

실시간 비디오 시퀀스로부터 형태학적 영역 병합에 기반 한 다중 객체 검출 및 추적

Multiple Object Detection and Tracking based on Morphological Region Merging from Real-time Video Sequences

박종현, 백승철, Nguyen Dinh Toan, 이귀상
전남대학교 전산학과

Jong-Hyun Park(jhpark@new21.com), Seung-Cheol Baek(backsc2@hanmail.com),
Nguyen Dinh Toan(toanhoian@gmail.com), Guee-Sang Lee(gslee@chonnam.ac.kr)

요약

본 논문에서는 카메라로부터 획득 되어진 비디오 시퀀스로부터 다중 움직임 객체와 배경을 분할하고 시공간 정보에 기반 한 객체 추적 방법을 제안한다. 제안한 방법은 3단계로 구성되어 있다. 먼저 입력 비디오 시퀀스로부터 프레임 사이의 차를 이용한 움직임 영역과 움직임이 존재하지 않는 영역을 구분하여 적응적 경계값을 추출한다. 두 번째는 참조 배경영상과 적응적 경계값을 이용하여 움직임이 존재하는 영역으로부터 개략적 객체 분할을 수행하며, 분할된 이진영상에 형태학적 영역 병합 알고리즘을 적용하여 객체 병합을 수행하였다. 마지막으로 분할된 객체에 시공간 정보를 이용하여 객체에 임의의 ID를 할당하여 추적하였다. 카메라로부터 획득되어진 비디오 시퀀스를 이용한 실험에서 객체들의 분할 및 추적의 효율성과 시스템의 유용성을 확인하였다.

■ 중심어 : | 비디오 감시 시스템 | 움직임 객체 검출 | 객체 분할 | 다중 객체 추적 |

Abstract

In this paper, we propose an efficient method for detecting and tracking multiple moving objects based on morphological region merging from real-time video sequences. The proposed approach consists of adaptive threshold extraction, morphological region merging and detecting and tracking of objects. Firstly, input frame is separated into moving regions and static regions using the difference of images between two consecutive frames. Secondly, objects are segmented with a reference background image and adaptive threshold values, then, the segmentation result is refined by morphological region merge algorithm. Lastly, each object segmented in a previous step is assigned a consistent identification over time, based on its spatio-temporal information. The experimental results show that a proposed method is efficient and useful in terms of real-time multiple objects detecting and tracking.

■ keyword : | Video Surveillance System | Moving Object Detection | Object Segmentation | Multiple Object Tracking |

I. 서론

움직임 객체들의 검출 및 추적은 컴퓨터 비전 응용 분야의 전 단계에서 폭넓게 사용되고 있으며, 비디오 감시 시스템, 로봇, 인증 시스템, 제스처에 의한 사용자 인터페이스 등에서 응용되고 있다. 또한 멀티미디어 시스템의 일반화로 비디오 감시 시스템과 휴먼 인터페이스 기술이 진화하고 있다.

일반적으로 움직임 객체의 검출과 추적을 위한 알고리즘들은 세 가지 방법론으로 분류될 수 있다. 이들은 배경영상의 차(difference) 기반 방법, 시각적 차이 기반 방법, 확률 기반 방법 등이다. Oliver et al.이 제안한 배경 차 알고리즘은 여러 정적 배경 영상들로부터 PCA (principal component analysis) 분석법에 의한 고유 배경(eigen-background)의 차이를 이용한 방법을 제시하였다[1]. [2]에서는 GMM (Gaussian mixture model)에 의한 배경을 모델링하는 방법을 시도하였으나 많은 계산량을 요구하는 문제점을 가지고 있었다. 따라서 Oliver 팀은 각 사람의 위치와 속도를 기반으로 사람을 추적하였으며, 칼만 필터와 확률 밀도 함수를 이용하여 사람 몸에 해당하는 개체 색상 성분들을 추정하는 방법을 채택하였다. Niyogi는 배경 차 영상열(background-subtracted image sequences)의 시공간에서의 Hough 변환을 기반 한 선 검출 알고리즘을 통하여 사람을 분할하였다[3]. Niyogi는 피사체가 일정한 속력으로 영상 내에서 움직인다고 가정하여 사람의 궤적을 추적하였다. Sato et al.은 영상내에서 측면으로 이동하는 두 사람을 독립된 객체로 분할 및 추적하는 연구를 수행하였다[4]. 이 시스템은 배경의 차이를 이용하여 사람 영상을 추출한 후 객체 분할 기술과 시간적 공간상 속도 변환(temporal spatio-velocity transform)을 수행하였다. [5]에서는 객체의 추적을 위한 첫 단계로 배경과 객체의 분할을 수행하였다. 움직임 분할은 현재 영상과 순열의 영상들 간의 차이를 경계값으로 사용하였다. 그러나 다중 객체의 추적 및 객체가 정지되어 있을 때에는 검출 및 추적의 문제점을 가지고 있어 비효율적인 결과를 보여주었다. Lipton은 시간적 차를 이용하여 움직임 객체들을 추출하였다[6]. Paragios et al.은 통계적 문제

와 확률을 고려하였으며, 영상 분할을 위한 확률적 기반의 분류 방정식을 얻기 위해 관찰된 정보를 이용하였다[7]. 이런 검출 모형은 완벽하게 움직이는 객체를 추적하고 검출할 수 있으나 높은 계산량을 요구하는 문제점을 가지고 있다. Davis et al.은 W^4 방법을 제안하였다[8]. 이 방법은 실시간으로 처리할 수 있게끔 확률 모형을 단순화하였다. 제안된 방법은 영상의 밝기를 이용하여 배경에 대한 학습을 수행하였으며, 움직임이 없는 객체들로만 이루어진 프레임의 모든 픽셀들의 변량 범위를 관측하기 위해 통계적 순환(statistical cycle)을 사용하였다. 그 이후에는 관측된 영상 정보를 토대로 배경 모형을 생성하였다. 그리고 현재 프레임과 배경 모형의 차이를 통하여 움직임이 존재하는 픽셀을 감지한다. 그러나 W^4 의 방법은 계산량 측면에서 효율적이지만 배경 모형을 종종 재생성하여야 하는 문제점이 있다.

또한 움직임 검출을 위한 접근 방법으로 영역기반 접근(region-based approach)과 경계 기반 접근(boundary-based approach) 방법으로 고려하였다 [9]. 영역 기반의 접근 방법은 배경 영상 차를 이용하는 방법과 optical flow를 이용한 방법이 있다. 배경 차에 의한 방법은 조명의 변화와 배경의 미세한 움직임에 매우 민감한 특징이 있다[10-13]. 이러한 문제점을 해결하기 위해 GMM[2]을, Barron et al.은 optical flow [14]를 적용하여 객체 검출을 시도하였다. 경계 기반의 움직임 객체 검출에는 에지기반의 optical flow [15-17]와 레벨셋[18]들이 있다. 그러나 이러한 방법들은 주위 화소들의 공간적인 정보의 사용 없이 계산되어지기 때문에 정확성이 떨어지는 단점을 가지고 있다. [19]는 움직임 객체의 초기 윤곽을 검출하기 위해 색상 성분을 이용한 분할을 하였다. 또한 Geodesic active contour와 레벨셋을 이용한 방법을 [20]에서 제안하였다. 그러나 제안한 방법들은 실시간 응용을 위한 처리에서 너무 많은 계산량을 요구하고 있다.

따라서 본 연구에서는 실시간 다중 객체의 검출 및 추적을 위한 효율적이고 효과적인 접근 방법을 제안한다. 또한 실험은 제한된 실내 공간에서 여러 사람의 움직임 객체를 검출하여 시공간적 특징 정보를 이용한 ID

할당과 연속적인 추적을 통하여 보안 및 검증 시스템에 응용하는데 목적이 있다. 제안한 방법은 정적인 배경 영상으로부터 정확한 객체의 분할과 추적을 위해 현재 프레임에서 움직이는 객체를 제외한 영역을 배경으로 생성하였으며, 형태학적 영역 병합에 기반한 객체 분할을 수행한다. 또한 검출된 다중 객체들은 색상정보와 시공간 위치정보를 이용하여 정해진 공간 내에서 실시간 추적을 수행한다.

본 논문의 구성은 다음과 같다. 먼저 II장에서는 입력 프레임으로부터 초기 참조 배경 영상 생성방법에 대하여 기술하며, III장에서는 적응적 경계값을 이용한 객체의 분할에 대하여 기술한다. 다음으로는 분할된 객체의 잡음제거와 고립영역의 병합을 위한 형태학적 병합 알고리즘에 대하여 IV장에서 기술한다. V장에서는 분할된 다중 객체들에 시공간 정보를 이용하여 객체에 ID 할당과 추적에 방법에 대하여 보여준다. VI장에서는 제안된 알고리즘의 효율성을 실시간 비디오 시퀀스에 적용하여 성능을 평가하였으며, VII에서 결론 및 향후 연구에 대하여 기술한다.

II. 참조 배경영상 생성

입력 프레임 $f_t(x,y)$ 로부터 배경 영상과 객체를 분할하는 부분에서의 출력 영상은 하나의 객체에 하나의 독립된 영역으로 분할된 이진 영상으로 정의할 수 있으며 다음과 같이 표시할 수 있다.

$$S_t(x,y) = \begin{cases} \text{if } f_t(x,y) \in \bigcup_{i=1}^N O_i^t, & 1 \\ \text{else} & , 0 \end{cases} \quad (1)$$

여기에서 O_i^t 는 t 번째 프레임에서 i 번째 객체에 대한 영역을 나타낸다. 배경과 객체를 분할하기 위해 입력 영상과 배경 영상의 차를 통하여 객체 분할을 하였다. 따라서 t 번째 프레임 f_t 에 대한 개략적 객체 분할은 다음 식 (2)와 같다.

$$S_t'(x,y) = \begin{cases} \text{if } |f_t(x,y) - B_{n-1}(x,y)| > TH_1, & 1 \\ \text{else} & , 0 \end{cases} \quad (2)$$

여기에서 B_{n-1} 은 $f_1 \dots f_{n-1}$ 로부터 얻어진 참조 배경영상이며, B_0 는 f_1 로 초기화 가능하며, $TH_1(x,y)$ 는 참조 배경영상 생성을 위한 경계 값을 나타낸다. 따라서 S_t' 에 대한 TH_1 은 다음 식 (3)과 같이 정의하였다.

$$TH_1 = \begin{cases} \text{if } |f_t(x,y) - f_{t-1}(x,y)| > TH, & TH - \Delta \\ \text{else} & , TH + \Delta \end{cases} \quad (3)$$

TH 는 두 프레임사이의 차에 의해 객체를 배경과 구분하기 위한 경계값으로 실험적 분석에 의해 20으로 설정하였다. TH_1 에 의해 참조배경영상 B_1 을 식 (4)에 의해 구할 수 있다.

$$B_n(x,y) = \begin{cases} \text{if } S_t'(x,y) = 0, & f_t(x,y) \\ \text{else} & , B_{n-1}(x,y) \end{cases} \quad (4)$$

[그림 1]은 참조 배경영상 생성에 대하여 보여주고 있다. $n > 1$ 인 경우에는 향상된 S_t' 영상을 사용하여 참조 배경 영상을 갱신할 수 있다. 이에 대한 자세한 설명은 V장에서 다루겠다.



(a) (b) (c)

그림 1. 참조 배경영상 생성: (a) 이전 참조 배경영상 B_{n-1} 프레임, (b) 현재 프레임 f_t , (c) (a)와 (b)로부터 추출한 참조 배경영상 B_n

III. 적응적 경계값 추출

객체 분할은 적응적 경계값에 의해 추출된 기준 배경 영상과 현재 프레임 사이의 차를 이용하여 수행할 수 있다. 일반적으로 동영상은 시간에 따른 조명의 변화 및 객체들의 특징 변화 등을 포함하고 있기 때문에 경

계값을 균일하게 적용한다면 정확한 객체 검출에 어려움을 제공할 수 있다. 따라서 각 픽셀마다 적절한 경계값을 객체의 특성에 따라 적용한다면 보다 더 정확하고 효율적으로 검출할 수 있을 것이다. 본 논문에서는 현재 프레임과 이전 프레임사이의 차를 이용하여 객체 기반의 적응적 경계값을 추정하는 방법을 제안한다. 비디오 영상열에서 각 프레임들은 몇 개의 균질 영역들로 구성되어 있다. 이는 객체의 이동에도 프레임 사이의 차가 객체 영역 내에서는 크지 않음을 나타낸다. 하지만 객체와 배경의 경계선 부분에서는 객체의 움직임 때문에 큰 차가 생긴다. 즉, 객체에 속한 균질 영역의 이동은 배경에 속한 다른 균질 영역과 겹쳐져서 나타나며, 서로 겹쳐진 부분과 그렇지 않은 부분에서 큰 특성의 변화를 가진다. 따라서 프레임 사이의 차가 큰 부분은 배경에서 객체로 객체에서 배경으로의 전이가 나타난 부분일 가능성이 높다는 것을 알 수 있다.

본 논문에서는 프레임 사이의 차 영상을 얻은 후에 경계값 TH_1 을 통하여 배경과 객체 영역 사이의 움직임이 일어났을 가능성이 높은 영역과 낮은 부분으로 영상을 분할하였다. 움직임이 일어났을 가능성이 높은 영역의 경우에는 이전 프레임에서 객체 영역이었던지를 판단하여, 만약 이전 프레임에서 객체 영역이었다면, 배경영역이 되기 쉽게 높은 경계값을 할당한다. 또한 이전 프레임에서 배경영역이었다면 객체 영역으로 검출 가능하게 낮은 경계값을 할당한다. 움직임이 발생할 가능성이 낮은 영역은 이전 프레임에서 객체 영역이었던지를 판단하여 만약 이전 프레임에서 객체 영역이었다면 객체 영역이 될 수 있도록 낮은 경계값을 할당한다. 이전 프레임에서 배경 영역이었다면 배경 영역이 되기 쉽게 높은 경계값을 할당하였다. 이 할당된 경계값들은 배경 영상과 프레임의 차 영상으로 객체 영역의 분할에 적용된다. 따라서 위에서 언급한 내용을 정리하면 다음과 같다. f_t 과 f_{t-1} 의 차를 이용하여 $I_t(x, y)$ 를 다음과 식 (5)와 같이 정의할 수 있다.

$$I_t(x, y) = \begin{cases} |f_t(x, y) - f_{t-1}(x, y)| > TH_1, & 1 \\ \text{else} & , 0 \end{cases} \quad (5)$$

$I_t(x, y)$ 에 의해 적응적 경계값 계산은 다음 식 (6)과 같다.

$$ATH_t(x, y) = \begin{cases} I_t(x, y) = 1 \text{ and } S'_{t-1}(x, y) = 1, & ATH_t^1 \\ \text{elseif } I_t(x, y) = 1 \text{ and } S'_{t-1}(x, y) = 0, & ATH_t^2 \\ \text{elseif } I_t(x, y) = 0 \text{ and } S'_{t-1}(x, y) = 1, & ATH_t^3 \\ \text{else} & , ATH_t^4 \end{cases} \quad (6)$$

여기서 $ATH_t^1 > ATH_t^2$ 또는 $ATH_t^3 < ATH_t^4$ 의 관계를 가지며 $ATH_t^2 = TH - \Delta$ 이며, $ATH_t^4 = TH + \Delta$ 이다. 이와 같은 사항을 적용하여, 식 (2)의 t 번째 프레임 f_t 에 대한 개략적 객체 분할을 수정하여 식 (7)로 만들 수 있다.

$$S'_t(x, y) = \begin{cases} \text{if } |f_t(x, y) - B_{t-1}(x, y)| > ATH_t(x, y), & 1 \\ \text{else} & , 0 \end{cases} \quad (7)$$

결론적으로, 식(6)에서 ATH_t 와 식 (7)을 통하여, 식 (2)에 비하여 향상된 S'_t 을 얻을 수 있다.

표 1. 형태학적 영역 병합을 위한 Pseudo Code ($SR_t(x, y)$ 값이 1이면 객체영역, 0이면 배경영역을 의미)

```

FUNCTION Merging( $S'$ )
FOR ALL  $(x, y)$ ,  $SR_0(x, y) = 1$ 
 $t = 0$ 
REPEAT
 $t = t + 1$ 
 $SR_t = SR_{t-1}$ ;
FOR  $x=1$  TO width
FOR ALL  $y$ ,  $SR_t(x, y) := \begin{cases} 1, & S'(x, y) = 1 \\ h_1(x, y), & S'(x, y) = 0 \end{cases}$ 
FOR  $x=width$  TO 1
FOR ALL  $y$ ,  $SR_t(x, y) := \begin{cases} 1, & S'(x, y) = 1 \\ h_2(x, y), & S'(x, y) = 0 \end{cases}$ 
FOR  $y=1$  TO height
FOR ALL  $x$ ,  $SR_t(x, y) := \begin{cases} 1, & S'(x, y) = 1 \\ h_3(x, y), & S'(x, y) = 0 \end{cases}$ 
FOR  $y=height$  TO 1
FOR ALL  $x$ ,  $SR_t(x, y) := \begin{cases} 1, & S'(x, y) = 1 \\ h_4(x, y), & S'(x, y) = 0 \end{cases}$ 
UNTIL  $(|SR_t - SR_{t-1}| < \epsilon)$ 
RETURN  $SR_t$ 
    
```

IV. 객체 분할을 위한 형태학적 영역 병합

1. 형태학적 영역 병합 알고리즘

우리는 입력된 실시간 동영상으로부터 움직임을 객체의 분할을 위해 배경과 객체는 잡음 또는 고립영역이 없는 하나의 영역으로 나타난다고 가정하였다. 전 단계에서 개략적으로 프레임 사이의 차를 이용하여 배경과 객체를 분할하였다. 그러나 분할된 객체는 많은 잡음과 고립영역들로 구성되어 있으며 하나의 객체는 여러 개의 영역들로 나누어져 있을 수 있다. 따라서 본 논문에서는 배경과 객체를 정확하게 분리하기 위해 프레임 전체 영역에 대하여 상하좌우 4 방향 스캔(scan)을 수행하여 배경 부분을 설정된 객체영역에서 제거한다. 이 방법을 4 방향으로 스캔을 하며 제거를 수행한 결과가 수렴할 때까지 반복적으로 수행한다. 제안한 객체 영역을 검출하기 위한 형태학적 영역 병합 알고리즘의 전반적 pseudo-code를 [표 1]에서 보여주고 있다. 여기에서 h_1, h_2, h_3 그리고 h_4 는 객체 또는 배경을 결정하기 위한 결정 함수(decision function)이며, $\epsilon = 0$ 이다. 이 결정함수에 대한 설명은 다음절에서 자세히 설명된다.

2. 객체 결정함수

비디오 프레임에서 움직임을 독립된 하나의 객체로 병합하기 위한 결정함수 $h(\cdot)$ 에 대하여 기술한다. 제안된 결정함수 $h(\cdot)$ 는 스캔 방향과 배경으로 조건이 되는 각도 $\theta(0^\circ < \theta < 90^\circ)$ 로 정의된다. 예를 들면, 조건 $\tan(\theta) = 1/2$ 이며, 왼쪽에서 오른쪽으로 스캔한다고 할 때, $h(\cdot)$ 는 그림 2와 같이 도식화 할 수 있다.

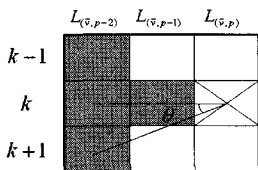


그림 2. $l=2, \tan(\theta)=1/2$ 인 h 에 대하여 도식화 (X는 현재의 위치, 회색 셀은 조건문에 사용되는 부분)

X는 스캔하는 위치이다. $h(\cdot)$ 는 모든 회색 셀들이 이미 배경으로 결정되어 있다면, X 또한 배경으로 결정한다.

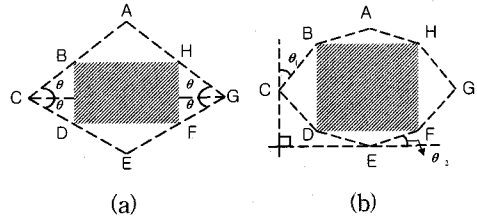


그림 3. 결정함수에 대한 형태학적 병합 알고리즘: (a) 좌우 스캔에 대한 h 의 θ 에 따른 병합 영역, (b) 상하 스캔에 대한 h 의 θ 인 θ_1 과 좌우 스캔에 대한 h 의 θ 인 θ_2 에 따른 병합 영역 (사선 부분은 검출된 영역)

[그림 3]은 결정함수 h 가 형태학적 영역 병합 알고리즘의 결과에 어떤 영향을 미치는지 보여준다. [그림 3]에서 회색 영역은 배경영상과 프레임의 차를 이용하여 분할된 영역으로 고려한다. 왼쪽에서 오른쪽 방향으로 스캔을 수행하면 $\triangle CDB$ 이 제거된다. 오른쪽에서 왼쪽 방향으로 깎기를 수행하면, $\triangle GHF$ 부분이 제거된다. 그리고 $\triangle BHA$, $\triangle DEF$ 그리고 $\square BDFH$ 부분은 남는다. 만약 이 상태에서 위에서 아래 방향으로 스캔을 수행하면, $\triangle BHA$ 가 제거되고, 아래에서 위 방향에 대한 스캔을 수행하면 $\triangle DEF$ 가 제거되어, 결국 $\square BDFH$ 부분만 남는다. 그러나 만약 $\triangle BHA$ 내에 하나의 점이 존재한다면, 그 점을 포함한 $\triangle BHA$ 의 일부분이 남아서 점과 $\square BDFH$ 이 연결된다. 결국 왼쪽에서 오른쪽으로 오른쪽에서 왼쪽으로 제거할 때 사용된 결정함수의 θ 가 상하로 놓인 연결된 영역들의 병합에 관계한다는 사실을 알 수 있다. 위와 같은 방법으로 각각의 방향에 대하여 스캔을 수행하면 [그림 3](b)에서 점선으로 이루어진 고유의 영역을 추출할 수가 있다. 이 영역에 다른 영역의 일부가 놓이면 두 영역은 하나로 병합이 된다. 위의 관찰에서 보듯이 연결된 영역에 대하여 다각형 ABCDEFGH와 같이 일종의 convex hull인 영역이 존재한다. 이 convex hull의 생성에 결정함수의 θ 값이 중요한 역할을 하는 것을 알 수 있다. 또한 상하 스캔에

관한 결정함수의 θ 을 θ_1 로 하고 좌우 스캔에 관한 결정 함수 θ 을 θ_2 로 할 때, θ_1 과 θ_2 의 관계가 병합 조건에 영향을 줄 수 있는데, 이에 대하여 기술하여 보면 아래와 같다.

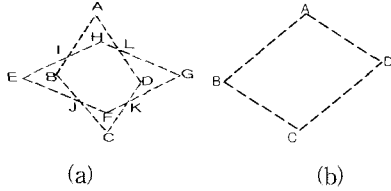


그림 4. 고립 영역 병합 도식화: (a) $\theta_1 \neq (\pi/2) - \theta_2$ 일 때의, (b) $\theta_1 = (\pi/2) - \theta_2$ 일 때의 클러스터 모양 도식화.

(a) $\theta_1 > (\pi/2) - \theta_2$ 이면, [그림 4](a)의 모양으로 도식화가 된다. 이 경우 다각형 BJFKDLHI내에 하나의 클러스터가 들어간다. 또한, $\angle IBJ = \angle LDK = 2((\pi/2) - \theta_1)$, $\angle IHL = \angle JFK = 2((\pi/2) - \theta_2)$, $\angle HIB = \angle BJF = \angle FKD = (\theta_1 + \theta_2 + (\pi/2))$ 가 성립된다.

(b) $\theta_1 < (\pi/2) - \theta_2$ 이면, [그림 4](a)의 모양으로 도식화가 된다. 이 경우에는 다각형 BJFKDLHI내에 하나의 클러스터가 들어가며, $\square AIHL$, $\square IEJB$, $\square JCKF$, $\square LDKG$ 에는 어떤 샘플도 존재해서는 안 된다. 또한, $\angle IBJ = \angle LDK = 2\theta_2$ 이고, $\angle IHL = \angle JFK = 2\theta_1$, $\angle HIB = \angle BJF = \angle FKD = ((3\pi/2) - \theta_1 - \theta_2)$ 가 성립된다.

(c) $\theta_1 = (\pi/2) - \theta_2$ 이면, [그림 4] (b)의 모양으로 도식화가 된다. 이 경우에는 $\square ABCD$ 내에 하나의 클러스터가 들어가게 된다. 또한, $\angle ABC = \angle BCD = \angle CDA = \angle DAB = 2\theta_1 = 2\theta_2$ 가 성립된다. θ_1 과 θ_2 의 관계와 무관하게 병합된 연결된 영역들은 일종의 convex hull 형태의 영역 내에 놓이게 된다. 이 convex hull 형태의 영역 내에는 다른 연결된 영역이 존재 할 수 없다. 따라서 형태학적 병합 알고리즘은 θ_1 과 θ_2 로 정의된 일종 convex hull 형태의 모양의 영역을 갖는 객체들을 추출할 수 있다. 본 논문에서는 추적할 객체의 모양이 세로로 긴 모양으로 가정하여, $\tan(\theta_1) = 1/2$,

$\theta_2 = (\pi/2) - \theta_1$ 로 결정하였다. [그림 5]는 입력 프레임에 제한한 형태학적 영역 병합 알고리즘을 적용하여 분할 및 병합을 수행한 결과이다.

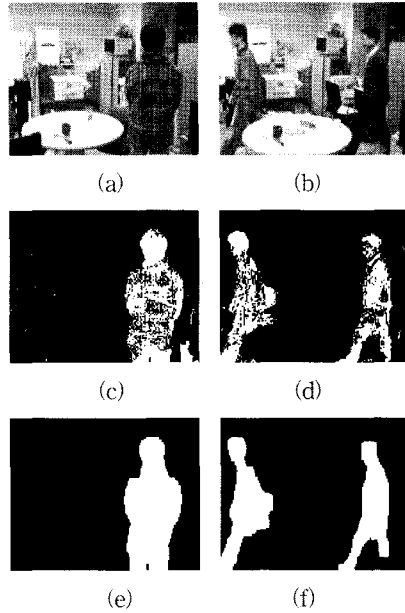


그림 5. 검출된 객체 영상: (a)와 (b) 입력 프레임, (c)와 (d) 적응적 경계값을 이용한 객체 분할, (e)와 (f) 형태학적 병합 결과

V. 시공간 특징 기반 객체 추적

본 절에서는 배경의 갱신 방법과 객체 분할 시에 나타나는 잡음의 제거를 기술한다. 객체 분할의 잡음은 곧 배경을 객체로 잘못 인식할 수 있다. 정지된 카메라로부터 획득된 비디오는 정적인 배경을 가지고 있지만 시간의 변화에 따라 배경영역의 물체 변화와 조명의 변화가 생길 수도 있다. 따라서 움직임이 있는 다중 객체 영역 이외의 영역은 모두 배경 영역으로 투영되어야 한다. 객체 주위에 나타나는 일부 잡음들은 객체와 연결되어 하나의 객체 성분으로 검출되는 경우가 있다. 잡음 부분의 가장자리가 객체와 연결된 경우에는 연결부위가 작기 때문에 opening 연산자로 분리 할 수 있다. 배경 내에서 변화하여 영역이 객체 영역으로 잘못 분할

된 잡음의 경우에는 객체 영역에 비해 면적이 작다고 가정을 세울 수 있다. 따라서 잡음 영역의 크기는 움직임 객체의 크기에 비해 작은 고립영역들로 나타난다.

본 논문에서는 잡음 영역의 면적은 A_1 보다 작다고 가정한다. 이를 이용하여 $n > 1$ 인 경우의 참조 배경영상 갱신에 대하여서는 식 (4)을 그대로 수행하지 않고, 잡음 영역을 제거한 후, 객체 영역의 경계 부분에서 생기는 잡음을 제거하기 위해 팽창 과정을 적용하였다. 객체의 이동에 따라서 일부 영역이 다른 사물들에 의하여 객체 영역에서 분리되어 새로운 객체 영역이 일시적으로 나타는 경우가 있다. 이와 같은 경우에 일시적으로 분리되는 영역이 다른 새로운 객체로 해석되는 것은 옳지 않다. 그래서 일시적으로 분리되는 영역을 제거할 필요가 있다. 일시적으로 분리되는 영역은 본 객체 영역에 비해 작은 고립영역으로 나타난다. 본 논문에서는 일시적으로 분리되는 영역의 면적은 A_2 보다 작다고 가정하였다. 따라서 다중 객체를 분할한 후, 검출된 객체의 식별과 추적을 위해 분할된 객체의 시공간 위치 정보와 색상정보를 기반으로 한 접근 방법을 제안한다.

객체 식별 방법은 객체의 특징 벡터를 정의하고 이 특징 벡터들의 거리를 정의하는 방법을 적용하였다. 현재 프레임 f_t 상의 i 번째 객체 O_t^i 의 특징 벡터와 이전 프레임 f_{t-1} 상의 객체들의 특징 벡터간의 거리를 구한 후, 최소 거리를 갖는 이전프레임 f_{t-1} 상의 j 번째 객체 O_{t-1}^j 의 ID를 객체 O_t^i 의 ID로 할당한다.

객체를 추적하기 위한 방법으로 크게 두개의 특징을 정의할 수 있다. 하나는 외형적 특징이고 다른 하나는 시공간 정보이다. 본 연구에서는 외형적 특징으로 RGB 성분을 이용한 히스토그램을 사용하였으며, 시공간 정보로는 객체의 중심 위치를 사용하였다. 다음 수식 (8)은 비디오 영상에 존재하는 다중 객체의 식별을 위한 것이다.

$$D(O_t^i, O_{t-1}^j) = w_1 \left[(x_t^i - x_{t-1}^j)^2 + (y_t^i - y_{t-1}^j)^2 \right]^{\frac{1}{2}} + w_2 \left[\sum_{i=1}^N (r_t^i - r_{t-1}^j)^2 + \sum_{i=1}^N (g_t^i - g_{t-1}^j)^2 + \sum_{i=1}^N (b_t^i - b_{t-1}^j)^2 \right] \quad (8)$$

여기서 w_1, w_2 는 가중치이며 본 연구에서는 가중치 값을 $w_1 = w_2 = 1$ 로 하였다. (x_t^i, y_t^i) 와 (x_{t-1}^j, y_{t-1}^j) 는 객체 O_t^i, O_{t-1}^j 의 중심에 대한 위치 값을 나타낸다. 그리고 r_t^i, g_t^i, b_t^i 와 $r_{t-1}^j, g_{t-1}^j, b_{t-1}^j$ 는 각각의 객체에 대한 red, green, blue의 색상 히스토그램을 나타내며 0.0-1.0 사이의 정규화된 값이다. 따라서 다중 객체에 대한 식별 ID를 할당하여 정해진 공간 내에서 객체들을 실시간으로 추적 가능하다.

VI. 실험 및 결과

본 논문에서는 제안된 알고리즘의 효율성 및 강건성 평가를 위해 실내에서 실시간으로 획득된 비디오 영상을 이용하였다. 또한 다중 객체의 추적을 위해 여러 사람의 공간 이동에 대하여 실험을 수행하였다. 실험에 사용된 시스템은 Logitech QuickCam 카메라와 Pentium 4, 1.8GHz, 512MB 메모리를 가지고 있다. 실험 영상은 실내 환경에 대한 비디오 영상이며, 사람의 움직임을 1-3명까지 증가하면서 각각의 상황에 대하여 객체의 검출 및 추적에 대한 성능 평가를 수행하였다. 또한 두 사람이 교차하여 겹침이 발생하여도 계속해서 할당된 ID의 불변성을 확인하였다. [그림 6]은 객체의 영역 분할을 위하여 기존의 방법과 제안된 방법에 대하여 효율성과 강건성의 성능을 비교 실험한 결과를 보여 주고 있다. [그림 6](b)는 [21]에서 제안된 벡터 미디언 방법을 적용하였으며, [그림 6](c)는 [22]에서 제안한 웨이블릿 기반의 객체 분할 결과, 그리고 [그림 6](d)는 제안된 방법에 의한 객체 분할 결과에 대하여 보여주고 있다. 벡터 미디언 방법은 작은 조명의 변화에도 매우 민감하게 반응하여 전체적인 공간에서 작은 잡음 영역들이 발생한다. 다음으로 웨이블릿 변환된 저주파 성분의 차를 이용한 방법으로 객체의 움직임 변화가 발생하더라도 저주파 성분의 변화는 매우 작게 나타나므로 하나의 객체영역들이 여러 개의 영역으로 분할되거나 정확한 영역 분할에 어려움이 있다. 그러나 제안된 방법에서는 기존 방법의 단점을 제거할 수 있었으며 객체와

배경에 대한 분할에 있어서 기존의 방법과 비교하여 강건하게 동작함을 볼 수 있다.

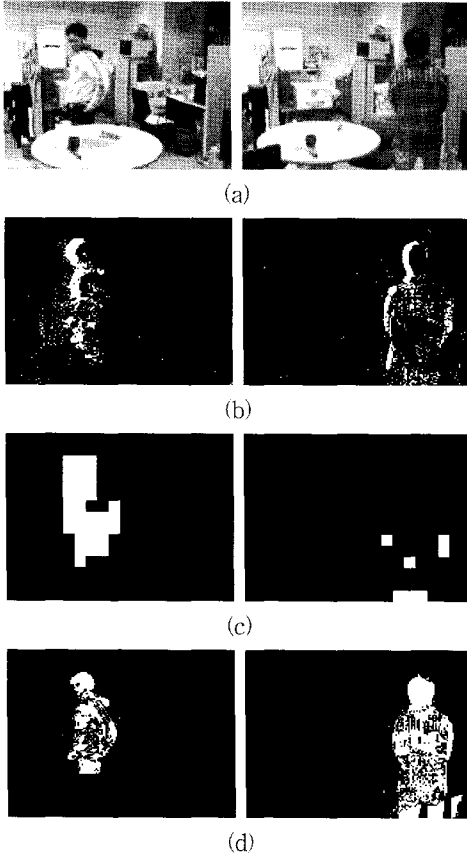


그림 6. 객체 분할을 위한 비교 실험: (a) 입력 프레임, (b) 벡터 미디언 방법에 의한 결과 영상, (c) 웨이블릿 기반 객체 분할 결과 영상, (d) 제안된 방법의 결과 영상

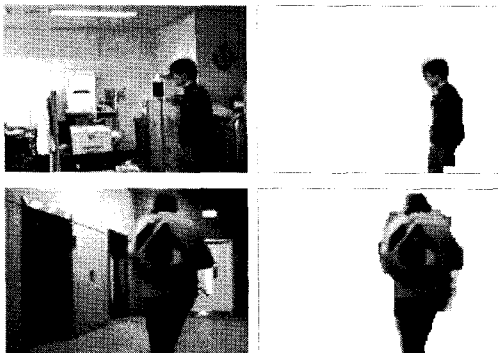


그림 7. 단일 객체 움직임에 대한 검출 결과:

[그림 7]은 단일 객체의 움직임 검출 결과를 보여주고 있다. 객체의 움직임에 대한 정확한 검출을 수행하였으며 검출된 객체는 배경과 완벽하게 분리 가능함을 보여준다.



그림 8. 다중 객체의 검출 및 추적 결과 (한 방향으로 이동한 경우)

[그림 8]은 두 사람이 일정한 방향(right->left)으로 이동하고 있는 경우로 각각의 사람에 대하여 추적 결과를 보여주고 있다. 배경과 객체의 검출에서 정확성을 보여주고 있으며 한방향의 이동으로 독립된 객체의 추적이 가능하였다. [그림 9]는 두 사람이 서로 교차하여 겹침이 발생한 경우로 겹침이 발생하기 전의 ID를 겹침 후에도 연속적으로 정해진 공간 내에서 객체의 추적을 효율적으로 수행하였다. 따라서 제안된 알고리즘은 실시간 다중객체의 검출과 추적에 있어서 효율성을 실험을 통하여 확인하였으며 다양한 분야의 비디오 감시 시스템에 응용할 수 있음을 확인하였다.



(a)

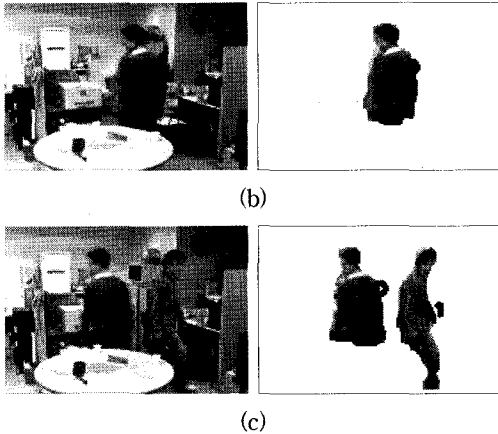


그림 9. 다중 객체에 대한 검출 및 추적 결과(교차하는 객체의 움직임):(a) 객체의 검침이 발생하기 전 프레임, (b) 두 객체의 검침이 발생했을 때의 프레임, (c) 검침 발생 후의 프레임

VII. 결론 및 향후 연구

본 논문에서는 실시간 비디오 시퀀스로부터 다중 움직임 객체의 검출 및 추적을 위한 새로운 접근 방법을 제안하였다. 제안된 방법은 프레임 차를 이용하여 적응적 경계값을 추출하였으며 이를 기반으로 참조 배경영상을 생성하였다. 그리고 움직임 객체를 배경으로부터 검출하여 다중 객체 각각에 식별 가능한 ID를 할당하여 정해진 공간내에서 객체의 추적을 수행하였다. 제안한 객체 검출을 위한 형태학적 영역 병합 알고리즘은 실시간 다중 객체 검출에 효율적임을 실험을 통하여 확인하였다. 또한 객체 추적에서는 색상정보와 시공간 위치정보를 병합하여 정해진 공간 내에서 움직임이 있는 객체에 임의의 ID를 할당하여 실시간 추적이 가능함을 보여주었다. 제안된 실시간 객체 검출 및 추적을 위한 실험에서는 알고리즘의 효율성과 다양한 분야에서 안정적으로 응용 가능함을 확인하였다. 제안한 객체 검출 및 추적 시스템은 제한적인 실내공간에서 실시간 사람의 이동경로 파악, 감시 및 보안, 또는 인증을 위한 분야에 응용 가능하다. 그러나 시간에 따른 조명의 변화가 급격하거나 배경 객체의 움직임이 존재할 경우는 정확한 배경영상 검출에 어려움이 있으며, 복잡한 실외 환경에

서 객체추적을 위한 연구가 요구된다. 향후 연구에서는 이러한 문제점을 해결하고 동적 카메라에 대한 객체 추적 연구가 수행되어야 한다.

참고문헌

- [1] N. Oliver, B. Rosario, and A. Pentland, "A Bayesian Computer Vision System for Modeling Human Interaction," Int. Conf. on Vision Systems, pp.255-272, 1999.
- [2] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "TM Pinder: Realtime Tracking of the Human Body," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No.7, pp.780-785, 1997.
- [3] S. Niyogi and E. Adelson, "Analyzing Gait Spatiotemporal Surface," IEEE Workshop, pp.64-69, 1994.
- [4] K. Sato and J. K. Aggarwal, "Tracking and Recognizing Two-person Interaction in Outdoor Image Sequences," IEEE Workshop on Multi-Object Tracking, pp.87-94, 2001.
- [5] M. Quming and J. K. Aggarwal, "Tracking and Classifying Moving Objects from Video," IEEE International Workshop on PETS, 2001.
- [6] A. Lipton, H. Fujiyoshi, and R. Patil, "Moving Target Classifying Moving Objects from Video," DARPA Image Understanding Workshop, 1998.
- [7] N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Object," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.3, pp.266-280, 2000.
- [8] I. Haritaoglu, D. Harwood, and L. S. Davis, "Real-time Surveillance of People and Their Activities," IEEE Transactions on Pattern

- Analysis and Machine Intelligence, Vol.22, No.3, pp.809-830, 2000.
- [9] M. Yokoyama and T. Poggio, "A Contour-based Moving Object Detection and Tracking," IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005.
- [10] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, "Foreground Object Detection in Changing Background based on Color co-occurrence Statistics," IEEE Workshop on Applications of Computer Vision, 2002.
- [11] C. Stauffer and W. E. L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," IEEE Conference Computer Vision and Pattern Recognition, 1999.
- [12] A. Mittal and N. Paragios, "Motion-based Background Subtraction using Adaptive Kernel Density Estimation," IEEE Conference Computer Vision and Pattern Recognition, Vol.2, pp.302-309, 2004.
- [13] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric Model for Background Subtraction," European Conference Computer Vision, Vol.II, pp.751-767, 2000.
- [14] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of Optical Flow Techniques," International Journal Computer Vision, Vol.12, No.1, pp.43-77, 1994.
- [15] J. H. Duncan and T. C. Chou, "On the Detection of Motion and the Computation of Optical Flow," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.14, No.3, pp.346-352, 1992.
- [16] V. Caselles and B. Coll, "Snakes in Movement," SIAM Journal Numerical Analysis, Vol.33, pp.2445-2456, 1996.
- [17] S. S. Beauchemin and J. L. Barron, "The Computation of Optical Flow," ACM Computing Surveys, 1995.
- [18] J. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge University Press, 1999.
- [19] L. Qiu and L. Li, "Contour Extraction of Moving Object," IEEE International Conference Pattern Recognition, Vol.2, pp.1427-1432, 1998.
- [20] N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking for Moving Objects," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.3, pp.266-280, 2000.
- [21] 김준철, 박은중, 이준환, "벡터 미디언을 이용한 비디오 영상의 온라인 배경 추출", 정보처리학회 논문지, 제13권-B권, 제5호, pp.512-524, 2006.
- [22] F. H. Cheng and Y. L. Chen, "Real Time Multiple Objects Tracking and Identification based on Discrete Wavelet Transform," Pattern Recognition, Vol.39, pp.1126-1139, 2006.

저자 소개

박종현(Jong-Hyun Park)

정회원



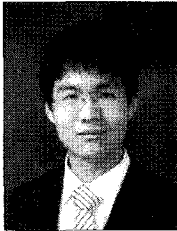
- 1994년 : 목포대학교 전자공학과 학사
- 1997년 : 목포대학교 전자공학과 석사
- 2001년 : 목포대학교 전자공학과 박사

- 2002년 ~ 2003년 : 전북대학교 Post-Doc.
- 2003년 ~ 2004년 : KOSEF 신진연구
- 2004년 ~ 2006년 : University of Southern California, USA, Post-Doc.
- 2006년 ~ 현재 : 전남대학교 BK21 교수

<관심분야> : 컴퓨터 비전, 패턴인식, 의료영상처리, 임베디드 시스템, 멀티미디어 통신

백 승 철(Seung-Cheol Baek)

준회원

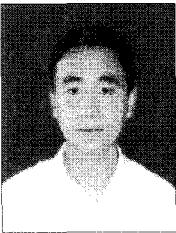


▪ 현재 : 전남대학교 컴퓨터공학과
(학사 과정)

<관심분야> : 영상처리 및 컴퓨
터 비전, 패턴인식

Nguyen Dinh Toan

준회원



▪ September, 2005 : Received the
B.S. degree in Information
Technology from Vietnam
National University Ho Chi
Minh City, University of
Natural Sciences

▪ Currently pursuing M.S. degree in Computer
Science at Chonnam National University in Korea

<관심분야> : Video coding, ubiquitous computing

이 귀 상(Guee-Sang Lee)

정회원



▪ 1980년 : 서울대학교 전기공학과
학사
▪ 1982년 : 서울대학교 전기계산기
공학과 석사
▪ 1991년 : Pennsylvania 주립대학
전산학 박사

▪ 1984년 ~ 현재 : 전남대학교 전자컴퓨터공학부 교수

<관심분야> : 멀티미디어 통신, 영상처리 및 컴퓨터
비전, 임베디드 시스템