

## 객체성 특징을 활용한 Correlation Filter 기반 열화상 객체 추적 기법

정현석 \*김원준

건국대학교

hsjeong5@konkuk.ac.kr \*wonjkim@konkuk.ac.kr

## Correlation Filter Based Thermal Object Tracking Using Objectness Feature

Jung, Hyeonseok \*Kim, Wonjun

Konkuk University

## 요약

객체 추적 기술은 컴퓨터 비전 분야에서 활발하게 연구되고 있는 분야로, 인간-컴퓨터 상호작용(HCI)이나 로보틱스, 그리고 자율주행 자동차와 같은 곳에 유용하게 사용될 수 있다. 그중에서도 열화상 객체 추적 기술은 빛이 전혀 없는 어두운 밤에도 적용 가능하기 때문에, 감시 시스템이나 사고 현장의 구조 상황, 또는 군사적 응용에 더욱 유용하다. 그러나 일반적으로 열화상 영상은 컬러 영상보다 해상도가 낮고, 객체의 경계가 흐릿하며, 텍스처가 거의 없다는 점 때문에 기존 추적 방법들의 성능 저하의 원인이 된다. 이에 본 논문은 기존 Correlation Filter 기반 추적기에 객체성 특징을 함께 이용하여, 열화상 영상에서의 객체 추적 성능을 향상 시키는 방법을 제안하였다. 또한 실험을 통해 기존 방법보다 열화상 영상에서의 객체 추적 성능이 향상된 것을 보였다.

## 1. 서론

객체 추적 기술은 동영상의 첫 프레임에서 객체의 위치가 주어졌을 때, 이어지는 프레임들에서 계속해서 객체의 위치를 예측하는 기술을 말한다. 이 기술은 지능형 감시 시스템, 인간-컴퓨터 상호작용(HCI), 자율주행 자동차와 같은 곳에 유용하게 사용될 수 있기 때문에 전통적으로 컴퓨터 비전 분야에서 활발하게 연구되고 있는 분야이다. 그중에서도 열화상에서의 객체 추적 기술은 빛이 전혀 없는 어두운 밤에도 적용 가능한 기술이기 때문에 야간 상황에서의 사용이 많이 요구되는 감시 시스템이나 사고 현장의 구조, 또는 군사적 목적에 더욱 유용하다. 열화상 영상의 이러한 장점에도 불구하고 일반적으로 열화상 영상은 컬러 영상보다 해상도가 낮고, 객체의 경계가 흐릿하며, 텍스처가 거의 없기 때문에 기존 객체 추적 방법들을 열화상 영상에 그대로 적용하였을 때에는 객체 추적 성능 저하의 원인이 된다.

이에 본 논문에서는 객체성 특징을 열화상 객체에서 추출하여 여러 가지 기존 추적 방법들 중 Correlation Filter 기반 추적기의 성능을 향상 시키는 방법을 제안하였다. 객체 추적 방법으로는 여러 가지 Correlation Filter 방식의 객체 추적 방법 중, [2][3]을 활용하였고, 객체성 특징으로는 영상의 그라디언트를 이용하는 방법인 [4]를 활용하였다.

## 2. 제안하는 방법

## 2.1 Correlation Filter 추적기

기존의 여러 추적 방법들 중, Correlation Filter 기반 추적기는

MOSSE[1]의 논문이 발표된 이후로 활발히 연구되고 있는 방법 중 하나이다. 그 중 CSK[2]는 추적 객체 주변의 학습 샘플들을 순환적으로 근사하여(Circularly approximated samples) 학습 데이터 행렬을 순환적으로 구성하였고, 이 학습 데이터 행렬에 커널 릿지 회귀(Kernelized ridge regression)을 적용하였다. 그리고 해당 논문에서는 순환 행렬을 Fourier 변환 행렬을 이용해 분해할 수 있다는 사실을 이용하여, Fourier 도메인에서 추적 객체를 매우 빠르게 학습하고 추적할 수 있다는 사실을 보였다. 또한 CSK[2]에서는 추적 객체의 회색조 픽셀 값을 단일 채널 특징맵으로 사용하였지만, KCF[3]에서는 CSK[2]의 방법을 확장하여 FHOG[7]와 같은 다중 채널 특징맵에서도 사용가능하도록 하였다. 본 논문에서는 KCF[3]의 추적 방법을 이용하였다.

## 2.2. 객체성 특징

객체성(Objectness)은 영상 내의 특정 윈도우가 객체를 포함하고 있는지를 나타내는 지표로 활용된다. 본 논문에서는 [4]에서 제안된 BING(Binarized normed gradient) 특징을 활용하였다. 이 특징은 일반적인 객체가 잘 정의된 경계(Well-defined boundary)를 가지고 있다는 관찰로부터 제안되었다. 따라서 이러한 관찰을 이용하기 위해 영상의 그라디언트를 구하고, 그 그라디언트의 크기를 구한 후 영상 내 관심 영역의 크기를  $8 \times 8$  로 조정하여 이것을 객체성을 나타낼 수 있는 64차원 특징으로 활용하였다. BING[4]에서는 연산의 효율성을 위하여 해당 특징의 이진화된 버전까지 제안하였지만, 본 논문에서는 그라디언트의 크기를 구하고  $8 \times 8$  로 크기를 조정하는 부분까지만을 활용하였다.

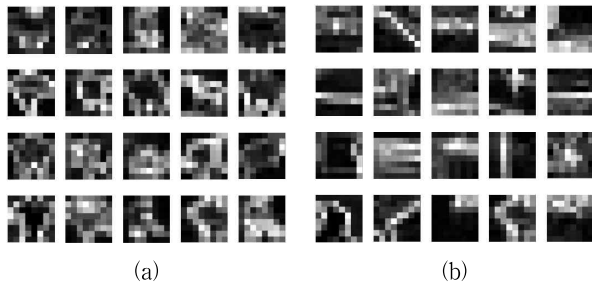


그림 1. 열화상 객체에 대한 객체성 특징 시각화. (a) 객체들에 대한 객체성 특징. (b) 객체가 아닌 것들에 대한 객체성 특징

LTIR v1.0 열화상 영상 데이터셋[5]에 있는 객체와 객체가 아닌 것들에 대해 해당 특징을 구한 후 시각화해보았고(그림 1.), 열화상 영상에서도 객체와 객체가 아닌 것을 구분하는 데에 이 특징이 잘 적용될 수 있다고 보았다.

### 2.3. 객체성 특징을 활용한 Correlation Filter 추적기

본 논문에서는 2.2.에서 설명한 객체성 특징을 열화상 객체에서 추출하여 기존 Correlation Filter 추적 방법[3]에 함께 사용하였다.

Correlation Filter가 더욱 일반적인 객체의 형태를 추적할 수 있도록 하기 위해, 열화상 영상에서 객체와 객체가 아닌 것들에 대해 객체성 특징을 추출하여 선형 분류기중 하나인 선형 서포트 벡터 머신 (Linear support vector machine)을 이용해 이진 분류를 학습하였다. 선형 분류기 학습 시에는, LTIR v1.0 열화상 영상 데이터셋[5]에서 추출한 20,000개의 샘플을 이용하였다. 이 과정에서 학습된 선형 가중치를 일반적인 객체의 형태로 활용하여 Correlation Filter 학습 시, FHOG[7] 특징맵의 채널에 결합하여 학습을 진행하였다. 객체를 추적하는 단계에서는 학습된 선형 가중치를 이용하지 않고, 객체에서 직접 객체성 특징을 추출하여 FHOG[7] 특징맵에 결합하여 추적을 하였다.

이와 같은 방법을 통해 Correlation Filter가 매 프레임에서 객체를 찾을 때, 더욱 객체일 확률이 높은 곳을 찾을 수 있게 된다.

### 3. 실험 결과

본 논문에서는 LTIR v1.0 열화상 영상 데이터셋[5]을 사용하여 실험을 하였고, FHOG[7] 특징과 객체성 특징을 결합할 때, 객체성 특징에 0.1을 곱하여 객체성 FHOG[7] 특징과 객체성 특징의 비율을 조절하였다. 성능 측정에는 VOT2014[6]에서 제안된 Accuracy와 Robustness(A-R plot)를 사용하였다. 그리고 측정된 성능을 기존 Correlation Filter 방법 중 대표적인 방법인 CSK[2], KCF[3]와 비교하였다. 표 1. 과 그림 2. 는 실험 결과를 그래프와 표로 나타낸 것이고, 제안한 방법이 기존 방법 대비 나은 성능을 보임을 확인하였다.

표 1. LTIR v1.0 열화상 영상 데이터셋에 대한 A-R 성능 비교

	Accuracy	Robustness
CSK[2]	0.47	33.50
KCF[3]	0.55	37.33
Proposed	<b>0.55</b>	<b>27.67</b>

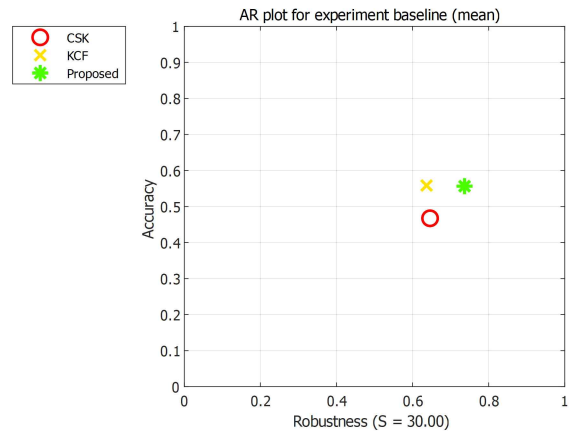


그림 2. 열화상 영상에 대한 제안하는 방법의 A-R plot

### 4. 결론

본 논문에서는 기존 Correlation Filter 기반의 추적기에 객체성 특징을 활용하여 열화상 영상에서의 객체 추적 성능을 향상 시키는 방법을 제안하였으며, 실험을 통해 실제로 성능이 향상되는 것을 확인하였다.

### 감사의 글

이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2017R1C1B2003044).

### 참고문헌

[1] D. S. Bolme, J. R. Beveridge, B. A. Drapper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters," in *Proc. Comput. Vis. Pattern Recognit.*, pp. 2544-2550, 2010.

[2] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in *Proc. Eur. Conf. Comput. Vis.*, vol. 7575, pp. 702-715, 2012.

[3] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583-596, Mar. 2015.

[4] M.-M. Cheng, Z. Zhang, W.-Y. Lin, and P. Torr, "BING: Binarized normed gradients for objectness estimation at 300fps," in *Proc. Comput. Vis. Pattern Recognit.*, pp. 3286-3293, 2014.

[5] A. Berg, J. Ahlberg, and M. Felsberg, "A Thermal Object Tracking Benchmark," in *Proc. Advanced Video and Signal Based Surveillance*, pp. 1-6, 2015.

[6] M. Kristan, R. Pflugfelder, A. Leonardis, J. Matas, L. Cehovin, G. Nebehay, T. Vojir, F. G. and et al., "The visual object tracking vot2014 challenge results," in *Proc. Eur. Conf. Comput. Vis. Workshop on Visual Object Tracking Challenge*, pp. 191-217, 2014

- [7] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627-1645, Sep. 2010.