

## 다수 객체 모델을 이용한 객체 추적 기법

김한울, 김창수  
고려대학교

hanulkim@mcl.korea.ac.kr, changsukim@korea.ac.kr

### Visual Object Tracking Using Multiple Object Models

Han-Ul Kim, Chang-Su Kim  
Korea University

#### 요 약

본 논문에서는 다수 객체 모델을 통해 잘못된 객체 모델 갱신의 영향을 줄이고 객체 추적의 정확도를 향상시키기 위한 방법을 제안한다. 객체 모델 집합은 다수의 객체 모델과 대응하는 가중치들로 정의된다. 각 모델과의 유사도 가중 합을 고려 하여 객체 위치를 추정하고, 각 모델의 추정 신뢰도를 계산하여 가중치를 갱신한다. 실험 결과를 통해 제안하는 기법이 오클루전, 밝기 변화로 인하여 객체 외형이 왜곡되었을 때 추적 성능을 크게 개선함을 보이고 기존 기법들보다 더 정확한 추적 결과를 제공함을 확인한다.

#### 1. 서론

최근 자율 주행 자동차, 무인 감시 시스템과 같은 다양한 응용 분야의 관심이 증가함에 따라, 기초 기술로서 객체 추적 기법에 대한 연구가 컴퓨터 비전 분야에서 활발히 진행 되고 있다[1]. 특히 기계학습 기법을 도입한 다양한 분류기 기반의 객체 추적 기법들[2,3,4]은 목표 객체와 배경을 효과적으로 구별함으로써 객체 추적의 정확성을 크게 향상 시켰다. 하지만 위 기법들은 분류기 학습을 위해 이전 프레임 결과를 이용하는데, 오클루전, 밝기변화 등으로 인하여 객체가 왜곡된 외형을 가질 경우, 객체 모델을 오염시켜 이후 프레임에서 추적 실패를 야기 하는 문제점을 가진다.

따라서 본 논문에서는 오염된 객체 모델에 대한 영향을 줄이고 추적 정확도를 향상시키기 위해 객체 모델 집합을 이용한 추적 기법을 제안한다. 객체 모델 집합은 다수의 객체 모델과 대응하는 가중치로 구성되며 가중치는 모델의 신뢰도에 따라 서로 다른 값을 가진다. 현재 프레임의 객체 위치는 후보 위치들 중 각 객체 모델의 가중 합을 최대화하는 위치로 결정한다. 이후 최종 위치와 개별 모델이 추정된 위치를 비교하여 모델의 가중치를 갱신한다. 실험 결과는 제안 기법이 단일 객체 모델 결과와 비교하였을 때, 오클루전, 밝기변화가 존재하는 상황에서 더 정확한 추적 결과를 제공하며 기존 기법들보다[2,3,4] 더 우수한 성능을 보임을 확인한다.

본 논문의 구성은 다음과 같다. 2 절에서 단일 객체 모델을 이용한 객체 추적 방법에 대해 설명한다. 3 절에서는 제안하는 객체 모델 집합을 통한 추적 기법을 설명하고 4 절에서 실험을 통해 제안 기법의 객체 추적 성능을 평가한다. 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

#### 2. 단일 객체 모델을 이용한 객체 추적

분류기 기반의 객체 추적 기법은 아래와 같이 분류기를 통해 학습한 객체 모델  $\mathbf{w}$  과 유사도를 최대로 만드는 위치를 현재 프레임  $\mathbf{x}_t$  에서의 객체 위치  $\mathbf{y}_t$  로 결정한다.

$$\mathbf{y}_t = \arg \max_{\mathbf{y}} \sum_{i=1}^N \mathbf{w}_i^T \phi(\mathbf{x}_t, \mathbf{y}) \quad (1)$$

위 식에서  $\mathbf{y}$  는 객체의 후보 위치이며  $\phi(\mathbf{x}_t, \mathbf{y})$  는 대응하는 특징벡터이다.

동영상 내 객체의 외형은 변화해 나가기 때문에 강인한 추적을 위해서는 이전 프레임 추적결과를 바탕으로 객체 모델을 지속적으로 갱신 할 필요가 있다. 따라서 본 논문에서는 객체 모델 갱신을 위해 학습데이터를 객체와 주변 영역 그리고 각 영역을 기술하는 특징벡터로 구성하고 이를 이용하여 객체 모델을 갱신한다.

객체와 배경을 명확하게 구별 해줄 수 있는 모델을 얻기 위해서는 식별력 있는 특징 벡터 사용이 필수적이다. 효과적인 영역 기술을 위해 본 논문에서는 영역의 색상, 경계, 그리고 지역적 특성을 이용한다. 구체적으로 영역의 지역적 특성을 반영하기 위해 영역을 64 개의 겹치지 않는 블록으로 분할하고 각 블록을 24 차원의 색상 히스토그램과, 8 차원의 경계 히스토그램으로 기술한다. 최종적으로 영역은 각 블록의 특징 벡터를 연결하여 얻은 2048 차원 특징 벡터로 표현한다.

현재 프레임의 객체 외형을 반영하는 새로운 객체 모델은 앞서 구성한 학습데이터에 대해 분류기를 학습시켜 얻는다. 본 논문에서는 객체 검출 및 추적 분야에서 자주 사용되는 구조화 서포트 벡터 머신(structured support vector machine) [2]을 분류기로 사용하였다.

#### 3. 다수 객체 모델을 이용한 객체 추적



(a)



(b)

그림 1. 왜곡된 객체 외형 예시 (a)오클루전, (b) 밝기 변화

객체 모델을 매 프레임 갱신하는 방법은 외형 변화에 효과적인 대응을 가능하게 하지만, 그림 1 과 같이 오클루전, 밝기 변화로 인하여 일시적으로 왜곡된 외형을 보일 때, 객체 모델을 오염시켜 추적 실패를 야기한다.

본 논문에서는 오염된 객체 모델의 영향을 줄이기 위해 다수 객체 모델과 대응하는 가중치들로 구성된 객체 모델 집합  $\mathcal{M} = \{w_1, \dots, w_N, \alpha_1, \dots, \alpha_N\}$  을 제안한다. 제안 기법에서 모델의 수는 실험적으로 결정하였으며 5 개의 모델을 고려한다. 객체 모델 집합  $\mathcal{M}$  이 주워졌을 때, 객체 위치는 각 객체 모델과 유사도의 가중 합을 최대화 하는 위치로 결정한다.

$$y_t = \arg \max_y \sum_{i=1}^N \alpha_i w_i^T \phi(x_t, y) \quad (2)$$

오염된 객체 모델의 영향을 줄이기 위해서는 오염된 모델의 가중치를 낮게 만들 필요가 있다. 위 목적을 위해 제안 기법은 오염된 모델은 정상적인 모델과 달리 잘못된 추정 결과를 내놓는 경우가 많다는 관찰 결과를 활용한다. 다시 말해 집합  $\mathcal{M}$  을 고려하여 찾은 위치  $y_t$  와 개별 모델  $w_i$  만을 고려하여 추정한 객체 위치  $y_{t,i}$  의 차이가 클 수록 모델  $w_i$  의 추정은 신뢰 할 수 없다고 판단하여 낮은 가중치를 부여한다. 구체적으로 가중치는 아래와 같이 정의 된다.

$$\alpha_i = \exp(-\gamma \|y_t - y_{t,i}\|^2) \quad (3)$$

여기에서  $\gamma$  는 0.5 를 사용한다.

가중치 갱신이 끝나고 가장 낮은 가중치를 가지는 모델을 집합에서 제외한다. 또한 변화하는 객체의 외형에 적응하기 위해 제안 기법은 새로 학습된 객체 모델을 집합에 추가한다. 새로운 모델의 가중치는 기존 모델의 평균 가중치를 할당한다.

### 3. 실험 결과

본 논문에서는 제안 기법을 정량적으로 평가하기 위하여 10 개의 실험 영상에 대해, 객체 추적 분야에서 자주 사용되는 평가 지표인 정밀도(precision)과 성공률(success rate)를 측정한다. 정밀도는 추정한 객체 위치와 실제 위치의 차가 20 픽셀 이내인 프레임의 비율을 의미하며 성공률은 추정한 영역과 실제 영역간의 교집합 넓이가 합집합 넓이의 절반 이상인 프레임의 비율을 나타낸다.

표 1. 객체 모델에 따른 비교 (정밀도/성공률)

	Single	Average	Proposed
--	--------	---------	----------

Bolt	1.00 / 0.87	0.99 / 0.99	1.00 / 0.98
Car4	0.72 / 0.37	0.63 / 0.37	0.78 / 0.37
David3	1.00 / 0.93	1.00 / 0.95	1.00 / 1.00
Football1	1.00 / 0.89	1.00 / 0.91	1.00 / 0.91
Girl	1.00 / 0.79	1.00 / 0.84	1.00 / 0.84
Lemming	0.27 / 0.27	0.94 / 0.89	0.94 / 0.89
Liquor	0.95 / 0.97	0.82 / 0.84	0.96 / 0.98
Matrix	0.06 / 0.06	0.11 / 0.08	0.48 / 0.30
Suv	0.87 / 0.87	0.86 / 0.86	0.90 / 0.90
Tiger1	0.61 / 0.77	0.69 / 0.85	0.68 / 0.82
Average	0.75 / 0.68	0.81 / 0.76	0.87 / 0.80

표 2. 기존 기법들 [2,3,4] 과의 비교 (정밀도/성공률)

	Struck[2]	TGPR[3]	KCF[4]	Proposed
Average	0.52/0.42	0.59/0.47	0.82/0.74	0.87/0.80

먼저 객체 모델 구성을 위해 단일 모델(Single), 평균 모델(Average), 제안 기법(Proposed)의 추적 성능을 표 1 을 통해 비교한다. 단일 모델은 한 개의 객체 모델만을 사용한 결과이며 평균 모델은 모든 가중치를 같은 값으로 고정하였을 때 결과이다. 실험 결과 제안 기법으로 인하여, 단일 모델 대비 16%, 평균 모델 대비 8.6%의 정밀도 향상이 있으며, 성공률 측면에서도 가장 높은 성능을 달성 하였다. 특히 “Lemming”, “Matrix”, “Suv” 와 같이 오클루전, 밝기 변화로 인해 왜곡된 객체 외형이 나타나는 영상에서 추적 성능을 크게 향상 시킨 것을 확인할 수 있다.

표 2 는 최근 객체 추적 벤치마크[1]에서 인상적인 성능을 보인 기존 기법[2,3,4]들과 제안 기법을 비교한다. 비교 결과 제안 기법이 기존 기법들보다 더 정확한 추적결과를 제공함을 확인할 수 있다.

### 5. 결론

본 논문에서는 왜곡된 객체 외형의 영향을 줄이기 위해 다수의 객체 모델과 가중치로 구성된 객체 모델 집합을 제안하였다. 객체 위치를 추정하기 위해 각 모델의 가중 합을 이용하였으며 추정 신뢰도를 통해 가중치를 갱신하였다. 실험결과를 제안하는 기법을 통해 오클루전, 밝기 변화가 존재하는 영상에서 추적 정확도를 크게 개선함을 보였고 기존 기법과 비교를 통해 제안 기법의 성능이 우수함을 확인하였다.

### 6. 참고문헌

[1] Y. Wu, J. Lim, and M.-H. Yang, “Online object tracking: A benchmark,” In *Proc. CVPR*, 2013.

[2] S. Hare, A. Saffari, and P. Torr, “Struck: structured output tracking with kernels,” In *Proc. ICCV*, 2011.

[3] J. Gao, H. Ling, W. Hu, and J. Xing, “Transfer learning based visual tracking with gaussian processes regression,” In *Proc. ECCV*, 2014.

[4] J. Henriques, R. Caseiro, P. Martins, and J. Batista, “High speed tracking with kernelized correlation filters,” *TPAMI*, 37(3):583–596, Mar. 2015.