

# 효율적인 화상회의 동영상 압축을 위한 블록기반 얼굴 검출 방식

정회원 김 기 주\*, 준회원 방 경 구\*\*, 문 정 미\*\*\*, 정회원 김 재 호\*\*\*

## A block-based face detection algorithm for the efficient video coding of a videophone

Ki Ju Kim\* *Regular Members*, Kyoung Gu bang\*\*, Jeong Mee Moon\*\*\*  
Jae Ho Kim\*\*\* *Regular Members*

### 요 약

비디오 화상회의 시스템을 위하여 동영상 압축 DCT 계수와 피부색정보를 이용하여 주파수 영역에서 정면 얼굴을 실시간으로 검출하는 알고리즘을 제안한다. 동영상 압축과정 중에 얻어지는 DCT계수값의 U와 V 색상정보로 피부색 범위를 추출하며 잡음성분제거를 위해 형태학적필터와 Labeling을 적용하고, 피부색블록의 방향성과 평탄도를 고려하여 검출하였다. 제안 알고리즘은, 배경에 피부색과 유사한 객체가 있는 경우와 배경이 단순한 경우에 대해 적응적으로 검출이 가능하도록 하였다 여러 인종에 대한 모의실험결과 제안 알고리즘이 약 94%의 검출 성공률을 보였다.

**Key Word** : Video coding, Face detection, Skin-color, DCT coefficient, Morphological filter

### ABSTRACT

We propose a new fast, algorithm which is used for detecting frontal face in the frequency domain based on human skin-color using DCT coefficient of dynamic image compression and skin color information. The region where each pixel has a value of skin-color were extracted from U and V value based on DCT coefficient obtained in the process of image compression using skin-color map in the Y, U, V color space. A morphological filter and labeling method are used to eliminate noise in the resulting image. We propose the algorithm to detect fastly human face that estimate the directional feature and variance of luminance block of human skin-color. Then Extraction of face was completed adaptively on both background have the object analogous to skin-color and background is simple in the proposed algorithm. The performance of face detection algorithm is illustrated by some simulation results carried out on various races. We confirmed that a success rate of 94% was achieved from the experimental results.

### 1. 서 론

최근에 정지 영상 및 동영상 압축 기술을 토대로 한 다양한 멀티미디어 응용 서비스들이 출현하고 있다. 그 중에서도 화상회의(Tele-Conference)나 비

디오폰(Video-Phone) 서비스는 대표적인 응용 분야로 인식되고 있다. 일반적으로 이들 응용 분야에 있어서 인간의 얼굴은 매우 중요한 객체(Object)로서 다루어진다. 따라서 열악한 통신 환경에서 보다 향상된 서비스를 제공하기 위해서는 얼굴 객체에 대

\* 구미기능대학 전자과(kkj@kopo.or.kr)

\*\*\* 부산대학교 전자공학과 (moonjmee@hanmail.net, jhkim@pusan.ac.kr)

논문번호 KICS2004-05-006, 접수일자 2004년 5월 11일

한 차별화된 처리가 요구된다. 이것은 객체 지향 압축 기법에 기반하여 얼굴 영역의 화질을 유지, 개선함으로써 해결될 수 있다. 그러나 이러한 방식은 무엇보다도 얼굴 영역에 대한 효율적인 인식과 검출을 전제로 하고 있다. 따라서 오늘날 동영상 압축 방식에 적합한 실시간 얼굴 검출 기법의 개발이 요구된다.

현재 디지털 영상에서 인간의 얼굴을 검출하기 위한 연구가 활발히 진행되고 있다. 지금까지 제시된 얼굴 영역 추출 방식들은 형판정합법(template matching method), 특징기반 접근법(feature-based approach), 신경회로망법(neural method), 그리고 피부색 기반 방식(skin color based approach)으로 분류할 수 있다[1]. 이 중에서 특징기반 접근법과 신경회로망법은 계산량이 많이 소모되어 동영상 압축 방식에 실시간으로 적용하기에 적합하지 않다. 그리고 형판정합법은 얼굴 모양의 변화에 대응이 곤란한 단점을 지니고 있어서 다양한 인종을 대상으로 하는 응용 분야에 부적합하다. 반면에 피부색 정보를 이용하여 얼굴 영역을 검출하는 피부색 기반 방식은 앞의 3가지 방식들에 비하여 구현이 간단하다. 또한 명암 성분과 색차 성분을 서로 분리하여 처리하는 기존의 동영상 압축 방식에서는 피부색 정보를 쉽게 추출, 활용할 수 있다. 이러한 사실을 고려할 때 따라서 피부색 기반 방식을 토대로 하여 동영상 압축 방식에 적합한 효율적인 얼굴 검출 방식을 개발하는 것이 타당할 것이다.

피부색 정보를 이용한 얼굴 검출 방식은 여러 컬러 모델을 기초로 제안되어 왔다. Yang[2]은 RGB 원 영상의 해상도를 감소시키면서 눈, 코, 입부분의 명암 특징을 이용하여 얼굴 영역을 추출하고 자 하였다. Herodotou[3]는 HSV 컬러 모델에서 피부색 구간의 스토그램을 측정하고 여기에 형태학적 연산자, 중간값 처리(median filtering), 퍼지 이론을 적용하여 얼굴 영역을 추출하였다. 그리고 Karin[4]는 HSV 컬러 모델에서 피부색 범위 안에 있는 화소들을 찾아내고 얼굴의 특징인 타원형이 생성되는 부분을 얼굴 영역으로 판단하였다. 그런데 이 방식들은 기존 동영상 압축 방식에서의 칼라 모델과는 다른 모델을 기반으로 있기 때문에 동영상 압축 방식에 적합하지 않다. 한편, Chai[5]는 화상회의용 영상에서 Y, Cb, Cr 모델을 사용하여 영상내에 후보 피부색 영역을 설정한 후 Y 성분에서의 표준편차를 구하여 최종 얼굴 영역을 검출하는 방식을 제안하였다. 제한된 조건하에서는 이 방식은 비교적 얼굴

영역을 잘 검출한다. 그러나 화면 내에서 얼굴 면적이 클 경우, 배경에 피부색이 많이 존재할 경우에 얼굴이 잘 검출되지 않는 단점을 지닌다. 또한 이 방식은 화소 단위 처리를 기본으로 하고 있기 때문에 실시간 처리를 수행하기에는 많은 계산량이 필요하고 아울러 블록 단위 처리를 기본으로 하는 동영상 압축 방식에 부적합한 문제를 지닌다. 따라서 Chai의 방식을 동영상 압축에 적용하기 위해서는 이러한 문제점들이 해결되어야 한다.

본 논문에서는 화상 회의 및 비디오폰 서비스에 있어서 객체 지향 압축 기법의 적용이 용이한 효율적인 블록 기반 얼굴 검출 방식을 제안한다. 제안 알고리즘은 DCT 계수값들을 이용한 블록 단위 얼굴 영역 검출을 기본으로 하고 있으며 검출의 효율성을 위하여 2단계로 구성된다. 색차 성분의 DCT 계수에 존재하는 피부색 정보를 이용하는 1단계에서는 비교적 단순한 영상들에 대한 효과적인 검출이 수행되어지며, 명암 성분에 대한 방향성, 평탄도를 이용하는 2단계에서는 복잡한 영상에서의 정확한 검출이 가능하다. 그리고 얼굴 후보 영역에 대한 모양 정보와 주파수 특성을 이용한 검증 과정이 각 단계에서 적용된다. 여러 인종에 대한 모의 실험 결과는 제안 방식이 적은 계산량으로 얼굴 영역을 효과적으로 검출함을 보여준다.

## II. 기존의 얼굴 영역 검출과 동영상 압축 방법의 적용

Chai[5]가 제안한 피부색 정보를 이용한 방식은 검출 가능한 얼굴의 크기가 48X48 이상이라는 전제 하에 화소 단위 처리를 통하여 정면 얼굴을 검출한다. 그림 1은 Chai가 제안한 얼굴 검출 방식에 대한 순서도이다.

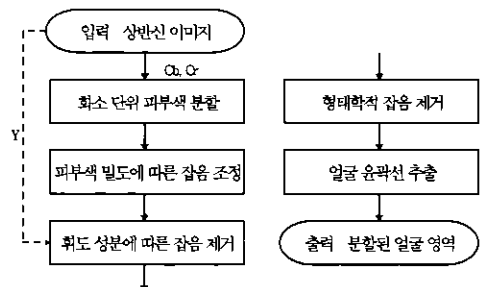


그림 1 기존 화소 단위 정면 얼굴 검출 방식의 순서도  
Fig 1 Flowchart of the conventional pel based face detection methods

그림 1에서 각 과정들은 다음과 같은 기능을 수행한다.

- 화소단위 피부색 분할 : Cb, Cr 프레임에서 피부색의 범위를 이용한 후보 얼굴영역 선정.
- 피부색 밀도에 따른 잡음 조정 : 4X4블록 단위로 순수 피부색 영역과 잡음 피부색 영역을 구분하고 팽창 및 침식 작용을 적용함으로써 잡음 피부색 영역을 제거.
- 휘도 성분에 따른 잡음 제거 : Y 프레임의 표준편차를 구하여 배경에 존재하는 피부색 영역 제거.
- 형태학적 잡음제거 : 얼굴의 형태학적 특성을 이용한 얼굴 영역의 객체를 추출.
- 얼굴윤곽선추출 : 블록단위의 얼굴객체 윤곽선을 화소단위로 매핑 시켜 윤곽선 추출.

이 방식은 평탄한 배경에서 얼굴의 크기가 크지 않을 경우에 검출이 잘 이루어진다. 그러나 얼굴의 크기가 점점 커져서 뺨, 이마 등 표준편차가 낮은 부분이 증가하게 될 경우 정확한 얼굴 검출이 이루어지지 못하고 배경에 얼굴색과 유사한 영역들이 존재할 경우 검출 오류가 발생하는 단점을 지니고 있다. 그림2는 기존 방식에 의하여 수행된 검출 결과이다. 그림 2의 (a)는 기존 방식에 의하여 검출이 양호하게 이루어짐을 보여준다. 그러나 그림 2의 (b),(c)에서는 검출 오류가 발생함을 확인할 수 있다. 따라서 기존 방식은 화상회의나 비디오 폰 서비스 분야에 기존 방식을 적용하기 위해서는 이러한 문제점에 대한 개선이 이루어져야 할 것이다.

일반적으로 동영상 압축 방식인 MPEG-4는 그림 3과 같이 영상을 압축하는 데 있어서 블록 단위 처리를 기본으로 하고 있다[6]. 즉, 시간적 중복성을 제거하기 위한 움직임 추정이나 공간적 중복성을 제거하기 위한 DCT 및 양자화는 일정한 크기의 블록 단위로 수행된다. 따라서 화소 단위 처리를 기반으로 하고 있는 기존의 검출 방식은 동영상 압축 방식에 적용되기에는 구조적으로 적합하지 않다. 또한 화소 단위 처리는 블록 단위 처리에 비하여 많은 계산량을 필요로 한다. 그런데 동영상 압축 방식을 이용하는 다양한 응용 분야들은 대부분 실시간 처리를 전제 조건으로 한다. 따라서 보다 고속의 얼굴 검출을 위해서도 블록 단위 처리가 요구되어진다. 이러한 사실들은 블록 단위 처리를 바탕으로 한

효율적인 얼굴 검출 방식이 개발되어야 함을 의미한다.

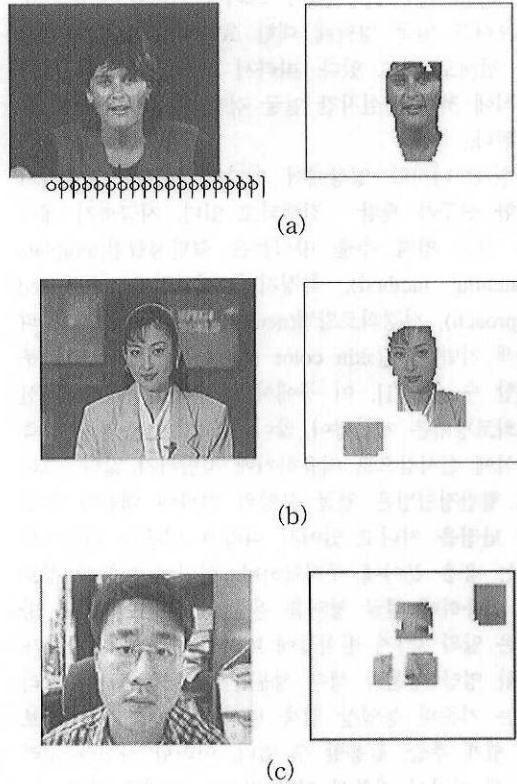


그림 2. 기존 방식에 의한 얼굴 검출 결과  
Fig.2. Results of the conventional detection method

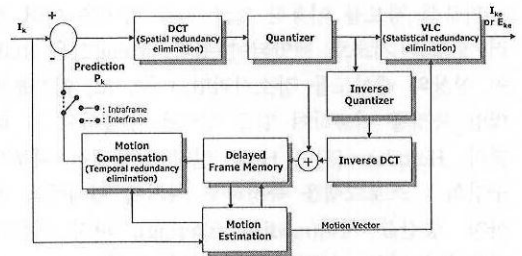


그림 3. MPEG-4 엔코더  
Fig.3 MPEG-4 encoder

### III. 제안 알고리즘

오늘날 보다 향상된 화상 회의나 비디오폰 서비스를 제공하기 위해서는 얼굴 객체에 대한 차별화된 처리를 통한 얼굴 영역의 화질을 유지, 개선하는 것이 필요하다. 이것은 동영상 압축 방식에 적합한 얼굴 검출 기법의 개발이 이루어질 때 가능한 일이다.

다. 그런데 이를 위해서는 새로운 검출 기법에 있어서 요구되어지는 전제 조건들이 무엇인지를 파악하는 것이 무엇보다도 중요하다. 기존의 얼굴 검출 방식[5]과 동영상 압축 방식을 분석, 고찰할 경우 아래의 사항들을 만족하는 얼굴 검출 방식이 개발되어야 함을 알 수 있다.

피부색 정보를 이용한 블록 기반 처리.

- 기존 압축 과정에서 생성되는 정보를 활용한 효율적인 처리.
- 고속 동작을 위한 불필요한 계산량 감소
- 기존 얼굴 검출 방식이 지니는 검출 오류 해결.

본 논문에서는 이상과 같은 조건들을 고려하여 동영상 압축 방식에 적용이 용이한 효율적인 얼굴 검출 알고리즘을 제안한다. 그림4 은 제안 알고리즘의 순서도이다.

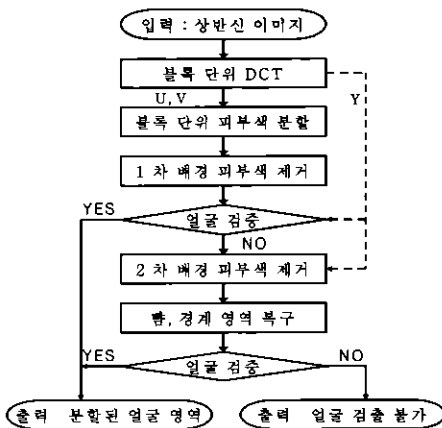


그림 4. 제안하는 얼굴 검출 방식의 순서도  
Fig.4 Flowchart of the proposed algorithm

그림 4에서 알 수 있듯이 제안 방식은 기본적으로 2과정으로 구성되며 기존 동영상 압축에서 얻어지는 DCT 계수값들에서 피부색 정보를 추출함으로써 얼굴 영역의 후보 블록들을 선정한다. 이것은 기존 동영상 압축에서 필수적으로 이용되는 DCT 과정을 얼굴 검출에 활용함으로써 블록 단위 처리를 수행하는 것과 동시에 불필요한 계산량을 감소시키는 장점을 지닌다. 그리고 얼굴 검출 단계를 통한 2 단계 검출 과정은 입력 영상의 복잡도에 따라 적응적으로 얼굴을 검출하여 불필요한 계산량을 감소시킴으로써 효율적인 처리를 수행하는 것을 목표로 한다. 이러한 특징을 지닌 제안 방식은 보다 효과적

인 얼굴 객체 코딩을 가능하게 할 것이다.

그림 4의 제안 방식에 있어서 각 단계별 상세한 처리 동작들은 다음과 같다.

### 1. 블록 단위 DCT 및 피부색 분할

동영상 압축 표준의 하나인 MPEG-4에서는 명암 성분을 나타내는 Y, 색차 성분을 나타내는 U, V 프레임은 입력으로 사용한다. 일반적으로 4:2:0 형식의 CIF 동영상에서 U, V는 22×18개의 블록으로 구성되어진다. 한편 기존의 동영상 압축에서 이 프레임들은 각각 8×8블록 단위로 DCT 된다. 따라서 기존 동영상 압축 과정에서 얻어지는 DCT 계수를 이용하여 얼굴 영역을 결정한다면 계산량 증가없이 블록 단위 처리가 가능해진다.

기존의 피부색 기반 얼굴 검출 방식들은 얼굴의 피부색에 대한 색차 성분의 값이 특정 범위내에 존재한다는 사실에 기인하고 있다. 그런데 이것은 각각의 화소값들에 대하여 적용되어지는 현상이다. 따라서 색차 성분의 DCT 계수들에 대하여 이러한 사실을 적용하기 위해서는 색차 성분의 DC 계수값에 대한 고찰이 요구된다. 즉, 얼굴 영역에 대한 색차 성분의 DC 계수값들이 화소값들의 경우와 같이 특정 범위를 구성하는 지를 살펴보아야 한다. 이를 위하여 본 논문에서는 다양한 인종에 대한 얼굴 영역에서의 색차 성분에 대한 DC 계수값들을 조사하였다. 그림 5는 황인종, 백인종, 흑인종 225샘플에서 얼굴 피부색에 대해 DCT변환된 8×8블록 DC값을 U, V 2차원 축에 표시한 것이다. 그림 5의 실험 결과로부터 인종에 관계없이 피부색이 일정 영역의 U, V 범위를 가진다는 특성을 확인 할 수 있다.

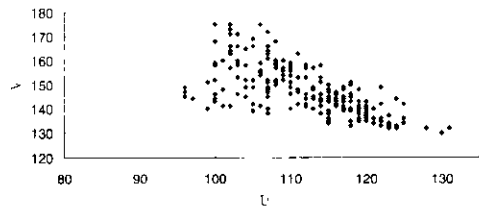


그림 5 얼굴에 대한 U, V성분의 DC분포  
Fig.5. DC distribution of the U and V components in the face area

본 연구에서는 이러한 사실을 토대로 2차원 좌표 상에서 DC값을 PLOT 하여 U, V값이 집중된 지역의 모양을 도형화해서 피부색 맵으로 정의한다. 그

림 6은 정의된 피부색 맵을 나타낸다.

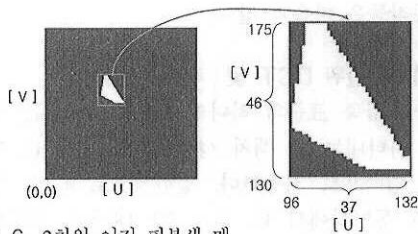


그림 6. 2차원 이진 피부색 맵  
Fig.6. 2D binary skin color map

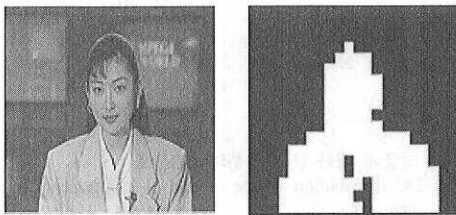
여기서 흰색으로 표현된 부분은 이진값 '1' 로서 피부색 영역임을 나타낸다. 그리고 가로, 세로축에 표시된 값은 U, V 성분에서의 DC 계수값의 범위이다.

이러한 사실을 토대로 제안 알고리즘에서는 기존 동영상 압축 과정에서 얻어지는 색차 성분의 DC 계수값을 그림 6의 피부색 맵에 적용함으로써 주어진 블록이 얼굴 영역에 속하는 지는 결정한다. 식 (1)은 이러한 과정을 수식화 한 것이다.

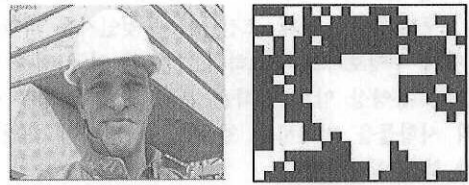
$$O_1(x,y) = \begin{cases} 1, & \text{if skin map}(UDC(x,y),VDC(x,y))=1 \\ 0, & \text{Otherwise} \end{cases}$$

(1)

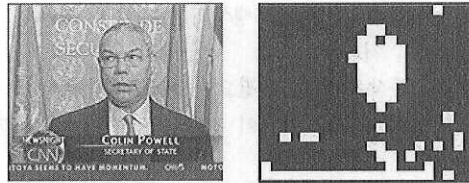
여기서  $O_1$ 은 피부색 분할 결과를 나타낸다. 그리고 UDC, VDC는 U, V 프레임에서의 각  $8 \times 8$ 화소의 평균값을 이다. 그림 7은 여러 인종에 대해 식 (1)과 같이 블록 단위로 피부색을 분할한 결과이다. 그림 7은 얼굴 영역에 속하는 블록들이 비교적 잘 검출됨을 보여준다. 그러나 얼굴에 속하지 않는 많은 블록들이 얼굴 영역 블록으로 존재함을 알 수 있다.



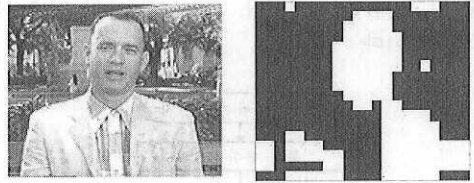
(a)



(b)



(c)



(d)

그림 7. 다양한 입력 영상에 대한 피부색 분할 ( $O_1$ )  
Fig.7 The skin color segmentation for the various Input image

## 2. 1 차 배경 피부색 제거

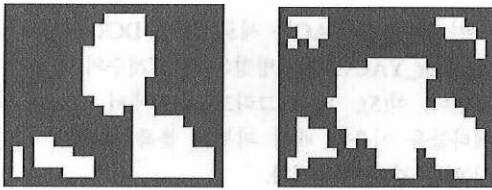
그림 7에서 볼 수 있듯이 단순한 피부색 분할은 조명 상태에 의한 잡음 및 배경 영역에 나타나는 피부색 블록을 얼굴 영역으로 결정한다. 그러나 이러한 잡음 피부색 블록들은 올바른 얼굴 검출을 위해서는 제거되어야 한다. 그리고 블록 내에서 눈, 눈썹, 입술 부위 등이 차지하는 면적이 클 경우 피부색 블록으로 검출되지 않는 경우가 발생한다. 이러한 블록은 다시 피부색 블록으로 환원시켜주어야 한다.

이러한 처리를 위해 본 논문에서는 먼저 화면 가장자리에 존재하는 피부색 블록을 잡음 블록으로 판단하여 제거한 후 식(2)와 식(3)의 침식 및 팽창 과정을 수행한다.

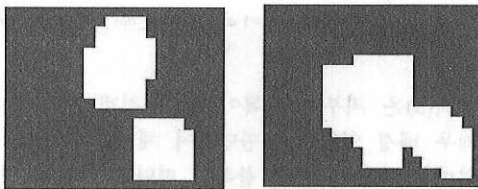
$$E(x,y) = \begin{cases} 1, & \text{if } \sum_{y=y-1}^{y+1} \sum_{x=x-1}^{x+1} O_1(x,y) - 1 > 2 \text{ and } O_1(x,y) = 1 \\ 0, & \text{Otherwise} \end{cases}$$

$$O_2(x,y) = \begin{cases} 1, & \text{if } \sum_{y=y-1}^{y+1} \sum_{x=x-1}^{x+1} O_1(x,y) > 4 \text{ and } E(x,y) = 0 \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

여기서 E는 침식작용의 결과를 O2는 팽창작용의 결과를 나타낸다. 그리고 얼굴의 형태학적인 특성 즉, 얼굴의 가로 대 세로 길이의 비율을 이용하여 필터링 작업을 수행한다. 본 논문에서는 얼굴의 가로 대 세로 비가 1:1.5이라는 실험적 사실을 바탕으로 검출의 최소 단위로 3×4 블록으로 가정하였다. 그리하여 제안 방식에서는 가로 방향으로 피부색 블록이 연속적으로 3개 미만인 블록과 세로 방향으로 피부색 블록이 4개 미만인 블록이 제거된다. 그림 8은 그림 7의 (b)와 (d)에 대하여 이 같은 과정을 적용하여 얻은 결과이다. 필터링 후 4방향 주변 성분 연결 분류화 알고리즘(4 neighborhood connected labeling algorithm)을 이용하여 후보 피부색 영역을 분할한다.



(O<sub>2</sub>) 침식 및 팽창



(O<sub>3</sub>) 형태학적 필터링

그림 8. 1차 배경 피부색 제거

Fig.8. The first removal of the skin color in the background

### 3. 얼굴 검출

앞에서 설명한 1차 배경 피부색 제거 단계를 적용할 경우 대부분의 단순한 영상들에서는 얼굴 영역이 잘 검출된다. 그러나 몇몇 영상들에서는 효과적인 검출이 이루어지지 않는다. 이러한 영상들을 분석해 보면 다음의 2가지 경우들로 구분할 수 있

다.

- 배경에 피부색이 폭넓게 존재하여 얼굴과 배경 영역이 잘 분리되지 못하는 경우
- 얼굴 외의 피부색 블록이 작게 존재하여 실제 얼굴 영역과 분리된 배경 영역이 얼굴영역으로 경우

이러한 영상들에 대한 효과적인 처리 방법은 후보 객체에 대하여 얼굴 검증 과정을 적용하여 얼굴로 검증되지 않은 경우에 대하여 보다 정밀한 과정을 적용하는 것이다. 이를 위해서는 무엇보다도 얼굴 검증 방식이 필요하다.

본 연구에서 제안하는 얼굴 검증 과정은 아래와 같이 크게 3가지 요소로 구성되며 주어진 객체가 3가지 요소를 모두 만족할 경우 얼굴 객체로 최종 결정한다.

- 후보 영역에 대한 외접 사각형의 가로와 세로의 비율
- 외접 사각형 내 피부색 블록의 비율
- 매크로블록 밝기 값의 평탄도

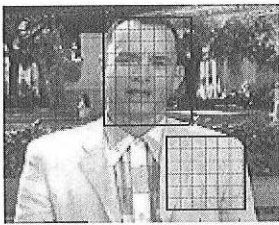
먼저 그림 9의 (a)와 같이 후보 영역의 외접 사각형들을 구한 다음 식(4)과 같이 후보 객체의 가로와 세로의 비율을 계산한다. 그리고 구해진 비율이 0.4와 1사이이면 얼굴로 인정한다.

$$Ratio = \frac{horizontal\_block\_number}{vertical\_block\_number} \quad (4)$$

여기서 horizontal\_block\_number는 외접 사각형의 가로 블록의 수를 vertical\_block\_number는 세로 블록의 수를 나타낸다. 다음 외접 사각형 내에 존재하는 피부색 블록의 개수를 조사하여 외접 사각형내의 전체 블록 개수에 대하여 80%이상일 경우 얼굴로 인정한다. 그리고 세 번째는 후보 영역들 중에서 얼굴이 아닌 부분에서의 밝기값은 균일하다는 특성을 이용한다. 즉, 후보 영역들중 얼굴 부분의 밝기 값 변화가 상대적으로 크다는 것이다. 이것은 외접 사각형 내에 존재하는 매크로 블록들에 대하여 Y성분의 DCT 계수들을 이용함으로써 예측할 수 있다. 본 논문에서는 이러한 사실을 바탕으로 매크로 블록의 밝기 평탄도를 식(5)와 같이 정의한다.

$$sum\_abs\_AC\_coef = \sum_{k=1}^4 \left\{ \sum_{y=1}^7 \sum_{x=1}^7 |YB_k(x,y)| \right\} \quad (5)$$

여기서, k는 8x8 크기의 부 매크로 블록을 나타내며 YBK는 DCT를 거친 부 매크로 블록을 나타낸다. 고주파 성분의 분포가 얼마나 되는지를 나타내는 sum\_abs\_AC\_coef는 그 값이 '0'일 때 고주파 성분이 존재하지 않고 블록이 매우 평탄함을 의미한다. 일반적으로 외접 사각형 내에 존재하는 평탄 블록의 개수는 얼굴의 크기에 따라 다르게 분포한다. 본 논문에서는 실험적 분석을 토대로 외접 사각형 내의 평탄 블록 비율이 35%이하일 경우 얼굴로 인정한다. 그림 9의 (b)에서 빨간색으로 표시된 블록은 평탄 블록을 나타낸 것이다.



(a) 후보 영역의 외접 사각형



(b) 후보 영역내의 평탄 블록

그림 9. 얼굴 검증의 예  
Fig.9 The example of the face verification

#### 4. 2 차 배경 피부색 제거

앞에서 설명한 얼굴 검증 방식을 적용하더라도 얼굴 영역이 제대로 검출되지 않을 경우에는 보다 정확한 검출 과정을 수행하여야 할 것이다. 이를 위하여 본 논문에서는 다음과 같은 사실을 이용하여 정확한 얼굴 영역을 검출하고자 한다.

- 얼굴 내부 영역에는 가로 및 복합 경계 성분이

존재한다.

- 얼굴 영역의 평탄도는 일정한 범위를 가진다.

이러한 특징들은 일반적으로 Y 성분의 DCT 계수들로부터 쉽게 추정할 수 있다. 먼저 얼굴 평탄도가 일정한 범위를 지닌다는 것은 식(5)에서 정의된 sum\_abs\_AC\_coef 를 이용할 경우 쉽게 확인할 수 있다. 그리고 가로 및 복합 경계 성분의 존재 유무는 블록내 영상의 가로 및 세로 방향의 성분이 얼마나 존재하는가로 추정할 수 있다. 그런데 이것은 주어진 매크로 블록의 Y 성분들에 대한 DCT 계수들로 쉽게 알 수 있다. 본 논문에서는 주어진 블록의 가로, 세로 경계 성분들을 식(6)과 식(7)으로 정의한다.

$$Ver\_YAC = \sum_{k=1}^4 \left\{ \sum_{u=1}^7 |YB_k(l,u)| \right\} \quad (6)$$

$$Hor\_YAC = \sum_{k=1}^4 \left\{ \sum_{v=1}^7 |YB_k(v,l)| \right\} \quad (7)$$

여기서 Ver\_YAC는 세로방향의 DCT계수의 정도를, Hor\_YAC는 가로방향의 DCT계수의 정도를 나타낸다. 식(5), 식(6) 그리고 식(7)에서 정의된 파라미터들을 이용한 배경 피부색 블록 제거는 식(8)과 식(9)와 같이 수행된다.

$$O_4(x,y) = \begin{cases} 1, & \text{if } O_3(x,y)=1 \text{ and } Ver\_YAC/Hor\_YAC < 3 \\ 0, & \text{Otherwise} \end{cases} \quad (8)$$

$$O_5(x,y) = \begin{cases} 1, & \text{if } O_4(x,y)=1 \text{ and } 2 < sum\_abs\_AC\_coef < 450 \\ 0, & \text{Otherwise} \end{cases} \quad (9)$$

식(8)은 피부색 블록이 세로 경계 성분이 강할 경우 배경 영역으로 판단하여 제거하는 것을 의미하며 식(9)는 피부색 블록의 평탄도가 얼굴 영역의 평탄도 범위를 벗어날 경우 배경 영역으로 판단하여 제거함을 나타낸다. 그림 10은 식(8)과 식(9)에 의하여 수행된 배경 피부색 영역 제거 결과이다.

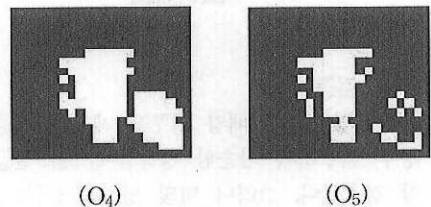


그림 10. 배경 피부색 영역 제거  
Fig.10. The second removal of the skin color in the background

한편 눈 주변의 블록들은 비교적 높은 분산값을 지니게 되어 식(8)과 식(9)의 과정에서 얼굴 영역이 아니라고 판정되어 제거될 수 있다. 이러한 블록들의 복구 및 잡음 블록을 제거를 위하여  $O_5$ 로 얻어진 결과에 앞에서 설명한 팽창 및 침식 작용과 형태학적 필터링을 적용한다. 그림 11은 팽창 및 침식의 결과  $O_6$ 와 형태학적 필터링을 거친  $O_7$ 의 결과를 보여준다.

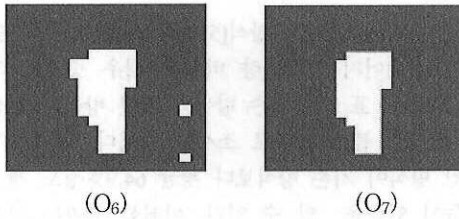


그림 11. 팽창, 침식 그리고 형태학적 필터링에 의한 검출영역

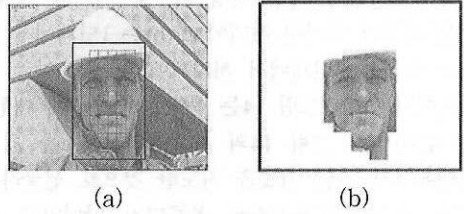
Fig.11. The detected region by the dilation, erosion, and morphological filtering

### 5. 얼굴 경계 영역 및 뺨 영역 복구

앞에서 설명한 과정들은 낮은 평탄도를 지니는 얼굴 영역 블록(뺨)과 높은 평탄도를 지니는 얼굴 영역 블록(얼굴 경계 부분)을 제거할 수 있다. 따라서 이러한 블록들은 복구되어야 한다. 이를 위하여 먼저 2차 배경 피부색 제거 단계에서 얻어진 후보 블록들을 포함하는 새로운 확장된 외접 사각형을 구한다. 그리고 이 외접 사각형 영역 내에서 식(10)을 적용하여 뺨 및 얼굴 경계 영역을 복구한다.

$$O_8(x,y) = \begin{cases} 1, & \text{if } \text{rect}(x,y)=0 \text{ and } O_7(x,y)=1 \\ & \text{and } (\text{sum\_abs\_AC\_coef}(x,y) \leq 2 \text{ or } \text{sum\_abs\_AC\_coef}(x,y) \geq 450) \\ 0, & \text{Otherwise} \end{cases} \quad (10)$$

여기서  $O_3$ 은 1단계 배경 피부색 제거 후의 결과를,  $\text{rect}(x,y)$ 는 확장된 외접 사각형을 나타내고 있다. 나타내고 있다. 즉 확장된 외접 사각형 내에서 피부색 블록을 다시 복구하는 것이다. 그림12는 복원된 얼굴 영역을 나타내고 있다.



(a) 확장된 새로운 외접 사각형 영역  
(b) 복원된 얼굴 영역

그림 12. 얼굴 경계영역 및 뺨 영역 복구(O8)  
Fig.12. Restoration of the face edge and chin region

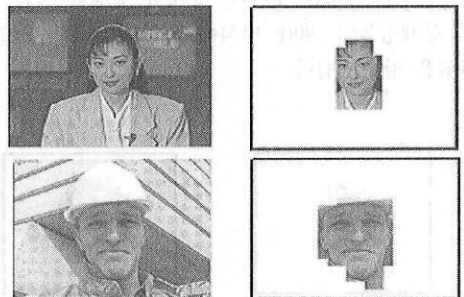
그리고 이렇게 복원된 결과는 앞에서 설명한 얼굴 검출 단계를 거쳐서 최종적으로 얼굴 검출을 완료한다.

## IV. 실험 결과 및 고찰

제안 방식의 효과를 검증하기 위하여 본 논문에서는 일반 CCD 카메라와 TV 캡처 영상으로 여러 인종에 대한 얼굴 영역의 검출을 조사하였다. 그림 13는 제안 알고리즘을 적용하여 얻어진 얼굴 검출 결과이다.



(a) 1 단계 얼굴 객체 검출

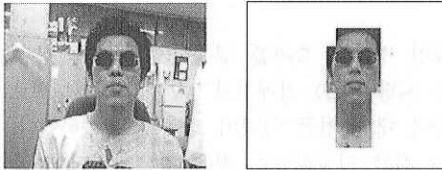


(b) 2 단계 얼굴 객체 검출

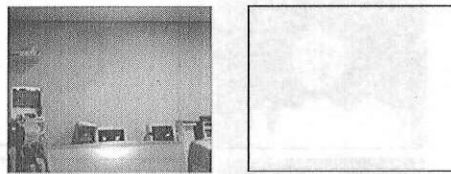
그림 13. 제안 알고리즘에 의한 얼굴 검출  
Fig.13. Face detection by the proposed algorithm



그림 13의 (a)는 1차 배경 피부색 제거와 얼굴 검증을 거쳐서 얻어진 결과이고 (b)는 1단계에서 처리되지 못하여 2단계에서 처리되어 얻어진 얼굴 검출 결과이다. 한편 그림 14는 특수한 경우들에 대한 검출 결과이다. 그림 14의 (a)는 선글라스를 착용한 상태에서의 얼굴 검출을 시도한 것으로 선글라스의 일부 영역이 피부색으로 분류되지 않더라도 팽창 작용에 의해 얼굴 영역이 양호하게 검출됨을 보여주고 있다. 그리고 기존 방식에서 배경영역이 피부색으로 검출되는 영상인 경우 얼굴 영역이 없음에도 불구하고 배경의 한 부분을 얼굴 영역으로 검출하고 있다. 그림 14의 (b)는 제안 방식을 이용하여 기존 방식의 이러한 단점을 효과적으로 제거함을 보여주고 있다.



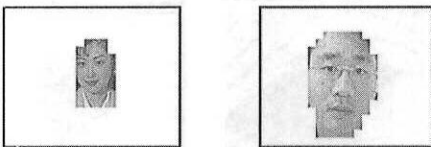
(a)



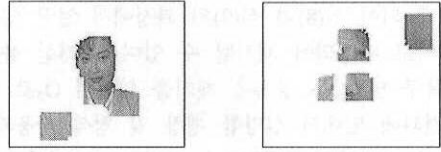
(b)

그림 14. 특수한 영상들에 대한 실험 결과  
Fig.14 Experiment result for the particular images

그림 15는 기존 방식[5]과 제안 방식의 얼굴 검출 결과를 비교한 것이다. 앞에서 설명한 기존 방식의 문제점들이 제안 방식에 의하여 효과적으로 해결됨을 알 수 있다.



(a) 제안 방식



(b) 기존 방식

그림 15. 기존 방식과 제안 방식의 결과 비교  
Fig.15. Comparison of the experimental results by the conventional and proposed methods

제안 방식과 기존 방식[5]간의 성능 평가에 있어서 사용되어지는 계산량 비교는 매우 중요한 의미를 지닌다. 표 1은 기존 방식과 제안 방식의 연산량을 매크로 블록 단위로 조사한 것이다. 표 1로부터 제안 방식이 기존 방식보다 평균 64.9%정도 계산량 이득이 있음을 알 수 있다. 이러한 결과는 실시간으로 영상이 압축되는 응용 분야에서 제안 알고리즘이 매우 효과적임을 암시한다.

표 1. 기존 방식과 제안 방식과의 연산량 비교  
Table1. Comparison of the computations for the conventional and the proposed methods

Sequence name	방법	Compare	Add	Sub	Mul	Total	비고
Akiyo	기존방식	222.3/MB	195.8/MB	3.4/MB	75.0/MB	496.5/MB	2단계
	제안방식	146.1/MB	88.6/MB	0.4/MB	0/MB	235.2/MB	
	이득	34.2 %	54.7 %	88.2 %	100 %	52.6 %	
Foreman	기존방식	224.0/MB	399.5/MB	8.2/MB	180.0/MB	814.6/MB	2단계
	제안방식	160.7/MB	104.2/MB	0.5/MB	0/MB	265.6/MB	
	이득	28.2 %	73.9 %	93.9 %	100 %	67.3 %	
Forest	기존방식	222.2/MB	175.8/MB	2.8/MB	64.6/MB	466.7/MB	1단계
	제안방식	39.2/MB	13.4/MB	0.1/MB	0/MB	52.9/MB	
	이득	82.3 %	92.3 %	96.4 %	100 %	88.6 %	
Paikha	기존방식	221.6/MB	103.3/MB	1.2/MB	27.2/MB	353.8/MB	1단계
	제안방식	36.0/MB	12.6/MB	0.1/MB	0/MB	48.9/MB	
	이득	83.7 %	63.6 %	91.6 %	100 %	86.1 %	
Average	기존방식	222.0/MB	162.2/MB	2.6/MB	57.2/MB	445.5/MB	100 sample
	제안방식	100.2/MB	55.3/MB	0.2/MB	0/MB	156/MB	
	이득	54.8 %	65.9 %	92.3 %	100 %	64.9 %	

표 2. 얼굴 영역 검출률  
Table 2. Detection rate of face region

Sequence type	Number of sequence	Correctly detected	Miss detected
Standard	3	3	0
Movie	50	47	3
CCD	13	13	0
Still image	34	31	3
Total	100 %	94 %	6 %

마지막으로 그림 16는 검출된 얼굴 영역에 객체 코딩을 적용한 결과를 보여주고 있다. 사용된 엔코

더는 MPEG4 VM을 사용하였으며 객체 코딩 시 얼굴 영역의 QP는 8, 얼굴 외 영역의 QP는 31을 적용하였으며, 객체 코딩을 하지 않을 시에는 모든 매크로 블록에 QP 20으로 적용하였다. 적용 결과 객체 코딩 미 적용시 비트율은 2,788Byte, 적용시는 2,733Byte에서 객체 코딩 적용시 얼굴 영역(ROI)의 PSNR이 5.74db증가하였다. 반면에 얼굴 영역이외의 부분에서의 화질 차이는 거의 느껴지지 않는다. 따라서 이러한 결과는 화상 회의를 진행할 시 실시간으로 얼굴 객체에 대한 객체 코딩을 적용함으로써 낮은 비트율에서 얼굴 영역의 화질을 보다 개선시킬 수 있음을 보여준다.



(a) 기존 코딩을 적용한 경우  
(Bitrate:2788byte, ROI PSNR:30.21dB)



(b) 객체 코딩을 적용한 경우  
(Bitrate:2733byte, ROI PSNR:35.95dB)

그림 16. 제안 방식에 기반한 객체 코딩 결과  
Fig.16. The experimental result of the object coding based on the proposed algorithm

## V. 결론

본 논문에서는 객체 지향 압축 기법의 적용에 적합한 효율적인 블록 기반 얼굴 검출 방식을 제안하였다. 제안 알고리즘에서는 여러 인종에 대한 얼굴의 피부색이 특정 범위 내에 존재한다는 사실을 기반으로 동영상 압축 과정에서 얻어지는 DCT 계수들을 이용하여 블록단위로 검출을 수행한다. 그리고 검출 과정을 1, 2단계로 분리함으로써 단순한 영상은 1단계에서, 복잡한 영상은 2단계에서 얼굴 영역

이 검출되게 하였다. 이것은 영상에 따른 적응적인 검출이 가능하여 계산량이 효율적으로 활용되게 한다.

영화 및 일반 CCD 카메라로부터 획득된 여러 인종에 대한 얼굴 영상에 대하여 제안 알고리즘을 적용한 결과, 평균적으로 94%의 검출율이 얻어졌다. 그리고 기존 방식과의 계산량 비교에서는 약 65%의 계산량 이득을 얻었다. 이것은 제안 알고리즘이 적응적으로 얼굴 영역을 검출하고 블록 단위로 처리를 수행하기 때문이다. 또한 제안 방식에 의하여 검출된 얼굴 영역에 대하여 효과적인 압축을 수행할 경우, 보다 우수한 주관적인 화질이 얻어짐을 확인할 수 있었다.

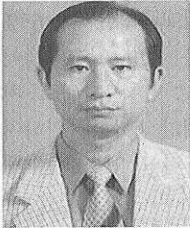
## 참고 문헌

- [1] M-H, Yang, D Kriegman, and N.Ahuja, "Detecting Faces in Images : A Survey," *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol.24, no.1, pp.34-58, Jan 2001.
- [2] G.Yang and T. S. Huang,"Human face Detection in a complex background,"*Pattern Recognition*, vol.27, no.1, pp.53-63, 1994.
- [3] N. Herodotou, K. N. Plataniotis and A. N. Venetsanopoulos, "Automatic Location and Tracking of the Facial Region in Color Video Sequences," *Signal Processing : Image Communication*, vol. 14, no. 5, pp.359-388, 1999.
- [4] K. Sobottka, I. Pitas, "Segmentation and Tracking of Face in color image," *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pp.236-241, 1996.
- [5] D.Chai and K.N.Ngan, "Face Segmentation Using SkinColor Map in Videophone Application" *IEEE Trans on Circuit and System for Video Tech.*, pp.551-564, June 1999.
- [6] ISO/IEC/SC29/WG11, "MPEG-4 Video Verification Model version 18.0," *ISO*, Jan 2001.
- [7] Y. S. Ho and A. Gersho, "Classified transform coding of images using vector quantization," *IEEE int. Conf. ASSP*, pp.1890-1893, May 1999
- [8] A. Murat Tekalp,

"Digital Video Processing," Prentice Hall, 1995.

김기주(Ki-Ju Kim)

정회원



1979년 2월:울산대학교 전자공학과 졸업(공학사)  
1981년 2월:부산대학교 전자공학과 졸업(공학석사)  
1985년 7월~1995년6월:한국 Sony전자 ISO추진실장  
1996년 3월~현재:부산대학교 메카트로닉스학과 박사과정

2004년1월~현재:구미기능대학 전자과 부교수  
<관심분야>영상신호처리, 동영상압축, 영상인식

문정미(Jeong-Mee Moon)

준회원



2001년 2월:홍익대학교 전기전자 컴퓨터공학부 (공학사)  
2001년9월~현재:부산대학교 전자공학과 석박사통합과

<관심분야> 신호처리, 동영상 압축, 영상처리

방경구(Kyoung-Gu Bang)

준회원



2002년 2월:경일대학교 전자정보공학과 (공학사)  
2004년 2월:부산대학교 전자공학과(공학석사)  
2004년 2월~현재 : LG전자 Digital Network Display Lab. 연구원

<관심분야> 동영상 압축, 영상처리, 영상인식

김재호(Jae-ho Kim)

정회원

한국통신학회논문지 제29권,제 1C지 참조