

# 영상 물체 탐지 및 추적 기술

이 석 호\*

## 1. 서 론

최근 테러 위협 및 강력 범죄 발생의 영향으로 CCTV를 비롯한 보안 감지 시스템에 대한 요구가 더욱 증가되고 있다. 특히 예방 차원에서의 위협 감지를 가능하게 하는 보안감지 시스템의 포괄적인 상황인지 능력이 필수적으로 요구되고 있다. 이런 상황인지 능력에는 기본적으로 물체탐지 또는 추적기술 등이 요구가 된다. 물체 탐지 및 추적 기술은 이외에도 컴퓨터 비전의 가장 기본적인 오늘날 물체탐지 또는 추적기술은 무수히 많이 개발되고 있으며 분류하기도 쉽지 않다. 여기서는 카메라가 정적 상태인지 동적 상태인지에 따라 물체 탐지 및 추적 기술 중 가장 기본적인 기술들에 대해서 분류하여 기술하고자 한다.

## 2. 배경분리 (Background Subtraction) 를 이용한 물체추적

배경분리를 이용한 물체 추적은 기본적으로 배경이 주어졌다고 가정하고, 각 프레임에 대하여 배경과의 차이를 낸 후, 차이가 크면 움직임이 있는 영역으로 판정한다. 여기서 배경과 물체 영역

을 어떻게 구분하는가에 따라 여러 가지 배경분리 방법들이 존재하며, 배경 분리와 관련하여 여러 가지 선/후 처리 기술들이 존재한다.

### 2.1 이웃 프레임 배경분리

이웃 프레임배경분리에서는 물체의 움직임이 없다고 판단되는 인접한 프레임을 배경으로 보고 현재 프레임을 배경 프레임과 차를 낸 후에 이의 절대값을 취하므로 차영상을 얻는다. 그 다음에 이 차영상에 대하여 임계값 필터링을 수행하여 이진화가 된 영상을 얻는다(그림 2.1). 이렇게 얻어진 이진화 영상에 대하여 다시 connected component analysis등을 수행하여 같이 뭉쳐 있는 화소들을 하나의 물체로 취급하여 각 물체를 구분한다. 그런데 영상잡음 등으로 잘못 생긴 영역들을 없애기 위해서는 차영상을 구한 후에 적당한 커널로 콘블루션을 수행한 후에 이진화하고, 경우에 따라서는 이진화된 영상에다가 다시 모폴로지 필터링을 하여 영상잡음의 영향력을 줄이기도 한다. 그림 2.2에서는 이렇게 하여 얻어진 이진화된 영상을 보여준다. 이런 후처리는 다른 배경분리기술에서도 사용이 된다.

이러한 간단한 배경분리의 단점은 다음과 같다. 먼저 적절한 임계값을 설정하기가 어렵다. 임계값이 너무 작으면 영상잡음이나, 조명의 세미한

\* 교신저자(Corresponding Author): 이석호, 주소: 부산시 주례2동 동서대학교 컴퓨터정보공학부, 전화: (051) 320-1744, E-mail: petrasuk@gmail.com

\* 동서대학교 컴퓨터정보공학부 전임강사

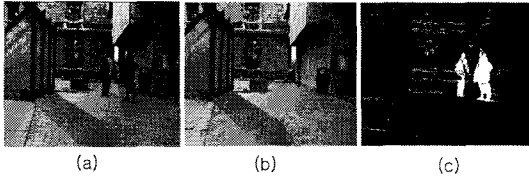


그림 2.1 (a)현재 프레임 (b) 배경 프레임 (c) 차영상을 임계치 필터링을 하여 얻은 이진화된 영상

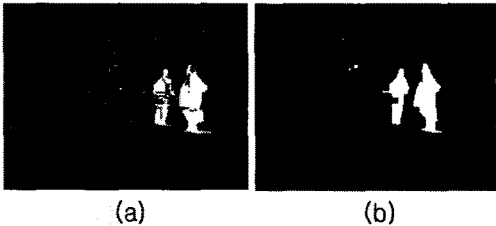


그림 2.2 (a) 가우시안 커널로 차영상을 콘볼루션을 수행한 후에 이진화한 영상. (b) (a)에다가 모폴로지 필터링을 수행하여 얻은 이진화된 영상

차이등에 의해 실제로는 움직임이 없는 영역이 움직임 영역으로 잘못 추출될 가능성이 있다. 반면 임계값이 너무 크다면 자칫 움직임이 있는 영역이 추출되지 못할 가능성이 있다. 또한 조명에 변화가 생기거나, 움직이는 나뭇잎과 같은 움직임이 있는 배경의 경우에 이를 배경으로 처리하기 어렵고 일시적으로 계속 정지한 물체는 배경으로 처리되기 때문에 검출이 어렵다.

### 2.2 이동 평균화 알고리즘

이동 평균화 알고리즘은 일정한 시간 동안의 프레임들을 모두 평균하여 배경을 얻어내는 방법이다. 이 방법을 이용하여 일시적으로 정지한 물체도 추출하여 낼 수 있고, 어느 정도 조명의 변화 등에도 적응적이다. 그러나 이도 이웃 프레임배경분리와 마찬가지로 물체와 배경을 이진화하기 위해 임계값을 정해야 문제가 있다. 또한 정확한 배경을 추출하기 위해서는 많은 프레임 수가 필요하다.

### 2.3 가우시안 모델

Wren 등은 [2]에서 시간에 따른 특정위치의 화소에서의 밝기값의 변화를 지켜보는 방법을 제안하였다. 일정 시간 동안 같은 화소 위치에서의 밝기값들을 관찰하여 관측된 값들로부터 가우시안 분포의 평균과 분산을 계산하여 모델링한다.

이렇게 해서 배경이 모델링이 된 이후에 들어오는 밝기값들에 대해서는 likelihood를 계산하여 배경모델로부터 많이 떨어진 밝기값들은 움직이는 물체로 본다.

### 2.4 Mixture of Gaussian (MOG) Model (Stauffer and Grimson 모델)

실외 환경에서는 나뭇잎의 움직임, 그림자 등과 같은 동적인 배경이 있을 수 있는데 이 경우 하나의 가우시안 분포만으로는 배경의 모델링이 부족하다(그림 2.3). 이때는 여러 개의 색상에 대한 다중 가우시안 분포를 가지고 배경을 모델링하는 것이 필요하다. 그림 2.4에서는 Stauffer와 Grimson이 제안한 다중 가우시안 분포로써 배경을 모델링하여 움직이는 물체를 추출한 예를 보여주고 있다[3].

배경분리에 기반한 추적 알고리즘들은 먼저 배경을 모델링하고, 이후 들어오는 시퀀스들이 모델링된 배경으로부터 얼마나 벗어났는가를 정량적

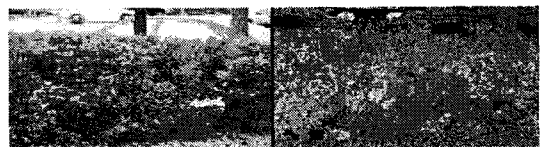


그림 2.3. 배경에 랜덤한 움직임을 가진 관목의 잎들에 의해서 오른쪽 그림과 같이 차영상이 나타나기 때문에 하나의 가우시안만으로는 배경 모델링이 어렵다.

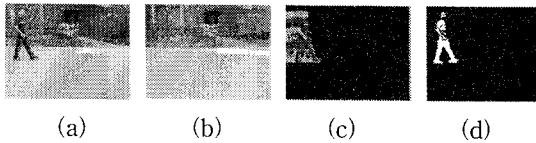


그림 2.4. (a) 현재 프레임 (b) 가장 높은 평균값을 가진 색상값들만을 보여줌. (c) 두번째로 높은 평균값을 가진 색상값들을 보여줌. (d) 배경분리를 하여 얻어진 전경(1).

으로 측정하는데, 일정 한도 이상 벗어난 영역들에 대해서는 움직임이 있는 물체라고 해석하고 이를 추적한다. 이들 알고리즘들은 배경에 대한 모델링이 정확하게 되었다고 가정하기 때문에 배경에 대한 모델링이 어려운 경우에는 제대로 동작하지 않는 한계를 지닌다.

감시 카메라 자체의 움직임(Pan/Tilt작동)으로 인해 배경이 움직이는 경우, 또는 배경에 움직임이 많은 사물이 있는 경우, 조도의 갑작스런 변화가 일어나는 경우 (실내환경에서 현광등의 깜빡거림등)에는 일반적으로 배경에 대한 모델링이 매우 어렵거나 불가능하다.

### 3. 동적 카메라 환경하의 물체추적

#### 3.1 Point Tracker에 기반한 물체추적

그림 3.1(a)나 (b)처럼 중심점이나 특징점들을 가지고 추적체를 표현하는 추적 알고리즘들을 Point Tracker라고 한다. 보통 이는 크기가 작은 물체를 표현하는데 사용하게 되며 물체의 크기가 커짐에 따라 추적의 정확도가 떨어지게 된다. 또 잡음이 있는 경우에 매우 약하며, 추적대상의 모양에 대한 정보를 제공해주지 못한다.

#### 3.2 템플릿 정합에 기반한 물체추적

템플릿 정합은 영상  $I_w$ 에 대하여 템플릿  $O_t$ 안

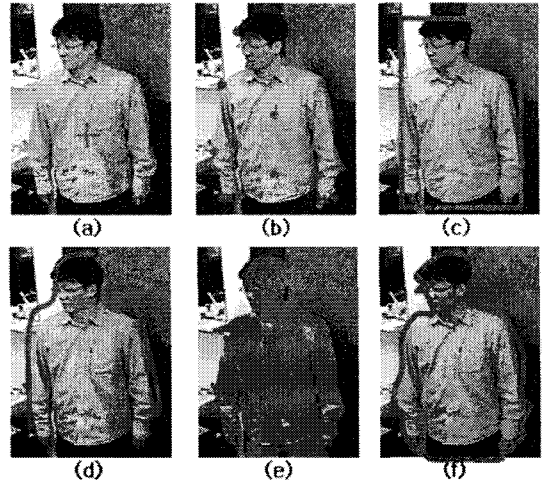


그림 3.1. (a) 중심점 (b) 특징점들 (c) 사각형 또는 그외의 쉬운 도형 (d) B-spline (e) 템플릿 (f) active contour

의 영역과 비슷한 영역을 찾는 방법이다. 비슷한 정도를 측정하는 측정방법으로는 예를 들어 cross-correlation이 사용되는데, 예를 들어 다음과 같은 수식이 사용될 수 있다:

$$\arg \max_{dx, dy} \frac{\sum_x \sum_y (O_t(x, y) \times I_w(x + dx, y + dy))}{\sqrt{\sum_x \sum_y O_t^2(x, y)}}$$

이렇게 해서 구해진 dx, dy이 다음 프레임 템플릿이 움직여간 변위를 나타낸다. 밝기값 이외에도 히스토그램이나, 가우시안 모델, 템플릿 내부의 색상의 평균값 등이 특징점들로 쓰일 수도 있다. 이외에도 밝기값이 조명의 변화에 민감하기 때문에 영상의 그레디언트값들이 특징점들로 쓰일 수도 있다. 간단한 도형을 쓰는 이외에도 추적 대상의 모양과 비슷한 템플릿을 만들어 사용하기도 한다. IBM의 W4 tracker등이 쓰고 있는 템플릿 정합에 기반한 알고리즘은 주로 정적 배경을 가진 경우에 사용되고 있으며, 이는 배경을 모델링한 후에 배경을 뺀 전경을 여러 프레임에 대해서 평균하여 먼저 움직이는 물체에 대한 템플릿을 만든

다 (그림 3.1 (e)). 이때 템플릿 내부에 구멍이 나 있는 곳 등은 모폴로지 필터 등의 기법을 이용해서 메꾼다. 이 템플릿을 기준으로 이후부터 들어오는 시퀀스에 대해서 템플릿과 정합해봄으로써 추적대상을 찾기도 한다.

그러나 템플릿 정합의 가장 큰 단점은 많은 계산량에 있다. 템플릿 안의 특징에 가장 잘 부합하는 영역을 찾기 위해 영상내를 검색하여야 하는데, 이때 템플릿의 위치를 바꿔가며 전 방위로 검색하여야 하기 때문에 계산량이 많아진다. 이 때문에 주로 전 프레임에서의 템플릿의 위치의 주변에서 검색하게 되는데, 여전히 계산량이 많은 편이다.

두 번째 단점은 간단한 도형을 템플릿으로 사용하였을 시에 배경의 특징들이 템플릿안에 포함이 된다는 점에 있다. 추적 물체를 원이나, 사각형 등과 같은 간단한 도형으로 표현하였을 때 추적하고자 하는 물체뿐만 아니라, 주변의 배경의 특징들도 함께 들어가게 되므로, 정확한 물체의 특징의 추출이나, correspondence를 구하지 못할 수도 있다. 그림 3.2는 추적대상을 사각형으로 표현하였을 때 사각형 내부의 히스토그램을 비교한 것이다. 사각형으로 추적대상을 정확하게 둘러치는 것

도 어렵지만 배경의 색상의 변화로 인한 히스토그램 분포의 차이를 볼 수 있다.

추적대상의 모양을 템플릿으로 사용하는 경우 추적대상의 모양이 거의 변하지 않는 Rigid motion에는 잘 적용되지만, nonrigid motion, 특히 갑작스러운 동작의 변화에는 잘 적용이 안 되는 한계점을 안고 있다.

이런 경우, spline등을 이용하여 추적물체에 대한 대략의 모양정보를 주고자 하는 추적 알고리즘들이 있다. 먼저 추적을 수행한 후에 추적의 결과 추출된 데이터를 가지고 모양정보를 모델링을 하는데 자칫 잘못된 모델링을 할 수 있는 경우가 많다는 한계점을 안고 있다. 그림 3.3은 Active shape tracker가 두 사람에 대한 부분 겹침이 일어날 때 다른 한 사람의 모양이 제대로 모델링이 되지 않은 경우를 보여주고 있다.

### 3.3 Mean Shift 추적기술

일반 템플릿 기술들은 전 영역을 검색하면서 템플릿 정합을 해야 하기 때문에 매우 소모적인 알고리즘들인데 반해 [4]에서는 전방위에 대하여 템플릿 정합을 수행하지 않고, Mean Shift 프로세

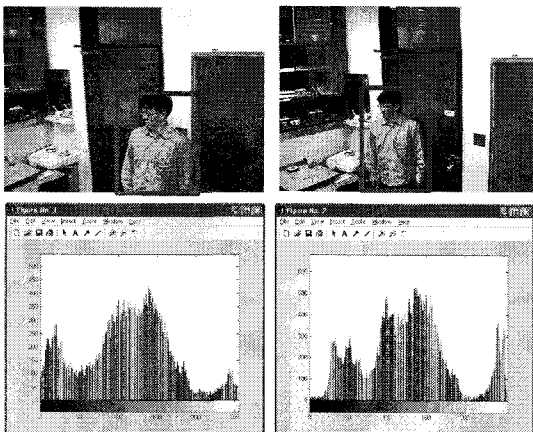


그림 3.2 단순도형을 사용하였을 때의 히스토그램



그림 3.3 Active Tracker에 의한 물체추적

스를 사용하여 템플릿이 위치를 찾아가도록 알고리즘을 제안하였다. 이 프로세스는 현재 타원형의 템플릿안의 히스토그램 분포와 가장 유사한 히스토그램 분포를 가진 영역으로 타원이 자동적으로 찾아가도록 제안된 알고리즘이다. 즉, 히스토그램의 분포의 동질성이 높아지도록 하는 Mean Shift 벡터를 계산하여 Mean Shift 벡터만큼 타원을 옮겨지게 한다. 이때 타원의 중심부에 있는 색상값들이 더 높은 가중치를 가지도록 설계되었다. Mean Shift 벡터는 다음의 식에 의해 계산된다.

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i \omega_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)}{\omega_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)} \quad (1)$$

여기서  $\hat{y}_0$ 은 이전 프레임에서 템플릿의 중심을,  $y_1$ 은 템플릿이 이동되어야 할 변위를 말하고,  $x_i$ 는 템플릿안에 있는 화소들의 위치벡터를,  $n_h$ 는 템플릿 내부의 화소들의 개수를 그리고  $h$ 는 적당한 상수를 말한다.

여기서  $g(x)$ 함수는 어떤 커널함수의 도함수인데, [4]에서는 커널함수를 다음과 같이 두었다.

$$k(x) = \begin{cases} 1-x & \text{if } \|x\| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

그러므로 이의 도함수는 다음과 같이 된다.

$$g(x) = -k(x) = \begin{cases} 1 & \text{if } \|x\| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

식 (1)이 의미하는 바는 가중치  $\omega_i$ 에 의해 가중된 무게중심을 향해 벡터  $y_1$ 가 생성된다는 것이다. 여기서 가중치 함수  $\omega_i$ 는 다음의 식에 의해서 계산되어진다.

$$\omega_i = \sum_{u=1}^m \sqrt{\frac{q_u}{\hat{p}_u(\hat{y}_0)}} \delta[b(x_i) - u] \quad (2)$$

여기서  $\hat{q}_u$ 는 전 프레임에서의 템플릿안의 히스

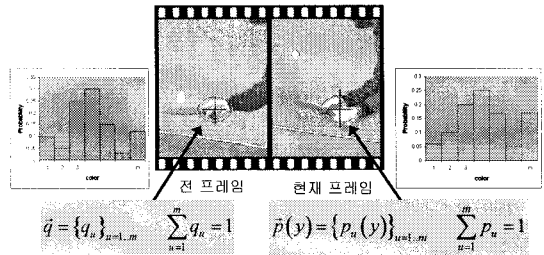


그림 3.4 프레임간의 히스토그램 비교

토그램 중  $u$ 번째 bin의 확률을 나타내며,  $\hat{p}_u(\hat{y}_0)$ 은 현재 프레임에서  $y$ 에 중심을 둔 템플릿안의 히스토그램중  $u$ 번째 bin의 확률을 나타낸다. (그림 3.4) 그러므로  $\omega_i$ 는  $i$ 번째 위치에서 발견되는 bin  $u$ 에 대한 함수가 된다.

그러므로 가중치함수가 뜻하는 바는 현재 프레임에서 초기 위치  $\hat{y}_0$ 에 중심을 둔 템플릿안에서 각 픽셀에서 발견되는 bin의 값이 전 프레임의 같은 픽셀에서 발견되는 bin의 값보다 작다면 그 bin에 해당하는 가중치값을 크게 하라는 뜻이다. 즉, 그 bin에 해당하는 색상이 많이 움직인 것으로 판단되었으므로, 그 bin의 값을 가진 화소들의 무게중심으로는 많이 shift되라는 뜻이다.

그림 3.5는 Mean Shift 알고리즘에 의해 타원의 template이 점차로 다음 프레임의 물체 주변으로 shift이 되는 과정을 보여주고 있다.

Mean Shift 알고리즘은 근본적으로 색상정보

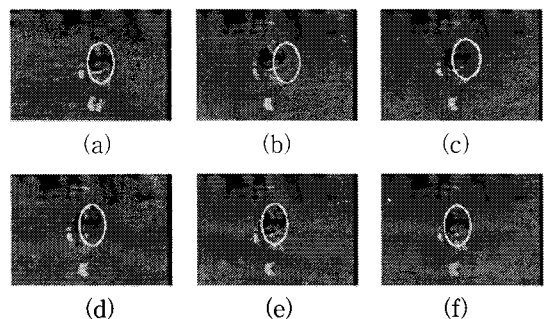


그림 3.5 Mean shift에 의한 물체추적 [1]

에 기초한 알고리즘으로 다음과 같은 한계점을 안고 있다.

1) 비슷한 색상값이 주변이 나타나면 잘못 shifting이 됨

주변에 물체의 색상과 비슷한 색상값이 배경에 나타나면 자칫 색상의 무게중심이 배경에 있을 수 있기 때문에 mean shift이 잘못된 방향으로 갈 수 있는 단점을 안고 있다.

2) 간단한 도형의 템플릿의 사용으로 인한 물체의 부정확한 표현

Mean Shift알고리즘에서는 타원, 사각형등의 간단한 도형을 사용하여 물체를 표현하기 때문에 물체가 완전하게 잡히지 않는다. 그렇기 때문에 도형을 크게 하여 잡으면 배경의 색상이 같이 포함되어 들어갈 수 밖에 없게 된다. 배경의 색상이

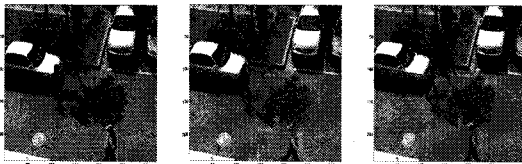


그림 3.6 mean shift에 의한 추적 실패 사례

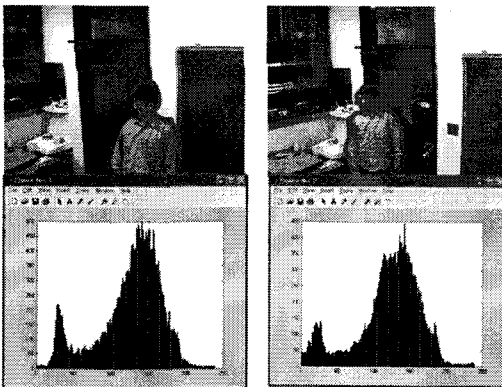


그림 3.7 Active contour를 이용한 추적대상의 보다 정확한 영역의 분할 및 active contour내부의 히스토그램의 표현



그림 3.8 기존의 Meanshift 추적 결과 (위)와 similarity를 달리한 Meanshift의 추적 결과(아래)의 비교

함께 들어가게 되면 1)의 한계점의 문제가 나타나게 된다. 그림 3.6은 물체의 주변에 너무 큰 도형을 씌우므로 추적에 실패한 예를 보여주고 있다.

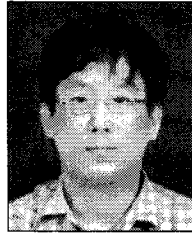
이와 같은 Mean shift의 단점을 극복하기 위해 Active contour기반의 영역분할을 병행하여 보다 정확한 추적대상의 히스토그램을 추출하거나 (그림 3.7), similarity함수에 거리, 위치등의 정보를 포함하여 mean shift알고리즘을 안정화시키는 알고리즘들이 제안되고 있다[5,6].

### 참 고 문 헌

- [1] Alper Yilmaz, "Object Tracking and Activity Recognition in Video Acquired Using Mobile Cameras," dissertation, 2004.
- [2] C.R. Wren, A. Azarbayejani, and A. Pentland. "Pfinder: Real-Time Tracking of the Human Body." IEEE Trans. on Pattern Analysis and Machine Intelligence, 19(7):780-785, 1997.
- [3] X. Gao, T.E. Boult, F. Coetzee, and V. Ramesh. "Error Analysis of Background Subtraction." in IEEE Int. Conf. on Computer Vision, 2000.
- [4] D. Comaniciu, v. Ramesh, and P. Meer. "Kernel-Based object tracking." IEEE Trans. on Pattern Analysis and Machine Intelligence, 25 :564-575, 2003.
- [5] Dong Xu, Yimin Wang, and Jinwen An: "Applying a New Spatial Color Histogram in

Mean-Shift Based Tracking Algorithm,”  
Image and Vision Computing New Zealand,  
2005.

- [6] C.J. Yang, R. Duraiswami, L.S. Davis,  
‘Efficient Mean-Shift Tracking via a New  
Similarity Measure’, CVPR05, 2005, 1, pp.  
176-183.



이 석 호

- 1993년 연세대학교 전자공학과 (공학사)
  - 2003년 연세대학교 전기전자공학과 (공학박사)
  - 2003년~2006년 연세대학교 수학과 BK21박사후연구원
  - 2006년~2008년 연세대학교 전기전자공학과 BK21 연구  
교수
  - 2008년~현재 동서대학교 컴퓨터정보공학부 교수
  - 관심분야 : 영상처리, 컴퓨터비전 등
- 
-