 모델에 비해 전체적으로 피부색 분포가 고르게 나타나는 것을 볼 수 있다. 이것은 조명의 밝기 변화뿐만 아니라 인종에 따라 피부색의 분포 영역에 차이가 있음을 보여준다. 이것은 백인의 모델에 의해 흑인의 피부색을 검출할 경우 피부색 검출 확률이 낮으면서 잡음이 피부색으로 포함될 확률이 높을 것이고, 그 반대의 경우도 마찬가지다.

다양한 인종을 고려한 피부색을 검출하기 위해 각 인종별로 칼라 공간을 세 가지(Skin, Skinness, Non-Skin) 영역으로 나누었다. 이렇게 생성된 피부색 모델의 샘플 데이터는 충분히 많은 영상으로부터 피부색 데이터를 추출하지 못하였기 때문에 위의 [그림 3]과 [그림 4]에서 보듯이 피부색과 인접한 영역에 홀(hole)들이 존재하고 있다. 또한 어떤 화소들은 다른 피

부색들과 일정 거리 이상 떨어진 고립된 상태로 존재하는 것도 있다. 일반적으로 피부색은 대체로 유사한 색으로 구성되므로 3차원 피부색 모델에서 인접 영역이 피부 화소로 둘러싸여 있는 홀(hole)들은 피부색일 확률이 매우 높기 때문에 피부색으로 지정하는 것이 합리적이다. 또한 다른 피부색 화소와 멀리 떨어진 고립된 화소는 피부색 샘플 데이터 추출 시 부주의로 인해 삽입된 잡음일 확률이 높으므로 제거해 줄 필요가 있다.

2.2 보간법에 의한 피부색 모델 확장

앞 절에서 작성된 초기 모델은 영상에서 나타날 수 있는 모든 피부색이 고려되지 못했으므로 훈련되지 못한 피부색이 존재할 수 있다. 또한 피부색 훈련 데이터 샘플을 선택할 때 부주의로 인한 잡음이 포함되어질 수 있다. 그러므로 이들에 대한 처리가 필요하다. 일반적으로 피부색은 일정한 범위 내에 존재 하므로 3차원 피부색 모델 내에서 인접한 화소의 정보에 의해 피부 화소 영역을 확대하거나 잡음 영역을 제거할 수 있다. 그러므로 제안된 초기 3차원 피부색 모델에서 높은 피부색 확률 값을 가진 화소들로 둘러싸인 홀은 피부색일 확률이 높으므로 보간법을 사용하여 피부색으로 포함시켜 모델의 피부색 화소를 확장시켰다. 보간법은 인접한 이웃 화소들을 값에 대한 거리의 비에 의해 홀 화소의 값이 결정되므로 확률 값에 의해 피부색 모델을 3개의 영역으로 나눌 때 각 영역이 공간적으로 구역화되는 이점을 가진다. 즉 구해질 화소 값이 공간적으로 자신의 구역의 값들에 의해 결정되기 때문이다. 제안 방법에서 보간법 수행은 피부 모델의 3차원 공간에서  $Y=z$ ,  $Cb=x$ ,  $Cr=y$ 로 지정하여 홀 화소  $f(x, y, z)$ 에 대해 다음의 [그림 5]에서 보여주는 6개의 인접 화소를 사용하여 식(3)과 같이 수행하였다.

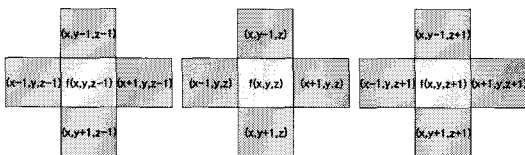


그림 5. 홀(hole) 보간법에 사용된 인접 화소

$$f(x,y,z) = \frac{HI + VI + DI}{(X_n + Y_n + Z_n)}$$

$$HI = \begin{cases} hi, & x_n = 2 \\ 0, & x_n < 2 \end{cases}, \quad VI = \begin{cases} vi, & y_n = 2 \\ 0, & y_n < 2 \end{cases}$$

$$DI = \begin{cases} di, & z_n = 2 \\ 0, & z_n < 2 \end{cases}$$

$$hi = (x-1,y,z)\alpha + (x+1,y,z)(1-\alpha)$$

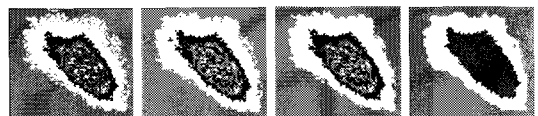
$$vi = (x,y-1,z)\beta + (x,y+1,z)(1-\beta)$$

$$di = (x,y,z-1)\gamma + (x,y,z+1)(1-\gamma)$$

$$X_n = \begin{cases} 1, & x_n = 2 \\ 0, & x_n < 2 \end{cases}, \quad Y_n = \begin{cases} 1, & y_n = 2 \\ 0, & y_n < 2 \end{cases}$$

$$Z_n = \begin{cases} 1, & z_n = 2 \\ 0, & z_n < 2 \end{cases}, \dots \dots \dots (3)$$

식(3)에서 홀 화소인  $f(x,y,z)$ 의 값은 HI, VI, DI의 평균값으로 채워진다. 위 식에서 HI, VI, DI 각각의 값은  $x_n, y_n, z_n$ 의 값에 따라  $hi(x_n=2$ 일 경우) 또는  $0(x_n<2$ 일 경우),  $vi(y_n=2$ 일 경우) 또는  $0(y_n<2$ 일 경우),  $di(z_n=2$ 일 경우) 또는  $0(z_n<2$ 일 경우)으로 지정하였다. 또한  $X_n, Y_n, Z_n$  각각의 값 역시  $x_n, y_n, z_n$ 의 값에 의해  $1(x_n=2, y_n=2, z_n=2$ 일 경우) 또는  $0(x_n<2, y_n<2, z_n<2$ 일 경우)으로 지정하였다. 위의 식에서  $\alpha$ 와  $\beta, \gamma$ 의 값은 모두 0.5를 사용하였다. 식(3)과 같이 보간법을 수행하면 피부색 모델은 자신의 영역 경계를 크게 벗어나지 않으면서 훈련되지 못한 피부색을 피부색 모델에 포함시키는 장점을 가진다. 또한 보간법에 의해 확장된 피부색 모델에서 다른 화소와 일정거리 이상 떨어진 고립된 화소 역시 제거하였다. 아래의 [그림 6]은 Y 값이 128일 때 보간법에 의해 피부모델을 확장시키는 과정을 보여주고 있다. (a)는 초기 모델을 나타내고, (b)는 1번, (c)는 3번, (d)는 10번의 피부색 모델 확장 연산을 적용한 결과이다.



(a) (b) (c) (d)  
그림 6. 보간법에 의해 확장된 피부색 모델

### 2.3 피부색 확률 값에 의한 피부색 모델의 분류

본 논문은 확장 연산을 수행한 피부색 모델을 피부색 확률 값  $p(x)$ 에 따라 세 영역으로 분류하였다. [그림 7]은 분류된 3 영역을 나타내고 있다. 각각의 영역은 피부색(Skin), 피부색 후보 영역(Skinness)과 아주 낮은 피부색 확률 값을 가지거나 확률 값이 0인 비 피부색(Non-Skin) 영역으로 나누어서 표현한 것이다. 피부색의 범위는 3D 피부색 모델에서 모든 밝기 값 Y에 대해 [그림 7]의 가로 축으로 표시된 Cb 범위는 77 - 133, 세로축으로 표시된 Cr 범위는 130 - 180까지의 사각형 영역으로 제한하였다. 이 범위는 D.Chai[2]가 식(4)와 같이 정의한 사람의 피부색의 Cb, Cr 범위를 확장한 것이다.

$$Skin\ Color \begin{cases} 77 \leq Cb \leq 127 \\ 133 \leq Cr \leq 173 \end{cases}, \dots (4)$$

본 논문에서 사용한 피부 샘플 훈련 데이터는 다양한 조명 변화와 다인종의 피부색을 가진 영상을 사용하였으므로 피부 화소의 분포가 D.Chai[14]가 정의한 피부색 영역의 Cb, Cr 범위보다 확장된 범위를 가질 수 있다. 본 논문에서 사용한 훈련 데이터의 피부색 분포 범위는 대부분(극소수를 제외한) 제안 방법에서 사용한 확장 범위 내에 밀집하여 분포하였다. 그러므로 제안한 3차원 피부색 모델의 Cb, Cr 값 역시 이 범위 내로 제한한다. Skin, Skinness, Non\_Skin 각각의 영역은 식(5)에 의해 분류 하였다.

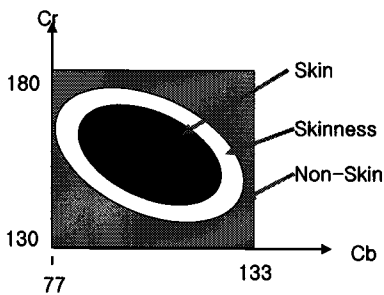


그림 7. 피부색 확률 값에 의한 분류

$$\begin{cases} Skin = P(YCbCr|Skin) \geq t_1 \\ Skinness = t_2 \leq P(YCbCr|Skin) < t_1, \dots (5) \\ Non - Skin = P(YCbCr|Skin) < t_2 \end{cases}$$

위의 식(5)에서  $P(YCbCr|Skin)$ 은 식(6)과 같다.

$$P(YCbCr|Skin) = \frac{h[Y][Cb][Cr]}{T_c}, \dots (6)$$

식(6)에서  $T_c$ 는 훈련에 사용된 모든 피부 샘플 화소들에 대한 히스토그램  $h[Y][Cb][Cr]$  값의 합을 나타낸 것이다. 식(6)에서  $h[Y][Cb][Cr]$ 의 값이 클수록 피부색일 확률이 높아지고,  $h[Y][Cb][Cr]$  값이 작을수록 피부색일 확률이 낮아진다. 그러므로  $t_1$ 과  $t_2$ 는 Skin, Skinness, Non-Skin을 결정하는 임계값으로 다음의 식(7)에 의해 구해진다.

$$t_1 = \frac{FDR}{FRR} \leq 1, \quad t_2 = \frac{FDR}{FRR} > 1, \dots (7)$$

식(7)에서 FRR은 피부 화소가 피부 화소로 검출되지 않은 미검출률(FRR : False Rejection Rate)을 나타내고, FDR은 배경이나 기타 객체가 피부색으로 검출되는 오검출률(FDR : False Detection Rate)을 나타낸다.  $t_1$ 은 FDR/FRR이 1보다 작거나 같은 것이므로 피부색 오검출률이 미검출률이 작거나 같은 것을 나타낸다.  $t_2$ 는 오검출률이 미검출률보다 큰 것을 나타낸다. 그러므로  $t_1$ 은  $t_2$ 보다 높은 피부색 확률 값을 가진다는 것을 알 수 있다. 본 논문에서는 FRR과 FDR을 구하기 위해 2.2절에서 기술한 확장된 3차원 피부 모델의 각 요소를 [표 1]에 나타낸 것과 같이 6개의 확률 값  $P(YCbCr|Skin)$ 에 대해 계산하였다.

FRR과 FDR 값은 피부색 샘플 추출을 위해 사용한 원래의 영상으로부터 피부 영역을 검출하고 그 결과 영상과 수동으로 피부 영역을 분할한 피부색 영상을 화소 대 화소의 1:1 매칭에 의해 식(8)과 식(9)와 같이 구하였다.

$$FDR = \sum_{i=1}^n \left( \frac{FD_i}{\text{count}(I_2|\text{nonskin})} \right) \times \frac{1}{n}$$

$$FRR = \sum_{i=1}^n \left( \frac{FR_i}{\text{count}(I_2|\text{skin})} \right) \times \frac{1}{n}, \dots\dots\dots(8)$$

$FD_i$ 와  $FR_i$ 는 다음과 같다

$$FD_i = 1, I_1(x, y) = \text{skin}, I_2(x, y) = \text{nonskin}$$

$$FR_i = 1, I_1(x, y) = \text{nonskin}, I_2(x, y) = \text{skin}, \dots(9)$$

위 식(8)에서  $\text{count}(I_1|\text{skin})$ 과  $\text{count}(I_2|\text{nonskin})$ 은 피부 화소 샘플링을 위해 수동으로 분할한 각 개별 영상에서 Skin 화소와 Non-Skin 화소의 개수를 나타내고,  $FD_i$ 와  $FR_i$ 는 각 개별 영상에 대한 오검출 화소의 개수와 미검출 화소의 개수를 나타낸 것이다. 또한  $I_1(x, y)$ 는 확장된 3차원 피부 모델을 사용한 피부 영역 분할 결과 영상에서 특정 화소를 나타낸 것이고,  $I_2(x, y)$ 는 피부색 샘플링을 위해 수동으로 피부색 영역을 분할한 영상에서 특정 화소를 나타낸 것이다. 다음의 [표 1]은  $t_1$ 과  $t_2$ 를 결정하기 위한  $P(YCbCr|\text{Skin})$ 의 값에 따른 FRR과 FDR의 값을 계산한 결과이다.

표 1. FRR과 FDR에 의해 결정된  $t_1$ 과  $t_2$  값

	$P(YCbCr \text{Skin})$	FRR	FDR
$t_2$	1%	2%	17%
	2%	3%	14%
	5%	5%	11%
$t_1$	10%	8%	8%
	30%	16%	5%
	50%	22%	3%

### 2.4 인종별 피부색 모델의 통합

[그림 3]과 [그림 4]에서 각 인종별 피부색 모델은 Skin, Skinness, Non-Skin 화소의 Cb, Cr 값의 분포 범위가 서로 다르다는 것을 확인했다. 이것은 조명뿐만 아니라 각 인종에 따라 피부색이 서로 다르기 때문이다. 그러므로 백인의 모델에 의해 흑인의 피부색을 검출할 경우 피부색 검출 확률이 낮으면서 잡음이 피부색으로 포함될 확률이 높을 것이고, 반대의 경우도 마찬

가지이다. 이런 문제를 해결하여 다양한 인종의 피부색을 검출하기 위해 이들 각 인종별 모델에서 공통적으로 피부색으로 분류되는 영역을 통합한 모델이 필요하다. 각 인종별 모델에 의해 통합 모델에서 Skin, Skinness, Non-Skin 영역의 분류는 인종별 모델의 세 영역(Skin, Skinness, Non-Skin)의 값을 사용하여 식(10)에 의해 구하였다.

$$f(X) = x + y + z$$

$$\begin{cases} \text{Skin} & 2 \leq f(X) \leq 3 \\ \text{Skinness} & -1 \leq f(X) \leq 1 \\ \text{Non-Skin} & -3 \leq f(X) \leq -2 \end{cases}, \dots(10)$$

식(10)에서  $x, y, z$ 는 [표 2]에 표시한 것과 같이 각각의 인종별 모델의 화소 값을 나타내는 것이고, 각 인종별 모델 화소의 값이 Skin, Skinness, Non-Skin에 따라서 +1, 0, -1로 할당하였다.

표 2. 각 인종별 모델과 세 영역 값 설정

각 인종별 모델	각 인종별 모델
백인 = x	Skin = 1
흑인 = y	Skinness = 0
황인 = z	Non-Skin = -1

[그림 8]은 [그림 4]에서 밝기 값  $Y=128$ 일 때, 각 인종별 모델에서 피부색 영역과 피부색 후보 영역을 사용하여 통합 모델의 피부색 영역과 피부색 후보 영역을 선택하는 방법을 보여주고 있다. 그림에서 (a)는 각 인종별 모델의 피부색(Skin) 영역을 이용하여 통합 모델의 피부색 영역을 선택한 것을, (b)는 피부색 후보(Skinness) 영역 선택을 보여주고 있다. (a), (b) 모두 빗금 친 영역을 각각 피부색(Skin)과 피부색 후보(Skinness) 영역으로 선택하였다. (c)는 제안된 방법에 의해 통합된 실제 모델이다. 아래의 그림 (a)에서 피부색 영역 선택 과정을 살펴보면 두 개 이상의 모델에서 피부색으로 결정된 영역이 통합 모델에서 피부색으로 선택하였으며, 하나의 모델에서만 피부색으로 결정된 영역은 통합 모델에서 피부색 후보 영역으로 선택하였다. 피부색 후보 영역 역시 두 개 이상의 모델에서 피부색 후보 영역으로 결정된 영역이 피부색 후보로 선택하

었고, 1개의 모델에서 피부색, 나머지 두 모델에서 비 피부색으로 결정된 영역은 통합 모델에서는 비 피부색으로 선택하였다.

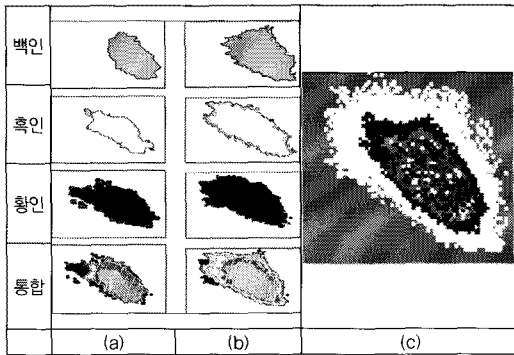


그림 8. 인종별 피부색 모델 통합 방법

### III. 다인종 피부색 검출

다양한 조명조건의 변화를 가진 다인종의 피부색이 포함된 RGB 컬러 영상을 본 논문에서 제안한 3차원 피부색 모델을 이용하여 피부 영역을 검출하는 단계다. 제안된 3차원 피부색 모델은 다양한 조명 조건과 다인종을 고려하기 위해 조명 요소를 포함한 Y, Cb, Cr 3차원 공간에서 피부 화소의 확률 값에 의해 3영역(Skin, Skinness, Non-Skin)으로 구분하여 작성하였다. 영상에서 피부 영역 검출은 영상의 화소가 피부색 모델의 Non-Skin과 일치하면 해당 화소를 제거하고, Skin과 일치하면 피부색으로 선택하였다. 그리고 영상의 화소 값이 피부색 모델의 Skinness와 일치하면 영상의 인접 화소의 값에 따라 Skin 또는 Non-Skin으로 지정하였다. 이것은 영상에서 대체로 피부색을 이루는 화소들은 서로 인접한 영역에 존재한다는 특성을 이용하여 피부색과 유사한 배경이 포함된 영상에서 효율적으로 피부색을 분할할 수 있는 장점을 가진다.

[그림 9]는 제안된 피부색 모델에서 입력 영상(a)를 이용하여 피부색을 검출하는 과정을 보여준다. (b)는 피부색 모델에서 Skin( $t_1=50\%$ ) 영역만으로 검출한 1차 결과이며, (c)는 Skinness( $t_2=5\%$ ) 영역으로 확대하여

검색한 결과이다. 아래의 그림에서 Skin 영역만으로 피부색을 검출한 결과(b)는 미검출률(FRR)이 높지만 Skinness 영역으로 확대하여 검출한 (c)는 영상의 인접 영역 화소를 고려하여 피부색 영역을 검출하였으므로 피부 영역 확대에 의해 발생하는 잡음 영역 추가되지 않고 없이 검출되는 것을 볼 수 있다.



그림 9. 피부 영역 검색 결과

[그림 10]은 각 인종별 모델과 통합 모델에 의해 검출된 결과이다. (a)는 입력 영상, (b)는 백인 모델에 의한 다인종 피부색 검출 결과, (c)는 황인 모델에 의한 결과 (d)는 흑인 모델에 의한 결과 (e)는 통합모델에 의한 결과이다. 아래의 [그림 10]에서 (b)의 경우 백인의 피부색 모델을 사용하였기 때문에 백인의 피부색이 포함된 영상에서 피부 영역 검출률이 높은 것을 볼 수 있으며, 반면 흑인의 피부를 포함한 영상에서 매우 낮은 피부 화소 검출률을 보이고 있다. 즉 각각의 인종별 3차원 피부색 모델을 사용한 경우 모델과 영상의 인종이 같은 경우 낮은 미검출률(FRR)과 낮은 잡음률(FDR)을 보여주고 있으며, 모델과 영상의 인종이 서로 다른 경우에는 미검출률(FRR)과 오검출률(FDR)이 모두 높은 것을 볼 수 있다. 그러나 다인종 통합 피부색 모델을 이용한 경우 미검출률(FRR)이 모델과 영상의 인종이 같은 경우와 비슷하게 나타나고 오검출률(FDR) 역시 비슷한 결과를 나타내고 있는 것을 볼 수 있다. 그러므로 제안된 통합 모델은 다양한 조명과 다인종이 포함된 영상에서 피부 영역 분할을 위한 좋은 모델임을 확인 할 수 있다.



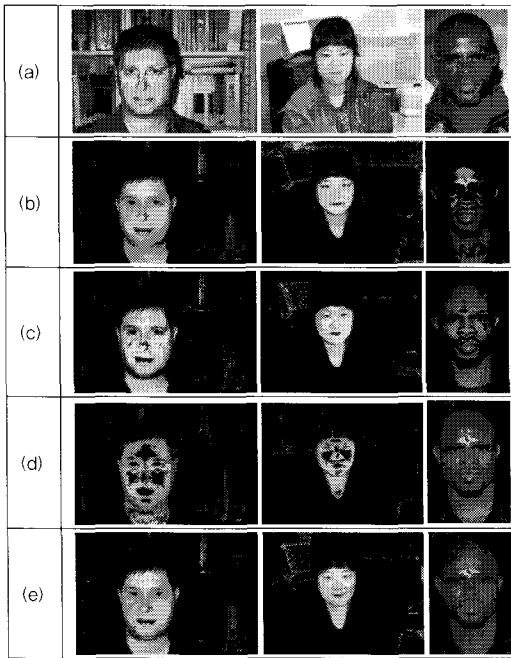


그림 10. 인종별 모델과 통합 모델에 의한 검출 결과



(a) 입력 영상



(b) 검출 결과

그림 11. 다인종 포함 영상에서 피부색 검출

[그림 11]은 다인종(흑인, 백인, 황인)의 피부색이 포함된 입력 영상으로부터 제안된 3차원 다인종 통합 피부색 모델을 이용하여 피부 화소를 분할한 결과를 보여주고 있다. 전반적으로 모든 인종의 얼굴과 팔 부분을 잘 검출해 내었다. 그러나 옷이나 나무의 일부분이 피부색으로 검출되는 오검출도 적은 부분에서 보인다.

#### IV. 실험 및 결과

본 논문에서 제안된 방법의 성능을 평가하고 측정하기 제안된 3차원 피부색 모델에 의해 피부 영역 검출 기법을 기존의 유사한 방법인 베이지언 방법과 비교하였다. 성능 평가 실험을 위해 3차원 피부색 모델 생성과 피부 영역 분류를 위해 사용한 데이터는 [표 3]의 내용과 같다.

표 3. 훈련 및 테스트 데이터

방법	훈련영상		테스트영상
	Skin Pixel	NonSkin Pixel	
베이지언	32,442,338	41,712,898	1,000장
제안방법	26,122,231	-	

제안된 알고리즘 실험을 위해 디지털 카메라로 촬영된 최소 크기 240\*320, 최대 크기 640\*768인 정면 얼굴 및 손 등 피부색이 포함된 칼라 영상 550장으로부터 26,122,231 화소(백인 : 9,433,823, 아시안 : 9,709,349, 흑인 : 6,320,107)을 추출하여 3차원 모델 작성하기 위한 학습 데이터로 사용하였다. 피부색 분류를 위한 성능 평가 및 테스트를 위해서 1,000장(디지털 카메라로 촬영된 영상 600장과 인터넷으로부터 획득한 400장)의 영상을 사용하였다. 각각의 영상들은 백인 300장, 아시안 400장, 흑인 180장, 기타 다인종이 함께 포함한 영상 120장으로 구성되어 있다. 본 논문에서 제안된 방법과 성능을 비교할 베이지언 분류 방법은 Skin 화소 훈련을 위해 32,442,338 화소를 사용하고, Non-Skin 화소 훈련을 위해 41,712,898 화소를 사용하였다. 베이지언 방법에서 Skin 화소 분류를 위한 확률밀도함수는 다음의 식 (11)을 사용한다.

$$\frac{p(x|skin)}{p(x|nonskin)} \geq T, \dots \dots \dots (11)$$

위 식(11)에서  $p(x|skin)$ 은 Skin 화소일 확률을 나타낸 것이고,  $p(x|nonskin)$ 는 Non-Skin 화소일 확률을 나타낸 것이다.

본 연구에서 피부색 검출 성능을 평가하기 위해 실제 피부색 검출률(True skin Detection Rate : TDR)과 피부색이 아닌 잡음이 피부색으로 잘못 검출되는 오검출률(False skin Detection Rate : FDR)을 평가하기 위해 피부화소 분류를 위한 테스트 영상을 수동으로 피부색 영역을 분할하여 제안방법에 의한 피부색 검출 결과와 비교하였다.

[그림 12]는 성능 평가를 위한 제안 방법과 베이지언 방법에 의한 일반 정지 영상에서 피부 화소를 검출한 결과를 나타낸 것이다. 그림에서 (a)는 피부 검출 성능 테스트를 위한 입력 영상을 나타내고, (b)는 본 논문에서 제안한 방법에 의한 피부 화소를 검출한 결과이며, (c)는 베이지언 방법에 의해 피부색을 검출한 결과를 나타낸 것이다.

영상으로부터 피부색 영역 분할은 피부색과 유사한 색의 배경이나 개체가 포함되어 있으면 오검출률(FDR)이 높아지고, 조명에 의해 피부색 일부가 왜곡된 경우에는 피부색 검출률(TDR)이 낮아져서 효과적이지 못하다.

실험 결과에서 피부색과 유사한 색을 배경에 포함하고 있거나 조명 조건의 변화에 의해 피부색 영역의 일부가 왜곡된 영상에서 제안 방법이 베이지언 방법에 비해 더 좋은 결과를 얻을 수 있었다. 즉 제안된 알고리즘은 베이지언 방법을 사용하여 얻어진 TDR과 유사한 값을 가질 때 더 낮은 오검출률(FDR)을 가지고 있는 것을 볼 수 있다. 이것은 우리가 제안한 3차원 피부색 모델이 Skin 영역과 Skinness 영역, Non-Skin 영역으로 3단계로 구성되어 있기 때문에 이들 영역을 이용하여 피부 화소를 동적으로 선택하여 검출의 효율을 높였기 때문이다.



(a) 입력영상 (b) 제안방법 (c) 베이지언방법

그림 12. 피부색 검출 결과 비교

먼저 영상의 화소가 모델에서 Skin 영역에 해당하면 피부색으로 선택하고, 영상의 화소 값이 모델의 Skinness 영역에 해당하면 영상의 이웃 화소의 정보를 사용하여 피부색으로 선택(인접한 화소가 Skin인 경우) 하거나 피부색에서 제거(인접한 이웃 화소가 Non\_Skin인 경우)하기 때문에, 피부색과 유사한 객체가 피부색과 인접하지 않고 일정거리 이상 떨어진 경우 피부색으로 선택되지 않아 기존의 방법보다 오검출률(FDR)이 낮아졌다. 또한 영상에서 피부색 영역의 일부가 왜곡된 경우 인접한 영역의 왜곡되지 않은 화소의 정보에 의해

피부색으로 선택되어 피부색 검출률이 높게 나타났다. 그러나 [그림 12]의 마지막 행의 그림과 같이 피부색과 유사한 색이 영상 전체에 고르게 분포되어 있는 경우, 제안 방법을 사용하여 검출 결과가 피부 화소 검출률(TDR)에 비례하여 오류 검출률(FDR)이 함께 높아져서 효율적이지 않다. 그러나 제안 방법은 대부분의 영상에서 베이지언 방법보다 더 좋은 결과를 얻을 수 있었다.

제안된 피부색 모델에 의해 영상에서 피부 화소 영역을 분할한 결과로 피부색 검출 성능 평가를 수행하였다. 성능 평가는 제안된 방법에 의한 생성된 모델을 사용하여 검출된 결과와 베이지언 방법에 의해 검출된 결과를 피부 영역을 수동으로 분할한 영상과 비교하였다. [그림 13]은 성능평가 결과를 베이지언 방법과 제안 방법의 FDR 기준으로 TDR 값을 나타낸 ROC 곡선을 그린 것이다. 아래의 그림에서도 제안된 방법이 베이지언 방법보다 오류 검출률(FDR)이 같은 경우 피부 화소 검출률(TDR)이 더 우수한 것을 알 수 있다.

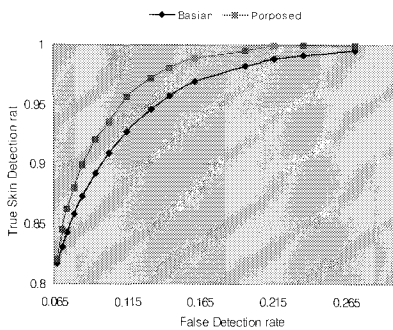


그림 13. 제안방법과 베이지언의 ROC 곡선

## V. 결론

본 논문에서 칼라 영상에서 다양한 조명 조건의 변화를 가진 복잡한 배경을 포함한 영상에서 다인종을 고려한 강건한 피부색 검출을 위한 3차원 피부색 모델 생성 방법을 제안하였다. 제안된 3차원 피부색 모델은 샘플 데이터의 학습에 의해 피부색 화소의 Y, Cb, Cr 값으로 3차원 모델을 작성하였다. 작성된 3차원 피부색 모델은

화소 밝기 값 Y에 대한 정보를 가지므로 조명의 변화에 강건하며, 또한 다양한 인종을 고려한 통합모델을 사용하여 다인종의 피부색이 포함된 영상에서도 강건하게 피부색을 검출하였다. 제안된 3차원 피부색 모델은 훈련 화소의 피부색 출현 빈도에 따라 누적된 값에 의해 모델을 Skin, Skinness, Non\_Skin 3영역으로 구분하였다. 이들 3개의 영역은 영상에서 피부색 검출할 때 Skin은 피부색으로, Skinness 영역은 영상의 이웃 화소 정보에 의해 Skin 또는 Non-Skin으로 동적으로 선택하는데 사용되었다. 그러므로 제안 방법은 조명이나 카메라 세팅에 의해 피부색이 왜곡되어 검출되는 못하는 미검출률을 낮추고, 또한 피부색과 유사한 배경이 피부색으로 검출되는 오검출률을 낮추어 피부 검출의 성능을 높여 주었다. 또한 3차원 피부 모델을 생성할 때 초기 훈련모델을 인접 화소 값들로 보간하여 편리하게 모델을 확장할 수 있는 장점을 가지고 있다.

## 참 고 문 헌

- [1] K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation," *Facial Feature Extraction and Tracking*, Signal Processing: ImageComm., Vol.12, No.3, pp.263-281, 1998.
- [2] D. Chai and A. Bouzerdoum, "A Bayesian approach to skin color classification in YCbCr color space," *IEEE TENCON00*, Vol.2, pp.421-424, 2000.
- [3] H. Wang and S. F. Chang, "A Highly Efficient System for Automatic Face Detection in Mpeg Video," *IEEE Trans. Circuits and Systems for VideoTechnology*, Vol.7, No.4, pp.615-628, 1997.
- [4] M. C. Shin, K. I. Chang, and L. V. Tsap, "Does Colorspace Transformation Make Any Difference on Skin Detection?," *Proc. IEEE Workshop Applications of Computer Vision*, pp.275-279, 2002.
- [5] M. J. Jones and J. M. Rehg, "Statistical Color

- Models with Application to Skin Detection," Int'l J. Computer Vision, Vol.46, No.1, pp.81-96, 2002.
- [6] J.-C. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images," Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition, pp.54-61, 2000.
- [7] J. Yang and A. Waibel, *A Real-Time Face Tracker*, Proc. IEEE Workshop Applications of Computer Vision, pp.142-147, 1996(12).
- [8] B. Menser and M. Wien, "Segmentation and Tracking of Facial Regions in Color Image Sequences," SPIE Visual Comm. and Image Processing 2000, Vol.4067, pp.731-740, 2000(6).
- [9] H. Greenspan, J. Goldberger, and I. Eshet, "Mixture Model for Face Color Modeling and Segmentation," Pattern Recognition Letters, Vol.22, pp.1525-1536, 2001(9).
- [10] M.-H. Yang and N. Ahuja, "Gaussian Mixture Model for Human Skin Color and Its Applications in Image and Video Databases," SPIE Storage and Retrieval for Image and Video Databases, Vol.3656, pp.45-466, 1999(1).
- [11] J. Karlekar, U. B. Desai, "Finding faces in color images using wavelet transform," International Conference on Image Analysis and Processing, pp.108/5, 1999.
- [12] S. L. Phung, D. Chai, and A. Bouzerdoum, "A universal and robust humanskin color model using neural networks," IJCNN01, 2001.
- [13] H. Sahbi and N. Boujemaa, "Coarse to fine face detection based on skin color adaptation," Workshop on Biometric Authentication, 2002, Lecture Notes in Computer Science, Vol.2359, pp.112-120, 2002.
- [14] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.27, No.1, 2005(1).
- [15] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," P. Kakumanu et al. / Pattern Recognition 40 1106 . 1122, 2007.
- [16] G. M. Park and Y. B. Kim, "3D Skin Color Model for Asian Skin Color Segmentation," MITA 2007 Track 4A-04, 2007.
- [17] S. Haykin, *Neural Networks: A Comprehensive Foundation, 2nd ed.*, Uppersaddle, N.J: Prentice-Hall, 1999.

저 자 소 개

박 경 미(Gyeong-Mi Park)

정회원



- 2000년 2월 : 한국방송통신대학교 컴퓨터학과(공학사)
- 2003년 8월 : 부경대학교 정보전산학과(공학석사)
- 현재 : 부경대학교 정보공학과 (박사과정)

<관심분야> : 컴퓨터그래픽스, 3D 애니메이션, 영상처리

김 영 봉(Young-Bong Kim)

정회원



- 1987년 : 서울대학교 계산통계학과(이학사)
- 1989년 ~ 1994년 : 한국과학기술원 전산학과(공학석사/공학박사)
- 1995년 ~ 현재 : 부경대학교 전자컴퓨터정보통신공학부 정교수

<관심분야> : 컴퓨터 그래픽스, 3D 컴퓨터 시뮬레이션