

논문 2004-41SP-6-8

HCr과 적응적 임계화에 의한 고속 얼굴 검출

(Fast Face Detection in Video Using The HCr and Adaptive Thresholding Method)

신 승 주*, 최 석 립**

(Seungju Shin and Seokrim Choi)

요 약

얼굴검출을 위한 다양한 연구가 행해지고 있으나 아직도 실시간성의 확보는 미진하다. 이에 본 연구는 연속영상에서 컬러와 움직임 정보를 이용한 실시간 얼굴검출 방법을 제안한다. 피부색 검출을 위한 컬러공간은 조명의 변화에 강인하고 피부색을 좁은 영역으로 정의할 수 있는 Hue와 Cr성분을 조합하여 재구성한 HCr을 사용한다. 배경참조영상 기반에서 밝기와 Cr 성분을 함께 사용하여 획득한 움직임 영역에서, HCr과 적응적 임계값을 이용해서 피부색 영역을 검출하고, 그 검출된 영역의 모양과 크기정보를 통해 얼굴 후보영역을 구한다. 이렇게 구해진 얼굴후보영역에서 G와 B성분의 차이, 밝기, Cr성분 값과 눈과 입의 위치 및 거리관계를 이용하여 눈과 입을 검출하여 얼굴을 확정한다. 실험결과 연속영상에서 실시간으로 얼굴을 검출 할 수 있었다.

Abstract

Recently, various techniques for face detection are studied, but most of them still have problems on processing in real-time. Therefore, in this paper, we propose novel techniques for real-time detection of human faces in sequential images using motion and chroma information. First, background model is used to find a moving area. In this procmoving area, edure, intensity values for reference images are averaged, then skin-color are detected in We use HCr color-space model and adaptive threshold method for detection. Second, binary image labeling is applied to acquire candidate regions for faces. Candidates for mouth and eyes on a face are obtained using differences between green(G) and blue(B), intensity(I) and chroma-red(Cr) value. We also considered distances between eye points and mouth on a face. Experimental results show effectiveness of real-time detection for human faces in sequential images.

Keywords : Face Detection, Color Space, HCr, Adative Threshold

I. 서 론

컴퓨터와 카메라를 비롯한 영상관련 하드웨어의 비약적인 발달과 첨단멀티미디어 시대로의 진행에 따라 수많은 영상정보들이 넘쳐나게 되었으며, 그 영상정보의 분석과 활용에 대한 관심과 노력 또한 갈수록 증대되고 있다. 특히 대용량의 데이터를 고속으로 처리할 수 있는 컴퓨터의 발달로 인해 컴퓨터 비전 기술은 이

제 추상적인 이론에서 벗어나 군사적인 목적을 가지는 시스템이나, 우리의 실생활에 유용한 여러 응용 시스템에 접목되고 있으며 활용분야도 더욱 넓어지고 있다.

컴퓨터 비전 기술이 일상생활에 사용된 예로는 다양한 생체인식시스템을 들 수 있다. 생체인식(Biometrics)이란 사람의 생체적 특성을 이용하여 개인을 식별하는 기술을 의미하는 것으로 지문인식, 얼굴인식, 홍채인식, 정맥인식, 망막인식 등이 있다. 이 중에서 지문인식이나 홍채인식, 정맥인식 등의 기술은 사용자의 협조를 가정해야 하기 때문에 사용자 편의성 측면에서 거부감을 일으키는 등의 단점을 가진다. 따라서 비록 정확성 측면에서 아직 성능이 떨어지지만 사용자의 협조 없이, 또는 전혀 의식하지 못하는 상황에서도 적용 가능한 얼굴

* 학생회원, ** 정회원, 세종대학교 전자공학과
(Dept. of Electronics Engineering, Sejong University)
※ 본 연구는 과학기술부 21세기 프론티어연구개발사업 중 유비쿼터스컴퓨팅 및 네트워크 원천기반기술 개발사업의 지원에 의한 것임.
접수일자: 2004년1월5일, 수정완료일: 2004년9월13일

인식 기술에 대한 요구가 확대되고 있다. 특히 미국의 9.11 테러이후 공항이나 공공기관시설 등에서의 보안에 대한 필요성은 폭발적으로 늘어나고 있으며, 이러한 장소에 적합한 보안시스템으로 평가받고 있는 얼굴인식시스템에 대한 관심이 더욱 높아지고 있다.

얼굴검출은 얼굴인식을 위한 사전단계의 기술로 연구되어왔으나 그 중요성이 인정되면서 점차 독립적인 분야로서 연구되고 있다^[1]. 얼굴검출 기술은 얼굴인식시스템 외에도 비디오 보안감시 시스템에서 DVR(Digital Video Recorder)기술과 접목되어 제3세대 비디오 보안감시 카메라 분야에서 중요한 역할을 하고 있다. 또한, 화상회의 시스템에서 얼굴 영역은 충실하게 압축하고 배경영역은 과감하게 압축하여 관심영역의 화질저하 없이 전체 압축률을 높여 실시간 전송을 보다 수월하게 할 수도 있으며, HCI(Human Computer Interaction)관련 시스템에서 보다 다양하고 편리한 인터페이스를 구성하는데 있어서도 필수적으로 요구되는 기술이다. 이처럼 얼굴 검출에 관련된 응용 분야는 다양하게 존재하며 앞으로 더욱 폭넓게 확대될 것이다.

임의의 영상에서 얼굴영역을 검출하는 것은 생각처럼 간단하지 않다. 영상에서의 얼굴은 다양한 크기를 가지며, 얼굴의 회전이나 촬영하는 카메라의 각도 또는 표정변화에 의한 얼굴형상의 변화도 있을 뿐만 아니라 조명의 변화에도 얼굴의 모양이나 특징들이 다양하게 나타나기 때문이다^[1,2].

얼굴검출 연구에서 가장 오래된 연구는 템플릿 정합(Template Matching)방법이다. 얼굴 전체 형상에 대한 템플릿 모델을 만들어 주어진 영상이 얼마나 그 템플릿 모델에 부합되는지를 따져서 판별해 내는 것이다. 이 방법은 배경이 복잡하거나 얼굴이 회전되어 있는 경우에 있어서 성능이 크게 떨어진다.

비교적 최근에는 모자이크 얼굴영상을 학습한 신경망(Neural Network)을 이용한 방법이나, SVM(Support Vector Machine)을 이용한 방법들이 많이 연구되고 있다. 이 방법들은 다른 방법들에 비해 정확성 측면에서 보다 우수한 성능을 보이고 있으나 계산복잡도가 너무 높아서 실시간 처리가 요구되는 시스템에서의 응용은 어렵다^[3-5].

이전부터 꾸준히 연구되고 있는 특징기반 방법은 눈이나 코, 입 등의 존재유무와 얼굴에서의 대칭성 타원모형, 그리고 피부색정보를 이용한 방법 등이 있다. 이 중에서 컬러를 이용한 방법은 계산량이 작기 때문에 실시간 처리에 적합하다. 그러나 컬러만을 이용할 경우에

는 배경에서 피부색이 존재한다거나 조명 상태가 변하는 경우에 효과적인 얼굴 검출이 어렵다^[6].

본 논문에서는 요즘은 대부분의 영상정보가 컬러동영상이라는 점과 동영상에기반한 응용시스템에서는 실시간성의 확보 여부가 중요하다는 판단아래, 컬러정보와 움직임 정보를 이용하여 전체적으로 계산 복잡도가 낮은 실시간 얼굴검출 방법을 제안하였다. 구성은 다음과 같다. 제II장에서 피부색 검출에 이용되는 컬러공간의 특성들을 살펴본뒤 본 연구에서 제안하여 사용한 컬러공간과 적응적 임계값 설정방법, 컬러와 위치관계를 이용하여 눈과 입을 검출하는 방법에 대해 설명한다. 그리고 제III장에서는 제안한 방법에 대한 실험결과를 보이고 제IV장에서 결론을 맺고 향후의 연구 과제를 제시한다.

II. 본 론

1. 컬러공간

피부색은 전체 컬러공간에서 아주 적은 영역을 차지한다. 이러한 특성은 영상에서 피부색과 기타 배경을 효과적으로 구분할 수 있도록 한다.

그러나 피부색의 영역 또는 분포는 컬러공간에 따라 각기 다르다. 따라서 어떠한 컬러공간을 선택하느냐에 따라 검출 성능에 상당한 영향을 미친다^[7].

디지털 영상을 얻기 위해 사용되는 대부분의 색채 카메라들은 RGB 포맷을 사용한다. 따라서 RGB 컬러공간을 사용하면 별도의 변환 과정 없이 하드웨어로부터 입력받아 바로 사용할 수 있는 장점이 있지만 조명에 따른 컬러 값의 변화가 심하다. 이러한 이유로 얼굴검출 연구와 관련하여 주로 사용되는 컬러공간은 조명의 영향이 적은 HSI^[8,9]와 YCbCr^[6,10,11], 그리고 Normalized RGB^{[12][13]} 등이다.

가. HSI 컬러공간

HSI는 색상(Hue), 채도(Saturation), 명도(Intensity)

표 1. HSI 컬러공간의 구성 성분
Table 1. HSI color components.

성분	범위	내용
Hue	[0, 360]	색상을 각도 θ 로 표시
Saturation	[0, 1]	원색의 희석정도
Intensity	[0, 1]	밝기의 강도

표 2. YCbCr 컬러공간의 구성 성분
Table 2. YCbCr color components.

성분	범위	내용
Y	[0, 255]	밝기의 강도
Cb	[-128, 127]	청색 성분
Cr	[-128, 127]	적색 성분

를 구성성분으로 하는 컬러공간으로서, 밝기 성분이 색 정보로부터 분리되고 사람이 색을 인지하는 방식과 가장 흡사하다는 특징을 가진다.

나. YCbCr 컬러공간)

밝기 값을 나타내는 휘도신호 Y와 두 개의 색차신호 Cb, Cr 성분으로 구성되어 있다. 비디오 코딩과 같은 영상압축 분야에서 주로 사용되는 컬러공간이다.

다. Normalized RGB 컬러공간)

조명의 영향을 줄여주기 위해 R, G, B 각각의 성분 값을 세 성분 값의 합으로 나누어 주는 정규화 과정을 거쳐 생성한다.

2. 컬러와 움직임 정보를 이용한 얼굴 검출

가. HCr 컬러공간.

피부색 영역을 정의하기 위해서는 컬러공간상에서 피부색의 분포 영역을 근사화 해야 한다. 그런데 피부색의 분포는 컬러공간에 따라 다르게 나타나고 그 특성들도 각기 다르다. 따라서 기존 연구들에서는 일반적인 컬러공간들 중에서 어떤 컬러공간을 선택하는 것이 피부색 검출에 더 우수한지를 가리는 것과, 피부색 영역을 더 정확히 근사화 시키는 방법에 대해서 주로 연구가 이루어졌다.

그러나 각 컬러공간을 구성하는 각각의 성분들은 조명에 따라 변화하는 정도나 표현할 수 있는 색의 영역이 다르다. 따라서 본 연구에서는 일반적으로 사용되는 컬러공간을 그대로 사용하지 않고, 컬러공간을 구성하는 성분들 중에서 얼굴피부색 검출에 보다 적합한 성분들을 찾아 새로운 컬러공간을 구성하여 사용한다.

피부색 검출에 사용되는 컬러공간의 조건은 크게 두 가지로 생각할 수 있다. 첫째로, 피부색을 규정짓는 영역이 협소하여야 한다. 즉 밀집도가 커야 한다. 피부색 이외의 다른 영역들이 동반 검출되는 것을 방지 할 수 있기 때문이다. 두 번째는 조명의 변화에 민감하지 않아야 한다는 것이다. 밝기에 민감하다면 동일 개체의

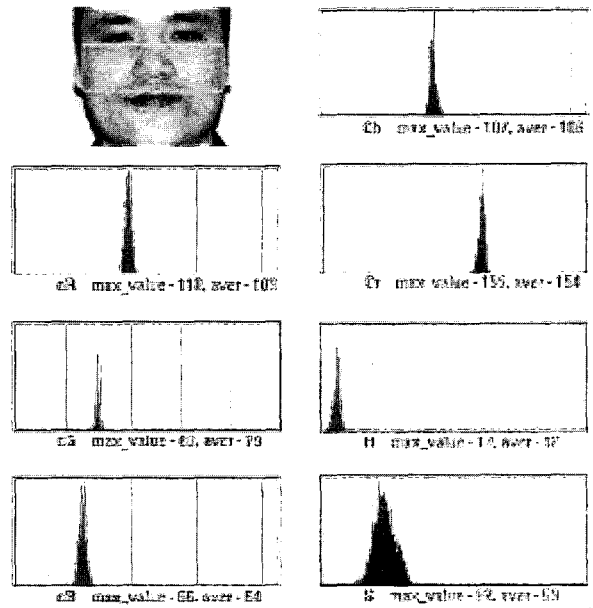


그림 1. 얼굴 샘플에 대한 히스토그램
Fig. 1. Histograms for components of facial sample.

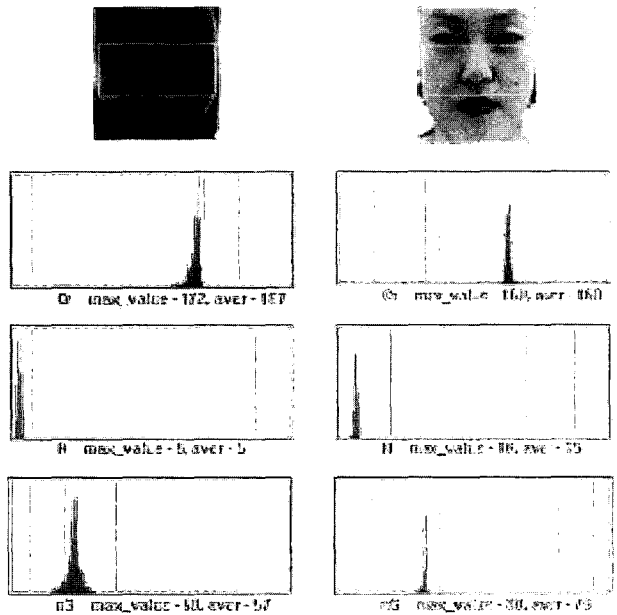


그림 2. 영상에 따른 히스토그램
Fig. 2. Histograms for different facial images.

색상 값이 조명에 따라 다르게 나타나게 되어 피부색을 정의하기가 어렵기 때문이다.

그림 1은 HSI 컬러공간에서 Hue(H)와 Saturation(S), YCbCr 컬러공간에서 Cb와 Cr 그리고 정규화 과정을 거친 RGB 컬러공간에서의 Normalized R(nR), Normalized G(nG), Normalized B(nB) 각 성분에 대한 얼굴 샘플의 히스토그램 분포를 보인 것이다. 구성성분 각각의 히스토그램들 중에서 S의 밀집도가 가장 떨어지는

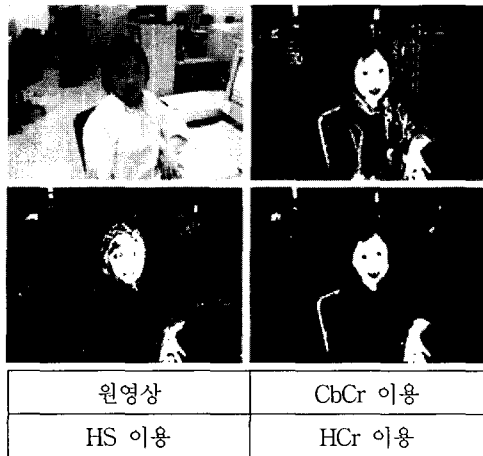


그림 3. 피부색 검출 영상 I
Fig. 3. Skin color detection I.

것을 볼 수 있고, 그에 반해 H, Cr, nG 성분이 우수한 밀집도를 갖는 것을 알 수 있다. 하지만 밀집도만 우수하다고 해서 피부색검출이 잘되는 것은 아니다.

그림 2는 영상에 따른 H, Cr, nG 성분의 히스토그램 분포의 변화 정도를 볼 수 있는 실험 영상이다. 왼쪽 영상은 실내조명은 비교적 어둡고 창문을 등쳐서 역광현상이 있는 상황에서 얼굴을 근접 촬영한 것이고, 오른쪽 영상은 일반적인 실내 조명환경에서 촬영한 것이다.

히스토그램을 보면 Cr, H, nG 각 성분의 변화 폭이 차이가 남을 볼 수 있다. Cr 성분의 경우 평균값을 비교 했을 때 7 크기의 변화폭을 보이고 있고, H 성분은 10 크기의 변화폭을, 그리고 nG 성분의 경우는 22 크기의 상대적으로 큰 변화폭을 보임을 알 수 있다.

영상에 따라서 히스토그램 분포의 변화폭이 크다는 것은 결국 피부색 영역을 정의하기가 어렵다는 것이고, 정의하더라도 좁게 정의하면 정의된 영역을 벗어나는 경우가 생기고, 넓게 정의하면 유사한 색의 영역까지 함께 검출된다는 것을 말한다. 따라서 얼굴피부색 검출을 위해 가장 적합한 성분은 밀집도와 향상성 면에서 모두 우수한 결과를 보인 H와 Cr 성분이다.

다음의 그림 3과 4는 피부색 검출을 위하여 기존의 HS와 CbCr 성분을 이용한 것과 HCr 성분을 이용한 결과 영상이다. 여기서 HS와 CbCr, 그리고 HCr의 각 성분들의 피부색에 해당하는 범위는 실험에 의한 통계치를 근거로 설정했다. 실험은 다양한(극단적이지 않은) 조명조건에서 촬영된 200장의 영상을 토대로, 앞의 그림 1과 2에서 보여 지는 눈과 입 사이의 샘플과 같이 피부영역을 제한하여 실시했다. 그리고 실험을 통해 얻은 각 성분의 중심(평균)값에서 ± 10 을 해주어서, 피부

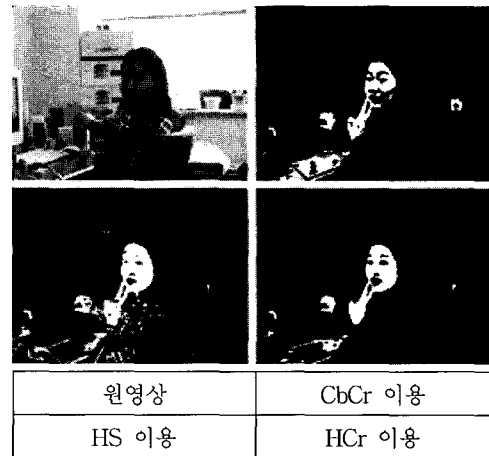


그림 4. 피부색 검출 영상 II
Fig. 4. Skin color detection II.

색 범위의 전체폭을 20으로 고정하여 검출한 결과이다.

앞의 그림 3을 보면 YCbCr 컬러공간의 CbCr이나 HSI 컬러공간의 HS 성분을 이용한 결과 영상에서는 원목책꽂이, 옷 또는 기타 배경부분이 피부색으로 검출됨을 볼 수 있다. 그러나 HCr을 이용한 결과영상은 얼굴과 손 부분은 확실하게 검출되면서도 다른 부분은 적게 검출되었고, 검출된 픽셀들의 응집정도도 우수함을 볼 수 있다.

그림 4는 창문을 통한 빛의 유입으로 역광효과가 있고 턱 부분이 어두운 영상에서의 결과이다.

그림 4의 CbCr을 이용한 영상을 보면 볼 아랫부분과 턱 부분의 어두운 영역이 제대로 검출되지 않았고, HS를 이용한 영상에서는 옷을 비롯한 주변 영역에서 불필요한 영역이 많이 검출된 반면에, HCr을 이용한 검출 영상은 얼굴과 두 손 외에 배경영역에서의 검출 영역이 현저하게 적음을 알 수 있다.

나. 적응적 임계영역 설정.

HCr 컬러공간이 다른 컬러공간에 비해 밀집도와 향상성 면에서 우수하지만 절대적이지는 않으며 사람에 따라 피부색의 개인차도 존재한다. 그러므로 고정된 임계영역을 이용하기 보단 입력된 개별영상에 따라 적응적으로 임계영역을 설정해 줄 필요가 있다.

적응적인 임계영역의 설정을 위해서, 우선 실험을 통해 구해진 H와 Cr 성분에 대한 고정 임계영역을 적용하여 조건에 만족하는 픽셀들만을 카운트하여 각 성분의 평균을 구한다. 이때 고정 임계영역의 범위를 정상적인 임계영역의 1/2정도로 엄격하게 적용한다.

위의 과정에서 구해진 평균값이 최종 임계 영역의 중

심 값이 되며 카운트한 픽셀 수에 따라서 임계 영역의 폭을 조정한다. 카운트한 픽셀 수가 9600보다 크면 식(1)을 이용하고, 9600 보다 작으면 식(2)를 이용해 최종 임계영역 값인 tH_{min} , tH_{max} , tCr_{min} , tCr_{max} 를 설정한다.

$$\begin{aligned} tH_{min} &= H_{min} + 2 \\ tH_{max} &= H_{max} - 2 \\ tCr_{min} &= Cr_{min} + 2 \\ tCr_{max} &= Cr_{max} - 2 \end{aligned} \tag{1}$$

$$W = \begin{cases} 15 & \text{if } Count < 480 \\ 10 & \text{if } 480 \leq Count < 9600 \end{cases}$$

$$\begin{aligned} tH_{min} &<- H_{ave} - W \\ tH_{max} &<- H_{ave} + W \\ tCr_{min} &<- Cr_{ave} - W \\ tCr_{max} &<- Cr_{ave} + W \end{aligned} \tag{2}$$

식(2)에서, W 는 임계영역의 폭을 결정하는 변수이다. 앞의 그림. 1에 해당하는 실험에서, 일반적으로 H 와 Cr 의 폭이 20 정도이면 대부분의 피부영역 픽셀이 검출 가능했으며, 조명이나 피부색의 개인차 등에 의한 변수에도 최대 30크기의 범위 폭이면 유효한 검출이 이루어졌다. 이러한 실험 결과를 토대로 식(2)의 10또는15라는 W 의 상수 값이 결정되었다. 480 이란 수치는 본 논문에서 얼굴영역을 찾기 위한 최소 제한 사이즈인 960 픽셀의 1/2에 해당하는 수치이고 9600 이란 수치는 320*240 크기 영상의 1/8 에 해당하는 크기이다. 1차 성분 조사를 위한 임계영역을 정상적인 임계영역 폭의 절반에 해당하도록 적용하였음을 감안할 때, 위에서의 두 수치는 결국 최소 크기와 최대 크기의 의미를 가진다.

식(1)의 경우, 카운트한 픽셀 수가 최대 크기를 넘어서는 값을 가지는 경우는 피부색이 작은 범위 내에 밀집해 있거나, 또는 피부 이외의 옷 등의 색이 피부색과 유사하여 함께 검출되었을 가능성이 크고, 이런 경우 오히려 1차 성분 조사에 의해 구해진 H_{ave} 와 Cr_{ave} 값이 왜곡되었을 수 있다. 따라서 1차 성분 조사를 무시하고 고정 임계영역을 사용하되 폭을 정상적인 임계 영역보다 보다 좁게(엄격하게) 설정한다.

다. HCr Map과 단계적 이진화.

앞에서 구한 tH_{min} , tH_{max} , tCr_{min} , tCr_{max} 값을 이용

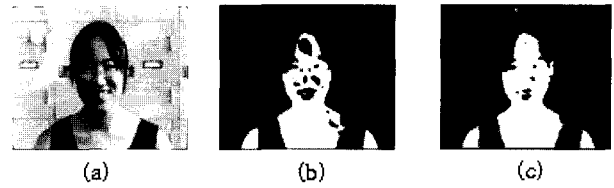


그림 5. 이진화 결과영상

(a) 원영상 (b) 임계값을 적용하여 한번에 이진화한 영상 (c) 단계적 이진화한 영상

Fig. 5. Results of Binary image.

(a) Original image (b) Binary image using fixed threshold (c) Binary image using adaptive threshold

하여 HCr Map을 구성한다.

H 값과 Cr 값은 임계영역에 해당하면 255, 임계영역을 벗어나는 정도가 ± 5 범위 안이면 128, 그 외는 0으로 설정한다. 미세한 차이로 영역을 벗어나는 픽셀들을 이후의 필터링 과정에서 다시 한번 판단이 이루어지게 하기 위해서 255와 0외에 중간 값을 갖도록 했다. 그리고 H 와 Cr 값을 단순히 and 연산을 수행하면 제거되는 픽셀이 많아지고, or 연산을 수행하면 검출 되는 픽셀이 과다해 질 수 있다. 따라서 이러한 경우 단순 and 또는 or 연산을 하지 않고 다음 식에 따라 값을 설정한다.

$$HCr_i = \begin{cases} 255 & \text{if } H_i == 255 \text{ and } Cr_i == 255 \\ 128 & \text{if } (H_i == 255 \text{ and } Cr_i == 128) \text{ or } \\ & (H_i == 128 \text{ and } Cr_i == 255) \\ 64 & \text{if } H_i == 128 \text{ and } Cr_i == 128 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

4가지 계층적 값을 가진 HCr Map 은 다시 필터링 과정을 통해 단계적으로 이진화 된다. 필터링 과정은 5*5 윈도우안의 25픽셀의 합을 기준으로 판단하여 값을 이진화 함과 동시에 영상의 경계 부분을 완만하게 하고 얇은 돌출 부위를 제거하기 위한 Opening 기능도 수행하게 된다.

그림 5는 임계값에 따라 한번에 이진화한 경우와 4단 HCr Map 구성 후 단계적으로 이진화한 경우의 결과 영상이다.

라. 입과 눈 검출

이진영상이 만들어 지면 레이블링(labeling)을 수행한다. 레이블링에 의해 검출 영역에 대한 정보를 얻게 되면 얼굴 후보 영역을 정한다. 960픽셀보다 작은 영역과 가로와 세로 크기가 최소 30픽셀보다 작은 영역은 얼굴 후보에서 제거한다. 얼굴 영역의 가로와 세로의 길이가 어느 한쪽이 일방적으로 크지 않은 점과, 앞에서의 얼

굴 최소크기가 960임을 감안해보면 가로나 세로 어느 쪽도 최소 30픽셀 이상의 크기를 가져야 한다는 추론이 가능하기 때문이다.

위의 과정을 거쳐 일단 얼굴 후보 영역이 정해지면 볼록 껍질(convex hull)를 이용해 얼굴후보영역에 대한 마스크(mask)를 구한다. 그리고 그 마스크영역 안에서 얼굴의 가장 큰 특징점인 눈과 입을 찾아 얼굴임을 확정한다. 얼굴후보 영역 안에서 눈과 입을 검출하기 위해서 얼굴영역에서의 눈과 입의 컬러정보와 밝기정보 그리고 위치 및 거리 정보를 이용한다.

(1) 컬러와 밝기 차이를 이용한 검출

얼굴영역에서 눈과 입을 제외한 피부 영역은 G성분이 B성분에 비해 절대적으로 크다. 그러나 눈과 입은 G와 B성분 간의 값의 차이가 작다. 따라서 얼굴후보영역에서 G와 B성분의 차이 값을 구하면 피부영역은 큰 값을 갖게 되고 눈과 입은 확연히 작은 값을 갖는다. G와 B성분간의 차이 값을 구하기도 용이하기 때문에 얼굴 피부영역으로부터 눈과 입을 구분하는데 효과적이다.

얼굴피부 영역으로부터 눈과 입이 구분 되면, 다시 눈과 입을 구분하기 위해서 Cr성분 값을 이용한다. Cr 성분은 조명의 영향에서 비교적 자유로우면서, 붉은 성분의 강도를 나타내기 때문에 눈과 입의 색 차이를 구분하기에 효과적이다. 또한 이전 과정에서 이미 값을 구했으므로 추가 연산이 필요하지 않다. 그리고 눈 검출의 정확성을 높이기 위해서 밝기 차이도 함께 이용한다. 입을 검출하고 눈의 후보 개체를 검출하는 과정은 다음과 같다.

1) 얼굴 후보 영역에 대해 밝기 값(이하 I), G와 B의 차이값(이하 GB), Cr 값을 각각 구한다.

2) 얼굴후보 영역에 한해서 Th_I와 Th_mouth GB, Th_eye GB, Th_mouth Cr, Th_eyeCr을 다음과 같이 구한다.

- Th_I = 밝기 값의 하위 50%를 구분 짓는 경계 값
- Th_mouthGB = GB 값의 하위 10%를 구분 짓는 경계 값
- Th_eyeGB = GB 값의 하위 5%를 구분 짓는 경계 값
- Th_mouthCr = Cr 값의 상위 10%를 구분 짓는 경계 값
- Th_eyeCr = Cr 값의 하위 10%를 구분 짓는 경계 값

3) 입을 검출하기 위해 식(4)을 이용하여 M_Map을

구한다.

$$M_Map_i = \begin{cases} 255 & \text{if } GB_i < Th_mouthGB \text{ and} \\ & Cr_i \geq Th_mouthCr \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

4) 3)에서 구한 M_Map에 대해 레이블링을 수행하여 개체 정보를 구한 뒤 가장 큰 개체를 입으로 검출한다.

5) 입이 검출된 얼굴후보 영역에 한하여 식(5)을 이용하여 E_Map을 구한다.

$$E_Map_i = \begin{cases} 255 & \text{if } I_i < Th_I \text{ and } GB_i < Th_eyeGB \\ & \text{and } Cr_i \leq Th_eyeCr \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

6) 5)에서 구한 E_Map에 대해 레이블링을 수행하여 검출된 개체수가 2보다 작으면 Th_eyeGB와 Th_eyeCr을 다음과 같이 다시 설정하여 5)의 과정부터 다시 수행한다.

- Th_eyeGB = Th_mouthGB
- Th_eyeCr = Th_mouthCr

위의 식에서 I_i 는 해당 픽셀의 밝기 값이고, GB_i 와 Cr_i 는 각각 해당 픽셀의 GB값과 Cr 값이다. 6)번은 처음에 임계값을 엄격하게 적용하여 개체를 검출하고 만약 개체검출 또는 눈 검출에 실패하면 임계값을 1차에서 보다 덜 엄격하게 적용하여 한번 더 눈 검출 과정을 수행하기 위한 것이다.

조명에 의해 얼굴의 일부분이 하이라이트 되면 본래 피부영역의 컬러특성이 왜곡되게 된다. 즉 GB 값이나 Cr 값이 원래 피부영역의 값과 다르게 나온다. 이것을 방지하기 위해 밝기 값의 하위 50%라는 Th_I의 조건을 주었다.

(2) 위치와 거리관계를 이용한 검증

앞에서 컬러와 밝기 차이를 통해서 입을 검출하고 눈의 후보 개체들을 검출했다. 그 눈 후보 개체들에 대해서, 이미 검출한 입의 위치정보를 기준으로 하여 눈 후보 개체와의 위치관계 및 거리정보를 이용하여 두 눈을 검출한다.

사람의 두 눈과 입이 이루는 일반적인 특징은 다음과 같다.

- 1) 입의 중심에서 왼쪽 눈과 오른쪽 눈까지의 거리가 동일하거나 거의 비슷하다.
- 2) 두 눈의 중간지점에서 입까지의 거리가 두 눈 사이의 거리보다는 크고 두 눈 사이의 거리의 2배 보다는

작다.

3) 두 눈을 잇는 선분과 입의 수평선과는 평행을 이루거나 평행에 가깝다.

이상 3개의 조건을 기준으로 하여 다수의 눈 후보 개체 중에서 가장 부합되는 개체를 눈으로 검출한다.

그림 6은 사람 얼굴에서 두 눈과 입이 이루는 위치 관계를 간략하게 표현한 것이다. 필수조건에서 2)번 조건은 코 구멍이 두 눈으로 잘못 검출 되는 경우를 방지하기 위한 것이고 4)번 조건은 두 눈을 잇는 수평선과 입 수평선의 평행정도를 약식으로 판단하기 위한 조건이다. 실험과정에서 한쪽 눈과 귀 부분의 안경다리 부분이 두 눈으로 잘못 검출 되는 경우가 있었는데 이 조건을 통해 방지할 수 있었다. 그리고 거리를 비교할 때 0.8과 1.2의 오차 폭을 준 이유는 카메라의 방향이나 얼굴의 회전 등으로 실제 거리와 조금씩 차이 나게 영상에 잡히기 때문이다.

다음의 조건에서 필수 조건은 반드시 만족해야만 하는 조건이고, 선택조건은 필수 조건을 만족하는 개체가 여러 개 존재 할 때 선택을 위한 조건이다. 위의 선택 조건에서 값이 작게 나올수록 선분 a와 b는 수직에 가깝고 이것은 사람의 두 눈과 입이 이루는 일반적인 특징 사항이다. 이상의 과정을 통해서 입과 눈을 검출하게 되면 그 검출된 영역을 얼굴 영역으로 최종 확정하게 된다.

마. 움직임 영역 정보

얼굴은 배경이 아닌 움직임 영역에 존재하므로 입력 되는 영상의 전 영역에 대해서 피부색을 검출하지 않고 움직임 영역에 한해서만 피부색을 검출하면 수행시간을 크게 단축할 수 있을 뿐만 아니라 배경에서의 오검출 가능성을 낮출 수 있다.

일반적으로 사용되는 움직임 영역 검출 방법은 차영상방법 (reference image)과 배경참조영상방법 (back-ground model)이 있다. 각각 장단점이 있는데, 본 논문에서는 순간적으로 움직임이 없거나 미세한 움직임에도 영역을 검출할 수 있는 배경 참조 영상 방법을 이용하여 움직임 영역을 검출한다.

배경참조영상 방법을 이용하기 위해서는 우선 움직임 물체가 없는 고정 배경 환경을 참조영상으로 만들어야 한다. 고정 배경이라 하더라도 조명의 미세한 변화나 카메라 잡음이 있을 수 있기 때문에 여러 장의 영상을 입력받아 평균한 값을 이용한다.

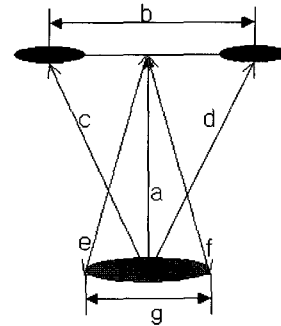


그림 6. 눈과 입의 위치관계
Fig. 6. Relations between eyes and mouth.

[필수 조건]

- 1) $a > b$ and $a < 2b$
- 2) $b > g$
- 3) $c > d*0.8$ and $c < d*1.2$
- 4) $e > f*0.8$ and $e < f*1.2$

[선택 조건]

$$\text{Min}(| a^2 + (b/2)^2 - c^2 |)$$

배경참조영상과 차후의 입력영상의 차 연산을 픽셀 단위로 처리하게 되면 조명이나 카메라의 잡음에 민감하게 되고, 계산양도 많아지게 된다. 이 문제를 해결하기 위해 전체 영상을 5*5 크기의 블록으로 분할하여 블록 단위로 처리를 수행한다. 블록을 대표하는 값은 5*5 윈도우 안의 25개 픽셀들의 평균값으로 한다.

차연산을 수행할 때 밝기 값만으로 움직임 영역 여부를 판단할 경우 간혹 현재 영상의 움직임 영역에 해당하는 블록의 값이 배경참조영상(이후 참조맵이라 한다)에서의 배경영역에 해당하는 블록 값과 비슷한 경우가 있다. 이런 경우 움직임 영역을 부분적으로 검출하지 못하게 되는데, 이를 방지하고자 밝기 값(Intensity)과 Cr 성분 값으로 하는 각각의 참조맵을 만든다. 전경체와 배경이 같은 밝기 값을 가지는 경우에 같은 컬러이기 보다는 다른 컬러일 가능성이 크기 때문에 두 성분 에 의한 움직임 영역 추정은 서로 상호 보완적이다.

참조맵이 만들어진 이후 입력되는 영상들과의 차 연산을 수행한다. 차 연산을 수행하여 블록간의 차이 값이 임계치를 넘으면 움직임 블록으로 그렇지 않으면 배경 영역으로 구분한다. 배경영역 블록은 참조맵을 갱신하는데 이용한다. 참조맵을 갱신하는 다양한 방법들이 존재하지만 처리시간에 대한 부담을 줄이기 위해 평균을 내는 방법을 이용한다. 참조맵 갱신을 통해 조명의

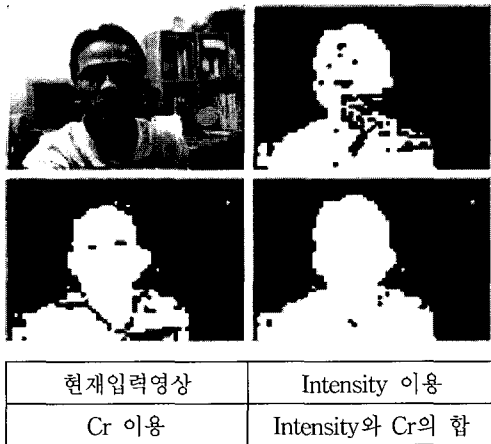


그림 7. 움직임 영역 검출
Fig. 7. Moving area detection.

점진적 변화에 의한 움직임 영역 오검출을 방지할 수 있다

밝기와 Cr 성분 각각에 대해 움직임을 구하고 논리합(OR 연산)을 통해 최종적인 움직임을 검출한다. 그림 7은 그 결과를 보여준다.

III. 실험

1. 실험 환경

조명의 변화가 심하지 않은 일반적인 연구실과 같은 실내에서, 소니 DSC-F717 디지털카메라에 의해 320*240 크기의 동영상상을 녹화한 뒤 같은 사이즈의 24bit 컬러 BMP영상으로 바꾸어 실험하였다.

VC++ 기반에서 구현하였으며, 배경영상을 초기화 하는 과정은 실행 시간에 포함시키지 않았다.

얼굴을 검증하기 위해서 눈을 검출하기 때문에 영상의 해상도와 영상 내에서 얼굴영역의 크기 제한을 가지지만 눈과 입은 얼굴인식이나 사용자 인터페이스 관련 응용에서 유용하게 사용되는 정보가 되므로 충분히 의미 있다고 판단된다.

2. 프로그램 실행결과.

본 실험에서 테스트영상 1, 2, 3 각 40개의 연속 영상, 즉 총 120개의 영상 중 107개의 영상에서 얼굴을 검출하였다. 따라서 성공률은 약 89%이며 소요시간은 아래 표와 같다.

실험에서 얼굴검출에 실패한 경우는 얼굴후보영역을 구한 뒤 얼굴검증 단계에서 눈을 감고 있거나 안경 렌즈가 조명에 하이라이트 되어있는 경우 또는 한쪽 눈

표 3. 프로그램 수행 소요시간
Table 3. Processing time.

시스템 사양	소요시간 (seconds/frames)
CPU : 550MHz RAM : 256M	8.89/120
CPU : 1.8GHz RAM : 512M	2.71/120
CPU : 2.4GHz RAM : 512M	1.86/120



그림 8. 피부색 검출 및 얼굴 검증
Fig. 8. Skin color detection and face verification.

영역이 안경의 검은 테 영역과 합쳐져 있어서 거리 및 위치관계를 이용한 조건검증에서 제한 조건을 벗어나는 경우 등이다. 따라서 눈 검출 또는 얼굴검증 방법을 보완하면 검출 성공률을 더욱 높일 수 있을 것으로 생각된다.

3. 얼굴검출 결과 영상.

얼굴검출 결과영상은 우선 정지영상에서 컬러를 이용한 피부색 검출 후 입과 눈을 검출하여 얼굴영역을 확정하는 과정과 이 방법을 이용해서 얻은 여러 결과영상을 보인다. 그리고 연속 영상에서 움직임 영역 검출 방법과 결합한 전체 알고리즘에 따른 단계별 결과영상과 연속 영상에서의 실험결과를 보인다.



그림 9. 다양한 영상에서의 얼굴검출 결과
Fig. 9. Face detection result of various images.

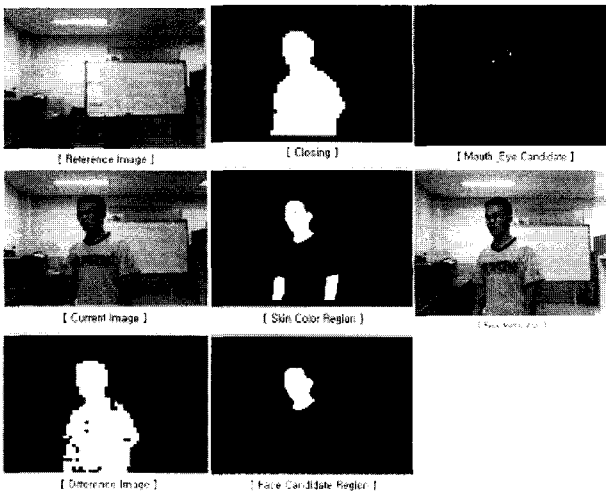


그림 10. 전체과정에 대한 결과영상
(a) 배경참조영상, (b) 현재 입력영상
(c) (a)와 (b)의 차영상, (d) 움직임영역
(e) 피부색 검출 영상, (f) 얼굴후보영역 마스크
(g) 눈과 입 후보 검출, (h) 최종결과영상

Fig. 10. Whole process for face detection.
(a) Reference image (b) Current image
(c) Difference image between (a) and (b)
(d) Moving area e) Skin color detected image
(f) Face mask for candidate (g) Detection of eyes and mouth (h) Final result for face detection

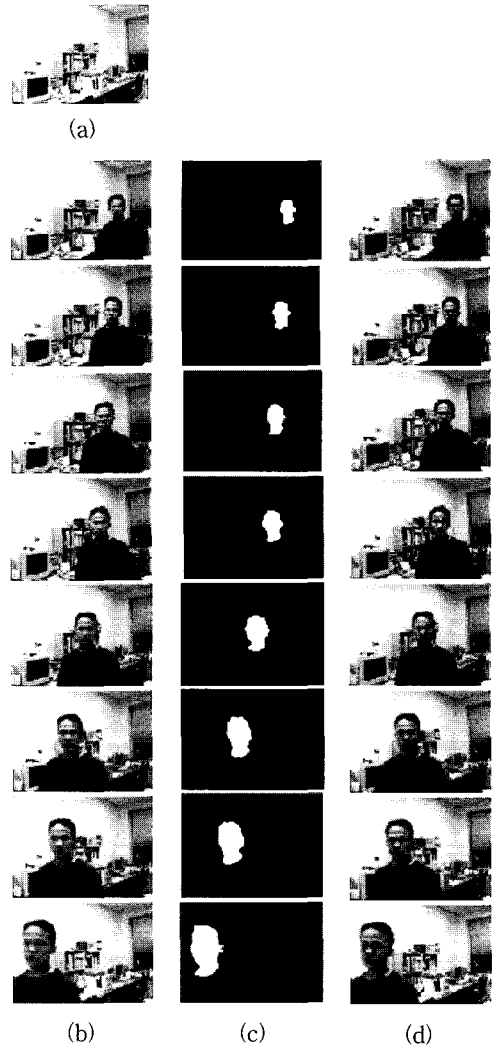


그림 11. 연속영상에서의 결과영상
(a)배경참조영상 (b)현재 입력영상
(c)얼굴후보영역 마스크 (d)얼굴검출

Fig. 11. Face detections for video sequences.
(a)Reference image (b)Current image
(c)Face mask for candidate (d)Face detection

IV. 결 론

본 논문에서는 연속된 입력영상에서의 움직임 정보와 컬러정보를 이용하여, 극단적이지 않은 조명환경에서 빠른 처리시간을 요하는 응용시스템에 이용 가능한 얼굴검출 시스템을 제안하였다.

피부색 검출에 있어서 조명의 변화에 보다 강인하고 피부색 특성에 알맞은 HCr 컬러공간을 사용하여 기존의 일반적인 컬러공간들을 사용할 때 보다 우수한 피부색 검출 결과를 얻었다.

HCr 컬러공간에서 영상에 따라 적응적으로 임계값을 설정하는 방법과 단계적인 이진화 방법을 통해 얼굴후

보 영역을 보다 안정적으로 획득할 수 있었다. 그리고 획득된 얼굴후보 영역에서 얼굴 특징 요소인 눈과 입을 검출함에 있어서 밝기 값, Cr 성분값, 그리고 G와 B성분의 차이 값을 함께 이용하여 얼굴피부영역에서 눈과 입을 보다 효과적으로 구분해 낼 수 있었다. 그리고 입과 눈의 위치 및 거리 관계를 이용하여 복잡하지 않은 방법으로 검증함으로써 처리시간의 부담 없이 눈과 입을 검출하였다.

또한 배경참조영상 기반에서 5*5 블록단위로 영상을 재구성하고, 밝기(Intensity)와 Cr 값을 함께 이용하여 움직임 영역을 검출함으로써 안정성을 높였다. 배경영역은 참조영상을 갱신하고 움직임 영역에 한해서 피부색을 검출하여 보다 효율적으로 피부색을 검출할 수 있었다.

본 논문에서 제안한 방법은 전체적으로 계산복잡도가 크지 않은 움직임 정보와 컬러정보만을 이용하여 얼굴 검출을 수행함으로써, 계산에 소요되는 시간을 많이 줄일 수 있었다. 따라서 실시간을 요하는 응용시스템에 유용할 것이다.

그러나, 본 연구를 통해 제안된 방법은 여전히 조명의 영향에서 자유롭지 못하며 눈을 검출하기위해서 화면상에서의 얼굴영역의 크기 제한이나 고해상도의 영상을 필요로 하는 제약조건이 있다. 따라서 계산량이 크지 않으면서도 조명의 변화에 따른 컬러변화를 보정할 수 있는 전처리 방법에 대한 연구와, 얼굴영역의 크기 및 영상의 해상도에 제한을 받지 않는 얼굴검출 방법에 대한 연구가 요구된다.

참 고 문 헌

- [1] M. Yang, D. Kreigman, and N. Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 No. 1, pp.34-58, Jan. 2002.
- [2] K. Sung, T. Poggio "Example-Based Learning for View-Based Human Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20 No:1, pp. 39-51, Jan.1998.
- [3] Henry A. Rowley, Shumeet Baluja, Takeo Kanade, "Neural Network-Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, January 1998.
- [4] Raphael Feraud, Olivier J. Bernier, Jean-Emmanuel Viallet, Michel Collobert, "A Fast and Accurate Face Detector Based on Neural Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 1, January 2001.
- [5] E. Osuna, R.Freund, F. Girosi, "Training support vector machines: An Application to face detection", IEEE Proc. of Int. Conf. On Computer Vision and Pattern Recognition, June, 1997.
- [6] C. Garcia, G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis", IEEE Transactions on Multimedia, Vol.1, No.3, September 1999.
- [7] M.J. Jones and James M.Rehg, "Statistical Color Models with Application to Skin Detection", Cambridge Research Laboratory Technical Report Series, 1998.
- [8] L. Jorado, M. Perrone, J. P. Costeria, "Active Face and Feature Tracking", in Processings of the 10th International Conference on Image Analysis and Processing, Kobe, Japen, Sep. 1999.
- [9] K. Sobottka, I. Pitas, "Extraction of Facial Regions and Features Using Color and Shape Information", IEEE International Conference on Pattern Recognition Vol. 3, August 1996.
- [10] D. Chai, 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, June 1999.
- [11] Rein-Lien Hsu, Mohamed Abdel-Mottaleb, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp.696-706, May 2002.
- [12] Jeonghee Park, Jung Seo, Dongun An and Seongjong Chung, "Detection of Human Faces using Skin Color and Eyes", in Proceeding of IEEE International Conference on Multimedia and Expo, vol.1, pp.133-136, 2000.
- [13] J. Yang, A. Waibel, "A Real-Time Face Tracker", IEEE Workshop on Applications of Computer Vision, Sarasota, Florida, 1996.

저 자 소 개



신 승 주(학생회원)
2001년 세종대학교 컴퓨터공학과
학사 졸업.
2004년 세종대학교 전자공학과
석사 졸업.
<주관심분야: 얼굴인식, 컴퓨터비
전, 영상처리>



최 석 림(정회원)
1981년 서울대학교 전자공학과
학사 졸업
1983년 서울대학교 전자공학과
석사 졸업
1992년 Syracuse University
전자공학 박사 졸업
현재 세종대학교 전자공학과 부교수
<주관심분야: 영상통신, MPEG, Multimedia>

