

화자(話者)와 스마트폰의 거리 측정을 위한 다중 색 좌표계와 다중 임계치 기반 실시간 얼굴검출

이재원^{*}, 권구락^{**}, 홍성훈^{***}

요 약

모바일기기가 발달함에 따라 핸드폰에 수많은 기능들이 탑재되고 있다. 영상 통화 기능도 그 중 하나이다. 본 논문에서는 화자와 스마트폰 사이의 거리를 측정하기 위한 다중 색좌표계와 다중 임계치를 사용하는 방법을 제시한다. 첫 번째로 피부색의 색상정보에 근거하여 얼굴영역을 검출한다. 두 번째로 검출된 얼굴영역의 크기를 이용하여 스마트폰과 화자 사이의 거리를 측정한다. 특히 본 논문에서 제시하는 얼굴영역 검출 알고리즘 개발에 있어 고려한 점은 스마트폰의 기본기능과 함께 실시간으로 처리가 가능할 정도로 연산량이 적어야 하며, 움직임이 많은 핸드폰의 특성상 프레임과 프레임 사이의 움직임과 조명 및 배경에 따라 검출된 얼굴 영역이 급격히 변화하는 문제를 해결할 수 있는 얼굴영역 검출 알고리즘을 개발하고자 하였다.

Real-Time Face-Detection Based on Multiple Color-Spaces and Multiple Thresholds for Distance Measurement Between Speaker and Smart-Phone

Jae-Won Lee^{*}, Goo-Rak Kwon^{**}, Sung-Hoon Hong^{***}

ABSTRACT

As the development of mobile devices, mobile phones are equipped with many features. Video-call feature is one of them. In this paper, we present distance measurement between speaker and smart-phone using multiple color spaces and multiple thresholds. first, detect face based on skin color information. and second, measure distance between speaker and smart-phone using the detected face region. Especially, the first considering point in the development of face area detection is real-time processing and the second point is robustness to solve the problems of face detection errors due to rapid change of object movement, lighting and background between adjacent frames.

Key words: Face-Detection(얼굴검출), Multiple Thresholding(다중 임계치), Mobile Phone(모바일), Distance Measurement(거리측정)

※ 교신저자(Corresponding Author): 홍성훈, 주소: 광주광역시 북구 용봉동 300번지 전남대학교 전자컴퓨터 공학부(500-757), 전화: 062)530-0820, FAX: 062)530-1759, E-mail: hsh@chonnam.ac.kr

접수일: 2010년 8월 25일, 수정일: 2011년 1월 25일

완료일: 2011년 2월 22일

^{*} 준회원, 전남대학교 전자컴퓨터공학과

(E-mail: 777kamja@naver.com)

^{**} 종신회원, 조선대학교 정보통신공학과

(E-mail: grkwon@chosun.ac.kr)

^{***} 정회원, 전남대학교 전자컴퓨터공학과, 정보통신연구소

※ 본 과제(결과물)는 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 광역경제권 선도산업 인재양성사업의 연구결과입니다.

1. 서 론

모바일 기기의 성능이 향상됨에 따라 핸드폰을 통해 영상통화가 가능하게 되었고, 음성인식 기능을 이용한 다양한 응용서비스가 제공되고 있다. 모바일 기기는 다양한 환경에서 사용되기 때문에 주위 잡음에 의해 통화 품질과 음성 인식률이 저하될 수 있다. 또한 화자와 모바일 기기의 거리에 따라서는 통화 품질과 음성 인식률이 좌우되기도 하는데, 화자와 모바일 기기의 거리가 너무 가깝거나 먼 경우 서비스 품질 저하가 발생한다. 따라서 모바일 기기와 화자 사이의 거리를 쉽게 인식할 수 있는 방법이 요구된다. 특히 모바일 기기의 연산 속도와 메모리 등의 제한과 실시간성을 고려하면 저 연산, 저 메모리를 요구하는 거리측정 방법의 개발이 필요하다. 본 논문에서는 영상통화에서의 통화품질을 보장하기 위한 최적의 거리를 화자가 인지할 수 있도록 돕는 얼굴영역 검출 기반 거리측정 방법을 제시한다.

기존의 얼굴검출 기술들은 지식 기반 방법, 특징 기반 방법, 형판 정합 방법, 외형 기반 방법 등의 분류로 나누어 연구되어지고 있다[1,2]. 지식 기반 방법은 사람의 얼굴이 눈썹, 눈, 코, 입 등을 포함하고 있고, 각각의 얼굴 성분들은 서로 일정한 거리와 위치 관계를 가지고 있다는 지식을 기반으로 얼굴을 검출하는 방법이고, 특징 기반 방법은 얼굴 특징 성분들의 크기와 모양, 상호 연관성, 얼굴의 색상과 질감 정보, 혹은 이런 특징들의 혼합 정보를 이용하여 얼굴 검출하는 방법을 말한다. 형판 정합 방법은 대상이 되는 모든 얼굴에 대한 표준 템플릿을 만든 후, 입력 영상과 비교하여 얼굴을 검출하는 방법이고, 외형 기반 방법은 학습 영상 집합에 의해 학습된 모델을 이용해서 얼굴을 검출하는 방법으로 PCA(Principal Component Analysis : 주성분 분석)에 의해 생성되는 고유 얼굴(eigen-face), LDA(Linear Discriminant Analysis : 선형판별식 해석), 신경망(Neural Network), SVM(Support Vector Machine)을 이용하는 방법이다. 학습이나 템플릿 매칭 등의 방법은 모바일에서의 실시간 연산에는 연산속도 문제로 그다지 적합하지 않다[3,4]. 따라서 저 연산으로 실시간 얼굴영역 검출을 수행하는 방법으로 특징기반 방법 중에서도 색 좌표계에 임계치를 적용하여 피부색영역을 검출하는 방식이 가장 많이 사용된다. 색 좌표계를 이용한

기존의 얼굴 검출에 대한 연구는 여러 색 좌표계에서 적용되어져 왔다. Terrillon 등[5]은 RGB 색 좌표계를 STV 칼색 좌표계로 변환하여 S와 T에 기반한 피부색에 대한 히스토그램 작성하여 얼굴영역 검출하였다. Dai 등[6]은 기존에 그레이 영상에 적용했던 SGLD matrix에 근거한 얼굴-텍스처 모델을 칼라에 적용하기 위해 YIQ 색 좌표계에서 I범위에 따른 얼굴영역을 추출하였다. Yang 등[7]은 사람의 피부색이 거의 같다는 점에 착안하여 RG값을 이용하여 얼굴을 검출하였다. Chai 등[8], Kwon 등[9]은 YCbCr 색좌표계에서 식(1)과 같이 Cb, Cr의 색상 값에 대한 단일 임계치를 적용하여 얼굴을 검출하였다.

$$Skin\ Region \begin{cases} 77 < Cb < 127 \\ 133 < Cr < 173 \end{cases} \quad (1)$$

위와 같이 일반적으로 색 좌표계에 임계치를 적용하여 피부색영역을 검출하는 기존 방식들은 단일 색 좌표계에 대해 단일 임계치 내에 포함되는 영역을 피부색으로 판단하는 방법을 사용한다. 그러나 얼굴의 색상이 사람마다 다르고, 조명환경에 따라 카메라를 통해 얻어진 살색영역의 색상이 변한다. 따라서 단일 색 좌표계에 대한 임계치만을 적용하는 기존 방식들은 조명 등 환경변화에 민감하여 얼굴영역 검출 결과의 신뢰성이 떨어진다. 또한 HSV 색 좌표계를 사용하는 방식[10,11]도 제안되었으나, 스마트폰 상에서 실시간 연산을 하기에는 색좌표계 변환 등을 위한 연산량이 많아 본 논문에서는 연산량이 적으면서 살색검출의 정확도가 높은 YCbCr과 YIQ 색 좌표계를 동시에 사용하고 임계치 또한 다단계를 사용하는 방법을 제안한다. 그리고 검출된 얼굴영역을 이용하여 통화품질이 보장되는 적정거리를 위하여 화자와 스마트폰 간의 거리를 화자에게 알려주는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서 제안하는 얼굴 검출 및 스마트폰과 화자간의 거리측정 방법에 대하여 설명하고, 3장에서는 실험 결과에 대한 분석을 한 후, 4장에서 결론을 맺는다.

2. 제안된 얼굴검출 및 거리측정 방법

2.1 얼굴 검출 알고리즘의 전체 구성

그림 1은 본 논문에서 제안하는 얼굴영역 검출 알

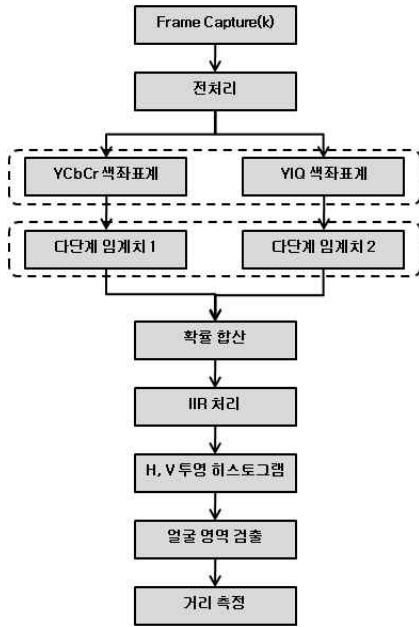


그림 1. 제안된 얼굴영역 검출 알고리즘의 전체 구성도

고리즘의 전체 구성도를 보여준다. 카메라로부터 얻어진 k번째 영상프레임 F(k)는 조명의 영향 등을 줄이는 전처리과정을 거친 후 색 좌표계 변환(Matrix)에 의해 살색검출을 위한 다수의 색 좌표계에 대응하는 색신호성분으로 변환된다. 색 좌표계 변환부의 출력인 색신호들은 다단계 임계치 처리과정을 거쳐 살색확률을 표현하는 확률 마스크를 출력한다. 본 논문에서는 2단계 임계치 처리과정을 수행하였다. 각 색 좌표계에 대한 확률 마스크는 확률 합산부에서 모두 더하여 누적된 확률 마스크가 구해지고, 이 누적확률 마스크는 시간 축 IIR(Infinite Impulse Response) 필터 처리되어 이전 프레임과의 상관성이 유지되도록 함으로써 선택된 얼굴영역이 프레임 단위로 급격히 변하지 않도록 하여 얼굴영역 검출의 신뢰도를 높인다. 이와 같이 구한 확률 마스크의 살색 확률을 수평 방향과 수직방향으로 투영하여 수평·수직 투영 히스토그램을 구하고 히스토그램 값이 일정 임계치보다 큰 영역을 얼굴영역으로 결정한다. 이와 같이 결정된 얼굴영역의 수평 크기를 기준으로 화자와 스마트폰 사이의 거리를 추정하게 된다.

2.2 전처리 과정

피부색을 이용한 얼굴검출에 있어서, 조명변화는

얼굴영역 검출의 성능에 크게 영향을 미친다. 특히 너무 어두운 경우 살색에 대응하는 색상정보가 불명확하게 된다. 본 논문에서는 조명변화에 강인하고 매우 어두운 조명 환경에서도 얼굴영역을 검출에 도움이 되는 색상 진폭 보정과 히스토그램 정규화를 이용한 역광보정의 선택적 전처리 방식을 수행한다. 색상 진폭 보정은 어두운 영역에서의 검출을, 히스토그램 정규화 보정은 역광에서의 검출을 위한 전처리 과정으로 입력된 영상 전체의 평균 밝기 값을 이용하여 전체적으로 어두운 영상과 역광이 나타난 영상을 구분하여 전처리 과정을 선택한다.

2.2.1 색상 진폭 보정

어두운 조명환경에서 스마트폰을 통해 획득한 영상의 색상을 정규화하기 위하여 휘도성분(Y) 중 밝은 화소 값들(reference white pixels)을 이용하여 R, G, B의 밝기 값을 조정하는 방식을 사용한다[12]. 스마트폰에서 획득된 R, G, B 신호로부터 휘도성분은 식 (2)에 따라 얻어진다[13]. 식 (2)는 연산을 줄이기 위해 소수점 연산을 곱하기와 bit-shift 연산만으로 변환한 것이다.

$$Y = ((76 * R + 150 * G + 29 * B + 128) \gg 8) \tag{2}$$

이와 같이 구한 휘도신호 중 밝기가 상위 6.25%에 속하는 화소에 대응하는 R, G, B 값들의 평균을 구하고 식 (3)과 같이 R, G, B의 밝기 값을 정규화하기 위한 비례 계수를 구한다. 여기서 6.25%는 실험에 의해 가장 적합하다고 판단된 값이다.

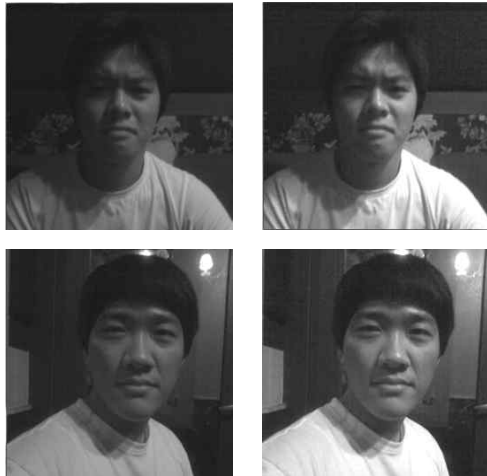
$$\begin{aligned} Ratio_R &= 255 / R_{avg} \\ Ratio_G &= 255 / G_{avg} \\ Ratio_B &= 255 / B_{avg} \end{aligned} \tag{3}$$

여기서 R_{avg} , G_{avg} , B_{avg} 는 각각 휘도신호의 상위 6.25%에 속하는 화소에 대응하는 R, G, B 값들의 평균이다. 식 (3)으로부터 구한 R, G, B 각각의 비례계수를 식 (4)와 같이 입력 R, G, B 신호에 곱하여 보정된 신호를 얻게 된다. 그림 2는 색상 진폭 보정 결과이다.

$$\begin{aligned} R_C &= Ratio_R * R \\ G_C &= Ratio_G * G \\ B_C &= Ratio_B * B \end{aligned} \tag{4}$$

2.2.2 히스토그램 정규화를 이용한 보정

얼굴검출에 있어 촬영 시 주변 환경에 따라 적정



(a) 보정 전 (b) 보정 후
그림 2. 색상 진폭 보정 전처리 수행 결과

량의 빛이 존재하지 않아 명암 대비 즉 콘트라스트가 낮은 영상을 취득하는 경우도 있는데, 이 경우 색 신호의 살색 특성이 많이 훼손된다. 특히 야외나 조명에 있는 공간에서 촬영할 때 피사체 보다 뒤쪽 배경이 밝아 피사체가 어둡게 되는 역광 영상을 취득하는 경우가 많아 이러한 영상을 개선할 수 있는 방법이 요구된다. 이를 해결하기 위한 전처리 방식으로 히스토그램 평활화(Histogram Equalization)를 이용한 방법을 사용한다. 식(5)와 같이 히스토그램 평활화 된 영상과 원영상을 합한 값을 이용한다[13]. 앞선 전처리 과정인 색상 진폭 보정이 적용된 경우는 이 과정을 생략하고, 그렇지 않은 경우에는 이 과정을 적용한다.

$$\begin{aligned} R_C &= \alpha * R_{EQ} + (1-\alpha) * R \\ G_C &= \alpha * G_{EQ} + (1-\alpha) * G \\ B_C &= \alpha * B_{EQ} + (1-\alpha) * B \end{aligned} \quad (5)$$

여기서 R_{EQ} , G_{EQ} , B_{EQ} 는 각각 히스토그램 평활화된 입력 R, G, B 신호에 대한 히스토그램 평활화된 신호이고, α 는 히스토그램 평활화된 입력신호에 대한 비례계수로 0~1사이의 값을 갖는다. 본 논문에서는 0.5로 설정하였는데, α 가 큰 경우 지나치게 평활화 결과가 영향을 미치게 된다.

그림 3은 히스토그램 정규화를 이용한 보정처리 결과를 나타낸 것으로 (a)와 (b), (c)와 (d)는 역광에 의해 화자의 얼굴영역이 매우 어둡게 촬영된 영상을 보정한 것이고, (e)와 (f)는 일반적인 조명에 적용시

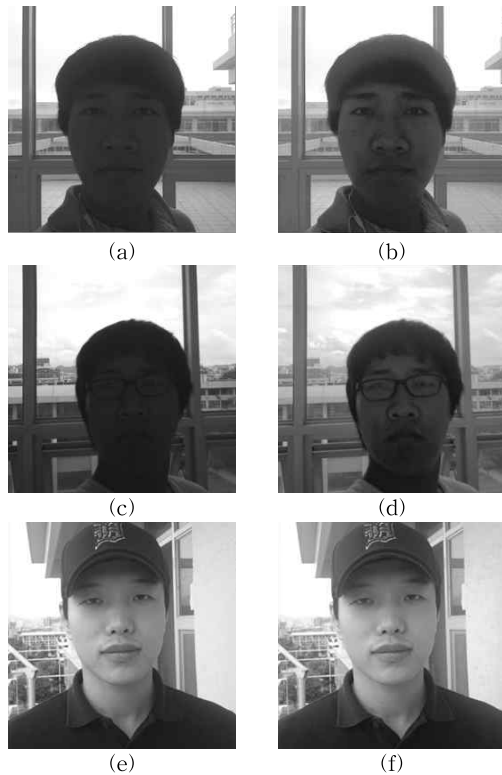


그림 3. 역광 보정 전처리 수행 결과

킨 결과이다. 그림 3을 통해 일반 조명에서의 영상에는 열화를 발생시키지 않으면서 역광에 의한 문제점을 일부 보정함을 알 수 있다.

2.3 다중 색좌표계 변환

전처리과정을 거친 k번째 영상프레임 F(k)는 다수의 색 좌표계에 대응하는 색신호성분으로 변환된다. 예를 들어 카메라로부터 얻어진 영상이 RGB 좌표계인 경우 색 좌표계 변환부에서 YCbCr 좌표계의 색신호 성분을 구하기 위해서 식 (6)의 연산을 수행하고, YIQ 좌표계의 색신호 성분을 구하기 위해서는 식 (7)의 연산을 수행한다[13]. 색신호 성분을 구하기 위한 식 (6), (7)도 앞서 휘도성분을 구한 식 (2)과 같이 연산을 줄이기 위해 소수점 연산을 곱하기와 bit-shift연산만으로 변환하였다.

$$\begin{aligned} Cb &= ((-43 * R - 84 * G + 128 * B + 128) \gg 8) + 128 \\ Cr &= ((128 * R - 107 * G - 20 * B + 128) \gg 8) + 128 \end{aligned} \quad (6)$$

$$\begin{aligned} I &= ((152 * R - 70 * G - 82 * B + 128) \gg 8) + 128 \\ Q &= ((54 * R - 133 * G + 79 * B + 128) \gg 8) + 128 \end{aligned} \quad (7)$$



그림 4. 좌표계 변환 처리 된 색신호 성분

얼굴영역 검출을 위한 살색검출에 적합한 색 좌표계는 YCbCr 좌표계와 YIQ 좌표계 및 HSV 좌표계 등으로 알려져 있다. 그러나 HSV 좌표계는 연산량이 많아 모바일 기기에서의 연산으로 적합하지 않고 판단되어 본 논문에서는 YCbCr 좌표계와 YIQ

좌표계만을 사용하였다. 그림 4는 YCbCr 좌표계와 YIQ 좌표계 변환 처리된 색신호 성분을 보여준다.

2.4 다 단계 임계치 처리과정

색 좌표계 변환부의 출력인 색신호들은 다단계 임계치 처리과정을 거쳐 살색확률을 표현하는 확률 마스크를 출력한다. 본 논문에서는 2단계 임계치 처리과정을 수행하였다. 색 좌표계 변환된 색신호들은 식 (8)과 같은 다단계 임계치 처리과정을 거쳐 살색 확률을 표현하는 확률 마스크를 출력한다.

$$\begin{aligned}
 & TL_c^i < c < TH_c^i \\
 & TL_c^{i+1} < TL_c^i \text{ and } TH_c^i < TH_c^{i+1}, i \geq 0
 \end{aligned}
 \tag{8}$$

여기서 TL_c^i 와 TH_c^i 는 각각 색 성분 c 에 대한 i 번째 단계의 하위 임계치와 상위 임계치를 의미한다. i 번째 단계 임계치 조건을 만족하는 색신호는 항상 $(i+1)$ 번째 단계 임계치 조건을 만족하는 화소보다 살색일 확률이 높은 화소임을 의미한다.

제안하는 다중 색 좌표를 이용하는 방법을 위해 다양한 조명과 환경에서의 영상으로부터 그림 5와 같이 황인, 백인, 흑인에 대해 각각 50명씩의 얼굴



그림 5. 황인, 백인, 흑인에 대한 얼굴색 샘플들 예

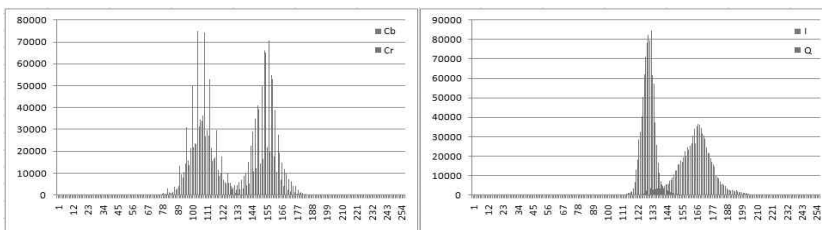
샘플들을 추출하였다. 그리고 각 인종별 Cb, Cr과 I, Q 색 좌표계에 대한 히스토그램을 얻어냈다. 그림 6은 인종별로 50개의 영상프레임에서 얼굴영역만을 추출하여 얻은 히스토그램 나타낸 것으로 (a)는 황인의 Cb와 Cr의 색신호와 I와 Q의 색신호에 대해 나타낸 것이고, (b)는 백인, (c)는 흑인, (d)는 3가지 인종의 히스토그램을 합친 전체 인종의 얼굴색 범위 히스토그램을 나타내는 것이다.

그림 7은 그림 6의 얼굴영역만 추출하여 얻은 히스토그램을 이용하여 2단계 임계치를 결정하는 예를 나타낸 것이다. Cb와 Cr 색신호, I와 Q의 색신호에 대해 각각 동일한 방법으로 미리 조사하여 놓은 범위를

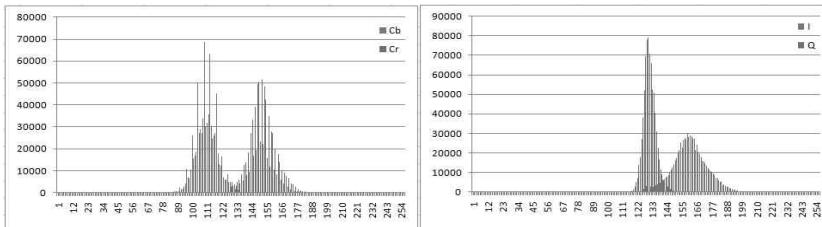
적용하여 얼굴 색신호에 대한 확률 값을 구할 수 있다. 표 1은 Cb, Cr, I, Q의 본 논문에서 적용한 피부색 2단계 임계치 값을 나타낸다. 여기서 2단계 임계값은 확실한 피부색 영역과 애매한 피부색 영역의

표 1. 피부색 2단계 임계치 값

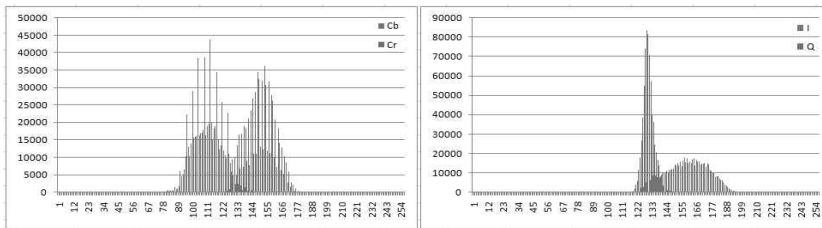
구분	2단계 Low	1단계 Low	1단계 High	2단계 High
Cb	90	100	115	125
Cr	135	145	160	170
I	140	150	170	180
Q	115	125	130	145



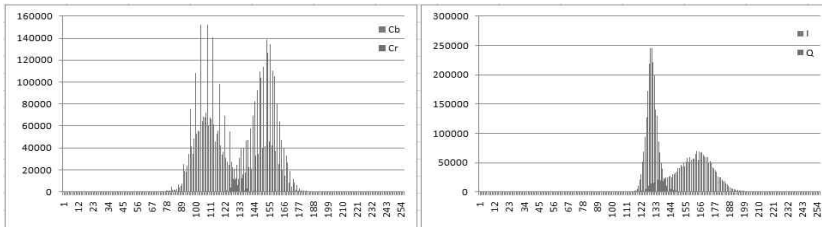
(a) 황인 얼굴 히스토그램



(b) 백인 얼굴 히스토그램



(c) 흑인 얼굴 히스토그램



(d) 황인, 백인, 흑인 전체 얼굴 히스토그램

그림 6. 인종별 얼굴 히스토그램(왼쪽 : CbCr 성분, 오른쪽 : IQ 성분)

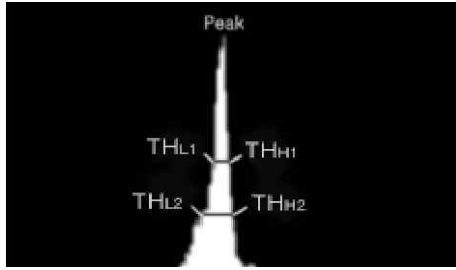
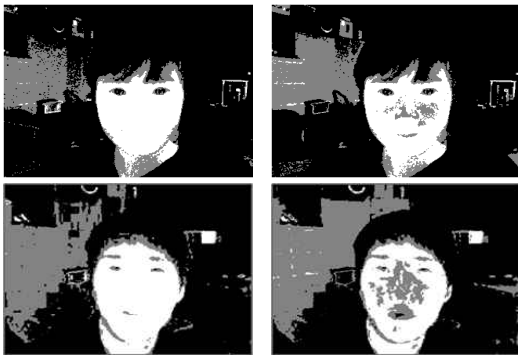


그림 7. 2단계 임계치 적용 예

구분이 가능하게 실험을 통해 정해진 값이다.

그림 8은 2단계 임계치를 적용하여 얻어진 확률 마스크 예를 나타낸 것으로 밝은 부분이 어두운 부분보다 살색 확률이 높은 영역임을 의미하는데, (a)는 Cb와 Cr 색신호에 대한 살색 확률 마스크의 예이고 (b)는 I와 Q 색신호에 대한 살색 확률 마스크 예이다. 확률 마스크는 2단계 임계치 처리가 되는 경우 0~2의 값을 갖는데, 큰 값은 살색 확률이 높음을 의미한다.



(a) Cb와 Cr (b) I와 Q

그림 8. 2단계 임계치를 적용하여 얻어진 확률 마스크 예

2.5 통합 확률 마스크 생성과 IIR 필터 처리

각 색 좌표계에 대한 확률 마스크는 모두 더해 통합 확률 마스크로 만들어진다. 그림 9는 그림 8에 나타난 Cb와 Cr 색신호와 I와 Q 색신호에 대한 개별 확률 마스크를 더해 얻은 통합 확률 마스크의 예를 나타낸 것으로 2단계 임계치 처리 방식을 적용할 경우 개별 확률 마스크가 가질 수 있는 값의 범위가 0~2 이므로 0~4 사이의 값을 갖는 통합 확률 마스크를 얻을 수 있다. 이와 같이 얻어진 통합 확률 마스크는 식 (9)와 같은 연산을 하는 시간 축 IIR 필터 처리되어 선택된 얼굴영역이 프레임 단위로 급격히

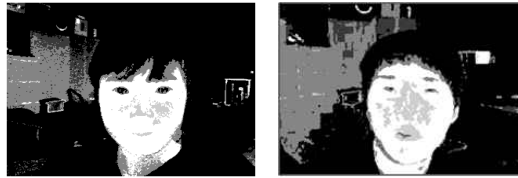


그림 9. 통합 확률 마스크 예

변하지 않고 이전 프레임과의 상관성이 유지되도록 함으로써 얼굴영역 검출의 신뢰도를 높인다.

$$M(k) = \alpha \cdot M(k) + (1-\alpha) \cdot M(k-1) \tag{9}$$

여기서 M(k)는 k번째 프레임에 대한 통합 확률 마스크이고, α는 가중치로 0~1 사이의 값을 갖는 상수이다. 만일 α = 1이면 시간 축 IIR 필터 처리 결과는 입력과 동일한 마스크가 된다. 그림 10은 IIR 필터 처리 결과를 보인 것으로 그림 9와 비교하면 모서리 또는 에지에 나타난 높은 확률 값이 낮아지고, 전반적으로 평활화되는 특징이 있다. 본 논문에서는 α값을 0.5로 적용하였다.

IIR 필터 처리는 단위 시간 동안 확률 마스크 값이 크게 움직일 수 없는 가정 하에 시간 축에 적용한다. 이는 단위 시간 내에 이전 얼굴 검출 부분과 동떨어진 곳에 얼굴 후보영역이 검출되면 이를 얼굴이 아닌 것으로 판단하는 과정으로 갑작스런 빛의 간섭에도 더욱 강한 검출을 할 수 있도록 한다. 그림 11은 IIR 필터를 적용한 것과 하지 않은 것의 비교이다. IIR 필터 적용으로 프레임 14의 갑자기 어두워 질 경우의 오 검출과, 프레임 132의 역광으로 인한 오 검출에 강인함을 확인할 수 있다.

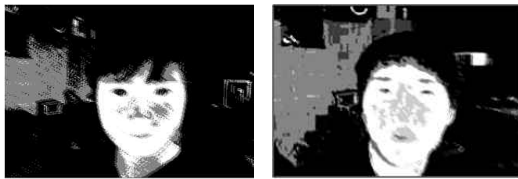


그림 10. IIR 필터 처리된 통합 확률 마스크 예

2.6 투영 히스토그램 생성과 얼굴 영역 검출

구해진 마스크의 점수가 높은 화소위치는 얼굴영역일 가능성이 높은 위치이므로 최종적으로 얼굴 영역을 선택하기 위해서 확률 마스크의 값(앞서 구한 통합 확률 마스크에서의 0~4 사이의 값)을 수평·수직 방향으로 투영하여 누적 값을 계산한다. 특히 살



그림 11. IIR 필터 적용과 미적용 비교

색 확률 마스크는 피부색과 유사한 배경이나 의상 같은 노이즈가 포함된 영역도 높은 확률을 가질 수 있기 때문에 노이즈는 제거하고 얼굴 영역만을 추출하기 위해서 확률 마스크의 살색 확률을 수평방향과 수직방향으로 투영하여 그림 12와 같은 수평·수직 투영 히스토그램을 얼굴영역 검출에 사용한다. 얼굴영역 검출은 수평·수직 투영 히스토그램에서 최대값을 구하고, 수평·수직 좌표의 히스토그램 값이 일

정 임계치보다 적어지는 위치를 얼굴영역 경계로 결정하여 그림 13과 같이 얼굴영역을 검출한다. 이때 2단계 임계치는 1단계 임계치보다 얼굴영역의 확률값을 더욱 높여 주고, 오검출을 줄여주기 때문에 수평·수직 투영 시 더욱 정확한 검출 결과를 보장한다. 여기서 임계치는 실험을 통해 투영 히스토그램 최대값의 1/3이 적합한 것으로 확인되었다.

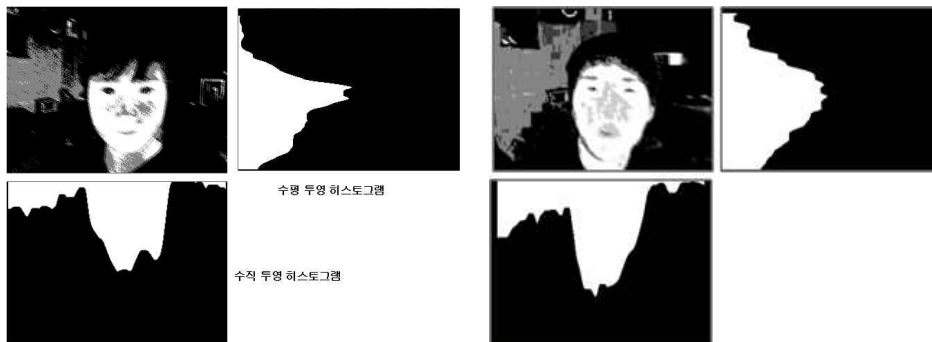


그림 12. 수평·수직 투영 히스토그램 예



그림 13. 수평·수직 투영 히스토그램을 통해 선택된 얼굴영역

2.7 화자와 스마트폰 사이의 거리측정

검출된 얼굴영역을 기반으로 화자와 스마트폰 사이의 거리 측정은 다음 3가지 방법으로 거리(d)를 측정할 수 있는데, 3가지 얼굴영역 파라미터(x) 모두 식 (10)과 같이 거리와 반비례관계를 갖는다.

- ① 검출된 얼굴영역의 면적(화소수)을 사용
- ② 검출된 얼굴영역의 수직길이(화소수)을 사용
- ③ 검출된 얼굴영역의 수평길이(화소수)을 사용

$$d = \frac{a}{x} + b \tag{10}$$

계수 a와 b는 카메라 종류에 따라 변하는데, 얼굴영역 파라미터와 거리에 대한 실험을 통해서 얻어진 데이터의 curve_fitting으로 구할 수 있다. 그림 14는 검출된 얼굴영역의 수평길이(화소수)와 사용자-폰 사이의 거리관계를 나타낸 것으로 3차례 실험을 통해 얻어진 데이터를 식 (10)로 fitting 한 예이다. 실험에 따르면 95% 신뢰구간에서 파라미터 a = 4630이고 b = 2.549이다.

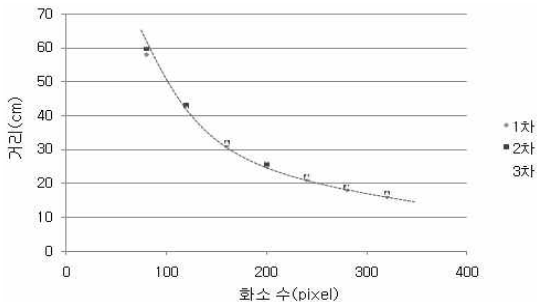


그림 14. 얼굴영역 수평방향 화소수와 화자-스마트폰 거리 관계

3. 실험 결과 및 분석

3.1 Bit-shift 연산에 의한 속도 향상

제안된 알고리즘의 성능 평가를 위해 CPU

806Mhz와 RAM 160MB 성능의 SAMSUNG SCH-M495 model을 가지고 다양한 조명과 배경 그리고 피부색에 대해 실험을 하였다. 영상 통화시를 기본으로 생각하여 영상에 포함되는 1명일 경우로 제한하였다. SAMSUNG SCH-M495 model은 기본적으로 320×240 영상 사이즈를 기준으로 15fps의 카메라를 제공하는데[14], 제안한 알고리즘의 전체 처리과정을 수행한 처리 속도는 14.1fps로 측정되었다.

일반적으로 모바일기는 부동소수점 연산에 대한 수치 프로세서(FPU: Floating Processing Unit)를 탑재하지 않은 모델들이 다수이기 때문에 일반적인 PC와는 다르게 수행속도의 보장을 받을 수 없다. 그렇기 때문에 얼굴 검출 과정에 있어서 발생하는 수많은 부동소수점 연산에 대한 별도의 처리를 해주지 않으면 수행속도가 매우 저하되게 된다[15]. 그렇기 때문에 본 논문에서는 부동 소수점 및 나눗셈 연산은 곱하기와 bit-shift연산으로 변환하여 수행함으로써 연산속도를 증가하였다. 수행 속도의 차이는 표 2와 같다.

표 2. 부동소수점과 나눗셈 연산의 사용비교

구 분	수행속도(fps)
모바일 기본 제공 카메라	15
부동소수점과 나눗셈 사용안함	14.1
부동소수점과 나눗셈 사용	3.9

3.2 얼굴 검출 결과 분석

본 논문에서는 얼굴 검출을 위해 YCbCr 색 좌표계와 YIQ 색좌표계에 대해 2단계 임계치를 적용하고 통합 확률 마스크 값에 대한 수직, 수평 프로젝션을 하여 얼굴 검출을 수행하였다. 그림 15는 본 논문에서 제안한 방법과 Chai 등[8]이 제안한 YCbCr 단일 색좌표계에 단일 임계치를 적용하여 얼굴을 검출한 결과를 비교한 것이다. 피부색에 단일 색좌표, 단일 임계치를 적용할 경우, 프로젝션을 통한 주변의 피부색과 유사한 배경과의 분리가 적정 임계치를 적용해도 쉽지 않음을 알 수 있지만, 본 논문에서 제안한 다중 색좌표계에 대해 다중 임계치를 적용하는 경우 피부색에 대한 각각의 확률 값을 계산하기 때문에 적정 임계치 적용이 더욱 쉬워지고 조금 더 정확한 피부색 영역 검출이 가능하다.

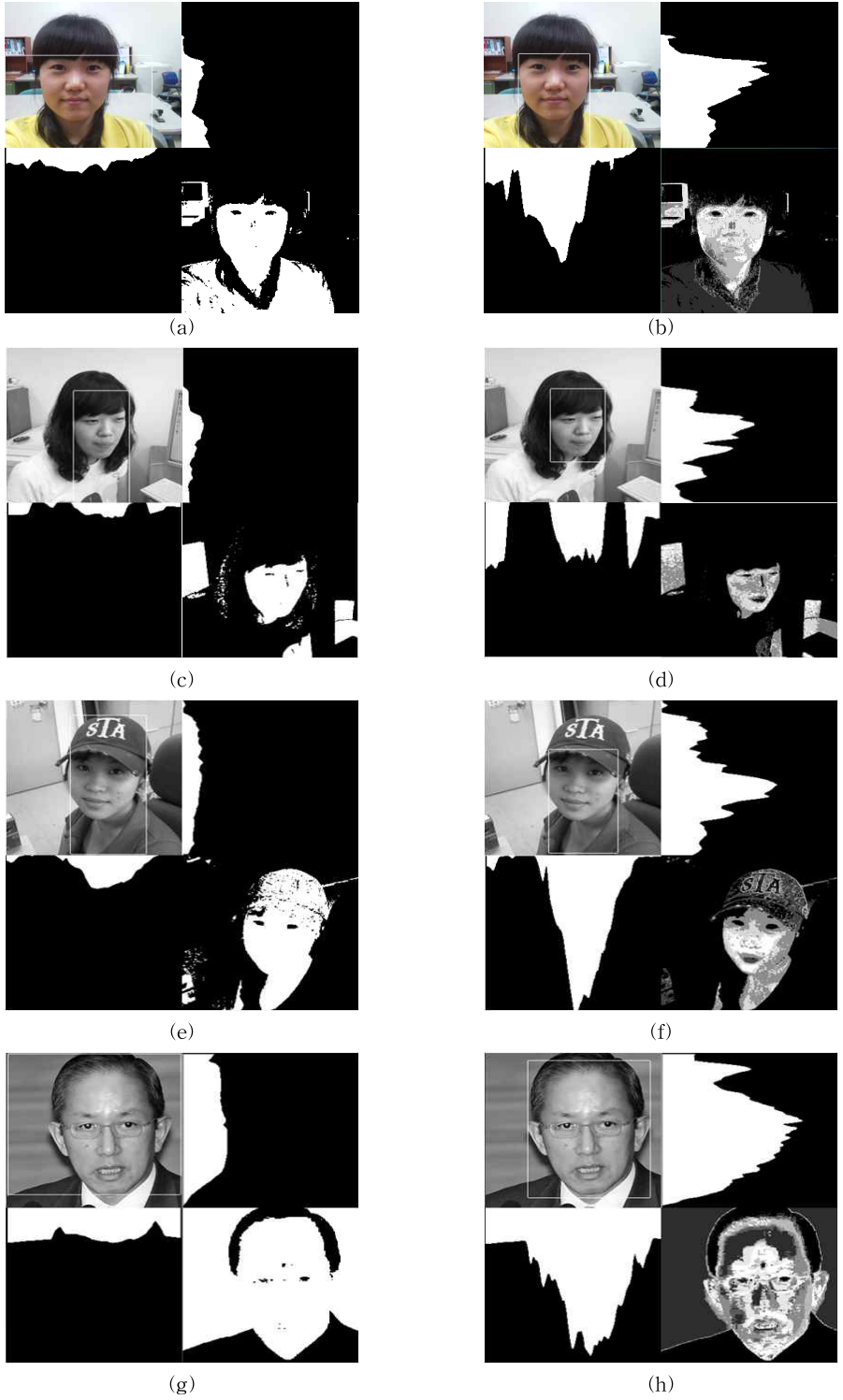


그림 15. 제안한 알고리즘을 이용한 얼굴 검출의 예
(a), (c), (e) CbCr에 1단계 임계치 적용
(b), (d), (f) CbCr, IQ에 2단계 임계치 적용

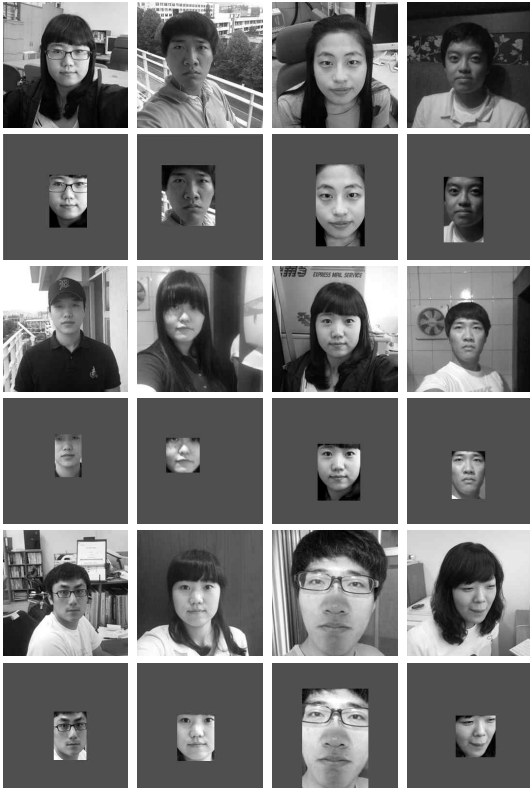


그림 16. 제안한 알고리즘을 이용한 얼굴 검출의 예

또한 얼굴 영역 검출에 대한 결과를 검증하기 위하여 환경에서 168 case에 대해 실험을 하였다. 그림 16은 실험영상에서의 얼굴검출 결과 중 일부를 보여준다. 다양한 환경과 인물, 조명등에서 얼굴 영역을 잘 검출한 것을 확인할 수 있다. 표 3은 168 case에 대한 얼굴 검출 실험 결과를 나타낸다. False Positive는 얼굴 영역을 찾았지만 얼굴 검출 박스에 얼굴 이외의 영역이 포함된 경우를 나타내고, False Negative는 얼굴 영역을 찾았지만 얼굴 검출 박스 안에 얼굴의 전체가 아닌 일부가 들어간 경우를

표 3. 얼굴 검출 실험 결과

Set No.	Correct	False Positive	False Negative	Detection Percentage
Set1	47/48	2	1	97.9%
Set2	46/48	1	2	95.8%
Set3	45/48	2	2	93.7%
Set4	24/24	0	1	100%
total	162/168	5	6	96.4%

나타내는 것이다.

본 논문에서는 얼굴 검출을 통해서 얻어진 얼굴의 수평 길이(화소 수)를 기준으로 사용자와 스마트폰과의 거리를 측정하는 방법을 제시하였다. 일반적으로 사람의 얼굴 크기가 12~16cm 내에 존재한다는 가정 하에 실험한 결과 표 4의 내용과 같이 측정 되었다. 14cm 정도의 얼굴크기에 대해서는 매우 근사한 값을 보이고 있으나, 12나 16cm일 때는 약간의 오차가 발생하였다. 스마트폰 단말의 이동성과 휴대성에 의한 사용자와 폰의 각도나 흔들림 등을 생각했을 때 허용 오차 범위로 판단된다.

그림 17은 본 논문에서 제안한 검출된 얼굴을 통한 화자와 스마트폰 간의 거리를 나타내어 화자가 스마트폰과의 거리를 쉽게 파악 할 수 있도록 한 결과를 보여준다. 거리에 대한 표현으로 절대적 거리(cm)와 상대적 거리(타원)를 5단계로 표현하여 화자에게 알림으로써 통화품질과 음성 인식률을 보장 받을 수 있는 미리 지정된 화자와 스마트폰간의 적정 거리를 알려준다. 결과를 살펴보면 실험영상에 대한 얼굴영역 선택결과와 이를 이용하여 스마트폰과 화자간의 거리를 나타낸 것으로 피부색과 유사한 배경에서도 얼굴 영역선택과 거리 측정을 수행함을 알 수 있다.

4. 결 론

본 논문에서는 모바일 장치에서의 얼굴검출과 이를 통한 화자와 단말간의 거리를 측정하기 위한 방법을 제시하였다. 이를 위하여 모바일 환경에 적합한 저 연산과 간단한 처리과정의 얼굴 검출 방법 및 거리 측정법을 보였다. 기존의 단일 피부색에 단일 임계치를 적용하는 방법에 비해 다중 색 좌표계와 다중 임계치를 적용함으로써 좀 더 환경과 조명에 강인한 색상을 이용한 얼굴 영역 검출 방법을 제안하였고, 이를 이용하여 모바일 폰을 이용하여 영상통화를 하는 화자와 단말 간의 적정거리를 측정 가능한 방법을 제안하였다. 실험 결과 일정 성능 이상의 거리 측정 결과를 보였지만, 이를 위해서는 좀 더 정확한 얼굴 검출 방법이 추가적으로 보완되어야지만 더욱 정확한 거리측정이 가능할 것이다. 그리고 본 논문의 방법은 모바일 환경에 적합하게 매우 연산량이 적어 립 리딩을 위한 얼굴 검출이나 모바일에서의 얼굴

표 4. 폰과 사용자간 실제거리와 측정거리 비교

실제얼굴크기(cm)	실제거리(cm)	20	25	30	35	40	45
12	측정얼굴크기(pixel)	209	182	143	129	110	98
	측정거리(cm)	24.70	27.98	34.92	38.44	45.02	49.79
14	측정얼굴크기(pixel)	216	194	151	138	119	106
	측정거리(cm)	23.98	26.41	33.21	36.09	41.45	46.22
16	측정얼굴크기(pixel)	228	206	163	149	130	116
	측정거리(cm)	22.85	25.02	30.95	33.62	38.16	42.46

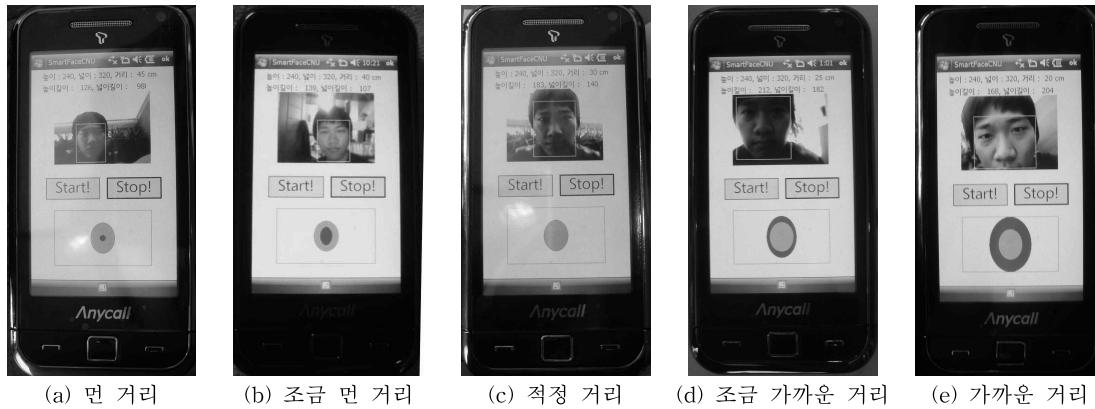


그림 17. 얼굴 검출 및 거리 측정 결과

인식 등의 기술과 접목한 기술로 발전시킬 수 있을 것으로 본다.

참 고 문 헌

[1] M-H Yang, D.J.Kriegman, and N.Ahuja, "Detecting Faces in Images: A Survey," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, Vol.24, No.1, pp. 34-58, 2002.

[2] W. Zhao, R. Chellappe, P.J. Philips, and A. Rosenfeld, "Face Recognition: A literature survey," *ACM Computing Surveys*, Vol.35, No.4, pp. 399-458, 2003

[3] R. Brunelli and T. Poggio, "Face Recognition : Features Versus Templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.15, pp. 1042-1052, 1993.

[4] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *Int. J. of Computer vision*, Vol.57, No.2, pp. 137-154, 2004.

[5] J.-C. Terrillon, M. David, and S. Akamatsu, "Automatic Detection of Human Faces in Natural Scene Images by Use of a Skin Color Model and of Invariant Moments," *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, Vol.1, No.1, pp. 112-117, 1998.

[6] Y.Dai and Y.Nakano, "Face-Texture Model-based on SGLD and its application in face Detection in a Color Scene," *Pattern Recognition*, Vol.29, No.6 pp. 1007-1017, 1996.

[7] J. Yang, W. Lu, and A. Waibel, "Skin-Color Modeling and Adaptation," *Proceedings of ACCV'98*, Vol.2. No.1. pp. 687-694. 1998.

[8] D.Chai and K.N.Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.9, No.4, pp. 551-564, 1999.

[9] 권혁봉, 권동진, 장언동, 윤영복, 안재형, "YCbCr 색공간에서 피부색과 윤곽선 정보를 이용한 얼굴 영역 검출," 한국멀티미디어학회지, 제7권, 제1호, pp. 27-34, 2004

[10] E. Weleh, R. Moorhead, and J. K. Owns, "Image Processing Using the HSI Color Space," In Proc. IEEE Conference Southeastcon'91, Williamsburg, Vol.2, pp. 722-725, 1991.

[11] Sanjay Kr. Sinhg, D. S. Chauhan, Mayank Vatsa, and Richa Singh, "A Robust Skin Color Based Face Detection Algorithm," *Tamkang Journal of Science and Engineering*, Vol.6, No.4, pp. 227-234, 2010

[12] R.L. Hsu, M.Abodel-Mottaleb, and A.K.Jain. "Face Detection In Colour Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.24, No.5, pp. 696-706, 2002.

[13] R. G. Gonzalez and R. E. Woods, Digital image processing, *Addison Wesley Publish*, Mar. 2009.

[14] SAMSUNG Mobile Innovator : <http://innovator.samsungmobile.com/>

[15] 윌리엄 스톨링스 저, 김종현 역, 컴퓨터 시스템 구조론, 사이텍 미디어, 2006.



이 재 원

2004년 3월~2009년 2월 전남대학교 전자컴퓨터공학부 학사
 2009년 3월~현재 전남대학교 전자컴퓨터공학과 석사과정
 관심분야 : 영상처리, 컴퓨터 비전, 멀티미디어시스템



권 구 락

1997년 경일대학교 전자공학과 학사
 1999년 성균관대학교 전기전자 및 컴퓨터공학부 석사
 2002년 3월~2007년 2월 고려대학교 메카트로닉스학과 공학박사
 2004년 5월~2007년 2월 DaliTech Co., Ltd. 책임연구원 및 대표이사
 2008년 3월~현재 조선대학교 정보통신공학과 조교수
 관심분야 : 미디어 신호처리, 정보보안, 컴퓨터 비전



홍 성 훈

1988년 2월 영남대학교 전자공학과 공학사
 1991년 2월 한국과학기술원 전기 및 전자공학과 공학석사
 1999년 2월 한국과학기술원 전기 및 전자공학과 공학박사
 1991년~2000년 7월 LG전자 DTV 연구소 책임연구원
 2007년~현재 전남대학교 전자컴퓨터공학부 교수
 관심분야 : 영상 처리, 영상 부호화, 멀티미디어 시스템