

컨볼루션 신경망의 특징맵을 사용한 객체 추적

임수창[†], 김도연^{**}

Object Tracking using Feature Map from Convolutional Neural Network

Suchang Lim[†], Do Yeon Kim^{**}

ABSTRACT

The conventional hand-crafted features used to track objects have limitations in object representation. Convolutional neural networks, which show good performance results in various areas of computer vision, are emerging as new ways to break through the limitations of feature extraction. CNN extracts the features of the image through layers of multiple layers, and learns the kernel used for feature extraction by itself. In this paper, we use the feature map extracted from the convolution layer of the convolution neural network to create an outline model of the object and use it for tracking. We propose a method to adaptively update the outline model to cope with various environment change factors affecting the tracking performance. The proposed algorithm evaluated the validity test based on the 11 environmental change attributes of the CVPR2013 tracking benchmark and showed excellent results in six attributes.

Key words: Object Tracking, Correlation Filter, CNN, Feature Map, Appearance Model

1. 서 론

최근 연산장치의 컴퓨팅 파워가 향상됨에 따라 많은 연산량을 요구하는 컴퓨터비전(computer vision)의 다양한 영역에서 향상된 연구결과들을 보여주었다.

컴퓨터비전에서 추적 분야는 가장 도전적이고 기본적인 문제로 여겨진다. 객체추적(object tracking)은 비디오 프레임의 추적대상 객체를 지속적으로 검출하고 추적하는 기술로서 주로 로봇비전, HCI, 보안, 국방 등 시각 처리를 필요로 하는 다양한 분야에서 응용되며 사용된다. 객체 추적은 비디오의 연속되

는 프레임에 존재하는 추적 대상에서 객체의 크기, 위치, 이동방향, 색상 분포, 엣지등 등 객체의 다양한 정보를 유추해 낸다. 여기서 가장 중요한 요소는 각각의 키 프레임에 나타난 추적대상 객체의 유사성을 비교하기 위한 분석 방법과 그 기준을 정하는 것인데, 일반적으로 추적대상 객체의 추상적 특징을 표현한 특징 모델을 사용한다. 특징 맵(feature map)이라고도 부르는 이 특징 모델은, 추적대상 객체를 나타내고 있는 각 픽셀들의 정보, 즉 특징 벡터들을 모델화 한 것이며, 현재도 객체추적을 위해 반드시 필요한 요소이다.

특징 맵은 추적대상 객체의 추상적 특징을 집약해

※ Corresponding Author : Do Yeon Kim, Address: (57922) Jungang-ro 235, Suncheon-si, Jeollanam-do, Korea, TEL : +82-61-750-3628, FAX : +82-61-750-3620, E-mail : dykim@sunchon.ac.kr

Receipt date : Dec. 14, 2016, Revision date : Jan. 9, 2017
Approval date : Jan. 13, 2017

[†] Dept. of Computer Science, Sunchon National University
(E-mail : suchangLim@sunchon.ac.kr)

^{**} Dept. of Computer Engineering, Sunchon National University

※ This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety(KOFONS), granted financial resource from the Nuclear Safety and Security Commission(NSSC), Republic of Korea (No.1403025)

놓은 모델이며, 특징 맵을 구성하고 있는 특징들이 얼마나 추적대상 객체를 잘 표현하고 있는지에 따라 객체추적 성능이 매우 달라진다. 전통적으로, 특징 맵 모델링은 사람의 수작업(hand-crafted)을 통해 수행되어 왔다. SIFT(scale invariant feature transform), SURF(speeded up robust features), HOG(histogram of oriented gradient)와 같은 알고리즘들은 영상에서 객체의 특징을 추출하기 위해 사용되는 알고리즘이며, 오늘날에도 특징 추출을 위해 많이 사용되고 있다. 그러나 이러한 방법을 통해 제작된 특징 맵은 몇 가지 한계점을 갖는다. 특정 객체를 표현하기 위한 특징 추출 알고리즘의 이상적인 파라미터가 존재하지 않으며, 특징 맵 모델링은 대상 객체에 대한 제작자의 이해 수준에 따라 다르게 진행될 수 있으며, 인간이 고려하기 어려운 다양한 특징 변수들이 분포되어 있기 때문이다. 이러한 문제점을 해결하기 위해서 본 논문에서는 최근 각광받고 있는 딥러닝의 한 부류인 CNN의 컨볼루션 레이어에서 특징맵을 추출하고 이를 수학적 모델링을 통해 객체 외형모델을 제작한다. 또한 추적 성능에 영향을 주는 다양한 환경변화에 강건히 대응할 수 있도록 외형 모델들만 매 프레임에서 적응적으로 업데이트된다.

논문의 구성은 다음과 같다. 2장에서 기존에 연구된 추적 알고리즘을 설명하고, 3장에서는 제안한 추적 알고리즘을 개발하기 위한 특징맵 추출, 외형모델 제작 방법 및 모델을 적응적으로 업데이트하기 위한 방법을 기술 하였다. 4장에서는 추적 알고리즘의 성능을 실험하고 검증 하였으며, 5장에서는 결론 및 향후 연구를 제시 하였다.

2. 관련연구

2.1 Particle Filter를 이용한 추적알고리즘

객체 추적에 사용되는 파티클 필터는 베이지안 확률 분포에 근거하고 있으며, 추적분야뿐만 아니라 다양한 분야에서 사용되고 있다. 이 필터가 가진 특성으로 선형(linear) 및 비선형(non-linear)환경에 적합하다. 파티클 필터를 이용한 추적단계는 파티클의 분포를 사용하여 사후확률을 예측하는 예측(prediction) 단계, 각 파티클들의 가중치를 갱신하는 업데이트 단계, 마지막으로 여러 파티클들을 중요 가중치를 기준으로 재 위치시키는 재샘플링(resampling) 단계

로 이루어져있다. 특히 파티클 필터에서 중요한 요소는 바로 측정모델(measurement model)이다. 이는 추적대상 객체를 표현하는 형태로서 기존에는 색상 분포를 이용한 히스토그램[1], HOG(Histogram of Oriented Gradients)[2], 평균이동(mean shift)와의 결합[3]을 통한 방법을 이용하였다. 이를 통해 생성된 측정모델에서 상태 변수를 통해 얻은 측정값인 likelihood를 이용해 상태를 갱신하다. 이들 방법은 가장 보편적으로 사용되는 방법으로서, 속도가 빠르고 많은 컴퓨팅과위를 요구하지 않기 때문에 실시간 응용분야에서 많이 사용되고 있다. 하지만 측정모델이 추적대상 객체를 표현하는 정도에 따라 추적의 성능에 영향을 준다. 특히 명암변화, 장애물에 의하여 객체가 가려지는 폐색현상(occlusion), 객체의 변형(deformation) 등이 대표적인 추적 성능에 영향을 주는 환경 변화 요소이다. 수작업(hand-crafted)특징들로 표현된 측정모델은 이러한 환경 요소들에 강건히 대응할 충분한 특징표현을 할 수 없다는 단점을 가지고 있다.

2.2 Correlation Filter 기반 추적알고리즘

주파수 도메인에서 FFT(Fast Fourier Transform)연산으로 얻을 수 있는 높은 계산 효율 때문에 객체추적 분야에서 많은 주목을 받고 있으며 다양한 알고리즘들이 연구되어왔다[4-6]. 코릴레이션 필터 기반 추적 알고리즘은 입력 프레임에서 객체 중심을 기준으로 서브 윈도우를 추출하고, 다음 프레임에서 추출된 서브윈도우와의 코릴레이션 연산을 진행한다. 이를 통해 획득한 코릴레이션 응답 맵에서 가장 큰 값을 추출하고 이 값을 목표의 위치로 추정하여 추적을 진행한다. 이후 추정된 객체 위치를 기준으로 코릴레이션 필터를 업데이트 한다.

코릴레이션 필터를 이용한 추적에서 가장 중요한 요소는 코릴레이션 필터의 제작, 즉 추적대상 객체를 강건히 표현하는 객체외형 모델을 제작하는 것이다. 이 외형모델이 객체를 표현하는 정도에 따라 추적의 성능은 달라진다.

기존의 연구들은 객체의 외형을 표현하는 외형모델 제작에 초점을 맞추고 있다. DS Bolme이 제안한 방법[4]은 코릴레이션연산 이용하여 최소제곱 에러합 기반의 MOSSE 필터를 제안하였다. MOSSE필터는 새로운 프레임의 추적대상 객체 외형을 적응적으

로 학습하고, 필터의 응답맵에서 가장 신호가 강한 PSR(Peak to Sidelobe Ration)를 이용해서 추적한다. JF Henriques 등이 제안한 KCF알고리즘[5]은 코릴레이션 필터를 이용해 신뢰맵(Confidence map)을 제작한다. 이 신뢰맵을 기반으로 베이지안 확률을 사용해 추적을 진행한다. 하지만 이 방법은 필터를 학습시키는 학습율이 고정되어 있다. Liu 등이 제안한 추적 방법[6]은 객체의 외형 전체를 이용하지 않고, 외형을 일정 영역으로 나눈 뒤 각 영역에서 코릴레이션 필터를 이용한 신뢰맵을 제작한다.

2.2 CNN 기반 추적 알고리즘

추적 알고리즘은 외형 모델을 강건히 유지하기 위한 실시간으로 업데이트 과정과 목표 객체의 위치를 찾는 두 가지 과정으로 주로 이루어져 있다. 최근의 CNN 기반의 추적 방법들은 상기 내용 중 외형 모델들을 제작 하는 것에 초점을 맞추고 있다. S. Hong[7]은 추적 대상 객체의 시각돌출맵(VIsual Saliency map)을 제작하기 위하여 CNN을 사용하였다. 제작된 시각 돌출맵을 사용하여 online으로 객체 추적을 진행하였다. 다른 방법은 FCN(Fully Convolutional network)을 보행자 추적에 사용한 연구 내용도 있었다[8]. FCN은 일반적인 CNN과 달리 모든 레이어가 컨볼루션 레이어로 이루어져 있으며 주로 의미적 분할(semantic segmentation)에 사용된다. 이 논문에서는 기존의 연구내용과 달리 단일 네트워크가 아닌 다중 네트워크를 사용하였으며, 입력된 이미지에서 feed forward과정을 거치며 객체의 영역을 분할한다. 이후 분할된 영역을 기준으로 추적을 진행한다.

3. 제안한 추적 알고리즘

본 논문에서 제안하는 추적알고리즘은 CNN의 컨

볼루션 레이어에서 추출되는 특징맵을 수학적 모델링을 통해 객체 외형을 제작하고 제작된 외형모델을 다양한 환경변화에서 강건히 대응할 수 있도록 적응적 업데이트 과정을 진행하며 객체 추적이 이루어진다.

3.1 컨볼루션 레이어를 이용한 특징맵 추출

동물의 시각처리 과정을 모방해서 만들어진 Convolutional Neural Network(CNN)은 2012년 ILVRC (Image Large Scale Visual Recognition Challenge)에서 AlexNet[9]이 우승함으로써 그 성능이 검증되었으며, 이를 기점으로 최근 컴퓨터 비전의 다양한 분야에서 이를 이용한 많은 연구들이 진행되고 있다.

CNN은 컨볼루션 레이어, 풀링 레이어, 풀리 커넥티드 레이어로 이루어진 형태로 구성되어있으며, 망이 깊어질수록 객체 고유의 특징들을 고수준으로 추출한다. 하위 계층의 레이어는 객체의 공간 (spatial detail)를 표현하는 특징들을 추출하고, 최상위 레이어로 이동할수록 의미 정보(semantic information)를 추출한다. 본 논문에서는 컨볼루션 레이어만 사용해서 특징을 추출한다. 풀링 레이어의 출력은 앞단의 컨볼루션 레이어의 출력되는 특징들 중에서 최대값을 추출하기 때문에 정보의 손실이 발생하고, 풀리 커넥티드 레이어의 출력은 1x1 크기의 공간 해상도로 구성되어있기 때문에 분류 목적으로는 유용하지만, 추적대상 객체의 위치를 추정하기 위한 충분하지 않은 정보로 구성되었기 때문에 사용하지 않는다.

본 논문에서는 100만장 이상의 이미지를 사용해서 사전에 학습된 19계층의 imagenet-vgg-verydeep19 네트워크[10]를 사용한다. Fig. 1에서 보여주는 이 네트워크는 16개의 컨볼루션 레이어를 포함하고 있으며, 전체 43계층의 레이어로 이루어진 거대한 네트워크이다. 특징맵은 VGG 네트워크의 각각의 컨볼루션 레이어에서 얻을 수 있다.

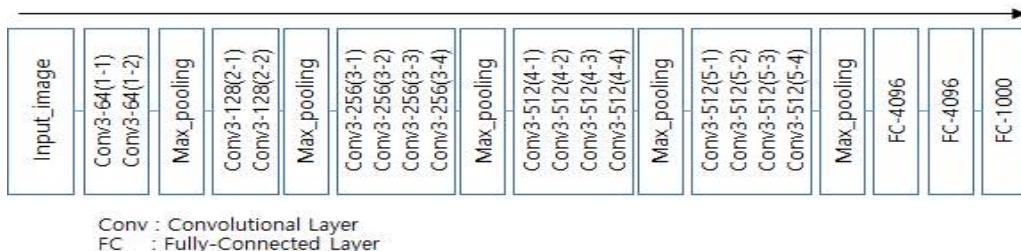


Fig. 1. vgg-verydeep19 Network configuration.

특징맵을 추출하기 위해서 공간 정보를 표현하는 Conv1-2 레이어와 의미 정보를 표현하는 Conv5-1, 5-2 레이어를 이용하였다. Fig. 1은 추출된 특징맵을 시각화한 결과를 보여주고 있다. 입력 이미지인 Fig. 2(a)는 계층별 컨볼루션 레이어를 거치며 공간 정보(Fig. 2(b))와 의미 정보(Fig. 2(c), (d))로 구성된 객체 고유의 특징들로 추출된다. 하지만 각 레이어에서 추출되는 특징맵의 크기가 다르기 때문에 이중 선형 보간법(bilinear interpolation)을 사용하여 동등한 크기로 업/다운 샘플링을 한다.

Fig. 2(b)의 특징 맵은 객체 영역뿐만 아니라 주변 배경 영역에서도 특징을 추출하는 것을 보여준다. 이를 통해 특징 맵에서 주된 관심영역은 중앙부에 분포된 객체의 영역이며, 가장자리로 갈수록 배경일 확률이 높다는 것을 알 수 있다. 따라서 객체의 영역에 집중할 수 있도록 cosine window[4]를 사용하여 가중치를 부여한다.

3.2 코릴레이션 필터를 이용한 객체 외형모델 제작

상관관계(correlation) 연산은 계산상의 이득을 위해 FFT(Fast Fourier Transform)을 사용해서 주파수 도메인으로 변환된다. 입력 이미지와 필터의 2D fourier transform은 식 (1)과 (2)로 표현이 가능하다.

$$F(u,v) = FFT(f(x,y)) \quad (1)$$

$$G(u,v) = FFT(g(x,y)) \quad (2)$$

회전(convolution) 정리의 상관관계(correlation) 연산은 주파수 도메인에서 $F(u,v)$ 와 $G(u,v)$ 의 곱셈과 같다. 이런 이론을 기반으로 상관관계 연산은 식 (3)로 표현 가능하다.

$$H(u,v) = F(u,v) \odot G^*(u,v) \quad (3)$$

식 (3)에서 H 는 상관관계 출력맵이고, $G^*(u,v)$ 는

$G(u,v)$ 의 켈레복소수(complex conjugate)를 나타낸다. 기호 \odot 는 아다마르(Hadamard) 곱이다.

컨볼루션 레이어에서 획득한 특징 맵들은 외형 모델을 제작하는 용도로 사용된다. 특징 맵은 $H \times W \times D$ 의 크기로 출력되는데 여기에서 H 는 높이, W 는 넓이, D 는 깊이를 의미한다. i 번째 레이어에서 출력되는 특징 맵을 x_i^d 로 나타낸다.

correlation 연산을 통해 얻어진 필터 H 는 특징 맵의 출력물 x_i 의 제곱합 에너지를 최소화하는 식 (4)의 minimization problem을 해결함으로써 얻어진다.

$$H^* = \underset{H}{\operatorname{argmin}} \sum_{w,h} \| H \cdot x_{w,h} - l(w,h) \|^2 + \lambda \| H \|^2_2 \quad (4)$$

여기서 λ 는 정규화 파라미터이고 벡터의 내적은 힐버트 공간(Hilbert space)에서 선형 커널로 나타낼 수 있다. 특징 맵 x_i 를 구성하는 $D(d \in \{1, \dots, D\})$ 개의 깊이의 주파수 도메인에서 학습된 필터는 식 (5)처럼 표현 가능하다[11].

$$H^d = \frac{L \odot X_i^{*d}}{\sum_i \sum_{d=1}^D X_i^d \odot X_i^{*d} + \lambda} \quad (5)$$

식 (5)에서 대문자는 주파수 변환된 값을 의미하고, $I(i \in \{1, \dots, I\})$ 는 특징 맵을 추출하는 컨볼루션 레이어의 개수를 나타낸다.

$$\max(f_i) = FFT^{-1} \left(\sum_{d=1}^D H^d \odot X_i^{*d} \right) \quad (6)$$

식 (5)를 이용해서 객체 이미지의 특징 맵에서 필터를 계산한다. 다음 프레임의 영상내에 객체의 위치를 추정하기 위해서 식 (6)을 이용한다. 계산을 통해 출력된 f_i 는 $w \times h$ 크기의 상관관계 응답맵이다. 목표 객체의 위치는 출력된 응답맵에 존재하는 최대값의 위치를 찾음으로 추정할 수 있다.

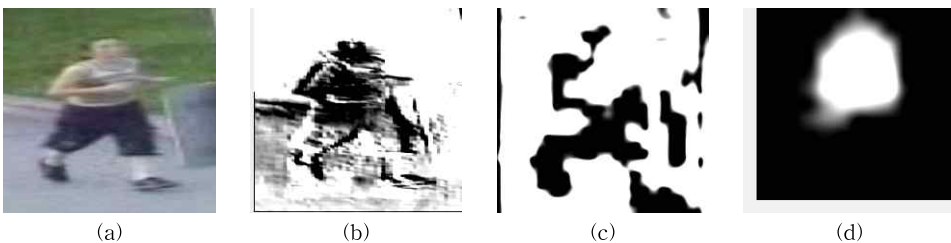


Fig. 2. Feature maps extracted from each convolution layer. (a) input image, (b) Conv 1-2, (c) Conv5-1, and (d) Conv 5-2.

3.3 외형모델의 적응적 업데이트

입력 영상에 첫 번째 프레임의 추적 대상 객체로 제작된 외형 모델은 모든 영상 시퀀스에 존재하는 객체를 추적하는 용도로 사용하기에 부적절하다. 그 이유는 각 프레임의 객체는 다양한 형태로 존재하여 초기의 외형 모델과 다음 프레임에서 제작된 외형 모델사이의 유사성이 상이하기 때문이다. 따라서 이러한 문제점을 해결하기 위해서 두 가지 사항을 고려하여야 한다. 첫 번째로 제작된 외형모델은 매 프레임에서 지속적으로 업데이트 되어야한다. 추적 알고리즘이 객체를 추적하는 동안 목표 객체는 회전, 크기 변화, 포즈 변화, 다양한 형태의 명암 상태 또는 비정형적인 변형 등으로 다양한 형태를 가진다. 따라서 객체 추적을 강건히 진행할 수 있도록 매 프레임마다 외형 모델의 실시간 업데이트를 수행 하여야 한다. i 번째 컨볼루션 레이어에서 출력되는 특징 맵을 이용해 제작된 가장 강건한 필터는 상관관계 연산을 통해 출력된 값을 최소자승법(Least Square Method)으로 업데이트가 가능하다[12].

두 번째로 외형모델이 잡음 등에 의해 오염되지 않도록 적응적으로 업데이트 되어야한다. 매 프레임에서 조건 없이 업데이트가 이루어진다면 추적기가 객체를 추적하는 도중에 객체를 잃어버리는 경우, 객체 영역에 위치한 배경을 기준으로 업데이트가 이루어지기 때문에 이후 프레임들에서 외형 모델은 오염된 채로 업데이트 된다. Fig. 3는 적응적 외형 모델 업데이트 알고리즘의 필요성을 보여준다. 이전 프레임의 객체이미지 Fig. 3(a)는 객체 원본 영역을 보여주지만, Fig. 3(b)는 객체의 대부분의 영역을 가리는 폐색 현상이 발생하였다. 일반적인 업데이트 알고리즘은 이런 상황이 발생했을 때, 업데이트 진행 여부를 고려하지 않고 무조건 업데이트를 진행하기 때문



Fig. 3. (A) the original image and (B) the image in which the occlusion occurred. (a) original image, (b) Images where occlusion occurred.

에 오염된 값으로 외형 모델 계수가 업데이트 된다.

본 논문에서는 이러한 문제점을 해결하고자 추적 모델에 대해서 적응적 업데이트 알고리즘을 제안 한다.

이전 프레임과 현재프레임의 외형 모델 사이의 유사도를 구하기 위해 코사인 유사도(cosine similarity)를 사용한다. 이 방법은 벡터공간(vector space)에 존재하는 두 벡터의 내적을 코사인 값을 이용해서 측정 가능하다. 측정된 결과 값은 코사인 함수의 특성에 의해 Θ 의 값이 0° 면 1, 90° 면 0의 값을 갖는다. 이 측정 방법을 통해 획득한 유사도가 임계값보다 작을 경우 폐색 현상이 발생한 것으로 판단하여 업데이트를 진행하지 않는다. 유사도가 1에 근접할 경우, 변화가 없는 것으로 판단하고 업데이트 과정의 계산량을 줄여주기 위해 업데이트를 진행하지 않는다. 유사도가 임계값 범위 내에 존재하면 객체 외형의 변화가 발생한 것으로 판단하고 획득한 가중치 값으로 업데이트를 진행한다. 이 과정을 적용하면 Fig. 2(b)처럼 폐색 현상이 발생한 영역을 업데이트하지 않고, 폐색 현상이 발생하지 않은 영역에 대해 업데이트를 진행한다.

$$A_i^d = \begin{cases} (1 - \eta w_{i-1}^d) A_{i-1}^d + \eta w_i^d (Y \odot X_i^{*d}) & w_i^d > threshold \\ A_{i-1}^d & otherwise \end{cases} \quad (7)$$

외형 모델은 식 (7)을 사용하여 적응적으로 업데이트 된다. 이 식에서 A 는 최종 업데이트된 외형 모델로서 첫 번째 프레임에서는 업데이트 이전의 출력 모델로 초기화 된다. 변수 i 는 프레임 인덱스이고 η 는 학습률이다. w_i^d 는 출력된 유사도 값에 따라 0.001 ~ 0.009 사이의 값으로 정규화된 값이다.

Fig. 4은 제안한 방법을 이용하여 객체를 추적한 결과를 보여준다. Fig. 4(a)은 객체를 정상적으로 추적하는 것을 보여주지만, Fig. 4(b)에서는 장애물로 인한 폐색 현상에 의해 객체를 놓친 결과를 보여준다. 하지만 제안한 적응적 업데이트 과정을 통해 객체를 놓친 영역(배경)을 업데이트 하지 않았기 때문에 Fig. 4(c), (d)에서 재 추적이 이루어지는 것을 볼 수 있다.

4. 추적 알고리즘의 성능검증

본 논문에서는 제안한 추적알고리즘의 성능 검증



Fig. 4. Tracking results using the proposed method.

을 위해 11개의 환경변화 속성으로 구성된 CVPR 2013 추적 벤치마크[13]와 TLD[14], SCM[15], LOT [16], CSK[17], VTD[18], ASLA[19] 등 6종의 추적 알고리즘을 사용하였다. 성능 검증을 실시한 실험 환경은 Table 1과 같다.

제안한 추적알고리즘의 정량적 평가를 위해서 매뉴얼로 지정된 바운딩박스와의 중첩율을 계산하는 Success plot를 사용하였다.

Fig. 5는 성능평가의 결과로서 제안한 알고리즘이 6개 항목에서 좋은 성능을 보여주었다. 이러한 결과를 보여주는 이유는 다음과 같은 두 가지 이유가 있다. 첫 번째로, 제안한 추적 알고리즘에서는 CNN 학습맵을 이용해서 추적 대상 객체의 고유의 특징들을

추출하고 이를 외형모델로 제작하여 추적에 사용하였기 때문이다. 이 알고리즘을 사용해서 추출된 특징 맵은 서로 형태가 다른 약 500개 이상의 고수준의 정보가 담긴 특징 맵으로 구성되어 있기 때문에 추적 대상 객체를 표현함에 있어 좋은 성능을 보여 준다. 두 번째로, 장애물 및 유사도 비교에 따른 적응적 업데이트 알고리즘을 이유로 볼 수 있다.

배경이 복잡하거나 장애물에 의한 폐색 현상이 발생한 상황에서 적응적 외형 모델 업데이트 알고리즘은 유사도 측정을 통해 업데이트를 진행하지 않고, 폐색 현상이 종료 되었을 때 업데이트 과정을 진행한다. 하지만 SV(Scale Variation)에서 성능상의 한계를 보여주었다. 제안한 알고리즘은 SV에서 추적에 실패하지 않고 객체를 추적하는 것은 다른 속성에서와 같았다. 하지만 SV의 속성으로 이루어진 영상에서, 추적 대상 객체의 크기가 지속적으로 작아지거나 커지는 경우 바운딩 박스가 변화된 객체 크기에 적응적으로 대응하지 못하고 고정된 크기로 추적을 진행하기 때문에 매뉴얼의 바운딩박스 영역과 넓이가 상이하여 낮은 성능 점수를 보여주었다.

Table 1. Experimental environment of tracking algorithm

	experimental environment
CPU/RAM	Intel I7-4790K 4.0GHz / DDR3 16G
GPU	Geforce Titan X 12GB
OS	Window 10 64bit
Development Tool	Matlab R2016b
Toolbox	MatConvNet Toolbox[20]

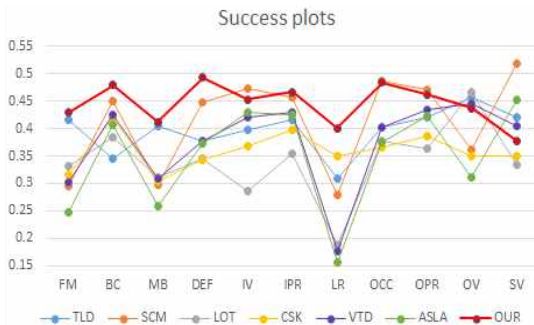


Fig. 5. Results graph of tracking performance evaluation.

5. 결 론

본 논문에서는 새로운 기술 트렌드로 떠오르고 있는 딥러닝 모델중 하나인 컨볼루션 신경망과 코릴레이션 필터를 결합한 형태의 강건한 객체 추적 알고리즘을 제안하였다. 기존의 알고리즘들은 hand-crafted 특징들을 사용해서 객체의 외형을 제작하였지만, 제안한 알고리즘은 Black-box 형태로 공간적 특징부터 의미론적 특징까지 추출해 주는 컨볼루션 레이어의 특징 맵을 이용하였다. 또한 적응적 코릴레이션 필터를 사용해서 다양한 요인들에 의해 변형되고 비정형적인 객체의 외형에 대응할 강건한 객체 외형 모델을 제작 하였다.

제한한 알고리즘의 성능 평가를 위해서 11개의 다양한 속성들로 구성된 50개의 비디오 벤치마크를 사용하였다. 정량적 평가의 결과로 6개 속성에서 우수한 성능을 보여주었지만, Scale Variation 속성에서 성능상의 한계를 보여주었다. 향후 연구과제로 낮은 성능을 보인 환경 변화속성에 강건히 대응하기 위한 추가적인 연구가 필요하다. 또한 제안한 알고리즘은 추적 대상 객체의 영역 주변을 패딩하는 형태로 위치를 추정하기 때문에 패딩영역을 벗어나는 객체에 대해서 추적이 실패한다. 따라서 이런 문제점을 해결하기 위해 이미지 전반에 걸쳐 객체의 위치를 추정하는 연구를 진행할 것이다.

REFERENCE

- [1] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based Object Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 5, pp. 564-577, 2003.
- [2] F. Xu and M. Gao, "Human Detection and Tracking Based on HOG and Particle Filter," *Proceeding of International Congress on Image and Signal Processing*, Vol. 3, pp. 1503-1507, 2010.
- [3] I.T. Whoang and K.N. Choi, "An Algorithm for Color Object Tracking," *Journal of Korea Multimedia Society*, Vol. 10, No.7, pp. 827-837, 2007.
- [4] D.S. Bolme, J.R. Beveridge, B.A. Draper, and Y.M. Lui, "Visual Object Tracking Using Adaptive Correlation Filters," *Proceeding of International Conference on Computer Vision and Pattern Recognition*, pp. 2544-2550, 2010.
- [5] J.F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-Speed Tracking with Kernelized Correlation Filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37, No. 3, pp. 583-596, 2015.
- [6] T. Liu, G. Wang, and Q. Yang, "Real-Time Part-Based Visual Tracking via Adaptive Correlation Filters," *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4902-4912, 2015.
- [7] S. Hong, T. You, S. Kwak, and B. Han, "Online Tracking by Learning Discriminative Saliency Map with Convolutional Neural Network," *Proceeding of the IEEE International Conference on Computer Vision*, pp. 1520-1528, 2015.
- [8] L.J. Wang, O.Y. Wanli, X.G. Wang, and H.C. Lu, "Visual Tracking with fully Convolutional Networks," *Proceeding of the IEEE International Conference on Computer Vision*, pp. 3119-3127, 2015.
- [9] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Proceeding of International Conference on Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012.
- [10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Proceeding of International Conference on Learning Representations*, pp. 1-14, 2015.
- [11] C. Ma, J.B. Huang, X. Yang, and M.H. Yang, "Hierarchical Convolutional Features for Visual Tracking," *Proceeding of the IEEE International Conference on Computer Vision*, pp. 3074-3082, 2015.
- [12] V.N. Boddeti, T. Kanade, and B.V.K. Vijaya Kumar, "Correlation Filters for Object Alignment," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2291-2298, 2013.
- [13] Y. Wu, J. Lim, and M.H. Yang, "Online Object Tracking: A benchmark," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2411-2418, 2013.
- [14] Z. Kalal, J. Matas, and K. Mikolajczyk, "P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 49-56, 2010.
- [15] W. Zhong, H. Lu, and M.H. Yang, "Robust

Object Tracking via Sparsity-based Collaborative Model," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1838-1845, 2012.

[16] S. Oron, A. Bar-Hillel, D. Levi, and S. Avidan, "Locally Orderless Tracking," *International Journal of Computer Vision* 111, No. 2, pp. 213-228, 2015.

[17] J.F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the Circulant Structure of Tracking-by-Detection with Kernels," *Proceedings of European Conference on Computer Vision*, pp. 702-715, 2012.

[18] J. Kwon and K.M. Lee, "Visual Tracking Decomposition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1269-1276, 2010.

[19] X. Jia, H. Lu, and M.H. Yang, "Visual Tracking via Adaptive Structural Local Sparse Appearance Model," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1822-1829, 2012.

[20] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional Neural Networks for Matlab," *Proceedings of the 23rd ACM International Conference on Multimedia*, pp. 689-692, 2015.



임 수 창

2015년 순천대학교 컴퓨터공학과 (공학사)
 2015년~현재 순천대학교 컴퓨터공학과 석사과정
 관심분야: 영상처리, 컴퓨터비전, 기계학습



김 도 연

1986년 충남대학교 계산통계학과 졸업(이학사)
 2000년 충남대학교 대학원 정보통신공학과 졸업(공학석사)
 2003년 충남대학교 대학원 컴퓨터공학과 졸업(공학박사)
 1986년~1996년 한국원자력연구원 선임연구원
 1997년~2008년 한국전력기술(주) 책임연구원
 2008년~현재 순천대학교 컴퓨터공학과 교수
 관심분야: 컴퓨터비전, 컴퓨터보안