

다중 랜덤 워커를 이용한 객체 추적 기법

문주혁, 김한울, 김창수
고려대학교

{jhmun, hanulkim, cskim}@mcl.korea.ac.kr

Visual Object Tracking Using Multiple Random Walkers

Juhyeok Mun, Han-Ul Kim, Chang-Su Kim
Korea University

요 약

본 논문에서는 다중 랜덤 워커(multiple random walkers)에 기반한 객체 추적 기법을 제안한다. 우선 서포트 벡터 머신(support vector machine)을 이용한 분류기 기반 객체 추적 기법을 소개한다. 다음으로 영상의 영역에 대한 특징 벡터 중 배경으로부터 추출된 특징 벡터를 억제하는 기법을 제안한다. 영역에서 배경 요소를 찾기 위해 다중 랜덤 워커를 이용한 전경 및 배경 추출 방법을 제시한다. 배경 요소를 억제하여 학습된 서포트 벡터 머신은 객체와 배경이 유사한 영상, 객체가 다른 물체에 의해 가려지는 영상 등에서 객체와 배경을 확실하게 구분하여 객체를 잃지 않고 추적할 수 있다. 마지막으로 실험을 통해 제안하는 기법이 기존 기법에 비해 우수한 추적 성능을 보임을 확인한다.

1. 서론

객체 추적은 동영상에서 객체의 초기 위치가 주어졌을 때 이후 위치들을 자동으로 추정하는 기술로 최근 연구가 활발하게 진행되고 있다 [2,3,4]. 특히 기계학습을 도입한 분류기 기반의 추적 기법들은 객체와 배경을 효과적으로 판별할 수 있는 객체 모델을 학습하여 객체 추적 알고리즘의 성능을 크게 향상시켰다.

분류기는 영상의 영역으로부터 추출된 특징 벡터를 학습한다. 학습 샘플의 후보 영역은 항상 박스 형태인데 반해 객체의 모양은 다양하므로 분류기 학습에 제공되는 샘플의 특징 벡터에는 배경 정보가 포함될 수 있다. 배경 정보로 학습된 분류기는 객체와 배경을 잘 구분하지 못하여 분류기 기반 추적 알고리즘의 성능은 현저히 떨어진다.

본 논문에서는 이러한 문제점을 해결하기 위해 다중 랜덤 워커(multiple random walkers) 기법[5]을 이용한다. 객체 박스를 그래프로 표현하기 위해 다수의 부분 영역으로 나누고, 각 부분 영역을 노드로 하는 다중 랜덤 워커 시뮬레이션을 수행한다. 이로부터 전경과 배경에 해당하는 요소를 추출하고 객체가 존재할 확률에 대한 가중치를 계산한다. 이 가중치를 특징 벡터에 곱해주어 분류기 학습에서 배경 요소가 미치는 영향을 최소화할 수 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 제안하는 기법에 대해 설명한다. 3 장에서는 실험을 통해 제안 기법이 우수함을

보이고, 마지막으로 4 장에서는 본 논문의 결론을 맺는다.

2. 제안하는 기법

2.1 분류기 기반 객체 추적

본 논문에서는 [4]와 같이 영역의 색상, 경계와 지역적 특성을 객체 추적에 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 2,048 차원 특징 벡터로 표현한다. 2,048 차원의 특징 벡터는 서포트 벡터 머신을 학습시키고, 학습된 서포트 벡터 머신의 분류기 모델은 다음 프레임에서 후보 영역 중에서 객체 영역을 찾아낸다. 분류기 모델은 후보 영역의 특징 벡터와 이전 프레임에서 학습된 특징 벡터들 사이의 유사도를 점수로 나타내어 가장 높은 점수를 얻는 후보 영역이 객체 영역으로 선정된다.

분류기 기반 객체 추적은 다음과 같은 문제가 있다. 객체 영역 안의 모든 블록이 객체만을 포함하지는 않기 때문에 2,048 차원의 특징 벡터는 배경으로부터 얻어진 특징 벡터가 많이 포함하고 있다. 분류기는 배경으로부터 객체를 분류해내어 추적을 수행하기 때문에 배경으로부터 얻어진 특징 벡터는 분류기를 오염시켜 추적 알고리즘의 성능을 현저히 떨어뜨린다.



그림 1. 부분 영역별 가중치 예시: 적색에 가까울수록 객체를 포함할 확률이 높다.

따라서 분류기가 높은 분류 성능을 보이기 위해서는 학습 단계에서 배경으로부터 얻어진 특징 벡터를 억제할 필요가 있다.

2.2 다중 랜덤 워커

본 알고리즘에서는 분류기를 오염시키는 배경의 특징 벡터를 억제하는 방법을 제안한다. 객체 영역 안 64 개의 블록 중 배경을 포함하고 있는 영역을 추출하기 위해 64 개의 블록을 노드로 하는 그래프를 구성하고 랜덤 워커 시뮬레이션을 수행한다. 영상의 전경과 배경에 해당하는 두 워커에 대해 랜덤 워커 기법을 동시 적용하여 확장한 다중 랜덤 워커 기법[5]을 이용한다. 전경과 배경 워커에 식 (2)을 적용하면 다음과 같다.

$$\mathbf{p}_f^{(t+1)} = (1 - \epsilon)\mathbf{A}\mathbf{p}_f^{(t)} + \epsilon\mathbf{r}_f^{(t)}, \quad (3)$$

$$\mathbf{p}_b^{(t+1)} = (1 - \epsilon)\mathbf{A}\mathbf{p}_b^{(t)} + \epsilon\mathbf{r}_b^{(t)}. \quad (4)$$

다중 랜덤 워커는 전경 확률과 배경 확률이 \mathbf{r} 에 영향을 줄 수 있다. 즉, 전경 확률이 높은 노드는 배경 확률이 낮게, 배경 확률이 높은 노드는 전경 확률이 낮게 나오도록 워커가 상호작용한다. 따라서 \mathbf{r} 은 매 반복 시행 t 마다 $\mathbf{p}_f^{(t)}$ 와 $\mathbf{p}_b^{(t)}$ 에 영향을 받아 새로운 $\mathbf{r}^{(t)}$ 로 갱신되어야 한다. $\mathbf{r}^{(t)}$ 는 사후 확률(posterior probability)를 이용하여 다음과 같이 구한다.

$$\mathbf{r}_k^{(t)} = (1 - \delta^t)\mathbf{r}_k^{(t-1)} + \delta^t\phi_k(\mathbf{P}^{(t)}), \quad \mathbf{P}^{(t)} = \{\mathbf{p}_k^{(t)}\}_{k=1}^K \quad (5)$$

$$\mathbf{p}_k^{(t)} = [p^{(t)}(x_1|\omega_k), \dots, p^{(t)}(x_N|\omega_k)]^T, \quad (6)$$

$$p^{(t)}(\omega_k|x_i) = \frac{p^{(t)}(x_i|\omega_k)p^{(t)}(\omega_k)}{\sum_l p^{(t)}(x_i|\omega_l)p^{(t)}(\omega_l)}, \quad (7)$$

$$\phi_{k,i}^{(t)} = \alpha \cdot p^{(t)}(\omega_k|x_i) \cdot p^{(t)}(x_i|\omega_k), \quad (8)$$

$$\phi_k^{(t)}(\mathbf{P}^{(t)}) = \alpha\mathbf{Q}_k^{(t)}\mathbf{p}_k^{(t)}. \quad (9)$$

$p^{(t)}(x_i|\omega_k)$ 는 워커 k 가 노드 i 에 분포할 확률을, $p^{(t)}(\omega_k|x_i)$ 는 노드 i 가 워커 k 에 의해 차지될 확률을 나타낸다. $\mathbf{Q}_k^{(t)}$ 는 각 대각 요소를 $\phi_{k,i}^{(t)}$ 인 대각 행렬이다. α 는 $\phi_k^{(t)}$ 를 확률 분포로 만들어주기 위한 정규화 상수이다. 식 (3)과 식(4)을 일정 T 만큼 시행하여 얻어진 전경 확률과 배경 확률을 얻는다. 전경과 배경확률을 이용하여 각 블록에서 객체가 존재할 확률은 식 (10)과 같이 구한다.

$$\Psi^{(t)} = \frac{1}{1 + e^{c(\mathbf{p}_f^{(t)} - \mathbf{p}_b^{(t)})}} \quad (10)$$

c 는 실험적으로 정한 상수이다. 블록에서 객체가 존재할 확률은 특징 벡터에 가중치로 곱해져 배경의 영향을 억제한다. 그림 1 은 식 (10)으로 얻어진 가중치를 도시한다.

3. 실험 결과

본 논문에서는 제안 기법을 정량적으로 평가하기 위하여 TB-100 벤치마크 영상[1]에 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. 기존 기법들[2,3,4]과의 비교 (정밀도/성공률)

	Struck[2]	KCF[3]	SOWP[4]	Proposed
Precision	0.635	0.692	0.825	0.836
Success	0.459	0.475	0.569	0.571

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

4. 결론

본 논문에서는 분류기 기반의 객체 추적 기법에서 배경이 분류기에 미치는 영향을 억제하기 위한 가중치 모델을 제안하였다. 억제할 배경 요소를 추출하기 위해 객체 박스를 다수의 블록으로 나누어 다중 랜덤 워커 시뮬레이션을 수행하였다. 각 블록에 객체가 존재할 확률을 해당 블록의 특징 벡터에 가중치로 곱하여 배경의 영향을 억제하였다. 실험결과 제안하는 기법을 통해 객체 추적 벤치마크 영상에서 추적 정확도를 크게 개선함을 보였고 기존 기법과 비교를 통해 제안 기법의 성능이 우수함을 확인하였다.

5. 참고 문헌

- [1] Y. Wu, J. Lim, and M.-H. Yang, "Object Tracking Benchmark," *TPAMI*, Sep. 2015.
- [2] S. Hare, A. Saffari, and P. Torr, "Struck: structured output tracking with kernels," In *Proc. ICCV*, 2011.
- [3] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Highspeed tracking with kernelized correlation filters," *TPAMI*, Mar. 2015.
- [4] H.U. Kim, D.Y. Lee, J.Y. Sim, and C.S. Kim, "SOWP: Spatially Ordered and Weighted Patch Descriptor for Visual Tracking," In *Proc. ICCV*, 2015.
- [5] C.W. Lee, W.D. Jang, J.Y. Sim, and C.S. Kim, "Multiple Random Walkers and Their Application to Image Cosegmentation," In *Proc. CVPR*, 2015.