

동영상에서 칼만 예측기와 블록 차영상을 이용한 얼굴영역 검출기법

A New Face Tracking Method Using Block Difference Image and Kalman Filter in Moving Picture

장희준, 고혜선*, 최영우*, 한영준, 한현수

Heejun Jang, Hyesun Ko*, Youngwoo Choi*, Youngjoon Han, Hernsoo Hahn

송실대학교 전자공학과, 숙명여자대학교 정보과학부*

Dept. of Electronic Engineering, Soongsil University, Computer Science,
Sookmyung Women's Univ.

요 약

복잡한 환경에서 이동하는 사람의 얼굴영역은 배경과 조명에 의해 확장, 축소 검출되기도 하고 잘못된 영역을 오검출하기도 한다. 본 논문에서는 동영상에서 얼굴을 추적하는데 있어서 확장 혹은 축소검출이나 오검출 문제를 해결하기 위해 블록차영상과 칼만예측기를 사용하는 방법을 제안한다. 블록차영상은 입력영상을 블록화하여 차영상을 얻는 방법으로 미세한 움직임까지 검출이 가능하여 영상에서 움직임이 작은 경우에도 검출이 가능하게 된다. 검출된 움직임영역에서 얼굴영역은 1차적으로 피부색을 이용하여 검출하며 피부색이 검출되지 않는 경우는 움직임영역의 경계선을 8이웃화소 창을 이용하여 부호화하고 머리 부분의 코드를 갖는 영역을 얼굴영역으로 추정하는 방법을 사용한다. 추정된 얼굴영역을 켈리분할하고 분할된 영역에서 피부색과 가장 가까운 색을 갖는 영역을 얼굴영역으로 판단한다. 얼굴영역은 최외각화소를 포함하는 4각형으로 표시하고 각 정점의 이동을 칼만예측기를 이용하여 추정하고 추정된 위치에서 얼굴영역을 검출하는 방법을 사용한다. 제안하는 방법은 동영상에서 얼굴영역검출의 정확도를 높이고 얼굴영역의 추적에서 얼굴영역검출에 소요되는 시간을 상당부분 감소시키는 효과를 거두게 됨을 실험을 통해 입증하였다.

Abstract

When tracking a human face in the moving pictures with complex background under irregular lighting conditions, the detected face can be larger including background or smaller including only a part of the face. Even background can be detected as a face area. To solve these problems, this paper proposes a new face tracking method using a block difference image and a Kalman estimator. The block difference image allows us to detect even a small motion of a human and the face area is selected using the skin color inside the detected motion area. If the pixels with skin color inside the detected motion area, the boundary of the area is represented by a code sequence using the 8-neighbor window and the head area is detected analysing this code. The pixels in the head area is segmented by colors and the region most similar with the skin color is considered as a face area. The detected face area is represented by a rectangle including the area and its four vertices are used as the states of the Kalman estimator to trace the motion of the face area. It is proved by the experiments that the proposed method increases the accuracy of face detection and reduces the face detection time significantly.

Key Words : Face Detection, Face Tracking, Kalman Filter, Surveillance System

1. 서 론

최근의 발전된 디지털 카메라와 영상처리기술은 동영상의 전송 및 저장, 처리를 손쉽게 함으로써 동영상의 활용분야를 오락이나 방송에서부터 의료와 보안 분야에 이르기까지 그 범위를 무한정 넓혀주고 있다. 특히 동영상을 이용한 무인감시 및 인증분야는 다양한 범죄를 예방한다는 차원과 재산권

을 보호한다는 차원에서 가장 많이 연구, 활용되고 있다[1]. 무인감시는 특정한 공간에 움직이는 물체가 접근하는 지를 판단하는 초보적인 단계에서 출발하여 지금은 많은 사람들이 출입하는 공간에 특정한 사람이 나타났는지를 판단하는 기술로 발전하고 있다. 또한 인증은 사람의 얼굴을 인식하여 특정한 정보에 접근할 권한이 있는지를 판단하는 기술로 이들 모두 얼굴을 인식하는 것이 기본적으로 가능해야 한다.

동영상에서 얼굴을 인식하는 기술은 Eigen-Face를 이용하는 방법[2], 템플릿을 이용하는 방법[3], 등과 같이 다양한 방법이 개발되어 사용되고 있다. 이들 기법은 모두 동영상에서 얼굴영역을 성공적으로 분리했다고 가정하거나 얼굴영역

접수일자 : 2004년 10월 15일

완료일자 : 2005년 2월 14일

이 영상에서 차지하는 위치와 크기를 일정하게 제한하고 있다. 따라서 대부분의 얼굴인식기법들이 얼굴의 크기와 위치, 그리고 영상을 획득하는 조명환경에 따라 성능이 크게 좌우되는 결과를 보여주고 있다. 이들 기법들이 다양한 환경에서도 일정한 수준이상의 성능을 갖게 하기 위해서는 조명환경과 얼굴의 크기나 위치에 강인한 얼굴영역의 획득방법이 사용되어야한다.

이러한 필요성에 의해 최근 동영상에서 얼굴영역을 검출하기 위한 다양한 방법들이 개발되고 있다[4-8]. 이러한 방법들은 사용 방법에 따라 배경의 복잡함, 얼굴 크기, 얼굴의 회전, 조명의 변화, 계산량 등에 따라 검출 성능이 좌우되는 문제점을 가지고 있다.

따라서 본 논문에서는 이러한 문제 요인들에 대해 강인한 검출을 하기 위해 불변특징 중의 하나인 얼굴색상정보와 얼굴영역 경계선의 특성을 활용하는 방법을 융합하여 얼굴영역을 검출하는 방법을 제안한다. 입력영상에서 얼굴영역의 크기는 이동하는 사람의 속도와 방향에 의해 좌우되므로 사람의 이동정보를 이용하여 얼굴영역의 이동위치와 크기를 예측하고 얼굴색상정보를 이용하여 얼굴영역을 확정한다. 얼굴영역은 입력영상에서 색상정보를 이용하여 측정하고 이를 포함하는 사각형의 꼭지 점으로 표현하며 이들의 이동정보는 칼만 예측기를 이용하여 예측한다. 예측된 얼굴영역은 차영상의 범위 내에 위치하는 것이 제약조건이 되며 카메라가 이동하는 경우 배경 영상이 변화되므로 차영상을 구할 때 장면분류기법을 적용한 다음, 처리속도를 향상시키고 미세한 움직임도 포함시키기 위해 블록단위의 영역검출기법을 적용하였다. 또한 얼굴영역의 측정단계에서 얼굴과 어귀의 형태 정보를 사용하여 조명의 변화에 강인하며 포즈의 변화에도 강인한 성능을 갖도록 하였다.

2 색상 및 움직임영역의 형태를 이용한 얼굴검출

얼굴영역 검출에 있어서 일반적으로 생각할 수 있는 방법은 피부색상을 이용하는 것이다. 피부색상을 이용하는 방법은 연산 시간이 비교적 빠르고 연산 과정이 간단하다는 장점으로 인해 널리 응용되고 있지만 배경색상이 피부색상과 유사한 경우 정확도가 떨어지고 조명의 변화에 민감하여 검출 성능이 조명의 변화에 따라 좌우되는 단점이 있다. 따라서 피부색상만을 이용한 얼굴검출의 오검출 위험을 감소시키기 위해 사람의 움직임 정보를 이용하여 배경이나 조명의 영향을 최소화하는 방법이 가장 많이 사용되고 있다. 본 논문에서는 색상정보와 함께 움직임정보에 포함된 얼굴영역을 정확히 검출하는 방법을 중점으로 다룬다. 본 논문에서 제시하는 얼굴검출 과정은 다음의 그림 1과 같다.

그림 1에서 보는 바와 같이 입력영상이 주어지면 이전영상과 차영상을 구하고 차영상은 8x8크기의 블록으로 표현되며 블록차영상에서 피부색상을 이용하여 얼굴영역을 검출한다. 피부색상을 가진 영역이 검출되지 않으면 움직임영역의 경계선 특성을 이용하여 머리영역을 찾고 머리영역에서 색상분할을 통해 피부색상과 가장 가까운 영역을 얼굴영역으로 검출하는 과정이다.

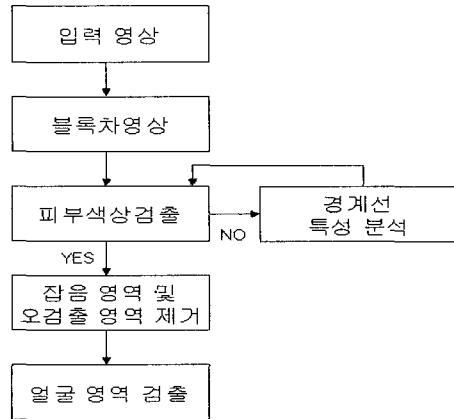


그림 1. 얼굴 영역 검출 과정 흐름도.

Figure 1. Flow chart of the face tracking algorithm.

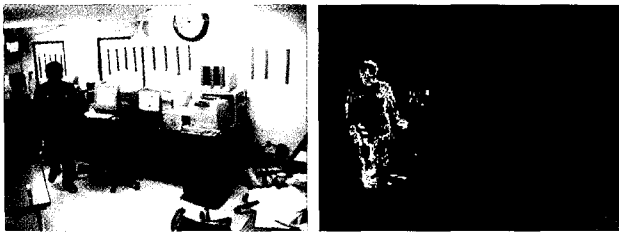
2.1 블록차영상

배경에서의 오검출과 조명의 영향을 최소화하기 위해 사용하는 차영상 기법은 카메라의 흔들림과 조명의 변화에 비교적 덜 민감한 장점이 있지만 객체의 움직임이 매우 미약할 때 실제 움직임 화소인지 조명에 대한 카메라의 잡음을 포함하는 화소인지의 구분이 매우 어렵고, 객체 전체를 나타내는데 어려움이 있다는 단점이 있다. 본 논문에서는 이러한 단점을 극복하기 위하여 영상을 8*8로 블록(block)화하고 차영상에서 블록내의 움직임화소가 일정 수 이상이면 전체 블록을 움직임 블록으로 판단하여 미세한 움직임도 검출될 수 있도록 하였다. 블록화 방법을 적용하기 위해서는 움직임 블록이 실제적인 움직임에 의한 블록인지 잡음에 의한 블록인지를 구분하는 방법이 필요하다. 이를 구분하기 위해 영상에서 카메라의 잡음이나 조명의 심한 변화에 대한 잡음은 임펄스(impulse) 형태로 존재한다는 점을 이용한다. 임펄스 잡음은 이웃 화소와 연관성이 없으며 일정 지역 내에 다중으로 발생하지 않고 전역적으로 발생하기 때문에 이러한 현상을 이용하여 블록 내에 움직임 정보량의 포함 정도로 실제 움직임 영역과 잡음영역을 구분할 수 있다. 블록 차영상은 현재 프레임(F0)과 이전 프레임(F-1)의 차이로 차영상을 얻은 후 8*8 블록 단위로 처리하게 된다. 각 프레임간의 노이즈로 인한 차영상 정보 발생을 극복하기 위하여 블록 내 일정 픽셀 이상의 차영상이 존재 할 때 그 블록을 1로, 그 이하 값을 가질 때 0값으로 채워지게 된다. 이렇게 하면 차영상이 존재하는 영역은 작은 것이라도 놓치지 않고 검출할 수 있다.

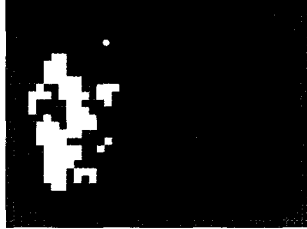
$$Diff(x, y) = \begin{cases} 1 & \text{if } (F^0(x, y) - F^{-1}(x, y)) > th_1 \\ 0 & \text{otherwise} \end{cases}$$

$$BDiff(x', y') = \begin{cases} 1 & \text{if } \sum_{x=0}^7 \sum_{y=0}^7 diff(x, y) > th_2 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

위의 식(1)에서 $Diff(x, y)$ 는 차영상의 이전영상을 $BDiff(x', y')$ 는 차영상의 블록영상을 의미한다. 그림 2는 현재 영상과 이전 영상으로부터 블록화된 차영상을 이용하여 움직임 객체를 검출하는 과정을 보여준다. 블록화된 차영상은 움직임 객체를 나타내는 동시에 블록의 변화 패턴을 이용하여 형태(Shape)정보를 나타낼 수 있다.



(a) 원영상 (b) 화소단위 차영상



(c) 블록화된 차영상

그림 2. 화소단위 차영상과 블록차영상.

Figure 2. Original difference image and blocked-based difference image.

2.2 피부색상을 이용한 얼굴 영역 검출

얼굴영역을 검출하기 위한 피부색상의 특성은 YIQ 색상 공간에서 정의하였다. YIQ 색상 표현법은 단순히 I와 Q값을 비교하여 검출하기 때문에 Y값의 변화에 강인한 검출 성능을 보인다. 식 (2)는 RGB로 입력된 영상을 YIQ로 변환한 후 피부색상의 범위를 실험적으로 구하여 정의하고 있다.

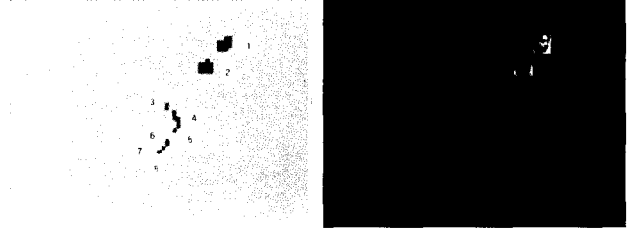
$$\begin{aligned} I &= 0.6R + 0.27G + 0.32B + 150 \\ Q &= 0.21R + 0.52G + 0.31B + 132 \\ \text{SkinRegion} &: (163 < I < 199), (123 < Q < 133) \end{aligned} \quad (2)$$

실시간으로 얼굴 검출을 하기위해서 본 논문에서는 입력 영상을 전체 영상의 1/16로 축소하여 검출하였으며, 움직임 영역이 있는 블록에서만 검색을 하였다. 배경이나 조명에 의해 오검출된 색상영역은 검출된 영역의 크기, 검출영역에 포함된 피부색상의 비율, 검출영역의 가로대 세로의 비율 이용하여 제거하였다. 검출영역의 크기가 12화소 이하일 경우(그림 3(c)에서 영역 3, 5, 6, 7, 8)는 배경 혹은 잡음에 의한 검출로 간주하여 제거하였고 검출영역에 피부색상이 60%이하로 포함된 경우(그림 3(c)에서 영역 4)와 검출영역의 가로와 세로의 비가 2/3이상일 때는 오검출로 판단하였다.

이와 같은 피부색상을 이용한 얼굴영역의 검출은 조명의 변화에 따라 검출성능이 좌우된다. 특히 역광으로 인한 얼굴의 실루엣 영상과 직접 광으로 인한 얼굴의 반사 영역이 크게 나타날 때, 검출 성능은 현저히 낮아지게 된다. 그림 4의 (a)와 (c)는 동일 장면을 역광과 전면조명을 이용하여 촬영한 것이고 그림 4의 (b)와 (d)는 각각의 영상에서 얼굴색상을 검출한 결과를 보여준다. 이러한 단점을 보완하기 위하여 본 논문에서는 블록차영상의 경계선 특성을 이용하여 얼굴을 검출하는 단계를 추가로 사용한다.



(a) 입력 영상 (b) 움직임 블록



(c) 움직임 블록내 Skin Color (d) 검출된 얼굴 영역

그림 3. 피부색상을 이용한 얼굴 검출 과정.

Figure 3. Segmentation and detection of the face region using the skin color.



(a) 역광조명 (b) (a)의 피부색상영역



(c) 전면조명 (d) (c)에서의 피부색상영역

그림 4. 조명에 따른 피부색상 영역의 검출결과.

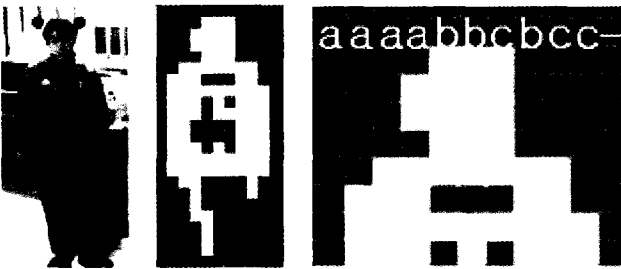
Figure 4. Detection of the face regions with the skin-color under irregular lighting conditions.

2.3 블록차영상의 경계선 특성을 이용한 얼굴영역 검출

조명의 영향에 의한 피부색상변화의 영향을 최소화하기 위해 블록차영상의 경계선을 분석하여 머리영역을 찾아내고 머리영역을 색상을 기반으로 분할하여 피부색상과 가장 유사한 영역을 얼굴영역으로 검출하는 기법을 사용한다. 블록차영상에서 검출된 영역은 사람의 움직임 블록이고 머리의 위치는 움직임 블록의 가장 위쪽에 있다. 따라서 차영상의 Y축에서 가장 큰 값을 갖는 블록이 사람의 머리영역일 가능성이 가장 크며 그 다음으로 큰 값의 블록은 어깨와 몸통일 가능성이 크다. 이와 같은 관점에서 블록차영상을 Y축 좌표 값으로 왼쪽에서 오른쪽으로 검색을 하면 양쪽 어깨선에서는 평활한 형태를 가지고 얼굴 영역에서는 상향 값을 가진다. 블록차영상의 형태를 쉽게 분석하기 위해 코드를 할당하였다. X축으로 왼쪽에서 오른쪽으로 검색하여 상향블록에 'a', 평활 블록에 'b', 하향블록에 'c' 코드를 할당하고, 블록 영상이 없는 곳은 '-'를 부여한다. 이와 같은 코드를 이용하면 사

림의 머리영역의 경계선은 어깨에서는 'b'를 어깨에서 얼굴로 올라갈 때는 'a'를, 머리의 상위부분에서는 'b'를, 다시 어깨로 내려올 때는 'c'를 갖게 됨으로 머리영역은 'abc'의 문자코드로 나타낼 수 있고 'b'의 숫자로 얼굴의 크기를 짐작할 수 있다. 식 (3)은 블록의 Y축 위치를 이용하여 코드를 할당하는 방법을 보여준다. 식 (3)에서 $Y_{max}(x, y)$ 는 블록 차영상의 (x,y) 점에서 최대 Y좌표 값을 나타내고 x^0 는 현재의 화소 위치이고 x^{-1} 은 이전 화소의 위치이다. 그림 5(a)의 입력영상에서 움직임영역이 그림 5(b)에 주어졌으며 그림 5(c)는 윗부분의 경계선 영역에 코드를 할당한 결과를 보여준다. 그림 5(a)의 입력영상에서 얼굴의 좌측과 오른쪽에 표시된 점은 각각 얼굴영역의 왼쪽과 오른쪽 경계를 나타낸다.

$$BC = \begin{cases} a & \text{if } Y_{\max}(x^0, y) > Y_{\max}(x^{-1}, y) \\ b & \text{if } Y_{\max}(x^0, y) = Y_{\max}(x^{-1}, y) \\ c & \text{if } Y_{\max}(x^0, y) < Y_{\max}(x^{-1}, y) \end{cases} \quad (3)$$



(a)입력영상 (b)움직임 영역 (c) 머리영역의 경계선코드
 그림 5. 검출된 움직임영역(b)과 머리영역의 경계선코드.
 Figure 5. Allocating the contour code to a head region.

블록차영상에서 경계선 코드를 이용하여 머리영역을 검출하는 경우 위치오차의 범위는 블록의 크기인 8화소가 된다. 또한 사람이 빠르게 이동할 때 차영상영역이 커져서 오차의 크기가 증가한다. 이러한 측정오차를 최소화하기 위해 머리영역이 예측되면 이 범위 내에서의 얼굴영역 검색은 화소단위로 이루어진다. 즉 검출된 머리영역블록을 화소단위 2진 차영상으로 표현하고 모폴로지 기법을 이용하여 머리영역내부의 화소들을 모두 1로 채워서 머리영역과 배경영역을 분리한다. 그림 6(a)는 그림 5에서 검출된 머리영역을 보여주고 그림 6(b)는 모폴로지 기법을 이용하여 머리영역과 배경영역을 분리한 결과를 보여준다. 그림 6(b)에서 얼굴영역을 검출하는 방법은 머리의 두께가 측면에서 관측한 경우를 제외하고는 항상 어깨영역의 두께보다 항상 좁다는 점에 착안한다. 따라서 얼굴영역을 구분하기 위해서는 머리영역의 상단부터 하단으로 내려가면서 그림 6(c)와 같이 사각형을 형성하며 사각형 내에 포함되는 머리영역의 비율의 변화를 관측하는 방법을 적용한다. 사각형은 머리영역의 좌우경계점을 폭으로 하고 머리영역의 최상단부터 수직으로 내려가는 화소의 길이를 높이로 한다. 그림 6(d)에서처럼 Y축은 영상의 높이를 X축은 사각형에 포함된 머리영역의 비율을 그려보면 높이가 머리에서 어깨로 변할 때 비율의 변화가 최대가 된다. 그림 6(c)의 작은 사각형이 검출된 경계선을 기반으로 구해진 얼굴영역이다. 다음의 식 (4)는 사각형에 포함된 머리영역의 비율(R_h)을 정의한다. 식 (4)에서 i 는 머리상단에서 영상화소단위로 검사하려는 영역까지의 높이를, a_i 와 b_i 는 각각 사각형의 왼쪽경계선과 오른쪽 경계선을 $F(i, j)$ 는 모폴로지 연산을 2진화된 영상을 의미한다.

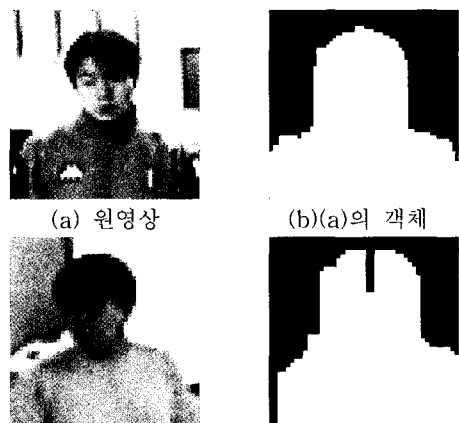
그림 6(b)의 얼굴영역을 검출하기 위해 사각형에 포함된 머리영역의 비율을 고려하는 이유는 그림 7(d)에서 보듯이 사람의 움직임이 작은 경우에 머리영역에서 구멍이 뚫린 부분이 발생하여 오검출 되는 경우가 발생하므로 폭의 변화율만으로는 얼굴영역과 어깨영역의 분리가 어렵기 때문이다.

$$R_h = \sum_{i=1}^l \sum_{j=a_i}^{b_i} F(i, j) / [i \times (b_i - a_i)] \quad (4)$$



(a) 원영상 (b) 배경과 사람의 분리
 (c) 사각형 검출창 (d) 사각형내의 머리영역비율
 (e) 색상기반 분할 영상 (f) 최종얼굴영역
 그림 6. 얼굴영역의 검출과정.

Figure 6. Detection process of a face region.



(a) 원영상 (b)(a)의 객체
 (c) 움직임이 적은 원영상 (d)(c)의 객체
 그림 7. 움직임 정보에 따라 각각 다르게 나타나는 객체의 형태.

Figure 7. Various object shapes of the moving head.

이와 같이 분리된 얼굴영역은 얼굴뿐 아니라 머리카락과 기타의 영역을 포함하게 되며 때로는 얼굴영역보다 일정부분

작거나 크게 분리될 수도 있다. 따라서 식 (4)를 이용하여 얼굴영역이 검출되면 정확한 얼굴영역을 결정하기 위해 미세한 조정 작업이 필요하다. 검출된 얼굴영역에서 얼굴을 결정하는 일은 검출영역 내에서 피부색상과 가장 가까운 색상을 갖는 영역을 찾는 것으로 정의한다. 얼굴영역은 조명의 변화가 있더라도 Red 성분이 Green 성분이나 Blue 성분보다 높기 때문에 관심 컬러를 Red 성분이 다른 성분보다 강한 색으로 정의하고 피부 색상과의 거리 정보를 이용하여 피부색상과 가장 가까운 영역을 군집화하면 검출이 가능하다. 그림 6(e)는 그림 6(c)에서 검출된 머리영역을 색상기반으로 분할하고 군집화한 영상을 보여준다. 각 분할영역은 대표색상으로 표현되며 이들 색상 중에서 식 (2)에 정의된 피부색상과 가장 가까운 색상을 갖는 분할영역이 얼굴로 결정된다. 그림 6(f)는 최종적으로 결정된 얼굴영역을 보여준다.

3. 칼만필터를 이용한 얼굴영역의 예측

동영상이 입력되는 동안 사람의 출입에 의해 조명이 영향을 받아서 피부색상으로 정의된 색상범위가 크게 바뀌기도 하고 움직임이 거의 없는 경우에도 얼굴의 검출이 어려워진다. 그러나 정지영상과는 달리 동영상에서 얼굴을 검출하는 일은 얼굴을 추적하는 일과 동일하여 이전 영상에서 검출된 얼굴영역에 대한 정보를 다음 영상에서 활용할 수 있다는 장점을 갖는다. 이와 같은 장점을 이용하여 동영상에서의 얼굴 검출의 정확도와 속도를 높이기 위해 본 논문에서는 식 (5), (6)과 같이 설계된 칼만필터를 이용하여 얼굴영역의 움직임을 예측하고 그 영역에서 얼굴을 찾는 방법을 사용한다. 식 (5)는 상태 모델은, 식(6)은 추정 모델을 나타낸다.

$$S(t) = \Phi(\Delta t)S(t - \Delta t) + w(t) \quad (5)$$

$$m(t) = H(t)S(t) + v(t) \quad (6)$$

식 (5)에서 상태벡터 $S(t)$ 는 시스템 상태를 나타내는 변수로, 본 논문에서는 검출된 얼굴영역을 포함하는 사각형의 중심 좌표점 (C_x, C_y), 너비 (L), 높이 (H), 중심 좌표점의 이동 속도 (V_x, V_y), 너비와 높이의 미분 값 (L', H'), 중심 좌표점의 가속도 (A_x, A_y), 너비와 높이의 이차 미분 값 (L'', H'')으로 구성되며 $S(t) = [C_x, C_y, W, H, V_x, V_y, W', H', A_x, A_y, W'', H'']^T$ 로 표현한다. 또한 t 는 이산시간, Δt 는 상태 추정의 샘플링 시간을 의미한다. Φ 는 상태전이 행렬로 다음 시점에서의 시스템 상태를 예측할 때 쓰이며, 본 논문에서는 식 (9)에서와 설계된다. 이는 상태예측 시 얼굴영역 중심점의 움직임 상태와 얼굴영역 크기 변화량을 반영해 주기 위해서 이다.

랜덤 변수 $w(t)$ 는 시스템 상태 모델과 실제 시스템의 차이를 모델링한 공정잡음(process noise)이며 $v(t)$ 는 추정시스템이 갖는 측정잡음(observation noise)을 나타낸다. 공정잡음과 측정잡음은 평균이 0이고, 공분산이 각각 $Q(t)$ 와 $R(t)$ 인 백색 가우시안 잡음으로 가정한다. 식(7), 식(8)은 공정잡음과 측정 잡음을 나타낸다.

$$p(w) \sim N(0, Q) \quad (7)$$

$$p(v) \sim N(0, R) \quad (8)$$

식 (6)에서 추정벡터 $m(t)$ 은 상태벡터의 구성요소 중 외부적으로 추정 가능한 요소들로 이뤄진 행렬로, 본 논문에서

는 얼굴영역 검출을 통해 얻은 얼굴영역의 중심점 좌표(C_x, C_y), 너비와 높이(W, H)로 구성되며 $m(t) = [C_x, C_y, W, H]^T$ 로 표현한다. $H(t)$ 는 관측행렬로 상태벡터를 측정벡터로 변환해주는 역할을 하고, 본 논문에서는 식 (10)에서와 같이 설계된다.

$$\Phi(\Delta t) = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (9)$$

$$\begin{pmatrix} C_x \\ C_y \\ W \\ H \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} S(t) + \begin{pmatrix} R(t) \\ R(t) \\ R_w(t) \\ R_h(t) \end{pmatrix} \quad (10)$$

칼만 필터의 적용과정은 크게 초기화, 상태예측, 측정갱신 이렇게 세 단계로 나뉘지며 초기화 과정을 거친 뒤 Δt 의 시간이 지날 때 마다 상태예측과 측정갱신이 재귀적으로 호출, 실행된다. 초기화 단계에서는 상태벡터와 예측오차 공분산 행렬을 초기화 한다. 상태벡터의 초기값 $S(0)$ 를 위해서 얼굴영역이 최소한 3번 연속으로 검출되어야 하는데, 이렇게 얻어진 얼굴영역들을 각각 $F1, F2, F3$ 라 한다면 다음의 식 (11)을 통해 상태벡터 초기값을 구할 수 있다. 예측오차의 공분산 행렬 초기값 $P(0)$ 은 단위행렬로 초기화한다.

$$\begin{aligned} V_x &= C_x(F3) - C_x(F2) \\ V_y &= C_y(F3) - C_y(F2) \\ A_x &= C_x(F3) - 2C_x(F2) + C_x(F1) \\ A_y &= C_y(F3) - 2C_y(F2) + C_y(F1) \\ W &= (W(F3) - W(F2)) / \Delta t \\ H &= (H(F3) - H(F2)) / \Delta t \\ W' &= ((W(F3) - W(F2)) / \Delta t - (W(F2) - W(F1)) / \Delta t) / \Delta t \\ H' &= ((H(F3) - H(F2)) / \Delta t - (H(F2) - H(F1)) / \Delta t) / \Delta t \\ S(0) &= [C_x(F3), C_y(F3), W(F3), H(F3), V_x, V_y, W', H', A_x, A_y, W'', H''] \end{aligned} \quad (11)$$

상태예측 단계에서는 식(12)를 통해 다음 시점에서의 상태벡터와 예측오차 공분산행렬 값을 예측한다. 식에서 '-'는 예측치를 '+'는 갱신치를 나타낸다.

$$\begin{aligned} S^-(t) &= \Phi(\Delta t)S^+(t - \Delta t) \\ P^-(t) &= \Phi(\Delta t)P^+(t - \Delta t)\Phi^T(\Delta t) + Q(t) \end{aligned} \quad (12)$$

상태벡터에 대한 예측치가 결정되면 다음 영상에서 얼굴영역을 찾기 위해 예측된 영역을 중심으로 얼굴검출 알고리즘을 적용한다. 알고리즘을 적용하는 범위는 예측된 얼굴영역을 2배 확장한 것으로 한다. 이는 예측한 영역에 얼굴영역이 포함되지 않을 수 있는 오차를 고려한 것으로 일반적으로 2배를 넘지 않으므로 2배까지만 확장한 영역을 검색한다.

얼굴검출 알고리즘을 적용하여 얼굴영역이 측정되면 이전 영상에서 검출한 얼굴영역과 동일한 얼굴영역인지 확인하기 위해 두 영역(i 와 $(i-1)$ 번째)의 유사도, $M(i)$ 를 식(13)을 이용하여 계산한다. 식(13)에서 μ_i 는 i 번째 영상에서 검출한 얼굴 영역 (N_1, N_2)의 평균 밝기를 나타내고 μ_{i-1} 는 $(i-1)$ 번째 영상에서 픽셀 좌표 (x, y) 의 밝기를 나타내며 $V_i(x, y)$ 는 i 번째 영상에서 각 좌표 (x, y) 가 Gray 레벨로 정규화된 분산을 나타낸다. 유사도 $M(i)$ 는 그림 8에서처럼 이전 영상에서 검출된 얼굴영역과 새로운 영상에서 검출된 얼굴영역을 같은 크기 (M_1, M_2)로 정규화한 뒤, 각 영역 밝기 값의 분산(그림 8(b),(c)) 비교를 통해 확인한다. 유사도가 만족되면 새로 검출된 얼굴영역의 정보를 이용하여 측정벡터 $m(t)$ 를 결정하게 된다. 만약 얼굴검출 알고리즘에 의한 얼굴검출이 실패한 경우는 형판정합의 방법으로 얼굴영역을 결정하고, 그 값을 가지고 측정벡터 $m(t)$ 를 결정한다.

$$\mu_i = \frac{1}{N_1 N_2} \sum_x \sum_y (b_i(x, y))$$

$$V_i(x, y) = (b_i(x, y) - \mu_i)^2 / 255 \quad (13)$$

$$M(i) = \sum_x \sum_y (V_i(x, y) - V_{i-1}(x, y))^2$$

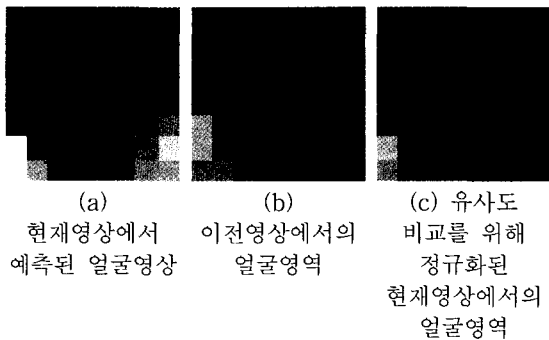


그림 8. 이전영상과 현재영상에서의 얼굴영역의 유사도 비교.
Figure 8. Comparing similarity between the previous and current face regions.

측정벡터 $m(t)$ 가 결정되면 측정갱신 단계를 수행하게 되고 이는 식 (14), (15), (16)을 통해 이뤄진다. 식 (14)를 이용하여 구한 칼만이득 $K(t)$ 과 측정벡터 $m(t)$ 를 식 (15)에 대입하여 상태벡터를 갱신한 후, 식 (16)에서 예측오차 공분산 행렬을 갱신한다.

$$K(t) = P^-(t)H^T(t)[H(t)P^-(t)H^T(t) + R(t)]^{-1} \quad (14)$$

$$S^+(t) = S^-(t) + K(t)[m(t) - H(t)S^-(t)] \quad (15)$$

$$P^+(t) = [I - K(t)H(t)]P^-(t) \quad (16)$$

실제 적용에 있어서는 이와 같이 칼만예측기를 이용하여 얼굴영역의 위치와 크기를 예측하고 예측된 영역을 확장해서 얼굴영역을 검출하고 상태벡터를 갱신하는 과정을 반복한다. 제안된 방법을 적용하면 순간적인 조명이나 배경의 변화에 의해 얼굴영역의 검출이 실패하는 경우에도 예측된 영역에서의 반복된 측정으로 검출이 가능해지며 얼굴영역을 검출할 영역을 예측하여 검색 영역의 크기를 최소화함으로써 얼굴영역 검출에 소요되는 시간을 줄일 수 있다. 또한 비 얼굴영역을 검출하는 오류를 최소화할 수 있으며 얼굴의 크기를 보다 정확하게 측정할 수 있는 효과를 얻을 수 있다.

4. 실험결과 및 분석

본 논문의 성능은 실험을 통해 검증하였다. 동영상은 특정 한 위치에 고정된 컬러 CCD 카메라를 이용하여 카메라로부터의 거리나 각도에 상관없이 무작위로 움직이는 사람(들)을 대상으로 획득하였다. 따라서 영상에 포함되는 사람의 얼굴은 카메라에 대해 다양한 각도와 크기를 갖는다. 실험 칼만 필터를 이용한 얼굴추적기법을 적용한 경우와 그렇지 않은 경우로 구분하여 실행하였고 그 결과를 비교하였다. 영상의 크기는 320x240화소이며 Intel Pentium 4, 3GHz의 CPU와 1G의 메모리를 갖는 PC를 사용하여 처리하였다. 실험에서는 입력되는 동영상의 블록화된 차영상에서 피부색상과 경계선 특성을 이용하는 얼굴영역 검출알고리즘을 기반으로 칼만예측기를 적용했을 경우에 얻을 수 있는 성능의 개선효과를 검증하였다.

얼굴영역 검출알고리즘의 첫 단계에서는 블록차영상에서 검출된 피부색상을 포함하는 후보영역을 선택한 후 후보영역의 가로와 세로비, 크기, 피부 색상 포함률을 기준으로 배경 및 잡음에 의한 영역을 제거한다. 그림 9(a)는 입력영상을, 그림 9(b)는 블록차영상에서 피부색상을 기준으로 선택한 후보영역을, 그림 9(c)는 배경 및 잡음 영역을 제거하여 최종적으로 얻어진 얼굴영역을 보여준다.

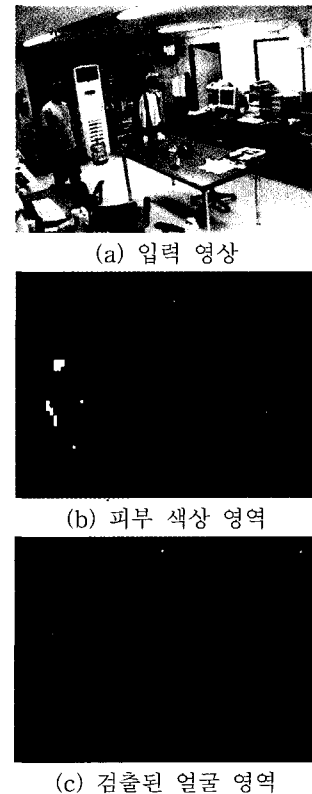


그림 9. 피부 색상으로 검출한 결과.

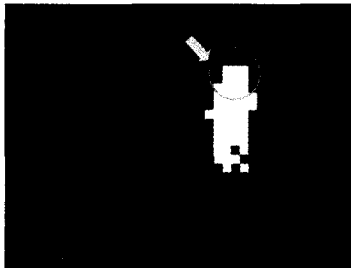
Figure 9. Segmentation of the face region using the skin color space.

블록차영상에서 피부색상영역이 검출되지 않는 경우는 경계선 형태를 이용해서 블록코드를 생성하여 코드로부터 얼굴영역의 검색 초기점을 검출하고 피부색상과의 유사도를 검출하여 가장 유사한 집합을 검출하는 방법을 적용한다. 그림

10(a)는 입력영상에서 사람의 얼굴색상이 조명에 의해서 피부색상으로 검출되지 않는 영상이고 그림 10(b)는 블록차영상에서 경계선이 머리영역을 나타내는 'abc'코드를 갖는 부분을 보여준다. 그림 10(c)는 검출된 머리영역을 시작으로 얼굴과 어깨 영역을 분리하기 위해 모폴로지 기법으로 배경과 객체를 분리한 후, 머리영역을 검출한 결과를 보여주며 그림 10(d)는 머리영역을 칼리 분할하여 피부색상과 가장 유사한 영역을 포함하는 부분을 얼굴영역으로 검출한 결과를 보여준다.



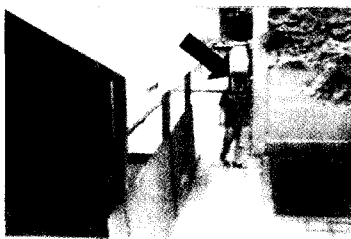
(a) 피부색상의 변화가 있는 입력영상



(b) 블록화된 차영상과 얼굴영역의 블록 코드점



(c) 얼굴과 어깨 영역의 분기점



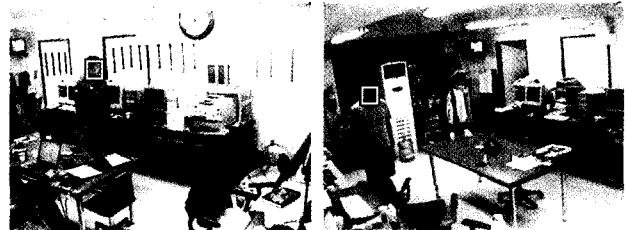
(d) 최종 검출 결과

그림 10. 블록차영상의 경계선 형태 분석을 통해 검출한 얼굴영역.

Figure 10. Face region detection using block-based difference image.

이와 같은 얼굴영역 검출알고리즘은 조명의 변화에도 강한 특성을 갖는 장점이 있지만 사람이 정지해 있거나 움직임이 매우 미약한 경우, 얼굴 이외의 영역에서 얼굴과 유사

한 가로 세로 비, 면적 비가 존재하는 경우 오검출이 나타날 수 있으며 모든 프레임마다 검출이 된다는 보장이 없다. 그림 11에서 보는 바와 같이 그림 11(a)~(f)는 검출이 성공적으로 이루어진 결과를 보이나 (g)와 (h)는 위에서 지적한 문제로 인해 검출이 실패한 영상을 보여준다. 따라서 이런 단점을 최소화하며 연산량을 최소화하기 위해 제안된 칼만예측기를 사용한 얼굴검출 실험을 실행하였다.



(a) 정상검출

(b) 정상검출



(c) 정상검출

(d) 정상검출



(e) 정상검출

(f) 정상검출



(g) 오검출

(h) 오검출

그림 11. 일반적인 얼굴 검출 방법 사용 결과.

Figure 11. Face detection using a conventional method.

그림 12(a)는 얼굴검출 알고리즘을 이용하여 검출하였을 때 움직임이 매우 미약하여 생긴 오검출된 영상과 그림 12(b)는 칼만예측기로 잘못된 얼굴 영역을 보정한 영상들, 그림 12(c)는 얼굴이 아닌 팔 영역에서 얼굴영역과 유사한 관계를 가져 오검출된 영상을 나타내며 그림 12(d)는 칼만 예측기로 얼굴 영역이 존재 가능한 위치를 예측하여 얼굴 이외의 영역을 제외하여 검출한 결과를 나타낸다.

움직임이 없는 경우는 블록차영상 만을 이용하면 기본적으로 얼굴검출이 불가능하다. 이러한 경우 칼만예측기는 얼굴영역을 예측하고 예측영역에서 이전에 검출된 얼굴영상과의 유사도를 비교하는 방법을 사용하기 때문에 얼굴영역 검출이 가능하다. 그림 13(a)는 사람의 움직임이 미약하여 차

영상이 얻어지지 않는 경우이고 그림 13(b)는 칼만예측기를 이용하여 보정한 결과를 보여준다.

또한 칼만예측기는 얼굴영역의 중심점과 더불어 너비와 세로 길이 정보를 예측하여 갱신하기 때문에 블록차영상을 이용한 얼굴검출방법에서 나타나는 매우 가변적인 길이의 얼굴검출을 보정하는 기능을 갖게 된다. 그림 14(a)는 블록차영상을 이용하여 검출한 얼굴영역으로 실제의 얼굴영역보다 크게 검출하고 있음을 보여주고 있으며 그림 14(b)는 칼만예측기를 이용하여 크기 오류를 보정한 결과를 보여준다. 칼만예측기를 이용하여 연속적으로 얼굴영역을 검출하는 경우 얼굴이 아닌 다른 영역에서의 검출을 사전에 방지를 하며 원하는 어느 순간에도 얼굴의 정보를 가질 수 있는 장점이 있다.

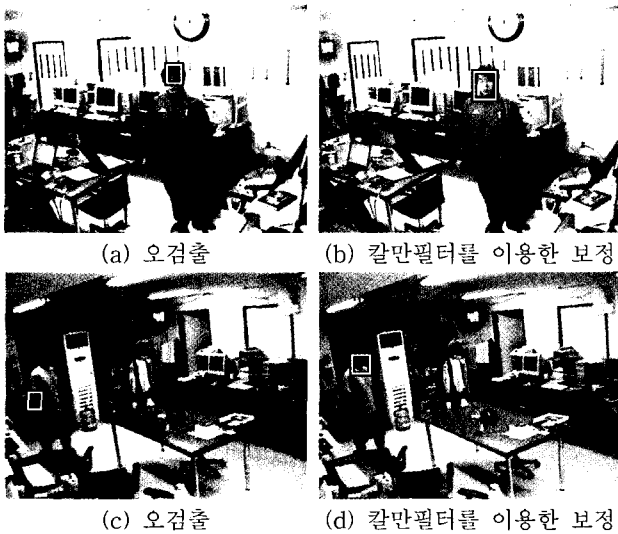


그림 12. 칼만 필터를 이용한 오검출 보정 결과.

Figure 12. Compensating a position error of the face region with Kalman filter.



그림 13. 칼만예측기를 이용한 연속적인 검출 보정 결과.

Figure 13. Compensating the position error of the face region in moving pictures with Kalman filter.

제안된 알고리즘의 성능을 분석해보기 위해서 제안된 알고리즘에서 칼만예측기를 사용하지 않은 경우와 사용한 경우에 동영상에서 몇 프레임정도 연속해서 얼굴영역을 정확하게 검출하는지를 표 1에서 비교하여 보았다. 두 경우 모두 피부색상만을 이용한 기존의 알고리즘에 비해 우수한 성능을 가지며 특히 칼만예측기를 사용한 경우는 현저한 성능향상을 보임을 알 수 있다.

얼굴검출의 정확도를 검증하기 위해 그림 15에서는 위의 표 1에서 비교한 세 가지 경우에서 대해 실제 얼굴영상의 크기와 측정된 크기와의 오차를 계산한 결과를 보여준다. 크기 오차는 세로 길이의 오차와 가로 길이의 오차의 제곱하여 합

으로 구성하였다. 가는 점선으로 표시된 것은 기존의 피부색상 영역과 차영상으로 검출을 하였을 때 크기 오차를 나타내며 평균 픽셀 오차의 제곱은 약 20으로 나타났으며 굵은 점선으로 표시된 것은 본 논문에서 사용한 일반적인 얼굴 검출 알고리즘의 결과를 나타내며 평균 오차는 약 10으로 나타났으며 실선으로 표시된 것은 칼만예측기를 사용했을 때의 결과를 나타내며 평균 오차는 약 8로 나타났다. 그림 15와 같이 기존의 피부색상과 차영상을 이용한 알고리즘의 경우와 일반적인 얼굴 검출 알고리즘만을 사용했을 때는 연속적인 검출이 보장되지 않았고 특히 고정된 피부 색상과 차영상만을 사용하는 기존의 알고리즘의 경우는 크기 오차가 매우 크게 나타났으며 본 논문에서 사용한 얼굴 검출 알고리즘으로 오차가 보정된 것을 알 수 있었으며 칼만예측기를 사용했을 때 연속적인 얼굴 영역의 검출이 보장되며 크기 오차가 대부분 줄어들어 드는 것을 알 수 있다.

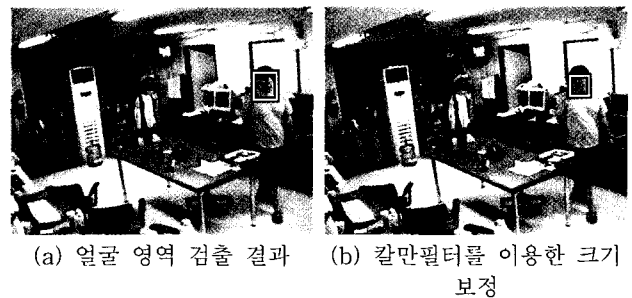


그림 14. 칼만 필터를 이용한 크기 보정 결과.

Figure 14. Compensating a size error of the face region with Kalman Filter.

표 1. 일반적인 얼굴 검출과 칼만 필터를 사용했을 때 검출 연속성 비교.

Table 1. Comparing the continuity of the face tracking between the conventional method and the proposed methods.

비교 항목	알고리즘 종류 기존의 피부색상 과 차영상기반 알고리즘	제한한 알고리즘	
		블록차영상 만 사용	칼만예측 기 사용
연속검출 최소프레임 수	1	1	17
연속검출 최대프레임 수	8	38	148
연속검출 평균프레임 수	3.2	5.7	49

또 다른 성능의 평가지수로 실시간성을 입증하기 위해 얼굴영역의 검출에 소요되는 시간을 그림 16에서 비교하였다. 그림에서 알 수 있듯이 칼만예측기를 사용하는 경우 프레임당 평균 연산시간은 0.015초로 이를 사용하지 않고 일반적인 검출 알고리즘으로 검출한 경우에 비해 약 40%의 연산시간이 절약되어 초당 30프레임 이상의 영상을 처리할 수 있음을 보여주고 있다. 그림 16에서 얼굴 검출 알고리즘을 매번 프레임당 검출한 시간과 칼만 필터를 이용하여 얼굴 영역을 예측하고 예측된 영역에서 얼굴 검출을 수행한 시간을 비교하였다. 가는 점선으로 표시된 것이 기존의 피부 색상을 이용한 검출 방법을 사용해 검출한 결과로 평균 프레임당 0.1초의 연산시간을 보였으며 굵은 점선으로 표시된 것이 블록 차영상을 이용한 얼굴 검출 알고리즘으로 검출한 결과로 평균

0.024초의 연산시간을 보였고 실선으로 표시된 것이 칼만 필터를 이용하였을 때의 결과로서 평균 0.015초로 감소된 연산 시간을 나타낸다.

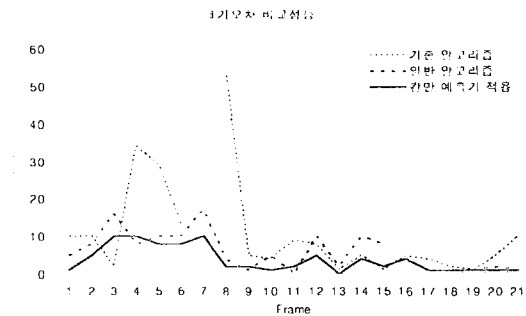


그림 15. 일반적인 얼굴 검출과 칼만 필터를 이용하였을 때 크기오차 비교: X축은 프레임 번호, Y축은 오차 화소수의 제곱 형태.

Figure 15. Comparing the size error of the face region between the conventional method and the proposed methods.

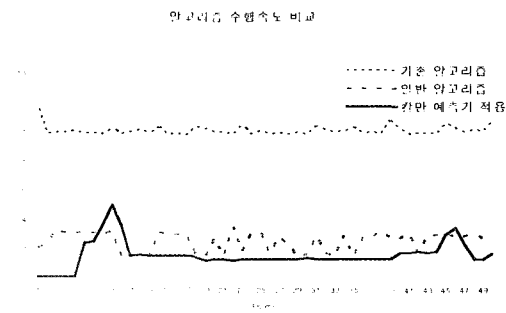


그림 16. 일반적인 얼굴 검출과 칼만 필터를 이용하였을 때 연산시간 비교: X축은 프레임번호, Y축은 시간(초).

Figure 16. Comparing the processing time between the conventional method and the methods.

5. 결론

본 논문은 복잡한 환경에서 이동하는 사람의 얼굴영역을 조명의 변화에 상관없이 정확하고 빠르게 검출하기 위해 블록차영상과 칼만예측기를 이용하는 방법을 제안하였다. 사람의 미세한 움직임도 검출 가능하도록 블록차영상을 생성하여 impulse 잡음을 제거함으로써 불필요한 전처리 과정을 피함으로써 연산량을 효과적으로 줄일 수 있었다. 또한 피부색상을 이용하여 얼굴영역을 검출함에 있어서 움직임 정보가 있는 영역에서의 피부색상만 고려를 하며 YIQ 컬러 공간에서 I와 Q성분만을 고려하기 때문에 배경과 조명의 변화에도 영향을 덜 받는다. 검출성능을 보였다. 추가적으로 조명의 영향으로 피부색상이 검출되지 않는 경우를 대비하여 블록차영상을 부호화하여 검출된 머리영역에서만 얼굴영역을 검색하는 방법을 사용함으로써 오검출의 가능성을 크게 낮추었다. 또한 본 논문에서는 동영상에서의 움직임을 평면적으로 연결된다는 점에 착안하여 다음 프레임에서 얼굴영역이 존재할 수 있는 영역을 칼만예측기를 이용하여 예측하고 그 범위 내에서

얼굴영역을 검색하는 방법을 적용하였다. 이와 같은 방법으로 지속적인 얼굴 검출이 가능하였고 오검출 가능성이 크게 줄어들어 정확한 얼굴영역의 검출이 가능함을 실험적으로 증명하였다. 실험결과 제안하는 방법은 기존의 방법에 비해 정확도에서는 평균 2.5화소 정도로 오차를 줄일 수 있었고 측정시간 면에서는 40%정도의 축소가 가능하여 초당 30프레임 이상의 속도에서도 적용 가능함을 입증하였다.

참 고 문 헌

- [1] Bowyer, K.W, "Face recognition technology: security versus privacy," *Technology and Society Magazine, IEEE*, Volume. 23, pp. 9-19, 2004.
- [2] M.Turk and A.Pentland, "Eigenfaces for Recognition," *J.Cognitive Neuroscience*, vol.3, no.1, pp. 71-86,1991.
- [3] A.Lanitis, C.J. Taylor, and T.F. Cootes, "An Automatic Face Identification System Using Flexible Appearance Models," *Image and Vision Computing*, vol. 13, no.5, pp. 393-401, 1995.
- [4] Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces in Images : A Survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions*, Vol.24, No 1, pp.34-58, 2002.
- [5] G. Yang and T.S. Huang, "Human Face Detection in Complex Background," *Pattern Recognition*, vol.27, no.1, pp. 53-63, 1994.
- [6] Rein-Lien Hsu, Abdel-Mottaleb, and M. Jain, A.K, "Face detection in color images," *Pattern Analysis and Machine Intelligence, IEEE Transactions, PAMI*, Vol.24, No 5, pp. 696-706, May 2002.
- [7] P. Sinha, "Object Recognition via Image Invariants : A Case Study," *Investigative Ophthalmology and Visual Science*, vol. 35, no. 4, pp. 1735-1740, 1994.
- [8] K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," *IEEE Trans. pattern Analysis and Machine Intelligence*, vol.20, no.1, pp.39-51, Jan. 1998.

저 자 소개



장희준(Heejun Jang)

2003년 숭실 대학교 전자정보통신 공학과 졸업

2005년 동대학 석사졸업

관심분야: 영상처리, MPEG, Vision System

E-mail : sweetga@visionlab.ssu.ac.kr



고혜선 (Hyesun Ko)

2003년 숙명여자대학교 정보과학부 학사 졸업
2005년 숙명여자대학교 컴퓨터과학과 석사 졸업
관심분야 : 영상처리, 컴퓨터비전

E-mail : hyesun99@sookmyung.ac.kr



한영준 (Youngjoon Han)

제 14권 7호(2004년 12월호) 참조



최영우 (Yeongwoo Choi)

1985년 연세대학교 전자공학 학사 졸업
1986년 University of Southern California
컴퓨터공학 석사 졸업
1994년 University of Southern California
컴퓨터공학 박사 졸업
1994년 11월-1997년 2월 LG전자기술원
선임연구원

1997년 3월-현재 숙명여자대학교 정보과학부 컴퓨터과학전
공 교수

관심분야 : 영상처리, 패턴인식, 문서인식

E-mail : ywchoi@sookmyung.ac.kr



한헌수 (Hensoo Hahn)

제 13권 4호(2003년 8월호) 참조