

딥러닝을 통한 움직이는 객체 검출 알고리즘 구현

이유경*·이용환**†

*백석문화대학교 스마트폰미디어학부, **†원광대학교 디지털콘텐츠공학과

Implementation of Moving Object Recognition based on Deep Learning

YuKyong Lee* and Yong-Hwan Lee**†

*Dept. of Smart Phone Media, BaekSeok Culture University

**† Dept. of Digital Contents, Wonkwang University

ABSTRACT

Object detection and tracking is an exciting and interesting research area in the field of computer vision, and its technologies have been widely used in various application systems such as surveillance, military, and augmented reality. This paper proposes and implements a novel and more robust object recognition and tracking system to localize and track multiple objects from input images, which estimates target state using the likelihoods obtained from multiple CNNs. As the experimental result, the proposed algorithm is effective to handle multi-modal target appearances and other exceptions.

Key Words : Object Detection and Tracking, Object Recognition, Moving Object, Deep Learning, CNN(Convolution Neural Network)

1. 서 론

실세계에서 객체를 인식하고 대상 객체의 움직임을 추적하는 기술은 컴퓨터 비전 분야에서 광범위하게 활용되는 응용 기술이다. 특히, CCTV를 포함한 감시 시스템, 자동 주행 자동차 시스템, 군용 시뮬레이션 시스템, 교통 통제 및 감시 시스템 등 매우 다양한 분야에서 객체 인식 및 추적 기술을 활용하고 있다[1]. 특히 컴퓨터 비전(Vision) 기반 기술을 활용하여 객체를 검출하고 추적하는 응용 분야로 증강현실 서비스가 주목받고 있으며[2], 이러한 기술은 산업현장에서도 많이 활용되고 있다. 주로 컴퓨터와 카메라를 이용해 산업 현장에서 사람의 눈을 대신하여 시각적으로 무엇인가를 인식하고 이를 응용하는 머신 비전을 특화하는 다양한 산업 기술이 발전되고 있다[3,4].

이러한 객체 검출과 추적의 주요 활용 목적은 입력되는 영상 시퀀스(Sequence)에서 객체의 이동 흐름과 동작을 획득하여 객체에 대한 다양한 형태의 분석 처리와 정보 제공이 가능하기 때문이다[5]. 산업 분야의 적용 가능성과 활용성을 고려하면 실세계에서 획득하는 입력 시퀀스를 기반한 컴퓨터 비전 기술은 사람과 컴퓨터의 상호 동작에서 매우 중요한 역할을 담당한다[6]. 이를 바탕으로, 감시 및 통제 시스템을 포함하여 객체 추적의 대표적인 응용 서비스 분야에서 효과적인 기법들이 제안되고 연구되고 있다[7-8]. 그러나 카메라로 투영되는 화면의 잡음, 센서 잡음, 이동하는 객체의 외곽선 변화, 조명의 불규칙적인 변화, 복잡한 배경과 객체의 상호 작용 등 다양한 환경과 배경에 따라, 객체의 검출과 추적은 여전히 어려운 문제점을 가지고 있다[9].

본 논문에서는 Multiple CNN(Convolution Neural Network)에서 얻은 객체 측정 여부를 통해 관심 대상 객체를 추정하

†E-mail: hwany1458@empal.com

고 다중 객체를 검출하는 알고리즘을 제시하고 이를 구현한다. 구현 시스템은 CNN 기반 트리를 구성하고 자동 갱신을 통해 검출 객체의 변화 이력을 추정하고 지속적인 객체 추적 여부를 검증한다.

본 논문의 구성은 다음과 같다. 2장에서 객체 검출 및 추적과 관련된 연구 동향을 살펴보고 3장에서 딥 러닝을 기초하여 입력 받는 사진 영상으로부터 관심 대상 객체를 검출하는 알고리즘을 제시하고 구현 시스템에 대한 평가를 수행하고 4장에서 결론으로 마무리한다.

2. 관련 연구

객체 검출 및 추적에서 대표적으로 Meanshift, Camshift, Kalman Filter 등의 알고리즘이 연구되고 있다[9]. Meanshift와 Camshift 알고리즘은 탐색 창을 통하여 추적하는 객체의 위치와 중심 영역을 계산한다[10-11]. Camshift는 Gary가 1988년에 처음으로 제안하였으며, Meanshift라는 확률 분포의 피크(Peak)를 찾기 위해 그라데이션(Gradient)을 동반한 비모수적(Non-parametric) 기법을 기반한다[12]. HSV 칼라 모델에서 색조(H)는 조명 영향을 받지 않기 때문에 칼라에서 색조 값을 활용하여 연산한다[13]. 채도(S)와 밝기(V)를 무시하기 때문에 문제가 발생할 수 있다. Camshift는 칼라 히스토그램을 대상 모델로 활용하여 색상 확률 분포 함수를 계산하고 탐색 창 중심점을 찾고 탐색 창 면적을 계산하여 객체 영역을 추출한다.

특징점 추출을 위한 알고리즘으로, 대표적으로 SIFT (Scale Invariant Feature Transform)과 SURF (Speed-Up Robust Features)이 있다. SIFT는 회전과 스케일에 영향을 받지 않는 이미지 특징 추출 알고리즘이다. 이미지 매칭 및 연결에서 많이 활용되지만 일반적으로 추출한 특징 정보와 계산 복잡도가 높다는 단점이 있다[14-15]. SURF는 SIFT를 보완한 알고리즘으로 성능이 보다 우수하고 처리시간이 빠르다[16]. SURF 특징 추출 연산은 키포인트 검출, 기술자 생성과 키포인트 매칭 과정으로 수행된다.

3. 구현 알고리즘

3.1 구현

본 논문의 객체 인식 및 추적 시스템의 흐름은 주기적인 상호 연결된 모듈 간의 집합으로 구성되며, 각 모듈은 적절한 데이터의 상호 전달을 통해 특정 관심 객체의 정보를 제시한다.

객체 검출을 위한 각 개별 구성 요소는 단일 신경망(Neural Network)을 통해 통합되며, 전체 이미지에서 객체 위치의 경계 영역을 추정하기 위해 콘볼루션 신경망(Convolutional Neural Network)의 모델을 적용한다. 입력 영상을 $S \times S$ 그리드 셀로 분할하고 객체의 중심이 격자 셀에 속하면 해당 격자 셀이 관심 객체를 감지한다. 신경망의 초기 콘볼루션 계층은 영상에서 특징을 추출하고 상호 연결된 계층에서 객체 여부 확률과 좌표를 추정하여 연결된 다음 계층의 연산으로 적용된다. 이미지 분류를 위한 네트 모델(Net Model)은 GoogLeNet 모델[17]을 참조하여 구성되며, 24개 콘볼루션 계층과 2개의 상호 연결된 레이어를 갖는다(Fig. 1).

3.2 실험

사진 학습을 위한 데이터셋으로 ImageNet에서 제공하는 1000-Class 이미지 데이터를 활용하여 20개 콘볼루션 계층을 사전 학습하였다. 딥 러닝 기법으로는 실시간 객체 탐지가 가능한 YOLO(You Only Look Once)을 사용하[18], 영상 처리 과정에서 객체 탐지 오류 여부를 검사하여 영역 박스 탐지 정보, 객체의 외곽선과 경계선 영역의 공간 정보를 활용하였다.

기술된 알고리즘 구현 환경으로, MacOS Sierra(10.12), 4GHz Intel Core i7 및 32GB DDR3 메모리를 장착한 iMac에서 개발 도구 Xcode 8.1을 이용하여 C++로 구현하였다. 실험에서는 대상 객체 모델링 및 훈련된 객체와의 일치에 초점을 맞추고 관심 대상 객체를 얼마나 잘 탐지하고 객체 모양 변형과 크기 변경 상황에서 대상 객체를 얼마나 잘 검

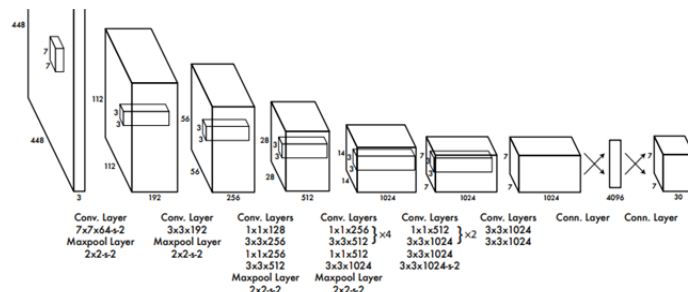


Fig. 1. Architecture of Detection Network between Interconnected Layers.



Fig. 2. Example of the detected and tracked object by input image.

출하는지에 집중적으로 실험한다. Fig. 2는 구현 알고리즘을 통해 관심 대상 객체를 검출하고 객체의 유형을 인식한 결과를 예시로 보여준다.

실험 결과, 중첩 정밀도의 관점에서 구현 시스템은 관심 대상의 검출 및 인식에서 좋은 성능을 보였다. 구현 시스템은 오버랩 정밀도를 효율적으로 증가시켜, 기존의 다른 방법과 비교하여, 적절한 시간 내에서 관심 대상을 추출하고 해당 객체가 무엇인지를 정확하게 인식함을 실험을 통해 검증하였다.

4. 결 론

본 논문에서는 다중 CNN에서 얻어진 객체 추정 가능성을 이용하여 목표 상태를 추정하고 다중 객체의 위치

를 파악하는 알고리즘을 제시하고 이를 구현하였다. 구현 시스템에서는 CNN을 기반으로 트리 구조를 유지하고, 트리 경로에 따라 자동으로 갱신한다. 각 트리 경로는 대상 객체의 변화에 대한 별도의 이력을 추적함으로써, 구현 시스템은 다중 모드 객체 운곽을 검증하고 추적 실패와 같은 예외 처리에서 효과적임을 실험을 통해 검증하였다. 이를 통해 새로운 영상 또는 화면을 입력 받아 관심 대상의 객체를 검출하는데 보다 효율적인 검출 결과를 얻을 수 있음을 실험을 통해 확인하였다.

향후 연구로, 복잡한 배경 화면에서 물체를 탐지하고 감시 또는 추적이 가능한 다양한 응용분야에 구현 시스템을 확장할 예정이다. 최근 증강현실은 산업계에서 많은 주목을 받고 있는 응용 서비스 분야이며, 실세계 환경에서 다양한 융합 응용 애플리케이션이 출시되고 있다. 특

히, 문화 유산에 적용한 문화 콘텐츠 증강현실은 우리나라에 좋은 활용 서비스의 예시가 될 수 있을 것으로 기대되며, 본 논문에서 구현한 시스템의 좋은 적용 사례가 될 것으로 판단한다. 또한 실세계 배경 화면이 상이한 물 속에서 관심 대상을 검출하는 실험에 적용할 예정이다.

감사의 글

본 연구는 2018년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업입과 제번호: 2018R1A2B6008255).

참고문헌

1. Yang H., Shao L., Zheng F., Wang L. and Song Z., "Recent Advances and Trends in Visual Tracking: A Review", *Neuro-computing*, vol.74, no.18, pp.3823-3831, (2011).
2. Yilmaz A., Javed O. and Shah M., "Object Tracking: A Survey", *ACM Computing Surveys*, vol.38, no.4, (2006)
3. Marcin Kuzanski, Anna Fabijanska and Dominik Sankowski, "Machine Vision – Automation of Selected Measurement Systems", *International Conference on Perspective Technologies and Methods in MEMS Design*, (2008).
4. Wai Lee, "3D Machine Vision in IoT for Factory and Building Automation", *International Symposium on Circuits and Systems*, (2017).
5. Klohofer W. and Kampel M., "Interest Point based Tracking", *International Conference on Pattern Recognition*, pp. 3549-3552, (2010).
6. Angela Zhou, "Cybernetics and Human-Computer Interaction: Case Studies of Modern Interface Design", *International Conference on Multidisciplinary in IT and Communication Science and Applications*, pp.1-6, (2016).
7. Kinjal A. Joshi and Darshak G. Thakore, "A Survey on Moving Object Detection and Tracking in Video Surveillance System", *International Journal of Soft Computing and Engineering*, vol.2, issue.3, pp.44-48, (2012).
8. Pawan Kumar Mishra and G. P. Saroha, "A Study on Video Surveillance System for Object Detection and Tracking", *International Conference on Computing for Sustainable Global Development*, (2016).
9. Yue Y., Gao Y. and Zhang X., "An Improved Camshift Algorithm Based on Dynamic Background", *International Conference on Information Science and Engineering*, pp.1141-1144, (2009).
10. Leichter I., Lindenbaum M. and Rivlin E., "Meanshift Tracking with Multiple Reference Color Histograms", *Computer Vision and Image Understanding*, 114(3), pp.400-408, (2010).
11. Ahn H., Lee Y., Lee J. and Cho H., "Research on Target Tracking based on CamShift Approach with Feature Matching", *International Conference on Convergence Technology*, pp.930-931, (2015).
12. G. R. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology Journal*, 2nd Quarter, (1998).
13. Woori Han, Youngseop Kim, Yong-Hwan Lee, "Multi-Object Tracking based on Keypoints", *Journal of the Semiconductor and Display Technology*, 14(3), pp.67-72, (2015).
14. Lowe D. G., "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, 60(2), pp.91-110, (2004).
15. Yong-Hwan Lee, Je-Ho Park, Youngseop Kim, "Comparative Analysis of the Performance of SIFT and SURF", *Journal of the Semiconductor and Display Technology*, 12(3), pp.59-64, (2013).
16. Bay H., Tuytelaars T. and Van Gool L., "SURF: Speeded-Up Robust Features", *International Conference on ECCV*, pp.404-417, (2006).
17. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke and Andrew Rabinovich, "Going Deeper with Convolutions", *IEEE Conference on Computer Vision and Pattern Recognition*, (2014).
18. Joseph Redmon and Santosh Divvala, "You Only Look Once: Unified, Real-time Object Detection", *Computer Vision and Pattern Recognition*, pp.1-10, (2016).

접수일: 2018년 6월 19일, 심사일: 2018년 6월 20일,
게재확정일: 2018년 6월 21일