

論文98-35C-11-6

# 무인 영상 감시 시스템을 위한 실시간 얼굴 영역 추출 알고리즘

(Real-time Face Localization for Video Monitoring)

朱 映 炫 \* , 李 定 勳 \*\* , 文 泳 植 \*\*\*

(Young Hyun Joo, Chung-Hoon Rhee, and Young Shik Moon)

## 요 약

본 논문에서는 영상 감시 시스템을 위한 이동물체 추적 및 얼굴 영역 추출 알고리즘을 제안한다. 제안된 방법은 두 단계로 구성되어 있으며, 첫 번째의 이동 물체 감지 단계에서는 각각의 입력 프레임에 대하여 화소값의 변화를 기반으로 한 세 종류의 특징값들을 추출한 다음, 그 값들에 의하여 입력 프레임을 다섯 종류의 클래스로 분류한다. 분류 결과 현재의 프레임이 이동물체를 포함하고 있다면 두 번째의 얼굴 영역 분리 단계에서 컬러 정보와 프레임간의 차이를 이용하여 그 얼굴 영역을 추출하게 된다. 제안된 알고리즘은 광류 등 계산시간이 많이 걸리는 특징값들에 의존하지 않으므로 실시간 처리에 적합하다. 또한 여러 테스트 영상 시퀀스에 대하여 실험을 수행한 결과 본 논문에서 제안된 알고리즘의 유용성이 입증되었다.

## Abstract

In this paper, a moving object detection and face region extraction algorithm which can be used in video monitoring systems is presented. The proposed algorithm is composed of two stages. In the first stage, each frame of an input video sequence is analyzed using three measures which are based on image pixel difference. If the current frame contains moving objects, their skin regions are extracted using color and frame difference information in the second stage. Since the proposed algorithm does not rely on computationally expensive features like optical flow, it is well suited for real-time applications. Experimental results tested on various sequences have shown the robustness of the proposed algorithm.

## I. 서 론

영상 감시 시스템은 교통량 감시, 출입자 확인, 그

\* 正會員, (株) 팬타그램

(Phantagram co. Ltd.)

\*\* 正會員, 漢陽大學校 電子工學科

(Department of Electronic Engineering, Hanyang University)

\*\*\* 正會員, 漢陽大學校 電子計算學科

(Department of Computer Science and Engineering, Hanyang University)

※ 이 논문은 1995년도 한국학술진흥재단의 대학부설

연구소 연구과제 연구비에 의하여 연구되었음

接受日字: 1998年4月20日, 수정완료일: 1998年10月29日

리고 공공 장소에서의 범죄 예방 등 많은 부분에서 인력을 보조하거나 대체하는데 유용하게 사용될 수 있다. 특히 전자기술의 급속한 발전에 따라 산업계의 공장 자동화나 빌딩 자동화 시스템에서의 영상 감시 시스템의 사용은 이미 일반화되었고, 그 외 일반 사회 생활의 여러 분야, 즉 아파트 관리 시스템이나 현금 입출기, 수퍼마켓, 편의점, 교차로 등의 무인 경비 및 관리 시스템에서의 영상 감시 시스템의 사용도 빠른 속도로 확산되어 가고 있다. 따라서 보다 효율적인 영상 감시 시스템을 위한 알고리즘들의 필요성은 점차 커지고 있다. 이러한 영상 감시 시스템이 사용되는 환경은 일반적으로 실험실에서처럼 잘 조정된 환경으로 만들기 어렵고 예외 상황 발생시 즉각적인 대응이 요구되는 경우가 많으므로, 영상 감시 시스템에 사용되는 알고리

즘은 환경의 변화에 대처하는 능력을 갖는 것이 요구되며 실시간으로 동작하는 것이 바람직하다. 본 논문에서는 이러한 영상 감시 시스템에 효과적으로 적용될 수 있는 효율적이며 환경 변화에 민감하지 않은 영상 분류 및 얼굴 영역 분리 알고리즘에 대해 기술한다. 제안된 알고리즘은 화소값에 기반한 특징값들의 시간에 따른 변화를 이용하여 배경 영상을 획득하고 그 배경을 참조 영상으로 사용해 각각의 입력 프레임들을 배경, 이동물체, 플래시, 배경변화, 조명변화의 다섯 종류로 구분한 다음 이동물체가 존재한다면 색상과 차영상에 근거한 특징값들을 이용하여 그 얼굴 영역을 추출한다. 또한 환경 변화에 따른 오류를 막기 위해 영상 분류 파라미터들은 필요한 경우 적응적으로 조정되며, 이 모든 과정이 실시간 처리가 가능한 시간범위 내에서 수행된다.

일반적인 연속영상 처리 알고리즘은 연속영상 중 각각의 프레임에서 특정한 특징값들을 추출하여 그 값의 연속된 프레임 사이에서의 변화의 패턴을 분석하여 각각의 작업을 처리하게 되는데, 이때 주로 사용하는 특징값에 따라 히스토그램 기반 방법<sup>[1]</sup>, 화소값에 기반한 방법<sup>[2]</sup>, 그리고 그 외의 특징값들을 이용하는 방법<sup>[3]</sup> 등이 있다. 히스토그램에 기반한 방법들은 각각의 프레임에 대해 밝기 히스토그램이나 색상 히스토그램을 계산하여 그 히스토그램의 프레임간 차이를 이용하는데, 화소값 기반 방법에 비해 카메라와 물체의 움직임에 비교적 덜 민감하여 장면 변화 검출 등의 응용 분야에 많이 사용되고 있으나 물체의 형태, 위치, 이동 방향 등 공간적인 정보를 얇게 되는 단점이 있어 일반적으로 영상 감시 분야에 사용하기에는 적절하지 않다. 화소값에 기반한 방법들은 여러 프레임 사이에서의 화소값들의 차이를 직접 이용하며, 물체에 관한 형태 등의 상대적으로 많은 정보를 얻을 수 있는 반면 짧음이나 카메라 움직임, 조명 변화 등에 매우 민감하게 반응하는 단점이 있다. 그 외에 많이 사용되는 특징값으로는 에지 정보, 광류, 색상 등이 있으며, 특히 광류는 연속 영상에서 얻을 수 있는 새로운 종류의 정보로 이동 물체 감지 및 추적을 비롯한 여러 연속 영상 처리에 있어 대단히 유용하게 사용될 수 있는 반면 계산 시간이 많이 걸려 실시간 처리를 위해서는 일반적으로 전용 하드웨어를 필요로 하는 단점을 갖고 있다.

채옥삼, 이정현 등<sup>[1]</sup>은 옥외에서의 사용을 고려한 차영상과 다단계 확장에 기반한 침입자 감지 및 추적

알고리즘을 개발하였다. 이 방법은 차영상에서 이동 물체의 마스크를 얻기 위해 차영상의 히스토그램을 분석하여 seed 임계값을 결정하고 이치화를 통해 seed 영상을 얻은 다음, 임계값을 점차적으로 낮추어 가면서 이동물체 영역을 확장시켜 나가는 방식을 취하였다. 이 방법에서 다단계 확장을 끝내기 위해서는 확장된 영역의 면적과 둘레, 밝기값에 따른 휴리스틱한 조건을 사용하고 있다. 이 방법은 차영상만을 사용하여 이동 물체가 배경과 비슷한 밝기 영역에 존재하는 경우도 처리할 수 있는 반면 이동 물체에 대한 면적, 둘레 등의 알려진 정보에 의존하며, 또한 다단계 확장 부분에서 면적과 둘레 등의 비교적 시간이 많이 걸리는 계산을 필요로 한다는 단점을 갖는다. Nagaya, Miyatake 등<sup>[2]</sup>은 옥외에서의 환경 변화를 고려한, 화소값에 기반한 이동물체 감지 방법을 개발하였다. 이 방법에서는 화소값들의 시간축 상에서의 변화를 이용한 time correlation, background correlation, normalized distance라는 세 종류의 특징값들과 판단 규칙을 사용하여 각각의 프레임을 분류해 배경 영상을 얻고 그 배경 영상과의 차를 이용해 이동물체의 존재 여부를 판단한다. 또한 그들의 연구에서는 이동 물체 영역의 분리를 위해 전체 영역을 다수의 같은 크기의 직사각형 영역들로 분할한 다음 각각의 영역에 대해 이 방법을 적용해 이동 물체가 감지된 영역들을 이동 물체 영역으로 인지하였다. 이때 분할하는 영역의 크기가 너무 작은 경우 이동물체 감지에 필요한 통계적 특성을 확보하지 못하며 너무 큰 경우 실제 이동물체의 영역과 거리가 있으며 하나의 영역 내부에서 오랫동안 움직이는 물체가 존재할 경우 이를 처리하지 못하므로 이 크기를 주의깊게 설정하여야 한다. 김희태, 김상철 등<sup>[3]</sup>은 누적 차영상의 특성과 다단계 영역 확장에 기초하고 카메라의 흔들림을 고려한 무인 감시 알고리즘을 제안하였다. 그들은 입력 영상에 대해 배경 영상에서의 차영상을 계산해 그 차영상의 평균과 표준편차로부터 각각의 프레임을 분류해 이동 물체를 감지한다. 이동 물체가 감지되면 다단계 영역 확장을 통해 이동 물체 영역을 추출한다. 이 방법은 [1]의 연구에서와 같은 면적이나 둘레 등의 특징값 대신 차영상의 평균과 표준 편차에 의거한 고정된 횟수의 영역 확장 방식을 취하여 보다 고속화를 추구하였다. 또한 입력 영상을 보다 낮은 해상도로 블록화해 이동 물체 영역을 구하고 다시 확대하여 마스크로 사용함으로

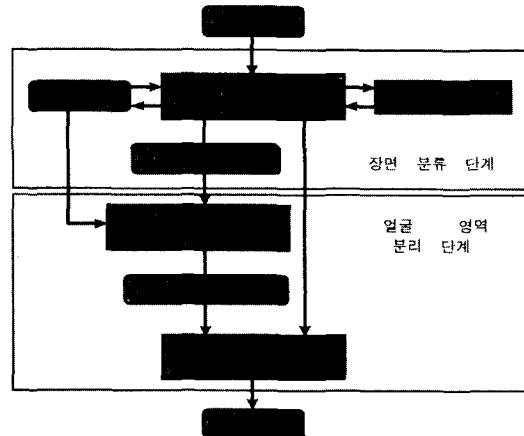
써 카메라의 흔들림을 흡수함과 동시에 고속화를 추구하였다. 이때 블록의 크기는 역시 카메라의 흔들림은 흡수되며 이동물체의 감지를 방해하지 않도록 주의깊게 설정해야 한다.

얼굴 영역의 판별 및 추출에 관해서도 최근 활발한 연구가 이루어지고 있다. Dai, Nakano<sup>[4,5]</sup>는 Space gray-level dependence(SGLD) matrix<sup>[6]</sup>를 이용한 얼굴-텍스처 모델과 색상 정보를 이용한 컬러 영상에서의 자동적인 얼굴 인식 방법을 제안하였다. 그들은 먼저 YIQ 색상 모델에 근거한 색상 정보를 사용하여 얼굴 후보 영역들을 얻은 다음 각각의 후보 영역에 대하여 SGLD matrix에 의한 판정으로 얼굴 영역들을 가려내었다. 특히 그들은 색상 정보의 활용으로 인하여 복잡한 배경 상에서의 얼굴 후보 영역 추출의 신뢰성이 높아졌으며 시간이 많이 걸리는 SGLD matrix 매칭의 대상 영역을 줄이고 필요한 경우에만 수행할 수 있게 되어 전체적인 시스템의 성능이 대폭 향상되었음을 보고하였다. Sobottka, Pitas<sup>[7]</sup>는 색상과 얼굴의 형태 정보를 이용한 얼굴 영역 판별 및 얼굴 특징 추출 방법을 제안하였다. 그들의 방법에서는 HSV 색상 모델을 사용하여 컬러 영상을 HSV 모델로 변환한 다음 H와 S의 범위에서 피부색 영역을 얻고 각각의 영역에 대하여 모멘트에 의한 best-fit ellipse를 구하여 적합한 얼굴 후보영역을 판정한 다음 그 후보영역에서 일반적인 투영(projection) 등의 방법으로 얼굴 특징들을 얻는 방식을 취하였다.

## II. 제안된 방법

본 논문에서는 감시 환경의 특성을 살려 실시간으로 적용될 수 있는 장면 분류 및 얼굴 영역 분리 방법을 제안한다. 제안하는 방법은 크게 장면 분류와 얼굴 영역 분리의 두 단계로 구성되어 있으며 각각 화소값의 차이와 색상 정보에 기반한 특징값들을 사용한다. 장면 분류 단계에서는 Nagaya 등의 연구<sup>[2]</sup>에서 도입된 두 종류의 특징값들과 장시간 움직이는 물체에 대한 인식 오류를 보완하기 위하여 본 논문에서 추가로 도입된 특징값을 포함한 세 종류의 화소값 기반 특징값들을 사용하며, 그 값들을 이용하여 각 입력 프레임들을 배경, 이동 물체, 플래시 조명, 배경 변화, 조명 변화의 다섯 클래스로 분류한다. 장면 분류 결과 배경 변화 또는 조명 변화가 감지되면 배경 영상을 그에 따

라 자동적으로 갱신한다. 또한 분류의 신뢰성을 높이기 위하여 장면 분류에 필요한 파라미터들 중 필요한 파라미터에 대하여 자동적인 조정을 수행한다. 장면 분류 단계에서 이동 물체가 감지된다면 다음 단계에서 그 이동 물체가 인간인 경우 배경 영상과의 RGB 차 영상과 YIQ 색상 모델의 피부색상 추출 효과를 이용하여 그 얼굴 영역의 마스크를 얻는다. 또한 얼굴 영역이 얻어진다면 그 위치의 변화로부터 적당한 키 프레임을 각각의 이동 물체 구간마다 하나씩 얻으며 얼굴 영역을 얻을 수 없는 경우 배경 영상과의 차이를 이용하여 키 프레임을 얻는다. 제안된 장면 분류 및 얼굴 영역 분리 방법의 전체적인 구조를 그림 1에 나타내었다.



을 갖는다. 그러나 배경 구간에서도 카메라 잡음 등의 영향으로 인해  $\delta(t)$ 가 정확히 0이 되는 경우는 거의 없으므로 0에 가까운 배경 임계값  $\theta_{bs}$ 와 비교하게 된다.  $\delta(t)$ 가  $\theta_{bs}$ 보다 클 경우는 배경이 변화했거나 이 동물체 또는 조명 등으로 인해 장면이 변화한 것을 뜻 한다.

① 배경 거리	$\delta(t) = \ \vec{r}(t) - \vec{\beta}\ _2$
② 프레임간 거리	$\delta_n(t) = \left\  \frac{\vec{r}(t)}{\ \vec{r}(t)\ } - \frac{\vec{\beta}}{\ \vec{\beta}\ } \right\ _2$
③ 정규화 거리	$\delta'(t) = \ \vec{r}(t) - \vec{r}(t-1)\ _2$
여기서	
$\vec{r}(t)$	: 시간 t에서의 영상
$\vec{\beta}$	: 배경 영상

그림 2. 사용된 특징값

Fig. 2. Features for the scene classification.

두 번째 특징값은 프레임간 거리로서, 현재의 프레임 영상  $\vec{r}(t)$ 과 직전의 프레임 영상  $\vec{r}(t-1)$  간의 유클리드 거리로 정의되며 역시 현재 프레임과 직전 프레임 사이의 차영상을 벡터로 나타내었을 때의 벡터 길이에 해당한다. 프레임간 거리  $\delta(t)$ 는 현재의 연속 영상에서 변화가 일어나지 않고 있으면 0에 가까운 값을, 변화가 일어나고 있다면 비교적 큰 값을 갖는다. 따라서 적당한 장면 변화 임계값  $\theta_{sc}$ 와 비교하여 현재 연속영상에서 물체 이동, 카메라 이동, 조명 변화 등 변화가 일어나고 있는지 여부를 알 수 있다. 본 논문에서는 보다 신뢰성을 높이고 잡음의 영향을 줄이기 위해 하나의 값을 사용하지 않고 1초 이상의 시간 구간 W를 정의하여 그 구간 동안의 프레임간 거리의 최대값을  $\theta_{sc}$ 와 비교한다.

세 번째 특징값인 정규화 거리는, 변화가 일어났을 경우 그 변화가 전체적인 조명 변화에 의한 것인가 아니면 부분적인 구조적 변화에 의한 것인가를 측정한다. 정규화 거리  $\delta_n(t)$ 는 배경 영상 벡터와 현재 프레임 벡터를 각각 그 길이로 나누어 단위원 상에 투영한 다음 그 단위 원에 투영된 두 벡터 사이의 유클리드 거리로 정의된다. 만약 전체적인 조명에 의한 변화가 발생하였다면 그 변화는 밝기값이 saturate되지 않았다

는 가정하에는 프레임 내의 모든 화소에 균등하게 영향을 끼칠 것이고 따라서 프레임 벡터의 길이에 주로 영향을 주며 방향에는 영향이 적을 것이다. 따라서 단위원 상에 투영된 벡터의 거리는 그다지 변화가 없으므로  $\delta_n(t)$ 값은 0에 가까운 값을 갖는다. 반면 이동물체에 의한 변화는 화소의 위치에 불균일한 변동을 가져오므로 벡터의 방향에 영향을 끼치고 따라서 투영된 벡터들 사이의 거리도 멀어질 것이다. 이 경우  $\delta_n(t)$ 값은 비교적 큰 값을 갖게 된다. 역시 본 논문에서는 보다 높은 신뢰성을 얻기 위해 변화가 일어나지 않고 있는 경우, 구간 W 동안의  $\delta_n(t)$ 의 변화량을 측정해 이를 조명변화 임계값  $\theta_{ii}$ 와 비교하여 그 결과에 따라 전체적인 조명 변화와 그 외의 경우를 구분한다. 장면 분류를 위한 알고리즘이 그림 3에 가상 코드 형태로 나타나 있다.

```

if  $\delta(t) \leq \theta_{bs}$  then /* 첫 번째 단계 */
[background period]
elseif  $\text{Max}_{u=t-W}^t (\delta'(u)) < \theta_{sc}$  then /* 두 번째 단계 */
    update background with current frame;
    if  $(\text{Max} - \text{Min})_{u=t-W}^t (\delta_n(u)) < \theta_{ii}$  then
        [background illumination change]
    else
        [background structure change]
    end;
else /* 세 번째 단계 */
    if  $\delta_n(t) \leq \theta_{ii}$  then
        [sudden illumination change]
    else
        [moving object period]
    end
end;

```

그림 3. 장면 분류 알고리즘

Fig. 3. Scene classification algorithm.

첫 번째 단계에서는 입력된 프레임을 현재의 배경 영상과 비교하여 그 프레임이 배경 구간에 있는지 여부를 알아보게 된다. 여기서 사용하는 특징값은 배경 거리  $\delta(t)$ 로, 이 값과 배경 변화 임계값  $\theta_{bs}$ 를 비교하여  $\delta(t)$ 값이  $\theta_{bs}$ 값 보다 작다면 이 프레임은 배경 구간에 있는 것으로 간주되어 분류가 끝나며  $\delta(t)$ 값이 크다면 이 프레임은 배경이 아닌 것으로 간주되어 다음 단계에서 보다 세부적인 분류를 수행한다. 이 단계에서는 현재 프레임에 대한  $\delta(t)$ 값을 직접 비교하므로 배경 구간과 다른 구간의 전환은 시간 지연 없이 즉각

적으로 일어난다.

두 번째 단계, 즉  $\delta(t) > \theta_{bs}$  인 경우는 현재 프레임이 배경 영상과 다른 경우인데, 이때는 현재 연속된 프레임들 사이에 변화가 급격히 일어나고 있는가 아니면 변화가 거의 없는 상태인가를 프레임간 거리  $\delta(t)$ 를 이용하여 판별한다. 이 단계에서는 현재의  $\delta(t)$ 값을 직접 사용하지 않고 임의의 W 프레임 길이의 측정 구간을 도입하여 그 구간 내에서의  $\delta(t)$ 의 최대값과 장면 변화 임계값  $\theta_{sc}$ 의 크기를 비교한다. 여기서 측정 구간을 사용하지 않고  $\delta(t)$ 값을 직접 사용한다면 이동 물체의 움직임이 짧은 시간동안 정지한 것처럼 보여 일시적으로  $\delta(t)$ 값이 작아진 경우에도 변화가 없는 상태로 판단할 것이다. 이 때는 배경 영상에 변화가 일어난 것으로 판단되어 현재의 프레임을 이후 새로운 배경 영상으로 갱신하게 되므로 이후 수십 프레임 이상에 걸친 연속적인 분류 에러를 야기할 것이다. 또한 연속적인 영상이 일어지는 간격은 대부분의 경우 0.1초 이하로 이러한 짧은 시간동안 일시적으로 물체의 이동이 정지하는 것은 드물지 않으며 이 때마다 배경 영상을 잘못 갱신하는 것은 바람직하지 않다. 이 문제를 해결하기 위하여 제안된 방법에서는 1초 이상의 비교적 긴 측정 구간을 도입하였다. 이 단계에서 측정 구간 동안의  $\delta(t)$ 의 최대값을 사용하므로 이 측정 구간 동안  $\delta(t)$ 값이  $\theta_{sc}$ 값 이하를 유지하면 점진적인 배경 변화가 일어난 것으로 간주하여 배경을 갱신한 후, 정규화 거리  $\delta_n(t)$ 와 조명 변화 임계값  $\theta_{il}$ 을 비교하여 조명의 변화인지 배경 자체의 변화인지를 판별한다.

세 번째 단계, 즉 일정 구간 동안  $\delta(t)$ 의 최대값이  $\theta_{sc}$ 값보다 큰 급격한 변화 구간에서는, 배경 영상과 차이가 생긴 것이 급격한 조명 변화에 의한 것인가 이동 물체에 의한 것인가를 정규화 거리  $\delta_n(t)$ 와  $\theta_{il}$ 을 비교하여 판별한다.  $\delta_n(t)$ 의 값이  $\theta_{il}$ 보다 작다면 카메라 플래시 등의 순간적인 조명에 의한 변화로 간주되어 플래시 조명 구간으로 분류되며,  $\delta_n(t)$ 의 값이  $\theta_{il}$ 보다 크다면 이동 물체에 의한 변화로 간주되어 이동 물체 구간으로 분류된다.

## 2. 파라미터 조정

앞에서 사용된 특징값들은 배경 거리  $\delta(t)$ , 프레임 간 거리  $\delta(t)$ , 그리고 정규화 거리  $\delta_n(t)$ 의 세 종류

로, 제안된 방법의 결과에 영향을 미치는 파라미터는 각각의 특징값들에 대응하는  $\theta_{bs}$ ,  $\theta_{sc}$ ,  $\theta_{il}$ 의 세 임계값들과 측정 구간의 길이 W이다. 이들 중 조명 변화 임계값  $\theta_{il}$ 과 배경 변화 임계값  $\theta_{bs}$ 는 화소의 수에만 영향을 받으므로 영상의 해상도를 알면 적당한 값을 구할 수 있으며, 또한 측정 구간의 길이 W는 극단적인 속도로 움직이는 물체의 존재를 배제한다면, 비디오 샘플링 속도에만 영향을 받으므로 역시 적당한 값을 구할 수 있다. 또한 이들 값들은 시스템이 동작하는 동안에는 거의 일정하게 유지되므로, 시스템이 설치될 때 조작자에 의해 설정되면 이 값들의 자동적인 조정은 필요하지 않다. 그러나, 프레임간 거리  $\delta(t)$ 는 이동 물체의 평균 속도나 크기, 영상의 평균 밝기, 비디오 샘플링 속도 등 많은 변수의 영향을 받으므로 장면 변화 임계값  $\theta_{sc}$ 는 자동적인 조정을 필요로 한다. 표 1은 여러 연속영상(student, game, door)에 대하여 각 파라미터들의 값을 바꾸어 가면서 장면 분류 결과를 측정한 것으로, 세 파라미터 중  $\theta_{bs}$ ,  $\theta_{il}$ 은 연속영상의 종류에 별로 영향을 받지 않고 비슷한 값에서 가장 좋은 결과를 보였으며,  $\theta_{sc}$ 는 최적의 결과가 연속영상에 따라 다르게 나타나 앞의 가정에서 벗어나지 않는 것을 볼 수 있다.

표 1. 파라미터 변화에 따른 장면 분류 결과  
(Test results with various parameter setting)  
(a)  $\theta_{bs}$  (b)  $\theta_{il}$  (c)  $\theta_{sc}$  (언급하지 않은 경우의  $\theta_{bs}$ ,  $\theta_{il}$ ,  $\theta_{sc}$ 의 값은 각각 2%, 5%, 8%임)

Table 1. Test results with various parameter setting.  
(a)  $\theta_{bs}$  (b)  $\theta_{il}$  (c)  $\theta_{sc}$

	student		game		door	
	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임
(a)						
$\theta_{bs} = 1\%$	457	43	71	29	452	48
$\theta_{bs} = 2\%$	471	29	71	29	452	48
$\theta_{bs} = 4\%$	470	30	58	42	427	73
$\theta_{bs} = 8\%$	450	50	45	55	393	107

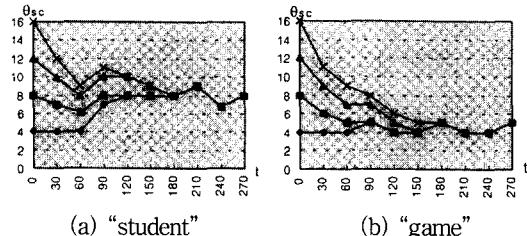
	student		game		door	
	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임
(b) $\theta_{il} = 2\%$	471	29	71	29	364	136
$\theta_{il} = 3\%$	471	29	71	29	452	48
$\theta_{il} = 5\%$	463	37	71	29	452	48
$\theta_{il} = 8\%$	449	51	71	29	452	48

	student		game		door	
	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임	바르게 분류된 프레임	잘못 분류된 프레임
(c) $\theta_{sc} = 4\%$	459	41	74	26	443	57
$\theta_{sc} = 8\%$	471	29	71	29	452	48
$\theta_{sc} = 12\%$	463	37	32	68	416	84
$\theta_{sc} = 16\%$	453	47	28	72	400	100

본 논문에서 사용하는 자동적인 파라미터 조정 방법은 통계적인 특성에 의거한 것으로, 어느 정도의 시간 동안 파라미터 변화를 기록하여 히스토그램을 계산한 다음 히스토그램 상에서 Otsu의 이치화 알고리즘<sup>[8]</sup>과 비슷한 방법으로 임계값을 선택하여 이전의 임계값과 조합해 새로운 임계값을 도출하게 된다.  $\delta(t)$ 는 동적적인 영역과 정적인 영역에서 서로 다른 통계적 특성을 가지므로, 그 히스토그램에 이치화 알고리즘을 적용한다. 여기서 히스토그램은 미리 정해진 간격(30 프레임)으로 계산하게 되며, 따라서 파라미터 조정도 그 간격으로 일어난다. 파라미터 조정은 이전의  $\theta_{sc}$ 값과 위에서 구해진 새로운 임계값에 각각 가중치를 가지고 더하여 다음 간격동안의  $\theta_{sc}$ 를 구하게 되는데, 여기서는 이전의 값에 0.7, 새로운 값에 0.3의 가중치를 주어 비교적 서서히 파라미터가 변화하도록 한다. 또한 시퀀스 자체가 대부분 동적인 구간으로 이루어진 경우는 정적인 영역이 거의 나타나지 않아  $\theta_{sc}$ 값이 비정상적으로 커질 우려가 있으므로  $\theta_{sc}$ 값이 가능한 범위의 16% 이상을 넘지 못하도록 제한하였다. 이상의 자동적인 파라미터 조정 방법을 두 연속영상에 대하여 적용 시킨 결과가 그림 4에 그래프로 나타나 있다. 각각 4종류의 서로 다른 초기  $\theta_{sc}$ 값에서 시간이 지남에 따라

$\theta_{sc}$ 값이 변화해 가는 것을 나타내었다. 초기값에 관계 없이 시간이 지남에 따라 같은 값으로 수렴함을 볼 수 있다.

그림 4. 시간에 따른  $\theta_{sc}$  값의 변화Fig. 4. Change of  $\theta_{sc}$  values with parameter tuning

### 3. 얼굴 영역 추출 방법

장면 분류 단계에서 이동물체의 존재가 감지된 프레임들에 대하여 이동물체 영역을 분리한다. 본 논문은 연속영상에서 이동하는 사람의 얼굴 영역을 올바르게 분리하는 것을 목적으로 하는데, 이는 이동물체 영역 추출과 얼굴 영역 추출의 두 과정으로 구성되며 이 두 과정에서 각각 얻어진 마스크들을 합성하여 최종적으로 목표하는 얼굴 영역 마스크를 얻는다. 이동물체 영역 추출 과정에서는 RGB 컬러 모델에서의 체스보드 거리 차영상을 사용하여, 구해진 차영상을 이치화하여 이동물체에 의한 변화 영역의 마스크를 추출한다. 컬러 성분들의 차이를 고려함으로써 밝기 성분의 차이만으로 변화 영역을 구하는 것에 비하여 이동물체에 의해 가려진 배경 부분과 이동물체의 밝기가 비슷한 경우 등에 더 나은 성능을 갖게 된다. 차영상의 이치화를 위한 임계값은 차영상의 평균값에 표준편차의 절반을 더 한 값을 사용하였다. 이동물체 영역 추출은 차영상에 기반한 것임에 반하여 얼굴 영역 추출은 색상 정보에 기반하고 있다. 앞에서 설명한 것과 같이 RGB 공간의 입력 프레임들을 YIQ 공간으로 변환한 다음 I 성분의 영상을 이치화하여 얼굴 영역의 마스크를 구한다. 이 과정에서의 이치화는 다른 여러 색상 정보를 사용한 얼굴 영역 추출 연구들<sup>[4,5,7]</sup>과 마찬가지로 인간의 피부색조를 대부분 포함하는 고정적인 값을 사용하였다.

앞의 두 과정에서 차영상에 기반한 마스크와 색상 정보에 기반한 마스크가 얻어지면 이 두 마스크를 논리 AND시켜 합성된 마스크 영상을 얻은 다음 모폴로지 연산자를 이용한 잡음 제거와 영역 확장 단계를 거쳐 최종적인 마스크를 얻는다. 이 단계에서 사용한 모

풀로지 연산은 이진 모풀로지 연산으로서, 고속 처리를 위하여 distance transform<sup>[9]</sup>을 이용하여 구현하였다. 이 방법은 기본적인 집합개념을 사용한 모풀로지 필터링 방법에 비하여 형태소 (structuring element)의 형태에 제한을 받으나 형태소의 크기에 무관한 수행속도를 가져 대단히 빠르게 처리된다는 장점을 갖는다. 본 논문에서 사용한 잡음 제거 필터와 영역 확장 필터는 다음과 같은 식으로 나타난다.

$$F(I) = \{(I \bullet S_1) \circ S_1\} \oplus S_2$$

여기서  $\bullet$ 은 불임(closing),  $\circ$ 은 열림(opening),  $\oplus$ 는 불립(dilation) 연산자이며,  $S_1$ 과  $S_2$ 는 각각 잡음 제거용과 영역 확장용  $7 \times 7$ 의 원형 형태소이다.

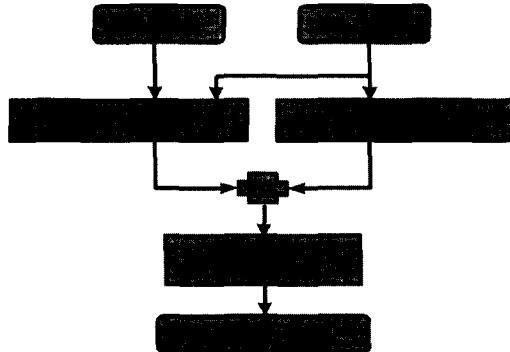


그림 5. 얼굴 영역 분리 과정

Fig. 5. Procedures for face localization.

### III. 실험 결과

제안된 알고리즘의 효과와 타당성을 입증하기 위하여 여러 연속영상들에 대하여 테스트하였다. 실험은 펜티엄 PC와 Windows 95 환경에서 Delphi 3.0을 사용하여 구현하였다. 입력영상은 흑백영상의 경우 Matrox사의 Meteor 보드를 이용하여 얻었으며 컬러 영상의 경우 별도의 시스템에서 얻어진 영상을 AVI 파일로 저장한 다음 MPEG 포맷으로 변환하여 사용하였다. 실험은 두 종류로 나누어 수행하였으며, 각각의 실험에서 장면 분류 부분의 성능과 전체적인 시스템의 성능을 측정하였다. 그림 6에 실험에 사용된 연속영상들에서 얻은 샘플 프레임들이 나타나 있다.

먼저, 제안된 방법의 장면 분류 알고리즘을 Nagaya 등의 장면 분류 방법<sup>[1]</sup>과 비교해 본 결과를 표 2에

나타내었다. 테스트한 모든 연속영상에서 제안된 방법 쪽이 보다 많은 이동물체 구간의 프레임들을 올바르게 분류하여 더 적은 에러율을 보임을 알 수 있다. 이 중 student 영상은 오랫동안 화면 내에서 움직이는 물체를 포함하는데 이때 Nagaya 등의 방법은 이를 배경 변화로 잘못 분류하여 이로 인한 연속적인 에러로 하나의 이동 물체 구간을 둘로 나누었고 또한 후속의 배경 구간을 또 다른 이동 물체 구간으로 인식하여 두 개의 오류가 나타난 것을 볼 수 있다. 실행 속도 면에서는 제안된 방법 쪽이 더 많은 특징값 사용으로 인하여 Nagaya 등의 방법에 비해 약간 느린 속도를 보였다.



그림 6. 실험 연속영상들 (a) student (b) game (c) door (d) 2man (e) sijong

Fig. 6. Test image sequences. (a) student (b) game (c) door (d) 2man (e) sijong

표 2. 장면 분류 실험 결과

Table 2. Test results for scene classification.

테스트 연속 영상	장면 분류 방법	이동 물체 프레임 수	이동 물체 구간 수			배경 변화			초당 프레임 수
			올바르 게 감지	missed	over detected	올바르 게 감지	missed	over detected	
student	제안된 방법	155/152	10	0	0	2	1	0	40.40
	Nagaya 등의 방법	169/159	10	0	1	2	1	1	42.12
game	제안된 방법	45/72	8	0	0	0	N/A	0	9.83
	Nagaya 등의 방법	44/72	8	0	0	0	N/A	0	10.21
door	제안된 방법	52/74	4	0	0	1	0	0	9.85
	Nagaya 등의 방법	50/74	4	0	0	1	0	0	10.27
2man	제안된 방법	229/225	5	1	0	1	0	0	15.89
	Nagaya 등의 방법	231/225	5	1	1	1	0	2	16.05
sijong	제안된 방법	70/87	4	0	0	0	N/A	0	15.92
	Nagaya 등의 방법	70/87	4	0	0	0	N/A	0	16.08

제안된 방법에 의한 얼굴 영역 추출 과정이 그림 7에 나타나 있다. 이동물체 영역과 얼굴 후보 영역으로부터 잡음 제거 및 영역 확장 과정을 거쳐 얼굴 영역이 추출되는 것을 볼 수 있다.

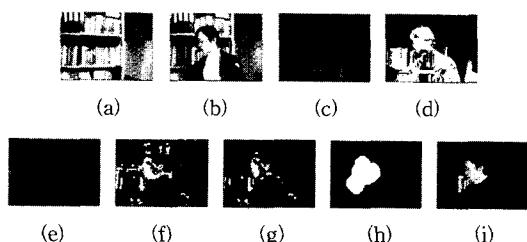


그림 7. 얼굴 영역 추출 과정

(a) 배경 영상 (b) 이동물체 영상 (c) a,b 의 차영상 (d) c 를 이치화한 마스크 (e) b 의 I-band 영상 (f) e 를 이치화한 마스크 (g) d ,f 를 논리 AND 한 초기 마스크 (h) g 에 잡음제거, 영역확대를 거친 최종 마스크 (i) 마스크 영역의 영상

Fig. 7. Steps for face localization.

(a) background image (b) image with moving objects (c) difference between a and b (d) binarized image of c (e) I-band image of b (f) binarized image of e (g) initial mask generated by ANDing d and f (h) facial mask obtained by dilation and noise suppression from g (i) localized facial region

제안된 알고리즘의 전체적인 테스트 결과를 표 3에 나타내었다. 두 번째의 테스트 연속영상에서 비교적 낮은 인식률을 보인 것은 사선 방향에서의 직접적인 조명으로 인하여 카메라의 자동 광량 조절 기능이 작동되어 피부색이 탈색되었기 때문이다. 카메라가 직접적인 조명에 노출되지 않은 다른 두 테스트 영상의 경우 90% 전후의 좋은 감지율을 보이고 있다. 또한 이 모든 과정이 실험 환경 내에서 실시간으로 수행되었다.

표 3. 전체적인 실험 결과

Table 3. Overall test results.

테스트 연속영 상	이동 물체 수	키 프레임 수	얼굴 영역				초당 처리 프레 임 수
			올바르게 감지	missed	over- detected	감지율	
2man	8	5	167	18	24	90.2%	5.98
sijong	4	4	37	12	0	75.5%	8.57
desk	1	1	99	11	20	90.0%	6.38

#### IV. 결 론

본 논문에서는 연속영상에서 실시간으로 이동물체의 존재를 감지하고 얼굴 영역을 추출하는 영상처리 알고리즘이 제안되었다. 제안된 방법은 크게 장면 분류와 얼굴 영역 분리의 두 부분으로 구성되어 있다. 장면 분류에서는 화소값의 변화에 근거한 특징값들을 이용하여 입력 영상 프레임을 다섯 클래스로 분류하고, 장면 분류의 결과 이동물체의 존재가 판명되면 컬러 차영상과 색상 정보를 이용하여 그 얼굴 영역 마스크를 찾는다. 제안된 방법은 계산량이 많지 않은 특징값들에 의존하므로 특별한 하드웨어 장비 없이 실시간 수행이 가능하며 환경변화에 적응적으로 파라미터가 조정되는 장점이 있다. 제안된 방법을 영상 감시 환경을 상정한 여러 테스트 영상들에 대하여 적용해본 결과 처리속도와 신뢰성 면에서 그 효율성을 확인할 수 있었다. 제안된 알고리즘은 영상 감시 환경과 같이 고정 카메라를 사용하는 분야에서의 실시간 처리에 적합하지만 카메라가 자유로이 움직이는 분야의 경우 장면 분류의 효율이 떨어진다는 점이 있으며, 이는 향후 연구되어야 할 과제이다.

#### 참 고 문 헌

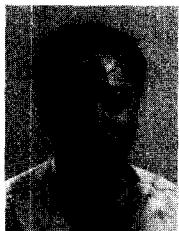
- [ 1 ] 채옥삼, 이정현, 안용학, 이성욱, “신경망을 이용한 침입자 감시 및 추적 시스템,” 제 8회 영상처리 Workshop 논문집, pp. 167-172, 1996.
- [ 2 ] S.Nagaya, T.Miyatake, T.Fujita, W.Ito and H.Ueda, “Moving Object Detection by Time-Correlation-Based Background Judgement,” Proc. of ACCV ’95, pp. II-717-722, 1995.
- [ 3 ] 김희태, 하영숙, 강경호, 박진권, 김상철, 임준석, 최태영, 김용득, “카메라의 혼들림을 고려한 무인 감시 알고리즘,” 1997년도 신호처리합동학술대회 논문집, pp. 127-130, 1997
- [ 4 ] Y.Dai and Y.Nakano, “Extraction of Facial Images from Complex Background using Color Information and SGLD Matrices,” International Workshop on Automatic Face and Gesture Recognition, pp. 238-242, 1995.
- [ 5 ] Y.Dai and Y.Nakano, “Face Texture Model Based SGLD and Its Application in Face Detection in a Color Scene,” Pattern

- Recognition, vol. 29 no. 6, pp. 1007-1017, 1996.
- [ 6 ] R.M.Haralick, "Textual Features for Image Classification," IEEE Trans. SMC, pp. 610-621, 1973.
- [ 7 ] K.Sobottka and I.Pitas, "Face Localization and Facial Feature Extraction Based on Shape and Color Information," Proceedings of ICIP, pp. 483-486, 1996.
- [ 8 ] N.Otsu, "A Threshold Selection Method from Gray-Level Histogram," IEEE Trans. SMC, pp. 62-66, 1979.
- [ 9 ] N.Otsu, "Distance Transforms in Digital Images." Computer Vision, Graphics, and Image Processing, no. 34, pp. 334-371, 1986.

## 저자소개

## 朱映炫(正會員)

1990년 3월 ~ 1996년 2월 한양대학교 전자계산학 학사.  
1996년 3월 ~ 1998년 2월 한양대학교 대학원 전자계산학과 석사. 1998년 1월 ~ 현재 (주) 팬타그램 연구원



## 李定勳(正會員)

1963년 3월 3일생. 1985년 5월 University of Southern California 전자공학사(BSEE). 1985년 8월 ~ 1987년 5월 University of Missouri 전자공학석사 (MSEE). 1990년 1월 ~ 1993년 5월 University of Missouri 전자공학박사 (Ph. D). 1994년 3월 ~ 1995년 8월 한국전자통신연구소 통신시스템 연구단 선임연구원. 1995년 9월 ~ 현재 한양대학교 전자공학과 조교수. 주관심분야는 컴퓨터비전, 패턴인식, 퍼지 시스템 및 신경회로망 모델링

## 文泳植(正會員) 第33卷 B編 第5號 參照

현재 한양대학교 전자계산학과 부교수