

논문 2014-51-9-16

# 이중계층구조 파티클 샘플링을 사용한 다중객체 검출 및 추적 ( Multi-Object Detection and Tracking Using Dual-Layer Particle Sampling )

정 경 원\*, 김 나 현\*, 이 승 원\*, 백 준 기\*\*

( Kyungwon Jeong, Nahyun Kim, Seungwon Lee, and Joonki Paik<sup>Ⓞ</sup> )

## 요 약

본 논문에서는 다중객체 검출과 동시에 추적을 수행하는 이중계층구조의 파티클 샘플링을 제안한다. 제안된 방법은 다중 객체 검출을 위한 상위 계층 파티클 샘플링과 검출된 객체의 추적을 위한 하위 계층 파티클 샘플링으로 구성된다. 상위 계층에서는 빠른 객체 검출을 위해 슬라이딩 윈도우 대신 움직임 추정 기반의 부모 파티클 (parent particles; PP) 윈도우를 사용하여, 이동 객체 주위로 리샘플링된 파티클을 통해 객체를 검출한다. 하위 계층에서는 상위 계층에서 검출한 객체의 객체영역에 자식 파티클 (child particles; CP)을 생성하여 해당 객체를 추적한다. 실험결과를 통해 비디오 시스템에서 기존 객체 검출 방법보다 빠른 검출이 가능하고, 다중 객체를 효과적으로 추적할 수 있음을 확인하였다.

## Abstract

In this paper, we present a novel method for simultaneous detection and tracking of multiple objects using dual-layer particle filtering. The proposed dual-layer particle sampling (DLPS) algorithm consists of parent-particles (PP) in the first layer for detecting multiple objects and child-particles (CP) in the second layer for tracking objects. In the first layer, PPs detect persons using a classifier trained by the intersection kernel support vector machine (IKSVM) at each particle under a randomly selected scale. If a certain PP detects a person, it generates CPs, and makes an object model in the detected object region for tracking the detected object. While PPs that have detected objects generate CPs for tracking, the rest of PPs still move for detecting objects. Experimental results show that the proposed method can automatically detect and track multiple objects, and efficiently reduce the processing time using the sampled particles based on motion distribution in video sequences.

**Keywords** : 파티클 필터 (particle filter), 객체 검출 (object detection), 객체 추적 (object tracking),  
파티클 윈도우 (particle window)

\* 학생회원, \*\* 평생회원, 중앙대학교 첨단영상대학원  
(Dept. of Image Engineering, Graduate School of  
Advanced Image Science, Multimedia, and Film,  
Chung-Ang University)

Ⓞ Corresponding Author (E-mail: paikj@cau.ac.kr)

※ 본 연구는 산업통상자원부 기술혁신사업 (SW융합형  
부품 기술개발사업 스마트 현장판단형 영상/음향 보  
안감시 SoC 및 핵심부품 개발, 10047788), 미래창조  
과학부 및 정보통신기술진흥센터의 정보통신·방송  
연구개발사업의 일환으로 수행하였고 14-824-09-  
002, (2세부) 대규모 실시간 비디오 분석에 의한 전  
역적 다중 관심객체 추적 및 상황 예측 기술 개발]  
과 미래창조과학부 및 정보통신산업진흥원의 대학  
IT연구센터 지원사업 (NIPA-2014- CAU)의 지원으  
로 수행되었습니다.

접수일자: 2014년6월09일, 수정일자: 2014년7월30일,  
수정완료: 2014년9월03일

## I. 서 론

최근 공공시설, 놀이공원, 백화점, 지하철 등 인구가  
밀집된 지역에서의 테러 위협이 증가하고 있으며, 유동  
인구가 적은 장소 및 통행량이 감소하는 시간대의 범죄  
및 사건·사고가 각종 미디어를 통해 끊임없이 보도되면  
서 보안 및 범죄 예방에 대한 관심이 증가하고 있으며  
감시 시스템의 연구가 활발히 이루어지고 있다<sup>[1~2]</sup>. 특  
히 카메라로부터 취득된 비디오 영상에서 관심영역을  
추출하는 객체 검출 분야와 관심 객체를 지속적으로 추  
적하는 객체 추적분야는 그 대상이 사람, 동물, 자동차,

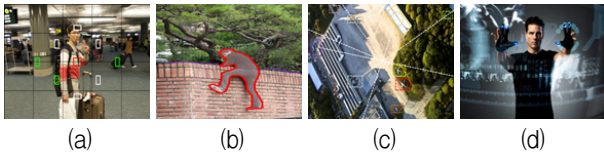


그림 1. 객체 검출과 객체 추적의 다양한 적용 분야; (a) 객체 자동초점, (b) 이상행위 검출, (c) 무인항공기, (d) 인간-컴퓨터 인터페이스

Fig. 1. Various applications of object detection and tracking; (a) object auto-focusing, (b) unusual behavior, (c) unmanned aerial vehicle, and (d) human-computer interaction.

및 기타 사물 등으로 확대되어 움직이는 객체의 자동초점, 교통상황 모니터링, 스포츠 전력 분석, 사람-컴퓨터 인터페이스, 증강현실과 같은 다양한 분야에 적용되고 있다<sup>[3-4]</sup>. 그림 1에서 객체의 검출 및 추적의 대표적인 응용 사례를 보인다.

기존의 객체 검출 및 추적 방법은 크게 확률론적인 방법과 결정론적인 방법으로 구분된다<sup>[5-6]</sup>. 현재 널리 사용되고 있는 Adaboost 방법은 결정론적인 방법의 한 종류로써 특정 객체에 대하여 학습된 분류기를 통해 객체를 검출하고, 검출된 객체의 위치들을 연속된 프레임에서 연결함으로써 객체를 추적한다. 그러나 Adaboost 기반의 방법은 객체의 왜곡이나 겹침현상이 발생하게 되면 검출 및 추적을 실패하는 단점을 가지고 있다.

대표적인 확률론적 접근법으로는 칼만 필터와 파티클 필터를 사용한 방법이 있다<sup>[7]</sup>. 칼만 필터와 파티클 필터는 관측된 객체 정보를 측정하여 동적 시스템에서 현재의 객체를 추정 및 보정하는 방법이다. 하지만 칼만 필터의 객체 정보 관측값은 확률적인 오차를 포함하고 있으며, 동적 시스템에서 객체의 상태가 가우시안 분포의 선형 모델을 가정하고 있기 때문에, 비선형 특성을 갖는 대부분의 실제 환경에서는 객체 추적을 실패하는 문제점을 가지고 있다. 이를 해결하기 위해 일반적으로 사용되는 파티클 필터는 무작위로 발생된 파티클들의 가중치를 통해 분포를 추정하기 때문에 다중 모드를 표현할 수 있기 때문에 비선형 특성을 가진 실제 환경에서도 적용이 가능하여 안정적인 객체 추적이 가능하다.

Okuma는 결정론적인 접근법인 Adaboost 기반의 객체 검출과 파티클 필터를 이용한 다중객체 검출 및 추적 방법을 제안하였다<sup>[8]</sup>. 그러나 이 방법은 영상 전체 영역에서 슬라이딩 윈도우를 사용해서 객체를 검출하기 때문에 계산량이 많은 문제점을 가지고 있다. Gualdi는

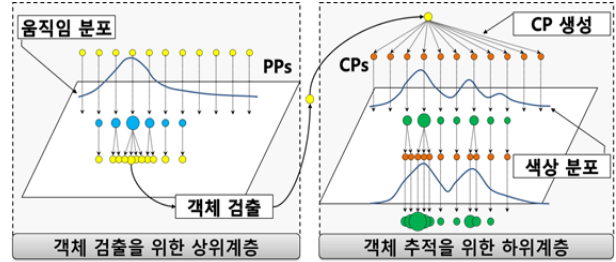


그림 2. 제안된 이중 계층 파티클 샘플링의 개념도  
Fig. 2. The concept of the proposed DLPS.

검출 단계에서 계산량을 줄이기 위해 단일 영상에서 파티클 필터를 이용한 효율적인 객체 검출 방법을 제안하였다<sup>[9]</sup>. 그러나 이 방법은 단일 영상을 위한 리샘플링 구조로 되어 있기 때문에 비디오에는 적용하기 어렵다.

기존 객체 추적방법은 기본적으로 객체의 초기 위치가 제공되고 초기 객체 위치로부터 객체 추적을 수행하게 된다. 그러나 지능형 감시시스템과 같은 환경에서는 감시하고자 하는 객체에 대하여 자동으로 객체의 위치를 제공해 주어야 하기 때문에, 초기 객체 정보 제공을 위한 객체 검출 방법이 필요하다. 본 논문에서는 이중계층의 파티클 샘플링 (dual layer particle sampling; DLPS)을 사용해서 객체의 검출과 추적을 동시에 수행할 수 있는 방법을 제안한다. 그림 2는 제안된 이중 계층 파티클 샘플링의 개념도를 보인다.

제안된 DLPS는 그림 2에서 보는 것과 같이 두 개의 계층으로 이루어져 있다. 상위 계층에서는 부모 파티클들 (parent-particles; PP)을 사용하여 객체를 검출하고, 하위 계층에서는 자식 파티클들 (child-particles; CP)을 사용하여 PP로부터 검출된 객체를 추적함으로써 객체 검출과 추적을 동시에 가능하게 한다.

객체 검출을 위한 기존의 검색 방법은 영상의 전체 영역에서 그림 3(a)와 같이 슬라이딩 윈도우를 사용한다. 이 경우 객체가 존재하지 않는 배경영역에서도 검색을 수행하기 때문에 연산량이 증가하는 단점이 있다. 제안된 방법에서는 파티클 샘플링의 상위 계층에서 확률론적 방법을 사용하여 그림 3(b)와 같이 PP들이 존재하는 영역에서만 객체 검색을 수행하여 계산량을 줄인다. 감시카메라 영상에서 객체는 움직임을 가지고 있다고 가정한 후, 영상에서 움직임을 검출하고 분포를 분석하여 PP의 리샘플링을 수행한다. 리샘플링된 부모 파티클에서 IKSVM (intersection kernel support vector machine)을 통해 특정 객체에 대하여 사전에 학습된 특

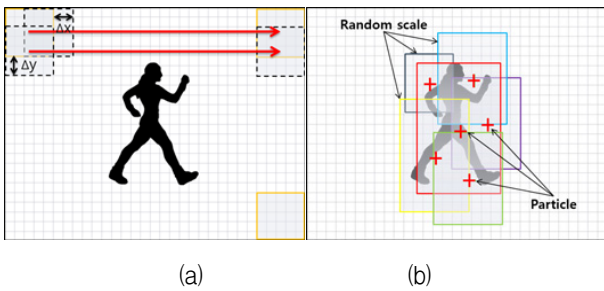


그림 3. 객체 검출을 위한 영상에서의 객체 검색 방법; (a) 슬라이딩 윈도우, (b) 파티클 윈도우  
Fig. 3. Searching method in the image for object detection; (a) sliding window, (b) particle window.

정을 사용하여 객체 검출을 수행한다<sup>[11]</sup>.

만약 상위 계층의 PP 중에서 임의의 PP가 객체로 판단되면, 해당 PP는 객체 추적을 위해 하위 계층을 위한 CP를 생성하게 된다. 객체로 판단된 PP는 초기 객체의 위치를 제공하고 CP는 객체 위치로부터 객체 정보를 추출하여 HSV(Hue-Saturation-Value) 기반 객체 모델을 생성한다. 생성된 CP는 해당 객체를 검출한 PP와 함께 지속적으로 객체를 추적하고, 객체가 영상 밖으로 사라지면 CP는 제거되고 해당 PP는 객체의 재검색을 위해서 상위계층으로 반환된다.

본 논문은 다음과 같이 구성된다. II장에서는 이중 계층 파티클 샘플링을 제안하고, III장에서는 실험결과를 제시한다. 그리고 IV장에서는 결론을 맺는다.

## II. 이중계층구조의 파티클 샘플링

본 논문에서 제안하는 DLPS는 다중 객체 검출을 위한 상위 계층 파티클 샘플링과 검출된 객체의 추적을 위한 하위 계층 파티클 샘플링으로 구성된다. 상위 계층에서는 부모 파티클 (parent particles; PP) 윈도우를 사용하여 움직임 추정 기반의 파티클 리샘플링을 통해 슬라이딩 윈도우 방식보다 빠른 객체 검출을 수행한다. 다음으로 하위 계층에서는 상위 계층에서 검출한 객체의 객체영역에 자식 파티클 (child particles; CP)을 생성하여 해당 객체를 추적함으로써 다중 객체 추적을 수행한다. 제안된 DLPS의 전체 블록도는 그림 4와 같으며, 상위 계층과 하위 계층의 두 개의 파티클 샘플링을 위한 계층으로 구분된다.

상위 계층은 객체 검출을 위해 인접 프레임 간의 움

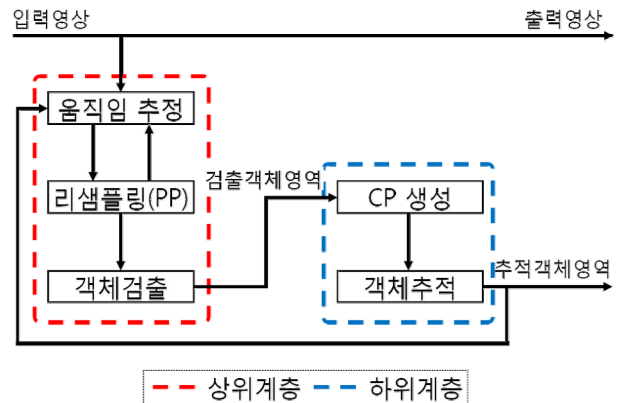


그림 4. 제안 방법의 블록도  
Fig. 4. Block diagram of the proposed method.

직임 정보를 기반으로 관측 모델을 생성하고 각 모델의 관측 우도(likelihood)를 계산하여 PP의 가중치를 갱신한다. 움직임분포에 따라 예측과 갱신단계를 거쳐 객체의 주위로 PP가 리샘플링되게 된다. 리샘플링 된 PP는 랜덤 스케일 기반의 객체 검출기를 사용하여 객체를 검출한다. 다음으로 임의의 PP가 객체를 검출하면, 해당 PP는 검출객체 정보를 하위계층으로 전달한다. 하위계층은 검출된 객체 주위에 존재하는 PP의 정보를 바탕으로 추적을 위한 CP를 생성한다. 생성된 CP는 검출된 객체정보를 초기 조건으로 사용하여 객체의 크기와 위치를 포함하는 객체 상태 모델과 HSV 색상 히스토그램을 기반으로 하는 객체의 색상 모델을 생성하여 지속적으로 검출된 객체를 추적하게 된다.

### 2.1. 객체 검출을 위한 상위 계층

고정형 카메라를 기반으로 하는 감시시스템 영상은 움직임 분포가 존재하는 전경(foreground) 영역과 분포가 없는 배경(background) 영역으로 구분 할 수 있다. 영상에서 객체를 검출하기 위해 배경 생성을 사용한 배경 차분 방법이 계산 복잡도가 낮아 많이 사용되지만 환경 변화와 잡음에 민감하기 때문에 객체 검출의 정확도가 떨어지고 객체 분류가 불가능한 단점을 가지고 있다. 반면에 학습기반 객체 검출 방법은 조명이나 그림자와 같은 환경 변화에도 객체 검출이 잘 되며 객체 분류가 가능한 장점을 가지고 있다. 하지만 객체의 크기 변화를 고려하기 위해 피라미드 영상을 생성한 후 슬라이딩 윈도우를 사용한 전역 검색(full search)을 통해 객체가 존재하지 않는 배경 영역에서도 객체 검출을 수행

하기 때문에 계산 복잡도가 높은 단점을 가지고 있다. 제안된 DLPS는 상위 계층에서 움직임이 발생하는 영역에서만 객체 검출을 위한 PP를 생성하고, 상태 예측, 갱신, 그리고 리샘플링 과정을 통해 객체를 검출하기 때문에 움직임이 없는 배경 영역에서 불필요한 검색을 수행하지 않으며, 외부 환경 변화에도 강인하다.

객체 검출 과정에서, 인접한 프레임간의 차이로부터 움직이는 객체의 거리, 방향, 속도를 연속적인 흐름으로 표현한 것을 광류흐름(optical flow)이라 한다. 영상 내에서 이동하는 객체의 광류흐름을 구하기 위해 대표적 방법으로 루카스-카나데 광류흐름 (lucas-kanade optical flow; LKOF) 추정 방법이 있다<sup>[12]</sup>.

이 방법은 영상에 가우시안 피라미드를 적용하여 인접한 프레임간의 이동 벡터를 구할 수 있으며 연산량이 적기 때문에 실시간으로 구현 가능하다는 장점이 있다. 구체적으로, 움직임벡터  $\mathbf{M}$ 은 광류 방정식에 의해 다음과 같이 최소평균제곱법에 의해 계산된다.

$$\mathbf{M} = \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} \sum_{\Omega} I_x(x,y)^2 & \sum_{\Omega} I_x(x,y)I_y(x,y) \\ \sum_{\Omega} I_x(x,y)I_y(x,y) & \sum_{\Omega} I_y(x,y)^2 \end{bmatrix}^{-1} \times \begin{bmatrix} -\sum_{\Omega} I_x(x,y)I_t(x,y) \\ -\sum_{\Omega} I_y(x,y)I_t(x,y) \end{bmatrix}. \quad (1)$$

여기서  $\Omega$ 는 윈도우 영역을 나타내고,  $x$ 와  $y$ 는 차례대로 가로 및 세로 방향의 좌표이다.  $I_x$ ,  $I_y$ ,  $I_t$ 는 차례대로  $x$ 축,  $y$ 축, 시간  $t$ 에 대한 미분을 나타낸다.

움직이는 객체를 중심으로 리샘플링을 하기 위해서 상태공간에서 파티클 필터의 샘플들의 상태모델을 고려한다. 즉, PP의 시간  $k$ 에 대한 상태벡터는  $s_k = [p_k^x, p_k^y, v_k^x, v_k^y]^T$ 로 나타낼 수 있으며 상태 전이 모델은 다음과 같이 정의된다.

$$\begin{aligned} p_k^x &= p_{k-1}^x + v_{k-1}^x \Delta k \\ p_k^y &= p_{k-1}^y + v_{k-1}^y \Delta k \\ v_k^x &= v_{k-1}^x \\ v_k^y &= v_{k-1}^y \end{aligned}, \quad (2)$$

여기서  $p_k^x$ 와  $p_k^y$ 는 샘플의  $x$ 축과  $y$ 축에서의 위치좌표이며,  $v_k^x$ 와  $v_k^y$ 는 각각 샘플의  $x$ 축과  $y$ 축에 대한 속도를 나타낸다.  $\Delta k = 1$ 을 가지는 이산적 상태 전이 방정식은 다음과 같이 정의된다.

$$s_k^- = A s_{k-1}^+ + \mathbf{n}_k, \quad (3)$$

여기서

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{n}_k = \begin{bmatrix} N(0, Q_{pos}) \\ N(0, Q_{pos}) \\ N(0, Q_{vel}) \\ N(0, Q_{vel}) \end{bmatrix} \quad (4)$$

를 나타내고,  $Q_{pos}$ ,  $Q_{vel}$ 는 각각 위치와 속도에 대한 시스템 잡음의 공분산을 나타낸다.  $s_k^-$ 는 시간  $k$ 에서 예측된 상태이고,  $s_{k-1}^+$ 는 시간  $k-1$ 에서의 상태이며,  $\mathbf{n}_k$ 는 잡음 모델이다. 관측 단계에서의 관측 모델  $\mathbf{y}_k$ 는 예측상태  $s_k^-$ 에 의해 다음과 같이 나타낼 수 있다.

$$\mathbf{y}_k = M(p_k^x, p_k^y) \quad (5)$$

여기서  $M(p_k^x, p_k^y) = \sqrt{(v_k^x)^2 + (v_k^y)^2}$ 는  $x$ ,  $y$ 위치에서의 움직임의 크기를 나타낸다.  $N$ 개의 후보 PP의 정규화된 가중치  $\mathbf{w}^i$ 는 다음과 같이 계산된다.

$$\mathbf{w}^i = \frac{1}{\sqrt{2\pi} R} \exp\left\{-\frac{\|\mathbf{y}_D^i\|^2}{2R^2}\right\}, \forall i = 1, \dots, N \quad (6)$$

여기서  $i$ 는 후보 PP의 인덱스이고,  $R$ 은 움직임 추정 잡음의 표준편차를 나타낸다.  $\|\mathbf{y}_D\| = \mathbf{y}_0 - \mathbf{y}_k$ 이고,  $\mathbf{y}_0$ 는 객체의 움직임 크기를 나타내며 실험을 통해서 8로 설정하였으며,  $M(x, y)$ 의 최대값을 8로 제한한다.  $\mathbf{w}^i$ 를 사용하여 로그 우도(log likelihood)  $L^i$ 은 다음과 같이 계산된다.

$$\begin{aligned} L^i &= \log(\mathbf{w}^i) \\ &= -\log(\sqrt{2\pi} R) - \frac{\|\mathbf{y}_D^i\|^2}{2R^2}, \forall i = 1, \dots, N. \end{aligned} \quad (7)$$

이렇게 계산된 로그 우도를 사용하여, 누적 값  $c_i$ 를 다음과 같이 구한다.

$$c_i = \sum_{m=1}^i \frac{L_m}{L_{total}}, \forall i = 1, \dots, N, \quad (8)$$

여기서  $L_{total}$ 은  $L$ 의 총 합을 나타낸다. 다음으로 정규 분포를 따르는 실수의 랜덤 변수  $u_i (i = 1, \dots, N)$ 를 생성하고,  $c_{j-1} < u_i$ 와  $c_j \geq u_i$  조건을 동시에 만족하는

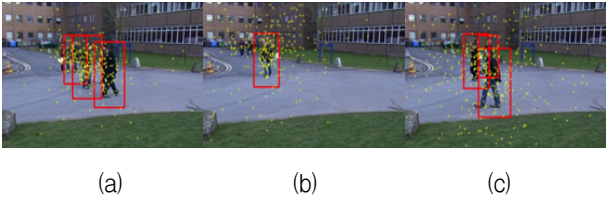


그림 5. 파티클 윈도우를 이용한 객체 검출 결과. (적색: 검출된 객체 영역, 노란색: PP); (a) 104번째 영상, (b) 205번째 영상, (c) 270번째 영상

Fig. 5. Experimental results of object detection using the particle window. (red: detected object region, yellow: PP); (a) 104th image, (b) 205th image, and (c) 270th image.

양의 정수  $j$ 가 결정되면  $s_{k,i}^+ \leftarrow s_{k,j}^-$ 으로 현재 상태를 갱신한다.

움직임 관측모델에 의해 리샘플링된 PP는 이동 객체 주위로 샘플들이 밀집되게 되고, 각 PP의 위치에서 랜덤으로 선택된 객체 스케일과 사전에 IK SVM을 사용하여 학습된 객체 검출 모델을 사용하여 객체를 검출한다<sup>[11]</sup>. 영상에서 움직임 관측모델에 의해 분포된 PP가 객체를 검출하면, 해당 PP의 위치  $(c_p^x, c_p^y)$ 에서 객체 영역의 너비( $w_p$ ), 높이( $h_p$ ), 내부의 PP 분포 정도( $\sigma_p$ )를 통해 CP의 초기 분포 모델을 결정한다. 이후로 해당 PP에서 검출된 객체 영역의 정보를 하위계층으로 전달하여 객체 추적을 위한 초기 정보로 이용한다. 반면에 객체를 검출하지 못한 PP는 검출된 객체를 제외한 영역에서 움직임 모델에 의해 다시 리샘플링을 수행하게 된다. 그림 5는 상위 계층에서 검출된 객체 주변 영역에 PP들이 모이게 되고, PP에 의해 객체가 검출되는 과정을 보인다.

## 2.2. 객체 추적을 위한 하위 계층

하위계층은 상위계층의 PP로부터 제공 받은 객체영역의 정보  $\mathbf{T}^T = [w_p, h_p, (c_p^x, c_p^y), \sigma_p]$ 를 사용하여 객체 추적을 위한 객체 모델의 초기조건을 생성하고, PP의 위치를 중심으로 추적을 수행하기 위한 CP를 생성한다. 하위계층에서 객체의 상태 벡터  $\mathbf{x}_k$ 는 객체 위치  $\mathbf{x}_k^p$ 와 객체 크기  $\mathbf{x}_k^s$ 로 구성되며, 관측값  $\mathbf{z}_k$ 는 조명에 의한 조도변화에 의한 영향을 줄이기 위해서 밝기와 채도 성분을 분리할 수 있는 HSV 색상 히스토그램을 사용한다<sup>[10]</sup>. 시간  $k-1$ 까지의 모든 관측값이 주어졌을 때 시간  $k$ 에서의 상태 확률밀도 함수를 알면, 채프먼-콜모고로

프(Chapman-Kolmogorov) 방정식을 사용하여 시간  $k$ 에서의 상태를 다음과 같이 예측할 수 있다.

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (9)$$

여기서  $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$ 는 사전확률을 나타내고,  $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ 은 2차 자동회귀 방법에 의해서 계산된 예측모델이며, 잡음 모델은 가우시안 분포를 갖는 것으로 한다. 시간  $k$ 에서 새로운 측정값  $\mathbf{z}_k$ 가 주어지면 사전확률은 베이시안 규칙에 의해 갱신되어 다음과 같이 사후확률을 얻는다.

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})}, \quad (10)$$

여기서  $p(\mathbf{z}_k | \mathbf{x}_k)$ 는 관측모델을 나타낸다. 파티클 필터에서 사후확률  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ 는  $N$ 개의 샘플과 그의 가중치  $w_k^i$ 로 근사화 될 수 있다. 후보 샘플  $\tilde{\mathbf{x}}_k^i$ 는 중요 분포  $g(\mathbf{x}_k | \mathbf{x}_{1:k-1}, \mathbf{z}_{1:k})$ 로부터 샘플링되며 독립적으로 계산된다. 샘플들의 가중치  $w_k^i$ 는 다음과 같이 정의된다.

$$w_k^i = \frac{p(\mathbf{z}_k | \tilde{\mathbf{x}}_k^i) p(\tilde{\mathbf{x}}_k^i | \mathbf{x}_{k-1}^i)}{g(\mathbf{x}_k | \mathbf{x}_{1:k-1}, \mathbf{z}_{1:k})}, \quad (11)$$

시간  $k$ 에서 색상분포  $\mathbf{q}_k(m; \mathbf{x}) = \{q_k(m; \mathbf{x})\}_{m=1, \dots, H}$ 는 다음과 같이 정의된다.

$$\mathbf{q}_k(m; \mathbf{x}) = K \sum_{l \in R(\mathbf{x})} \delta[h_k(l) - m], \quad (12)$$

여기서  $h_k(l)$ 는 시간  $k$ 일 때  $l$ 위치에서의 빈(bin) 인덱스로 정의되며,  $\delta$ 는 크로네커(Kronecker) 델타 함수,  $K$ 는 정규화상수,  $R(\mathbf{x})$ 는 샘플링된 객체 영역을 나타낸다. 관측모델  $\mathbf{q}_k$ 와 참조모델의 유사도는 바타차야(Bhattacharyya) 거리 측정 방법을 사용하여 다음과 같이 계산된다.

$$\mathbf{B}[\mathbf{q}^*, \mathbf{q}_k(\mathbf{x})] = \left[ 1 - \sum_{m=1}^H \sqrt{q^*(m) q_k(m; \mathbf{x})} \right]^{\frac{1}{2}}, \quad (13)$$

여기서  $\mathbf{q}^*$ 는 참조 색상 모델로 참조모델의 분포는 상위계층에서 PP가 객체를 검출 한 초기 시간  $k_0$ 일 때 객체 영역에서 생성된다. 관측모델의 유사도 측정에 의한 샘플들의 가중치  $w_k^i$ 는 다음과 같이 식 (13)에 의해 재



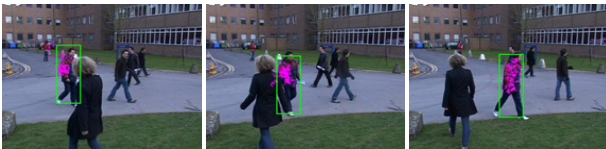
정의 될 수 있다.

$$w_k^i \propto e^{-\lambda B^0[\mathbf{q}_k^*, \mathbf{q}_k^i(x)]}, \quad (14)$$

여기서  $\lambda$ 는 관측모델 유사도의 가중치 상수로 본 논문에서는 15로 설정하였다. 시간  $k$ 에서의 최대 사후확률 (maximum a posterior; MAP) 추정 값은  $N$ 개의 샘플로부터 다음과 같이 정의 된다.

$$x_k^{MAP} = \underset{x_k^i}{\operatorname{argmax}} w_k^i, \quad \forall i = 1, \dots, N. \quad (15)$$

$x_k^{MAP}$  값은 객체의 현재 시간에서의 최종 상태를 결정하고, 하위계층에서는 CP에 의해 객체의 상태  $x_k^{MAP}$ 가 지속적으로 갱신되며 객체를 추적한다. 그러나 CP에 의해 추적 도중에 객체의 움직임이 상위 계층에 반영되어 지속적으로 관측되면, 추적되는 객체에 PP가 집중되는 현상이 발생된다. 이러한 현상은 새로운 객체가 나타나도 PP의 샘플들이 새로운 객체에 집중되지 않고, 이미 CP가 할당된 객체 영역에 중복되어 CP를 생성하는 문제를 발생시킨다. 이러한 문제를 해결하기 위하여 추적 객체 영역 정보  $[w_c, h_c, (\mu_c^x, \mu_c^y)]$ 를 상위 계층으로 반환하고, 추적되는 객체 영역의 움직임 크기를 0으로 설정함으로써 CP에 의해 추적되고 있는 객체 영역에서는 PP가 리샘플링 되는 것을 방지한다. 마지막으로  $x_k^{MAP}$ 를 가지는 샘플의 가중치  $w_k$ 가 일정한 값 이하로 떨어지게 되면 CP가 추적하고 있는 객체가 사라지거나 객체를 놓친 것으로 판단하여 해당 객체를 위한 CP를 제거한 후, 그의 부모 샘플인 PP는 상위 계층에서 리샘플링 되어 객체 검출을 위한 PP의 역할을 다



(a) (b) (c)

그림 6. 객체 추적 결과. (녹색: 추적 객체 영역, 분홍색: CP); (a) 740번째 영상, (b) 744번째 영상, (c) 748번째 영상

Fig. 6. Experimental results of object tracking using the HSV-based particle filter. (green: tracked object region, magenta: CP); (a) 740<sup>th</sup> image, (b) 744<sup>th</sup> image, and (c) 748<sup>th</sup> image.

시 수행하게 된다. 그림 6은 하위계층에서 하나의 객체를 지속적으로 추적하고 있는 과정을 보인다.

### III. 실험 및 분석결과

제안된 방법의 성능 평가를 보이기 위해 PETS-2009, i-LIDS AVSS-2007, CAVIAR-2004의 공용 데이터 셋과 SONY의 CMOS 카메라를 사용하여 직접 촬영한 영상을 사용하였다. 제안된 방법은 인텔코어 i7 환경에서 MATLAB 2013b을 사용하여 구현되었다.

그림 7은  $720 \times 576$  해상도를 갖는 PETS 2009 영상에서 제안된 방법을 사용하여 객체를 검출하고 동시에 객체를 추적한 결과 이다. 그림 7(a)는 객체 움직임을 기반으로 PP들이 객체 주위로 리샘플링 되고, 그림 7(b)는 파티클에 의해 객체를 검출하는 것을 확인할 수 있다. 그림 7(c)는 객체를 검출 후 CP를 통해 검출된 객체의 추적을 수행하게 되고 추적되는 객체영역의 움직임 분포가 0이 되고 그 외의 영역에서 움직임이 존재하지 않아 PP들이 영상에 전반적으로 리샘플링되는 것을 확인할 수 있다. 그림 7(d), 7(e), 7(f)는 다중 객체를 안정적으로 추적하고 있는 것을 보여주며 추적되는 객체 영역의 움직임 분포는 0이므로 PP가 추적객체 주위로 리샘플링 되지 않는 것을 확인할 수 있다.



(a) (b) (c)



(d) (e) (f)

그림 7. PETS 2009영상에서의 제안된 방법의 실험결과 (적색: 검출된 객체 영역, 녹색: 추적된 객체 영역, 노란색: PP, 자색: CP); (a) 800번째 영상, (b) 804번째 영상, (c) 805번째 영상, (d) 842번째 영상, (e) 866번째 영상, (f) 900번째 영상

Fig. 7. Experimental results of the proposed method for PETS 2009 sequence (red: detected object region, yellow: PP, green: tracked object region, magenta: CP); (a) 800<sup>th</sup> frame, (b) 804<sup>th</sup> frame, (c) 805<sup>th</sup> frame, (d) 842<sup>th</sup> frame, (e) 866<sup>th</sup> frame, and (f) 900<sup>th</sup> frame.

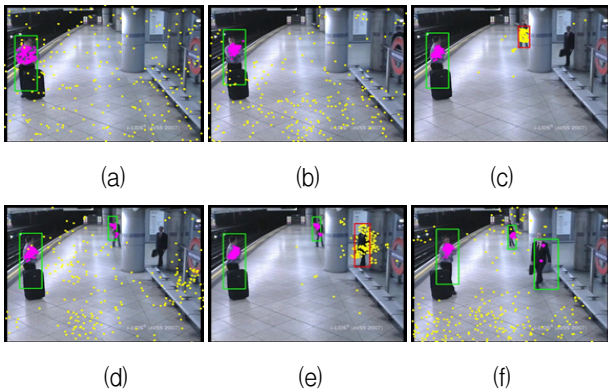


그림 8. AVSS 2007 영상에서의 제안된 방법의 실험결과 (적색: 검출 객체 영역, 녹색: 추적 객체 영역, 노란색: PP, 자색: CP); (a) 1720번째 영상, (b) 1870번째 영상, (c) 1906번째 영상, (d) 1908번째 영상, (e) 1909번째 영상, (f) 1967번째 영상

Fig. 8. Experimental results of the proposed method for AVSS 2007 (red: detected object region, yellow: PP, green: tracked object region, magenta: CP); (a) 1720<sup>th</sup> frame, (b) 1870<sup>th</sup> frame, (c) 1906<sup>th</sup> frame, (d) 1908<sup>th</sup> frame, (e) 1909<sup>th</sup> frame, and (f) 1967<sup>th</sup> frame.

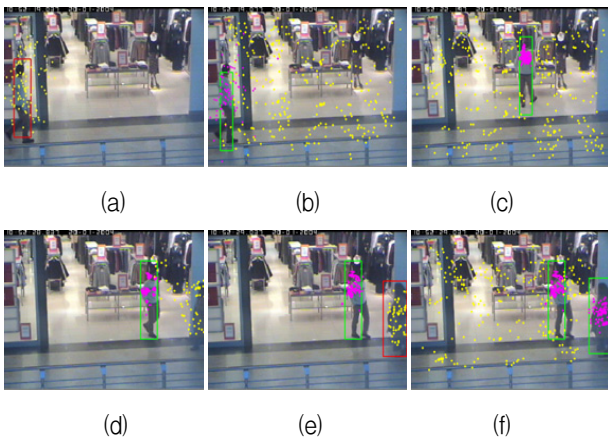


그림 9. CAVIAR 2004의 영상에서의 제안된 방법의 실험결과 (적색: 검출 객체 영역, 녹색: 추적 객체 영역, 노란색: PP, 자색: CP); (a) 471번째 영상, (b) 472번째 영상, (c) 530번째 영상, (d) 612번째 영상, (e) 619번째 영상, (f) 623번째 영상

Fig. 9. Experimental results of the proposed method for CAVIAR-2004 sequence (red: detected object region, yellow: PP, green: tracked object region, magenta: CP); (a) 471<sup>th</sup> frame, (b) 472<sup>th</sup> frame, (c) 530<sup>th</sup> frame, (d) 612<sup>th</sup> frame, (e) 619<sup>th</sup> frame, and (f) 623<sup>th</sup> frame.

그림 8은 720×576해상도를 갖는 AVSS 2007 영상에 제안된 방법을 적용한 결과이다. 그림 8(a), 8(b)는 추적

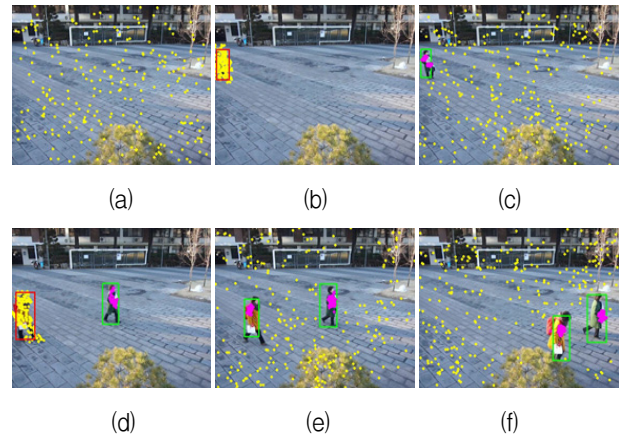


그림 10. 직접 촬영한 영상에서의 제안된 방법의 실험결과 (적색: 검출 객체 영역, 녹색: 추적 객체 영역, 노란색: PP, 자색: CP); (a) 10번째 영상, (b) 34번째 영상, (c) 42번째 영상, (d) 197번째 영상, (e) 215번째 영상, (f) 275번째 영상

Fig. 10. Experimental results of the proposed method for the sequence acquired by ourselves (red: detected object region, yellow: PP, green: tracked object region, magenta: CP); (a) 10<sup>th</sup> frame, (b) 34<sup>th</sup> frame, (c) 42<sup>th</sup> frame, (d) 197<sup>th</sup> frame, (e) 215<sup>th</sup> frame, and (f) 275<sup>th</sup> frame.

객체의 움직임이 작을 때의 시퀀스로 CP는 색상 분포에 의해 리샘플링되므로 움직임 분포에 영향을 받지 않고 독립적으로 수행되어 효과적으로 객체가 추적되고 있는 것을 확인할 수 있다. 그림 8(c)-8(f)는 새로 등장한 객체 주위로 PP가 리샘플링되어 객체를 검출 후 CP에 의해 추적되는 것을 보여준다.

그림 9는 CAVIAR-2004 영상에 적용한 결과로 384×288해상도를 갖는 저해상도 영상에도 강건하게 동작하는 것을 확인할 수 있다. 그림 10은 640×480해상도, 30fps를 갖는 영상으로 SONY의 CMOS 카메라를 사용하여 직접 촬영한 영상에서의 제안된 방법의 실험 결과이다. 그림 10(e), 10(f)과 같이 객체가 그늘진 영역에서 조명이 밝은 영역으로 이동해도 CP들을 통해 강건하게 추적되고 있는 것을 확인할 수 있다.

표 1은 슬라이딩 검색에 의해 IKSVM[11]을 사용하여 객체 검출을 수행하는 방법과 제안된 파티클 샘플링을 적용한 IKSVM을 사용하여 객체 검출을 수행하는 방법의 객체 검출 시간을 비교한 것이다. 각 데이터 셋에 대해 동일하게 200프레임을 사용하였으며, 검출에 사용된 파티클(PP)의 300개로 설정하였다. 실험 영상의 크기에 따라 검출 시간의 감소 비율은 다르지만 기존의

표 1. 제안된 방법의 객체 검출과 기존 방법의 객체 검출 수행 시간비교

Table 1. Comparison of computational time between the proposed and the conventional methods for object detection.

실험 영상	영상 해상도	평균검출시간 [s]		시간감소율 [%]
		IKSVM <sup>[11]</sup>	제안 방법	
PETS-2009	720×480	3.716[s]	2.299[s]	38.13
AVSS-2007	720×576	4.562[s]	2.397[s]	47.46
CAVIAR-2004	384×288	0.832[s]	0.588[s]	29.33
CAU-(SONY)	640×480	3.344[s]	2.034[s]	39.17

검출시간에 비해 30%~40% 정도의 계산 시간이 크게 감소되었음을 보이며, 특히 영상의 크기가 커질수록 효율이 증가함을 표 1을 통해 확인할 수 있다.

#### IV. 결 론

본 논문에서는 감시 시스템 영상에서 이계층구조의 파티클 샘플링을 사용한 다중 객체 추적 방법을 제안하였다. 본 연구의 결과를 정리하면 다음과 같다. 첫째, 파티클 윈도우를 이용한 다중 검출 방법을 제안하였고, 움직이는 객체를 검출한다. 둘째, 객체검출에 의한 추적 접근방식을 이용함으로써 객체영역을 이용한 자동 다중 객체 추적을 수행한다. 셋째, 색상기반의 객체 추적을 사용하여 겹침현상에 대해 강건한 객체 추적을 수행한다.

제안된 방법은 일반 영상 환경에서 기존의 파티클 필터와 같이 사용자에게 의해 지정된 영역에서 추적을 수행하는 것이 아닌, PP로 객체를 검출하고, 검출된 PP로부터 생성된 CP를 통해 독립적인 객체 추적을 하여 다중 객체를 추적 할 수 있었다. 또한 영상 전체에서 슬라이딩 윈도우를 사용하여 객체를 검출하는 것이 아닌 연속된 시퀀스에서 움직임을 통해 리샘플링되는 PP의 위치에서 객체검출을 수행하여 검출을 위한 연산 시간을 효율적으로 줄일 수 있었다. 실험 결과를 통해 일반 비디오 환경에서 강인한 다중 객체의 검출과 추적이 가능하며, 지능형 감시시스템에 적용 가능성을 보인다.

향후에는 동적환경 영상에서의 객체 검출 및 추적에 대한 알고리즘을 개발하고자 한다.

#### REFERENCES

- [1] S. Lee, T. Kim, J. Yoo, and J. Paik, "Abnormal Behavior Detection Based on Adaptive Background Generation for Intelligent Video Analysis", *Journal of The Institute of Electronics Engineers of Korea*, vol. 48SP, no. 1, pp. 111-121, January 2011.
- [2] S. Moon, and S. Shin, "Implementation of Intelligent Image Surveillance System based Context", *Journal of The Institute of Electronics Engineers of Korea*, vol. 47SP, no. 3, pp. 11-22, May 2010.
- [3] J. Kim, D. Yeom, and Y. Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems," *IEEE Trans. Consumer Electronics*, vol. 57, no. 3, pp. 1165-1170, August 2011.
- [4] Y. Chai, S. Shin, K. Chang, and T. Kim, "Real-time user interface using particle filter with integral histogram," *IEEE Trans. Consumer Electronics*, vol. 56, no. 2, pp. 510-515, May 2010.
- [5] Y. Rui and Y. Chen, "Better proposal distributions: object tracking using unscented particle filter," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 786-793, 2001.
- [6] H. Grabner, M. Grabner, and H. Bischof, "Real-time tracking via on-line boosting," *Proc. British Machine Vision Conference*, vol. 1, pp. 47-56, 2006.
- [7] Z. Chen "Bayesian filtering: From Kalman Filters to particle filters, and beyond," *Technical Report McMasters University, Hamilton*, 2003.
- [8] K. Okuma, A. Taleghani, N. Freitas, J. Little, and D. Lowe, "A boosted particle filter: multitarget detection and tracking", *Proc. European Conference on Computer Vision*, vol. 3021, pp. 28-39, 2004.
- [9] G. Gualdi, A. Prati, R. Cucchiara "Multi-stage particle windows for fast and accurate object detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.34, no. 8, pp. 1589-1604, August 2012.
- [10] P. Peerez, C. Hue, J. Vermaak, and M. Gangnet, "Color based probabilistic tracking," *Proc. European Conference on Computer Vision*, pp. 661-675, 2002.



- [11] S. Maji, A. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [12] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. DARPA Image Understanding Workshop*, pp. 121-130, 1981.

— 저 자 소 개 —



정 경 원(학생회원)  
2013년 서울과학기술대학교 전자공학과 학사 졸업.  
2014년 현재 중앙대학교 첨단영상대학원 석사과정 재학중.  
<주관심분야 : 영상처리, 객체 추적, 지능형 영상 보안 및 감시>



김 나 현(학생회원)  
2011년 수원대학교 전자정보통신공학과 학사 졸업.  
2014년 중앙대학교 첨단영상대학원 석사 졸업.  
<주관심분야 : 영상처리, 영상 개선>



이 승 원(학생회원)  
2009년 서울과학기술대학교 전자공학과 학사 졸업.  
2011년 중앙대학교 첨단영상대학원 석사 졸업.  
2014년 현재 중앙대학교 첨단영상대학원 박사과정 재학중.  
<주관심분야 : 영상처리, 객체 추적, 지능형 영상 보안 및 감시 >



백 준 기(평생회원)-교신저자  
1984년 서울대학교 제어계측공학과 학사 졸업.  
1987년 노스웨스턴대학교 전기 및 컴퓨터 공학과 석사 졸업.  
1990년 노스웨스턴대학교 전기 및 컴퓨터 공학과 박사 졸업.  
2014년 현재 중앙대학교 첨단영상대학원 영상공학과 교수.  
<주관심분야 : 영상복원, 신호처리, 반도체>