

Scalable Re-detection for Correlation Filter in Visual Tracking

Kayoung Park*

*Researcher, Agency for Defense Development (ADD), Daejeon, Korea

[Abstract]

In this paper, we propose an scalable re-detection for correlation filter in visual tracking. In real world, there are lots of target disappearances and reappearances during tracking, thus failure detection and re-detection methods are needed. One of the important point for re-detection is that a searching area must be large enough to find the missing target. For robust visual tracking, we adopt kernelized correlation filter as a baseline. Correlation filters have been extensively studied for visual object tracking in recent years. However conventional correlation filters detect the target in the same size area with the trained filter which is only 2 to 3 times larger than the target. When the target is disappeared for a long time, we need to search a wide area to re-detect the target. Proposed algorithm can search the target in a scalable area, hence the searching area is expanded by 2% in every frame from the target loss. Four datasets are used for experiments and both qualitative and quantitative results are shown in this paper. Our algorithm succeed the target re-detection in challenging datasets while conventional correlation filter fails.

▶ **Key words:** Visual Tracking, Correlation Filter, Searching Area, Re-detection, Long-term Tracking

[요 약]

본 논문에서는 상관필터를 이용한 영상 추적에서 탐색 영역의 크기 조절이 가능한 재탐지 방법을 제안한다. 실제 장비를 통해 영상 추적 기능을 실행할 때에는 표적이 특정 물체에 가리고 다시 나타나는 일이 빈번하게 일어나는데, 따라서 표적의 소실 판단과 재탐지 방법이 필요하다. 본 알고리즘은 강인한 추적을 위해 커널 상관필터를 사용한다. 일반적인 상관필터를 활용한 영상 추적 알고리즘에서는 표적을 탐지하는 범위가 학습된 필터의 크기에 국한된다. 하지만 표적의 가림이 오랜 시간 지속될수록 표적의 위치는 예측된 위치에서 벗어날 가능성이 커지고, 따라서 충분히 큰 범위에서 표적의 탐색이 이루어져야 한다. 제안하는 방법은 매 프레임 2%씩 탐색 범위를 넓히며 재탐지를 시도하여 성공률을 높인다. 실험은 항공에서 촬영된 4가지 영상을 활용하였고, 제안한 알고리즘은 재탐지가 어려운 데이터셋에서도 성공적인 결과를 보였다.

▶ **주제어:** 영상 추적, 상관 필터, 탐색 영역, 재탐지, 장기간 추적

-
- First Author: Kayoung Park, Corresponding Author: Kayoung Park
 - *Kayoung Park (kayoung@add.re.kr), Agency for Defense Development (ADD)
 - Received: 2020. 06. 01, Revised: 2020. 07. 13, Accepted: 2020. 07. 15.

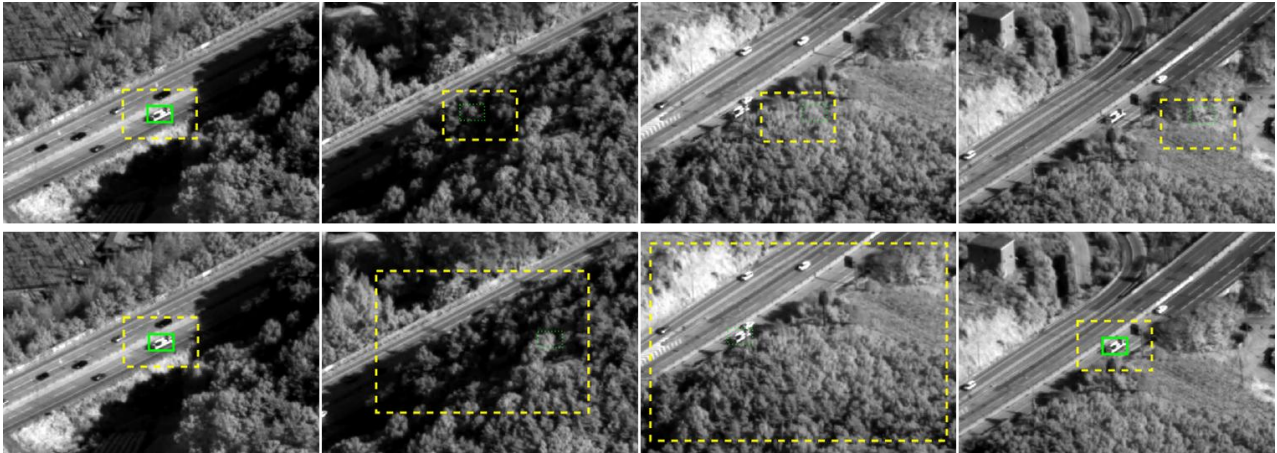


Fig. 1. Visual tracking with target disappearances. Top row : conventional correlation filter with fixed searching area, bottom row : Our approach with scalable searching area.

I. Introduction

최근 하드웨어의 발전과 스마트 기기의 보편화로 영상 정보를 사용한 다양한 자동화 기능의 중요성이 점점 대두되고 있다. 그중 동영상에서 자동으로 특정 사물의 위치를 지속적으로 찾아내는 영상 추적은 컴퓨터 비전 분야에서 깊은 연구가 이루어지고 있는 주요한 분야 중 하나이다.

대다수의 영상 추적 알고리즘 연구는 표적이 영상 안에 존재 하는 짧은 시간 동안 정확한 위치를 찾는 것에 초점을 맞추고 있다[1-3]. 하지만 실제 시스템에 영상 추적 기능을 적용한 경우, 추적하고자 하는 표적이 영상 안에서 사라졌다가 다시 나타나는 일이 빈번하게 일어난다. 이동 중인 자동차가 다리 밑 또는 터널을 지나가거나, 카메라 앞에 다른 물체가 지나갈 가능성도 있다. 카메라 시선이 고정이거나 무작위로 움직일 경우 목표하는 물체가 화면 밖으로 나갔다가 다시 나타나는 경우도 존재한다. 따라서 실제 장비에 적용하고자 하는 영상 추적 알고리즘은 추적 결과로 나타나는 위치의 정확도뿐만 아니라 영상 내 표적의 소실을 판단하고 다시 나타났을 때 재탐지하는 성능 또한 중요한 요소가 된다.

상관 필터(Correlation Filter)를 사용한 추적 알고리즘은 실시간 환경에서도 높은 성능을 얻을 수 있어 활발히 연구되고 있는 방법이다[1-6]. 주파수 도메인에서 빠른 연산이 가능한 장점을 활용하는 일반적인 상관 필터[1-4]는 학습된 필터의 크기를 변경하지 못하고 같은 크기의 영역에서 표적을 탐지하는 특징이 있다. 표적을 지속적으로 추적하고 있을 때에는 연속된 프레임 간 표적의 위치 변화가 크지 않아 높은 탐지 성능을 낼 수 있다. 반면 표적이 영상 내에서 사라졌다가 다시 나타나 위치가 크게 변화한 경우,

필터의 크기에 국한되는 탐색 범위를 벗어나 탐지에 실패할 수 있다.

본 논문에서는 탐색 범위를 키워 재탐지 확률을 높이기 위해, 학습된 필터의 크기와 상관없이 탐지 영역의 크기를 상황에 따라 조절 가능한 알고리즘을 제안한다. 넓은 영역에서 표적의 위치를 찾는 별도의 재탐지 알고리즘 적용[4, 8] 없이 표적 소실 전 학습된 필터를 그대로 사용하여 영상 추적 시스템의 효율성을 높였다. Fig. 1은 탐지 영역의 크기가 고정인 일반적인 상관 필터와 제안하는 알고리즘의 추적 결과를 보여준다. 사용된 영상은 표적이 산 뒤로 가려졌다가 다시 나타나는 영상이다. 우측 상단에서 고정 탐색 범위 밖에 표적이 나타나 재탐지에 실패한 영상을 볼 수 있고, 제안한 알고리즘으로 넓어진 탐색 범위 안에 표적이 재출현하여 탐지에 성공한 영상은 우측 하단에 나타나 있다.

본 논문의 구성은 다음과 같다. II장에서 상관 필터를 이용한 영상 추적 알고리즘을 요약하고, III장에서 제안하는 알고리즘을 설명한다. 적용하고자 하는 시스템의 영상으로 구성된 데이터셋과 이를 이용한 실험 결과를 IV장에서 소개한다. 마지막으로 V장에서 본 논문의 결론을 기술한다.

II. Correlation Filters

상관필터는 표적 주변을 포함한 윈도우 내에서 표적과의 상관도를 계산하는 방식인데, 주파수 도메인을 활용하여 빠른 계산이 가능하다. 영상 추적에서는 이전 프레임에서 표적 주변을 포함하는 윈도우 영역을 학습하여 표적과 배경을 구분하는 상관 필터를 구하고, 이를 다음 프레임에 적용하여 윈도우 내 가장 높은 응답(Response)이 나오는 위치를 표적

으로 판단한다. 작은 메모리를 필요로 하고 비교적 간단한 계산에도 불구하고 표적의 밝기, 포즈, 형상 변화 등에 강인한 성능을 보여 많은 연구가 이루어지고 있다[1-4].

상관 필터 결과의 오차 제곱의 합을 최소화 시키는 방향으로 학습한 알고리즘[1]은 영상의 밝기 정보를 사용하여 1초에 수백장을 처리하는 속도를 달성하였다. 보다 성능을 높인 커널 상관필터[2]는 영상에서 다차원 특징 벡터 (Feature)인 HoG(Histogram of Gradient)[7]를 계산하여 사용하였다. 리지 회귀(Ridge Regression)의 목표 함수를 주파수 도메인에서 계산하고, 쌍대 공간(Dual Space)을 활용해 선형 회귀뿐만 아니라 비선형 커널을 효과적으로 계산하는 방법을 제안하였다. [3]에서는 다차원 특징 벡터를 사용하면서 표적의 크기 변화를 학습하는 별도의 과정이 추가된 알고리즘이 제안되었다.

상관필터를 사용한 영상 추적에서 필터의 크기를 변경한 방법은 [5][6]에서 연구되었다. 학습 시 사용한 필터의 크기를 확대하여 더 많은 정보를 학습해 추적기의 성능을 높였지만, 탐지 단계에서 학습 시 사용한 필터의 크기를 그대로 적용한다. 상관필터를 사용한 장기간 추적 방법도 제안되었는데[4], 추적 중 응답 값으로 표적의 가림을 판단하지만 재탐지는 별도의 분류기를 도입하였다.

본 논문에서 제안하는 알고리즘은 커널 상관필터(KCF, Kernelized Correlation Filter)[2]를 기반으로 연구되었다. 커널 상관필터의 세부내용은 다음절에서 간략하게 설명한다.

1. Kernelized Correlation Filter Overview

상관필터의 목적은 표적과 배경을 분리하는 필터를 학습하는 것이다. 탐지를 위해 필요한 입력 영상 데이터가 z 일 때, 표적과 유사할수록 높은 응답을 도출하는 함수 $f(z) = \mathbf{w}^T \mathbf{z}$ 를 계산하기 위한 리지 회귀 목적식은

$$\min_w \sum_i (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2 \quad (1)$$

로 나타낸다. \mathbf{x} 는 이전 프레임에서 표적의 주변을 포함하는 영역인 학습 데이터, y 는 이상적인 응답, λ 는 정규화 계수이다. 위 식에서 \mathbf{w} 의 빠른 계산을 위해 주파수 도메인에서 계산하면 아래와 같다.



Fig. 2. Patch examples for circular correlation.

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda} \quad (2)$$

위 식에서 $\hat{\mathbf{w}} = F(\mathbf{w})$ 는 이산 푸리에 변환(DFT, Discrete Fourier Transform)을 사용한 주파수 도메인에서의 매트릭스를 의미하고, \odot 는 매트릭스의 대수 곱, 분모는 매트릭스의 대수 나눗셈을 나타낸다. 선형 회귀 대신 분별력을 높이기 위해 비선형 커널을 적용하려면 쌍대 공간에서의 계산이 용이한데, 여기서 응답을 내는 목적함수는 다음과 같이 변경된다.

$$\hat{\mathbf{f}}(z) = \hat{\mathbf{k}}^{zz} \odot \hat{\alpha} \quad (3)$$

$$\hat{\alpha} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda} \quad (4)$$

커널 매트릭스인 \mathbf{k} 는 커널의 종류에 따라 다르게 계산될 수 있다. 가우시안 커널을 사용할 경우 커널 매트릭스는 다음과 같다.

$$\hat{\mathbf{k}}^{zz} = \exp\left(-\frac{1}{\alpha^2}(\|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - 2F^{-1}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{z}}))\right) \quad (5)$$

위 식을 식 (3)에 대입하면 새로운 프레임에서의 응답 지도(Response map)를 얻게 되고, 응답이 가장 높은 위치를 표적의 위치로 판단하게 된다.

III. Scalable Searching Correlation Filter

1. Scalable Searching Correlation Filter

본 논문에서 제안하는 알고리즘은 상황에 따라 탐색 영역의 크기를 조절할 수 있는 상관필터 알고리즘이다. 학습된 상관 필터보다 큰 영역에서 탐지하기 위해서는 상관 필터의 크기를 원하는 탐지 영역에 맞추어 주어야 주파수 도메인에서 대수 곱이 가능해진다.

가장 손쉽게 크기를 맞추는 방법은 공간 도메인에서 상관 필터의 주변에 0을 채우는 제로 패딩(Zero padding)이다. 하지만 탐지 단계에서 제로 패딩을 적용하여 단순히 크기만 맞추면 표적 위치에서 가장 높은 응답이 나오지 않는다. 상관필터 알고리즘에서 활용하는 주파수 특성은, 주파수 도메인에서 켈레와의 대수 곱이 공간 도메인에서 순환 상관(Circular correlation)과 대응한다는 것이다. 그 예시가 Fig. 2에 나타나 있다. 패치는 추적을 위해 표적을 포함하는 윈도우, 즉 상관 필터 크기 또는 탐색 영역 크기 만큼 추출한 영상의 특징 벡터를 의미한다. 가우시안으로 정의한 이상적인 응답인 y 를 대입해 보면 좌상단의 표적이 중심에 있는 패치는 응답 1, 나머지 패치는 0에 가까운 응답으로 학습될 것이다. Fig. 2에서 보이는 순환(Circular) 특성 때문에 작은 상관필터에서 학습된 정보는 넓은 영역에서의 순환 상관과 상이할 수밖에 없다.

공간 도메인에서 윈도우 크기의 절반 이상의 크기로 사방에 제로 패딩을 적용하면, 윈도우 영역 내 응답은 선형 상관(Linear correlation)과 유사하게 학습된다. 본 논문에서는 학습 및 탐지 시 순환 특성을 해결하기 위한 최소값인 윈도우 크기의 절반 크기만큼 제로 패딩을 적용한다. 재탐지 시에는 넓어진 탐색 영역에 학습한 상관 필터에서 적용한 크기로 제로 패딩을 적용시켜 탐색영역 내 표적을 찾아낼 수 있다.

순환 특성이 파생하는 또 하나의 문제는 가장자리 효과(Boundary effect)이다. 순환 상관이 적용되는 효과로 영상의 가장자리 부분이 맞게 되는데 갑자기 변하는 부분에서 의도하지 않은 정보가 생성되기 때문이다. [2]에서는 공간 도메인의 특징 벡터에 코사인 윈도우(Cosine window, Hann window)를 곱해주어 간단한 방법으로 가장자리 효과 문제를 해결하였다. 본 알고리즘에서는 [2]에서 사용한 코사인 윈도우를 제로 패딩 전에 적용하였다. Fig. 2에서 나타내고 있는 예시는 코사인 윈도우가 적용된 패치이다. 이 식을 그대로 증가한 탐색 영역의 특징 벡터에 적용하면 영상의 가장자리로 갈수록 정보 손실이 커, 탐색 영역의 본래 목적인 중심부에서 떨어진 표적 탐지가 어려워진다. 따라서 큰 탐지 영역의 가장자리

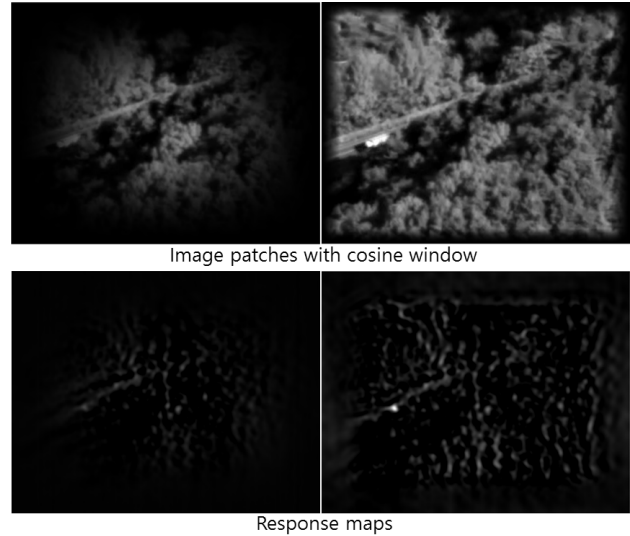


Fig. 3. Comparison with conventional cosine window (left) and a proposed cosine window (right).

부분에서는 추적 시 상관 필터에 적용한 곡률을 그대로 적용하여 가장자리 효과를 감소시키면서 그 외 부분은 1로 채워 지나친 정보 손실을 방지하였다.

$$C = \begin{cases} \frac{1}{2}(1 - \cos \frac{2\pi i}{l}), & 1 \leq i \leq \frac{l}{2} \\ 1, & \frac{l}{2} < i \leq L - \frac{l}{2} \\ \frac{1}{2}(1 - \cos \frac{2\pi(i - L + l)}{l}), & L - \frac{l}{2} < i \leq L \end{cases} \quad (6)$$

l 은 추적 중 상관 필터의 크기, L 은 증가된 탐색 영역의 크기, 즉 x 또는 y 방향 길이이다. i 는 픽셀 인덱스를 나타낸다. 재탐지용 코사인 윈도우는 탐색 영역의 x , y 축 길이로 각각 구한 두 벡터 C_x, C_y^T 로 계산된다.

기존의 코사인 윈도우와 제안된 코사인 윈도우를 Fig. 3에서 비교하였다. 상단은 각 코사인 윈도우를 적용한 패치인데, 기존 방법인 좌측 영상에서는 표적인 버스가 거의 보이지 않을 정도로 가장자리의 정보가 손실된 것을 확인할 수 있다. 반면 제안된 방법인 우측은 가장자리 효과를 감소시킬 수 있도록 가장자리 값이 서서히 작아지지만 탐색 영역 내 대부분의 정보를 보존하였다. 뿐만 아니라 가장자리에서 감소되는 정도가 확대 전 상관 필터에서 사용한 코사인 윈도우와 유사하여 표적 탐지에도 유리하다. Fig. 3 하단에서 실험 결과 응답 지도를 확인할 수 있다. 좌측에서는 눈에 띄는 응답이 없지만, 우측 영상의 표적 위치에서 높은 응답을 확인할 수 있다.

2. Re-detection

표적이 다른 물체에 가려져 소실되거나 다시 나타났을 때 추적기는 이를 판단하는 기준이 필요하다. [1]에서는 표적에 대한 추적의 실패를 판단하는 방법(Failure Detection)으로 PSR(Peak to Sidelobe Ratio)을 제안하였다.

$$PSR = \frac{r_{\max} - \mu_{sl}}{\sigma_{sl}} \quad (7)$$

r_{\max} 는 상관필터 적용 결과 계산된 응답 지도의 최대값, μ_{sl} 은 배경 영역(Sidelobe)에서 응답 평균, σ_{sl} 은 배경 영역에서 응답의 표준편차를 의미한다. 여기서 배경 영역이란 응답 지도에서 최대값, 즉 표적이라고 판단되는 픽셀 위치를 중심으로 11x11 픽셀을 제외하고 나머지 부분을 의미한다[1]. 하지만 고정된 크기의 표적 부분을 제외하게 되면 표적이 너무 크거나 작을 때 PSR 값의 수준이 달라질 우려가 있다. 본 논문에서는 이상적인 응답(y)의 분산을 표적의 크기와 비례하게 정의하였고, 그 결과로 도출되는 응답 지도에서 PSR을 계산할 때 또한 표적의 크기와 비례한 영역을 제외하였다.

정상 추적 중에는 매 프레임 응답 지도를 도출한 후, PSR을 계산하여 특정 값 이하가 되면 표적의 소실로 판단한다. 본 논문에서 목표로 하는 실제 장비의 경우, 표적의 가림을 판단한 순간부터 표적을 따라 움직이는 카메라의 시선이 최근 이동하던 방향으로 계속 이동하게 된다. 따라서 영상의 중심에 탐색 영역을 설정한다. 반면 구동이 없는 카메라로 찍은 데이터셋으로 시뮬레이션할 경우, 추적시 이동하던 방향으로 예측한 위치에 탐색 영역을 설정하고 재탐지를 시도할 수 있다. 재탐지를 위한 탐색 영역에서도 추적과 같은 방법으로 응답 지도를 구하고, PSR이 특정 값 이상이 되면 재탐지로 판단하게 된다.

표적이 소실되고 난 후 시간이 지날수록 표적이 다시 출현할 것이라고 예측한 위치가 부정확해 질 것이다. 예를 들어 표적이 신호등이나 고가도로 등 작은 물체에 가려졌을 때는 금방 비슷한 위치에서 다시 나타나기 때문에 작은 탐색 영역에서도 표적을 쉽게 찾을 수 있다. 반면 산이나 터널 등 큰 물체에 의해 가려졌을 경우 비교적 긴 시간 후에 다시 출현하게 되고, 그 때 표적의 위치가 예측하지 못한 방향으로 변화했을 가능성이 더 커지게 된다. 본 알고리즘은 앞에서 설명한 것과 같이 고정된 크기의 상관 필터를 사용하여 다양한 크기의 탐색 영역에서 상관도를 계산하는 것이 가능하다. 이를 활용하여 표적이 손실된 후에는

탐색 영역의 크기를 매 프레임 2%씩 증가시키며 재탐지를 시도한다. Fig. 1에 표적의 가림 시간에 따른 탐색 영역의 크기 변화가 나타나 있다. 이 방법은 특히 표적의 소실이 오래 지속되는 경우 재탐지의 성공 유무에서 기존의 알고리즘과 큰 차이를 보인다.

탐색 영역을 증가시킬 때 발생할 수 있는 문제점은 크게 두 가지가 있다. 첫 번째, 계산해야 하는 정보가 많아짐에 따라 처리 시간이 증가한다. 하지만 실제 추적기가 적용된 실시간 시스템에서 재탐지를 시도할 경우, 실시간성 보다는 재탐지의 성공 유무가 시스템의 실용성 측면에서 더 중요하다. 모든 프레임에서 표적의 재출현을 검사하여 재탐지 시점을 조금이라도 앞당기는 데에 초점을 맞추기 보다는, 탐색 영역을 확대해 처리 시간이 수 프레임 걸리더라도 표적을 다시 찾아내야 추적을 유지할 수 있기 때문이다. 때문에 탐색 영역의 크기는 최대 4배까지 증가시킨다.

두 번째 문제로, 주변에 표적과 유사한 오답(Distractor)이 존재할 때 응답이 표적보다 높게 나올 우려가 있는데, 탐색 영역이 증가할수록 오답이 포함될 가능성이 높아진다. 이를 해결하기 위해 재탐지를 위한 응답 지도에 가우시안 가중치(Gaussian weight)를 적용하여 최대값의 위치를 찾아내었다. 탐색 영역 내 비슷한 물체가 여러 개 존재할 때 예측한 위치의 중심에 가까운 물체가 선택될 확률을 높인 것이다. 이렇게 결정한 표적 후보의 위치에서 PSR 계산 시에는 보정되지 않은 응답 지도를 그대로 사용하여 가우시안 가중치가 재탐지 판단에 주는 영향을 제한하였다.

IV. Experiments

제안하는 알고리즘과 KCF[2]의 비교를 위한 실험 결과는 다음과 같다.

Table 1. Dataset List

Sequence name	Length [frame]	Target size [pixel]	Camera LOS
Truck Bridge	150	77x79	Stop
Car Bridge	200	25x24	Moving
Truck Mountain	800	36x28	Tracing
Bus Mountain	638	47x31	Tracing

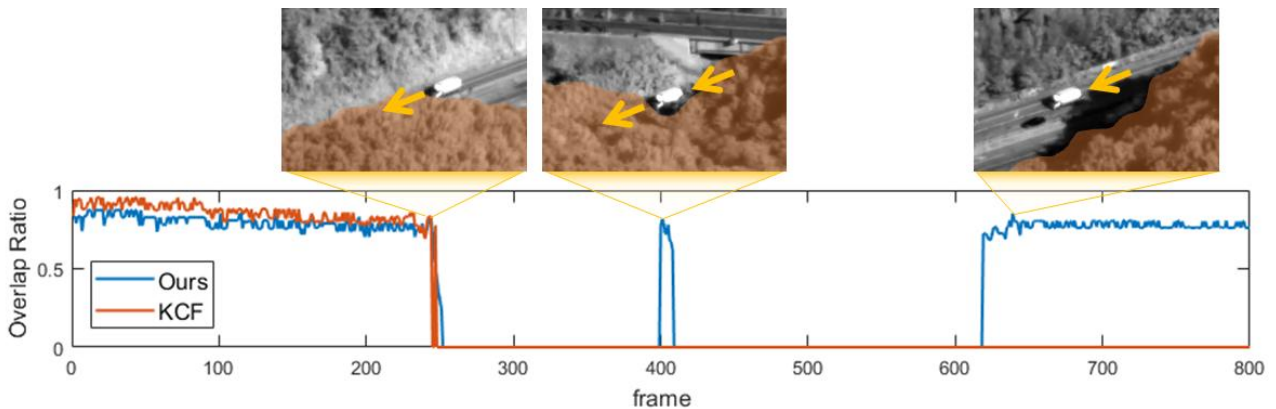


Fig. 4. Overlap ratio graph for a frame number. Our algorithm (blue line) succeed to re-detect the target twice while original KCF fails (orange line).

1. Experimental Details

본 알고리즘에서 학습 시 표적 주변을 포함하도록 표적의 크기보다 3배 큰 윈도우를 사용하였고, 표적의 소실 후에는 탐색 영역의 크기를 매 프레임 2%씩 증가시키며 재탐지를 시도하였다. 학습과 탐지에 사용된 특징 벡터는 HoG[7] (Orientation=9, cell size=1)와, 밝기 정보 (Grayscale), 총 10채널이다. 상관필터, 즉 α 와 \mathbf{x} 로 표현되는 학습 모델은 매 프레임 0.02의 학습율(Learning rate)로 갱신된다. 이상적인 응답(\mathbf{y})은 가우시안 분포로 정의하였는데, 여기서 표준편차(σ_y)는 표적의 크기에 0.1을 곱하여 사용하였다. 표적의 상태를 결정하는 기준으로는 앞장에서 설명한 PSR을 사용하였고, 이 때 배경 영역은 표적 위치를 중심으로부터 $3\sigma_y$ 크기를 제외한 나머지 부분으로 정의하였다. PSR이 12.0 이하이면 표적 소실, 9.5 이상이면 표적의 재탐지로 판단하였다.

모든 실험은 Intel i7 3.2GHz CPU와 16GM RAM 환경에서 수행되었다. 특징 벡터를 계산하는 부분은 C언어, 그 외 부분은 MATLAB에서 구현한 코드가 사용되었다.

2. Datasets

실험은 항공에서 촬영한 4개의 데이터셋으로 진행하였다. 재탐지 성능을 확인하기 위해 모두 표적의 소실과 재출현이 한 번 이상 일어나는 영상이고, 정량적인 결과 분석을 위한 정답지(Ground truth)는 바운딩 박스(Bounding box)로 정의되어 있다. 트럭과 승용차가 다리 밑을 지나갈 때 짧은 시간 가리는 영상 2개, 트럭과 버스가 진행 방향에 있는 산에 비교적 길게 가렸다가 나타나는 영상 2개로 구성되어 있다. 데이터셋에 대한 자세한 내용은 Table 1.에서 확인할 수 있다. 표에서 'Camera LOS(Line of Sight)'란 카메라의 시선을 의미한다. 앞 절에서 설명한 바와 표적의 소실 후 카메라의 시선이 이전

이동 방향으로 계속 이동하며 촬영한 경우 'Tracing'이라고 표현하였다. 'Stop'은 카메라의 시선이 고정된 경우, 'Moving'은 표적과 무관하게 시선이 등속 이동하는 경우이다. 이때 간략하게 'Tracing'을 시뮬레이션하기 위해 표적 손실 시점에서 최근 30 프레임 동안 표적의 이동 방향을 평균 내어 탐색 영역의 위치에 적용하였다.

3. Evaluation

실험 결과 다리에 짧은 시간 가리는 영상의 경우 표적의 위치가 많이 변하지 않아 예상할 수 있듯 기존의 상관필터 알고리즘과 제안한 알고리즘 모두 표적의 재탐지에 성공하였다. 하지만 산에 가리는 영상의 경우 표적이 다시 나타났을 때 기존 상관필터 크기의 영역에서 벗어나 확대한 탐색 영역에서만 재탐지가 가능하였다.

앞에서 확인한 Fig. 1은 'Bus Mountain' 데이터셋의 정성적인 결과이다. 'Truck Mountain'의 결과는 Fig. 4.에 나타나 있다. 정답지와 추적 결과로 나오는 사각형의 추적창을 비교하여 전체 영역 대비 겹치는 부분의 비율(Overlap ratio)을 그래프로 나타내었다. 파란색 그래프는 제안한 알고리즘, 주황색 그래프는 탐색 영역의 크기 변화가 없는 기존 KCF를 나타낸다. 영상에서 표적은 산 뒤로 움직이며 사라졌다가 다시 나타나고, 곧 다시 가렸다가 마지막에 다시 출현하며 끝나는 구성이다. 표적이 영상 내 존재하지 않는 구간에서는 정답지의 바운딩 박스도 존재하지 않고 겹치는 비율은 0으로 계산된다. 그래프에서 변화가 일어나는 부분은 그래프 상단에서 영상으로 확인할 수 있다.

모든 데이터셋 실험 결과는 정답지와 비교하여 정량적 결과를 정리하였다. Fig. 5에 각 데이터셋의 실험 결과를 그래프(Success plot)로 나타내었다. 모든 프레임에서 정답지와 추적 결과를 비교하여 겹치는 비율을 계산하고 기준 값(Overlap threshold) 이상일 때 성공으로 정의한다.

기준 값을 변화시키며 전체 프레임에서 성공한 프레임의 비율을 나타낸 것이다. 표적 소실 상태인 프레임은 합산에서 제외하였다. 파란색은 제안한 알고리즘, 주황색은 KCF, 노란색은 KCF에서 탐색 영역만 키운 알고리즘 결과를 나타낸다. 상관 필터의 학습 시에는 제로 패딩을 적용하지 않은 경우 학습된 필터의 특성이 달라 탐색 영역의 크기를 조절하는 순간부터 표적을 찾지 못하게 됨을 확인할 수 있다. (a)와 (b)에서는 짧은 가림 이후 KCF의 고정된 탐색 영역에서도 재탐지에 성공하여 제안한 알고리즘과 유사한 결과를 보이고 있다. 반면 (c)와 (d)에서는 제안된 알고리즘이 탐색 영역을 키워 오랜 시간 후 재출현 표적의 재탐지에 성공하면서 그래프의 양상이 달라졌다. KCF도 첫 번째 가림부터 동영상이 끝날 때 까지 표적 탐지에 실패하기 때문에 주황색과 노란색이 거의 같은 모양의 그래프를 그리고 있다. 본 실험을 통해 표적이 일정 시간 이상 소실되었다가 다시 나타나는 경우 제안한 알고리즘의 효과가 두드러짐을 확인할 수 있다.

V. Conclusions

본 논문에서는 탐색 영역의 크기를 조절 가능한 상관 필터 추적 방법을 제안하였다. 추적 시 상관필터에 제로 패딩을 적용하여 더 넓은 영역에서 재탐지가 가능해졌고, 가장자리 효과를 줄이면서 정보 손실을 최소화하는 코사인 윈도우를 사용하였다. 추적 실패와 표적의 재탐지를 판단하는 PSR은 표적의 크기에 적응적인 배경 영역(Sidelobe)을 정의하여 계산하였다. 표적의 가림과 재출현이 잦은 실제 상황에 적용 가능하도록 표적 소실 후 일정 비율씩 탐색 영역을 증가시켰고, 증가한 탐색 영역에서 오답(Distractor)의 영향을 줄이기 위해 중심과 가까울수록 높은 가중치를 적용하였다. 본 알고리즘은 일반적인 상관필터 추적기가 실패하기 쉬운 도전적인 데이터셋에서도 표적의 재탐지에 성공한 결과를 보였다. 제안된 알고리즘은 다양하게 응용될 수 있다. 항상 영상 전체에서 탐지를 시도할 필요는 없기 때문에 상황에 맞게 매 프레임 탐색 영역을 조절할 수 있고, 표적의 속도가 매우 빠르거나 프레임 갱신률(Frame rate)이 낮아 프레임 간 표적의 위치 차이가 클 때는 추적 중에도 탐색 영역을 증가시키며 진행이 가능하다.

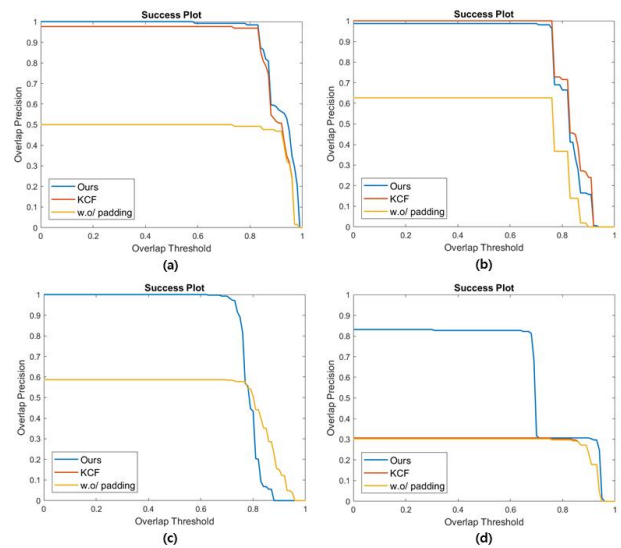


Fig. 5. Success plot for (a) Truck Bridge, (b) Car Bridge, (c) Truck Mountain, (d) Bus Mountain.

REFERENCES

- [1] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters.", IEEE Conference on Computer Vision and Pattern Recognition, 2010.
- [2] 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, 2015.
- [3] M. Danelljan, G. Hager, F. S. Khan, M. Felsberg, "Accurate scale estimation for robust visual tracking.", British Machine Vision Conference, 2014.
- [4] C. Ma, X. Yang, C. Zhang and M. Yang, "Long-term correlation tracking", IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [5] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, "Learning spatially regularized correlation filters for visual tracking.", IEEE International Conference on Computer Vision, 2015.
- [6] H. K. Galoogahi, T. Sim, and S. Lucey, "Correlation filters with limited boundaries.", IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection.", IEEE Conference on Computer Vision and Pattern Recognition, 2005.
- [8] N. Wang, W. Zhou and H. Li, "Reliable re-detection for long-term tracking.", IEEE Transactions on Circuits and Systems for Video Technology, 2019.

Authors



Kayoung Park received the B.S.(2013) degrees in Electronis Engineering from Kyungpook National University and M.S.(2015) degrees in Creative IT Engineering from POSTECH (Pohang University of Science and

Technology), KOREA. Kayoung Park is currently a researcher in Agency for Defense Development (ADD) since 2015. She is interested in visual tracking.