

# 컨벌루션 특징맵과 코릴레이션 필터를 이용한 물체 추적에 관한 연구

임수창\*, 김도연\*\*

\*순천대학교 컴퓨터학과

\*\*순천대학교 컴퓨터공학과

e-mail:suchangLim@sunshon.ac.kr

## A Study on Object Tracking using Convolution Feature Map and Correlation Filter

Su-Chang Lim\*, Do-Yeon Kim\*\*

\*Dept of Computer Science, Sun-Chon National University

\*\*Dept of Computer Engineering, Sun-Chon National University

### 요 약

컴퓨터비전의 한 분야인 추적은 다양한 방법론들에 근거하여 활발히 연구되어온 분야이다. 추적알고리즘은 연속되는 영상 시퀀스의 객체를 지속적으로 추적하는 방법으로, 객체의 외형 변형, 이동, 회전, 폐색등 복잡한 환경에서도 강건히 추적하는 것에 초점이 맞춰져 있다. 본 논문에서는 딥러닝의 한 부류인 CNN의 컨벌루션 레이어에서 출력되는 특징맵과 변화되는 객체에 적응적으로 대응하는 코릴레이션 필터를 결합하여 복잡한 환경에서도 객체를 추적하는 방법을 제안한다.

### 1. 서론

컴퓨터비전에서 한 분야인 추적(tracking)은 최근까지도 활발히 연구 되어온 분야이다[1]. 추적알고리즘은 영상의 처음 프레임에 존재하는 객체의 정보를 기준으로, 이후 연속되는 프레임에 존재하는 객체를 추적하는 것을 말한다. 이러한 추적 알고리즘은 영상에 존재하는 목표 객체의 선형, 비선형적인 이동, 변형, 회전, 밝기의 변화, 복잡한 배경과 같은 복잡한 환경에서 객체를 강건히 추적하기 위해 다양한 방법론에 기반 하여 제작되고 있다. 객체의 밀집된 특징분포 집합을 찾는 것에 기반 한 mena-shift 추적기와, 베이지안 확률 분포 기반인 파티클 필터를 이용한 추적기, 주파수 도메인에서 FFT(Fast Furier transform)를 사용하여 연산성능에서 빠르고, 효과적인 성능을 가진 적응적 Correlation Filter를 사용한 추적기[2], 최근 새롭게 떠오르는 Deep Neural Network의 한 부류인 CNN(Convolution Neural Network)의 Convolution Layer에서 추출되는 추적 객체의 feature map을 이용한 추적기[3]가 연구되고 있다. 본 논문에서는 이런 추적 방법들 중, 코릴레이션 필터와 CNN을 결합하여 객체를 추적하기 위한 알고리즘을 제안한다.

코릴레이션 필터는 오래전부터 연구되어 왔으며, 필터의 연산방법이 가진 특징으로 신호처리, 매칭등 다양한 분야에서 사용되고 있다. 특히 영상처리분야에서 이 필터는 주파수 공간에서 연산이 이루어지므로 기준 되는 객체의 회전, 변형, 폐색등에 좋은 성능을 보여준다.

다양한 이미지의 고수준의 고유 특징들로 이루어진

CNN은 최근 이미지 분류, 객체 검출, 지역화 문제에서 뛰어난 성능을 보여주고 있다. 기존의 특징들을 추출하는 알고리즘들과 달리 CNN은 각각의 레이어를 거치며 이미지가 가진 높은 수준의 정보들로 구성된 특징맵을 추출하고, 이러한 특징들을 사용하여 여러 부류의 객체를 분류한다.

본 논문에서는 사전에 학습된 학습모델을 이용해 특징맵을 여러 계층의 Convolution Layer에서 추출하고, 생성된 특징맵을 기반으로 Correlation Filter와 결합하여 지속적으로 객체를 추적하는 알고리즘을 제안한다.

### 2. 제안하는 추적 알고리즘

#### 2.1 특징맵 추출을 위한 학습 네트워크 제작

CNN기반의 기존 공개된 추적기에서 사용되는 네트워크 모델은 대량의 이미지를 분류하기 위해 심층 레이어로 구성되어 있으며, 각 레이어에서 출력되는 특징맵이 거대하기 때문에 연산 속도가 느리다. 하지만 CNN의 특징으로서 레이어가 깊어지면 높은 수준의 특징들을 추출할 수 있다. 따라서 본 논문에서는 속도에서 효과적인 성능을 얻기 위해 분류의 목적이 아닌 추적되는 객체의 효과적인 특징맵을 추출하기 위한 학습 네트워크를 제작한다. 각 레이어에서 출력되는 결과물을 시각화하여 올바르게 학습이 진행되고 있는지, 충분한 특징들을 추출하는지 판단한다.

#### 2.2 Convolution Layer 특징맵

본 논문에서는 추적 객체의 외형을 묘사하기 위해 기존의 HOG, SIFT와 같은 hand-craft 특징 추출기를 사용하

지 않고, CNN의 컨볼루션 레이어를 통해 획득한 특징맵을 사용한다. 각 레이어의 컨볼루션 연산을 통해 획득한 특징맵은 입력된 이미지의 고 수준의 특징들로 이루어져 있으며, 각 레이어의 출력결과에 따라 다양한 형태의 특징맵을 가지고 있다. 레이어가 깊어지면 점진적으로 출력 맵의 공간 해상도는 적어지며 마지막 fully-connected layer에서는 1x1의 해상도로 출력이 된다. 하지만 이런 저 해상도에서는 추적 객체의 정확한 위치를 판단하기 위한 정보가 충분하지 않기 때문에 이용하지 않고, 상위의 컨볼루션 레이어만 이용 한다.

### 2.3 추적을 위한 코릴레이션 필터

추적에 사용되는 필터는 CNN을 통해 획득한 특징맵을 기반으로 제작된다. 사이즈가 고정된 필터를 사용하지 않고, 출력되는 특징맵의 사이즈(width×height)와 유사하게 선언을 하고, 사용되는 필터는 추적 객체의 변화에 강건히 대응하기 위해 단일 필터가 아닌 복수 개의 필터를 이용한다. 최초 필터의 형태는 처음 프레임에서 객체의 영역을 지정하여 입력 객체로서 사용되며, 영상 시퀀스가 진행 될 때마다 객체의 추적과 필터의 계수는 적응적으로 갱신된다. 추적 객체는 이전 프레임의 특징맵을 기반으로 갱신된 필터와 코릴레이션 연산을 하고, 연산을 통해 획득한 최대 값에 상응하는 위치를 추적객체의 새로운 위치로 판단한다. 이후 새로 위치 추정된 객체를 기준으로 필터는 갱신된다. 이런 방법은 연산량이 많아 연산 속도에 많은 영향을 주므로, 코릴레이션 연산은 주파수 도메인에서 FFT(Fast Fourier transform) 형태로 계산되어 성능을 높여준다.

### 2.4 코릴레이션 필터 갱신

필터를 갱신하기 위한 방법은 식 (1)과 같이 Vishnu . et. al [4]이 제안한 최소평균제곱오차 방법을 기반으로 갱신을 한다. 파라미터  $n$ 은 컨볼루션 레이어에서 출력되는  $n$ 개의 특징맵이며,  $x$ 는 출력되는 각  $i$ 번째 특징맵,  $y$ 는  $i$ 번째의 입력 이미지의 값,  $\lambda$ 는 정규화 계수,  $W$ 는 최소 값을 가지는 필터이며,  $w$ 는 각 특징별 필터이다.

$$W = \underset{w}{\operatorname{argmin}} \sum_i^n \| w \cdot x_i - y_i \|^2 + \lambda \| w \|^2 \quad (1)$$

하지만 본 논문에서 필터 갱신을 위해 사용되는 학습 이미지는 컨볼루션 레이어에서 추출된 특징맵으로 앞 단의 레이어에서 출력되는 특징맵을 이용할 경우 출력 맵의 크기  $w(\text{width}) \times h(\text{height})$ 와 출력되는 맵의 깊이  $d(\text{depth})$ 가 크기 때문에 연산량이 많아진다. 또한 강건한 필터를 결정하기 위해 각 필터가 어느 정도의 범위로 갱신이 이루어져야 하는지 결정해야한다.

### 2.5 제안 알고리즘

본 논문에서 제안하는 방법으로 구성된 추적방법을 다

음과 같은 알고리즘 형태로 제안 한다

- ① 첫 번째 프레임에서 추적 객체의 영역을 지정하여 입력 이미지로 받는다.
- ② 입력 이미지는 추적 목적으로 직접 제작한 CNN의 각 Convolution Layer계층을 지나며 특징맵 형태로 출력된 고유의 외관모델로 변환된다.
- ③ ②를 통해 획득한 특징맵은 다음프레임에서 추적 객체를 찾기 위한 코릴레이션 맵으로 사용된다.
- ④ 다음 프레임의 영역에서 객체의 위치를 추정하기 위해 주파수 도메인환경에서 위상 상관관계를 계산한 후 최대 값이 나온 위치  $x, y$ 를 찾는다.
- ⑤ ④를 통해 획득한 위치에서 새로운 추적 객체의 이미지를 구하고, Convolution Layer를 통해 특징맵을 구한다.
- ⑥ 획득한 특징맵을 이용하여 코릴레이션 필터의 계수들을 갱신한다.
- ⑦ 이후 ②~⑥의 과정을 비디오 시퀀스의 마지막까지 반복한다.

## 3. 결론

본 논문에서는 컨볼루션 레이어에서 추출되는 특징맵과 적응적 코릴레이션 필터를 사용하여 온라인 학습을 통해 강건하게 객체를 추적하는 알고리즘을 제안하였다. 추후 이론적인 알고리즘을 실제 구현하고, 검증 영상을 통해 추적알고리즘을 실험하고 검증하고자 한다.

## 사사의 글

본 연구는 원자력안전위원회의 재원으로 한국원자력안전재단의 지원을 받아 수행한 원자력안전연구개발사업의 연구결과입니다. (No. 1403025)

## 참고문헌

- [1] Smeulders, Arnold WM, et al. "Visual tracking: An experimental survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1442-1468, 2014.
- [2] Bolme, David S., et al. "Visual object tracking using adaptive correlation filters," Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, pp. 2544-2550, 2010.
- [3] Wang, Lijun, et al. "Visual tracking with fully convolutional networks," Proceedings of the IEEE International Conference on Computer Vision, pp. 3119-3127, 2015.
- [4] Naresh Boddeti, et al. "Correlation filters for object alignment," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2291-2298 2013.