

Robust Visual Tracking using Search Area Estimation and Multi-channel Local Edge Pattern

Eun-Joon Kim*

Abstract

Recently, correlation filter based trackers have shown excellent tracking performance and computational efficiency. In order to enhance tracking performance in the correlation filter based tracker, search area which is image patch for finding target must include target. In this paper, two methods to discriminatively represent target in the search area are proposed. Firstly, search area location is estimated using pyramidal Lucas-Kanade algorithm. By estimating search area location before filtering, fast motion target can be included in the search area. Secondly, we investigate multi-channel Local Edge Pattern(LEP) which is insensitive to illumination and noise variation. Qualitative and quantitative experiments are performed with eight dataset, which includes ground truth. In comparison with method without search area estimation, our approach retain tracking for the fast motion target. Additionally, the proposed multi-channel LEP improves discriminative performance compare to existing features.

▶ Keyword: Visual tracking, Correlation filter, Search area estimation, Local edge pattern

I. Introduction

최근 영상 추적은 컴퓨터 비전 분야에서 가장 핵심적이면서 도전적인 연구 분야 중 하나이다.[1] 영상 추적은 연속되는 영상 프레임에서 표적의 궤적을 예측하는 것이다. 일반적인 영상 추적은 첫 번째 프레임에서 표적의 초기 위치만 알고 있는 상태에서 추적을 수행한다. 이러한 제한된 조건과 연속되는 프레임에서 나타나는 표적의 크기 및 양상변화, 부분적/완전 가림, 밝기 변화, 유사 객체의 존재, 그리고 급격한 위치변화와 같은 요인들과 더불어 안정적인 추적을 어렵게 하고 있다.[2]

많은 영상 추적 방법에서 표적을 분별하기 위해 표적의 모델을 구성한다.[3] 생성한 표적 모델은 새로운 프레임의 탐색 영역에서 표적의 위치를 탐색하고 예측하기 위해 사용한다. 표적의 급격한 변화나 카메라의 시선 변화로 인해 영상에서 표적이 빠르게 이동하면 표적의 양상이 변화하게 된다. 따라서 급격히 위치와 양상이 변하는 표적에 대해 추적을 유지함으로써 전반적인 추적 정확도를 높일 수 있다. 본 논문에서는 이동 변위

가 큰 표적에 대한 탐색영역을 예측하고, 변화하는 양상에 대해 분별력 있는 모델을 구성함으로써 전반적인 추적 정확도를 높일 수 있는 방법에 대해 기술한다.

매년 영상 추적 분야에서 다양한 알고리즘이 제안되고 있고, 경연을 통해 성능을 비교 검증하고 있다.[4][5] 최근 높은 추적 정확도와 실시간성으로 주목받는 상관필터 기반 추적 방식에 대해 활발한 연구가 진행되고 있다.[6][7] 상관필터 기반 추적은 표적을 배경이나 유사표적으로부터 분별할 수 있는 필터를 구성하여 표적의 위치를 예측하는 추적 방식이다.

기본적으로 상관필터 기반 추적에서 표적을 추적하기 위해서는 탐색영역 내부에 표적이 포함되어야 한다. 카메라 시선의 이동이나 표적의 빠른 이동 속도로 인해 영상에서 표적의 이동 변위가 크더라도 탐색영역 내부에 표적이 위치해야 상관필터를 적용해 표적위치를 예측할 수 있다. 또한 탐색영역 내부의 표적을 유사표적이나 배경으로부터 구분하기 위해서는 분별력 있는 모델을 바

*First Author: Eun-Joon Kim, Corresponding Author: Eun-Joon Kim
*Eun-Joon Kim (hseunjun@add.re.kr), Agency for Defense Development (ADD)
• Received: 2017. 05. 29, Revised: 2017. 06. 19, Accepted: 2017. 07. 23.

탕으로 상관필터를 구성해야 한다. 이를 위해 분별력 있는 표적의 특징으로부터 상관필터를 생성하고 업데이트해야 한다.

본 논문에서는 상관필터 기반 추적 방식에서 양상과 위치가 급격히 변화하는 표적을 추적하면서 전반적인 추적 정확도를 높이기 위해 두 가지 향상된 방법을 제안한다. 먼저 피라미드 Lucas-Kanade 알고리즘을 이용한 탐색영역 예측방법을 고안하여 급격히 위치가 변하는 표적이 탐색영역 내부에 포함될 수 있도록 한다. 또한 상관필터 기반 추적 방식에서 탐색영역 중심의 표적을 강조하고 배경의 영향을 줄이기 위해 코사인 윈도우를 사용한다. 하지만 추적 시 항상 표적을 탐색영역 중심에 위치시킬 수 없고, 표적이 탐색영역의 가장자리에 위치하면 표적 정보가 감쇄하여 상관필터의 분별력이 저하된다. 이에 상관필터로 표적의 위치를 예측하기 전에 탐색영역 예측 방법을 수행하여 표적이 탐색영역 중심에 위치하도록 하여 특정 정보가 코사인 윈도우에 의해 감쇄되지 않도록 하여 전반적인 추적 성능을 높인다.

또한 기존 모서리 성분을 이진화한 패턴인 Local Directional Pattern (LDP) 특징[8]을 상관필터 기반 추적 방식에 적용할 수 있도록 변형한 다중채널 Local Edge Pattern (LEP) 특징 추출 방법을 제안한다. 다중채널 LEP는 표적의 모서리 강도를 표현한 특징이다. 여덟 방향의 모서리를 구할 수 있는 Kirsch 윈도우를 적용해 얻은 모서리 강도 값을 획득하는 단순하면서 효율적인 특징추출 방식이다. 본 방식은 별도의 여덟 모서리 방향으로 분리해 다중채널로 구성함으로써 상관필터 기반 추적 시 적용해서 사용할 수 있다. 모서리 성분은 기존의 명암도나 경사도(Gradient) 기반 특징 대비 밝기 변화나 잡음이 섞인 영상에서도 분별력을 높일 수 있는 성분으로 이를 이용해 개선된 특징 추출 방법을 제안한다.

제안한 두 가지 방법의 효용성을 검증하기 위해 가시선(EO, Electric Optics) 뿐만 아니라 열화상(IR, InfraRed) 영상이 포함된 여덟 개의 실험 영상으로 정성적, 정량적 시험을 수행한다. 제안한 탐색영역 예측 방법을 적용하여 급격히 위치가 변하는 표적이 탐색영역 내부에 위치하도록 하여 표적이 지속적으로 탐색되어 위치를 예측할 수 있음을 보였다. 또한 다중채널 LEP를 추출함으로써 여덟 개의 실험 영상에서 기존의 명암도나 경사도 기반 특징 대비 높은 분별력을 가진 추적 성능을 보임을 검증하였다.

본 논문은 다음과 같이 구성된다. 2장에서는 기존 연구에 대해 기술하고, 3장에서는 본 논문의 기본 추적 방식인 다중채널상관필터에 대해 설명한다. 4장에서 급격히 변하는 표적을 추적하기 위해 제안하는 알고리즘인 탐색영역 예측과 다중채널 LEP에 대해 설명하고, 5장에서는 다양한 실험을 통해 제안하는 방법을 검증한다. 마지막으로 6장에서 본 논문의 결론을 도출한다.

II. Related Works

최근 영상 추적 시 상관필터 기반 추적 방식이 빈번히 적용되고 있고, 여러 논문에서 높은 추적 성능을 보고하고 있다.[5] 본 방식은 상관필터 구성 및 위치 예측 시 주파수 도메인에서 합성곱 연산을 하여 효율적인 계산을 통해 실시간 작동을 하면서 주요 벤치마크 영상에 대해 높은 추적 성능을 보여주고 있는 것이 특징이다.[7] 본 장에서는 상관필터 기반 추적 방법 중 빠르게 이동하는 표적을 추적하면서 전반적인 추적 성능을 개선한 관련 연구들을 기술한다.

카메라 시선이 움직이거나 표적이 빠르게 이동하는 경우 영상에서 표적의 이동 변위가 크게 나타난다. 상관필터 기반 추적에서 표적의 위치를 예측하기 위해서 표적이 탐색영역 내부에 포함되어야 한다. 영상에서 이동 변위가 큰 표적을 탐색영역에 지속적으로 포함시키기 위한 기본적인 방법은 표적의 탐색영역을 확장하는 것이다.[3] 하지만 탐색영역을 확장하게 되면 표적 탐색 영역이 증가하게 되므로 계산량이 증가한다. 계산량 증가는 실시간 구현을 요구하는 추적 알고리즘 개발을 어렵게 하는 요인이다. 따라서 영상에서 빠르게 움직이는 표적의 움직임을 예측해서 현재 표적의 위치에 맞게 탐색영역을 설정하는 것이 필요하다.

또한 표적 모델의 분별력을 높여 탐색영역 내부에 포함된 표적의 위치 예측 정확도를 높이기 위한 연구도 수행되고 있다. 초기 방식은 표적의 명암도 특징을 활용한다.[9] 표적의 명암도와 기대응답 간 픽셀 오차가 최소화 되는 필터를 생성하여 표적의 위치를 예측한다. 초기 부족한 표적 샘플의 한계를 극복하기 위해 상관필터 생성 시 표적의 회전, 평행이동 및 닳음변환을 한 다수의 샘플을 이용하여 표적의 분별력을 높이고 있다. 최근 연구에서는 명암도 단일 특징뿐만 아니라 경사도(Gradient) 기반의 다중 차원의 특징을 이용하고 있다. 경사도의 방향에 따라 특징을 다중 차원으로 확장하고, 각 차원에 맞게 상관필터를 구성하여 표적의 특징을 추출한다. 경사도 기반 특징으로 HOG(Histogram of Oriented Gradient)를 다중차원으로 분할하여 사용한 경우가 대표적이다. [6][7] 명암도와 경사도 기반 특징은 표적의 양상과 변화량을 표현하지만 밝기변화와 잡음에 민감하다. 따라서 밝기변화와 잡음의 영향을 줄이고 배경으로부터 표적의 분별력을 높일 수 있는 특징을 통해 모델을 구성하고 필터를 생성해야 한다.

본 논문에서는 상관필터 기반 추적에서 빠르게 움직이는 표적의 탐색영역을 예측하고, 표적과 배경의 분별력을 높여 전반적인 추적 정확도를 개선하는 방법을 제안한다. 제안하는 방법의 효용성을 검증하기 위해 탐색영역 예측 방법을 통해 빠르게 이동하는 표적을 지속적으로 추적하고, 탐색영역 확장 방법과 비교해 계산 효율적임을 보인다. 또한 제안하는 모서리 기반 특징인 다중채널 LEP를 적용하여 기존 명암도나 경사도 기반 특징 대비 표적의 분별력을 높여 전반적인 추적 정확도가 개선되었음을 증명한다. 총 여덟 개의 ground truth를 포함한 실험영상을 통해 본 방법의 효용성을 검증하여 결과를 제시하였다.

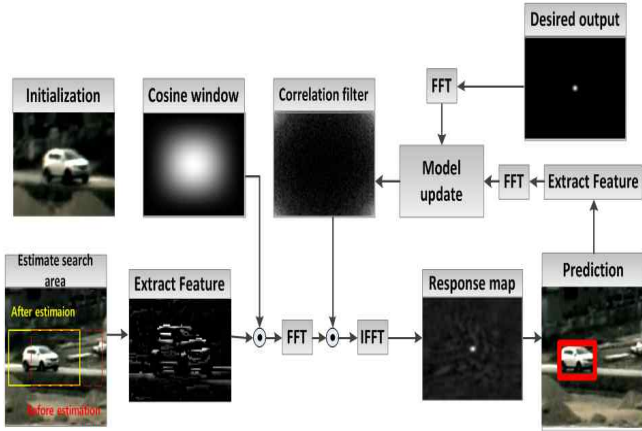


Fig. 1. Overall process of proposed tracking algorithm

III. Multichannel Correlation Filter Tracking

본 논문에서 추적은 상관필터에 기반을 두고 있다. 상관필터 기반 추적은 첫 프레임에서 표적의 양상이 주어졌을 때, 새로운 프레임에서 표적의 위치를 예측할 수 있는 최적의 상관필터를 도출하는 것을 목적으로 한다.[9] 상관필터 추적은 다중 채널의 특징정보를 활용한 다중채널상관필터로 확장되어 사용될 수 있다.[6][7]

다중채널상관필터 추적은 단일채널상관필터를 다중으로 확장한 것이다. 상관필터를 d 차원의 채널로 확장하면 추적의 대상이 되는 표적이 포함된 패치 f 의 d 채널 특징정보는 f^l 로 표현되며 $l \in \{1, \dots, d\}$ 이다. h^l 로 구성된 다중채널상관필터 h 를 얻기 위한 목적함수는 식 (1)의 비용함수 ϵ 을 최소화 하는 식과 같다.

$$\epsilon = \left\| g - \sum_{l=1}^d h^l \cdot f^l \right\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2, \quad (1)$$

식 (1)의 g 는 표적의 위치에서 가장 높은 값을 가지는 가우시안 함수이고, \cdot 는 순환 컨볼루션 연산을 나타낸다. λ 는 정규화 항의 영향을 조절하는 파라미터이다.

식 (1)은 선형 최소제곱법 문제이며 본 식을 주파수 도메인으로 변환하여 풀면 식 (2)와 같이 단순화된 식을 도출할 수 있다.

$$H^l = \frac{\overline{GF^l}}{\sum_{k=1}^d \overline{F^k} F^k + \lambda}, \quad l = 1, \dots, d, \quad (2)$$

H, G, F 는 각각 h, g, f 의 이산 푸리에 변환(DFT)을 의미하고, \overline{G} 는 G 의 켈레복소수를 나타낸다. 식 (2)는 단일 표적 샘플 f 에 대해 도출한 최적의 다중채널상관필터 h 를 이산 푸리에 변환한 식이다.

연속되는 프레임에서 표적과 배경의 분별력을 지속적으로

유지하기 위해서는 새로운 표적 패치에 대한 다중채널상관필터의 업데이트가 수행되어야 한다. t 번째 프레임의 새로운 표적 패치 f_t 에 대한 필터 H_t^l 의 업데이트는 식 (3)과 같이 분자 A_t^l 와 분모 B_t^l 에 대해 별도로 이루어진다. 분모, 분자 각각 새로운 표적패치의 업데이트는 기존 $t-1$ 번째 필터에 대해 새로운 표적패치를 학습률 η 을 가중합 하여 얻는다.

$$A_t^l = (1-\eta)A_{t-1}^l + \eta \overline{GF_t^l}, \quad l = 1, \dots, d$$

$$B_t^l = (1-\eta)B_{t-1}^l + \eta \sum_{k=1}^d \overline{F_t^k} F_t^k. \quad (3)$$

t 번째 프레임에서 필터를 적용하기 위해, 프레임에서 표적이 존재할 것으로 예상되는 지점에 표적 패치 z^t 가 추출된다. 일반적으로 z^t 는 표적이 중심영역에 존하며, 코사인 윈도우에 의해 패치의 중심영역에 가중치가 적용되며, 다중채널상관필터가 적용되어 표적의 위치를 예측할 수 있는 탐색영역을 의미한다. 새로운 표적 패치 z^t 에서 다중채널상관필터를 적용하는 방법은 식 (4)와 같다.

$$Y_t = \frac{\sum_{l=1}^d \overline{A_{t-1}^l} Z_t^l}{B_{t-1} + \lambda}, \quad (4)$$

Z_t^l 와 Y_t 는 각각 표적이 포함된 패치 z_t 와 상관도 y_t 의 이산푸리에변환을 나타낸다. 상관도 y_t 가 가장 높게 나타나는 픽셀 위치가 실제 예측된 표적의 위치가 된다.

IV. The Proposed Scheme

제안하는 추적 알고리즘의 과정은 그림 1과 같다. 본 방식은 초기화(Initialization), 탐색영역 예측(Estimate search area), 특징추출 및 표적 위치 예측(Extract feature & Prediction) 그리고 상관필터 업데이트(Model update)로 구성된다.

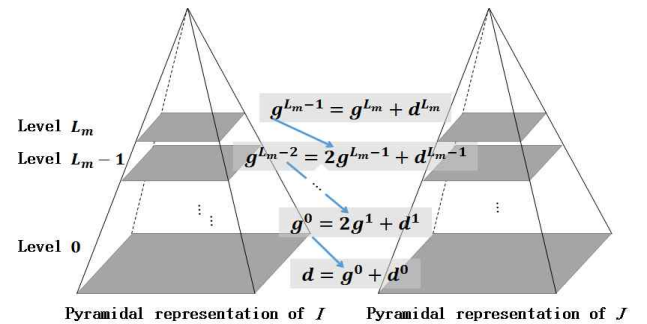


Fig. 2. Pyramidal representation of Lucas-Kanade algorithm

현재 표적 위치를 중심으로 탐색영역을 설정하는 기존 방식에서는 표적의 이동 변위가 탐색영역의 너비 보다 클 경우 탐색영역을 벗어나 표적을 놓치게 된다. 또한 탐색영역을 확장하는 경우 탐색영역의 증가로 계산량이 증가하게 된다. 따라서 기존 상관필터 기반 추적 과정에 탐색영역을 예측 과정을 추가하여 이동 변위가 큰 표적에 대해서도 추적을 유지할 수 있도록 하였다. 탐색영역 예측은 피라미드 Lucas-Kanade 알고리즘을 사용해 예측한 표적의 광학흐름 변위만큼 탐색영역을 평행이동한다. 탐색영역 예측을 통해서 표적 탐색을 위한 계산량의 증가 없이 이동 변위가 큰 표적을 탐색영역에 포함시킬 수 있다.

또한 탐색영역에 포함된 표적을 배경이나 유사표적으로부터 분별력을 높이기 위해 모서리 기반 특징인 다중채널 LEP 특징 정보 추출 방법을 특징 추출 시 적용한다. 모서리 기반 특징은 기존 명암도나 경사도 특징 대비 밝기 변화나 잡음의 영향을 줄일 수 있다. 또한 모서리 강도를 방향에 따라 다중채널로 분리하여 구성함으로써 다중채널상관필터 기반 추적 방식에 적용할 수 있도록 한다.

1. Pyramidal Lucas-Kanade based search area estimation

다중채널상관필터 추적에서 표적 패치는 새로운 프레임에서 필터를 적용하여 표적의 중심위치를 찾을 수 있는 영역을 의미한다. 영상에서 프레임 간 표적의 위치변화가 작을 때 표적 패치는 일반적으로 전 프레임의 표적 중심위치를 기준으로 획득한다. 반면 표적의 위치가 급격히 변할 때는 프레임 간 표적의 거리 차이가 크기 때문에 표적 패치의 크기를 크게 설정하여 표적 패치 내부에 표적이 포함되도록 해야 한다. 하지만 표적 패치의 크기가 증가하면 필터의 크기가 커지기 때문에 계산량이 증가하고, 유사 표적이 포함될 확률이 증가하기 때문에 추적 성능에도 영향을 줄 수 있다.

본 논문에서는 표적이 포함되어 있고, 필터를 적용하여 표적 위치를 예측하기 위한 영역인 표적 패치를 탐색영역으로 정의한다. 탐색영역의 크기를 적절하게 유지하면서 표적이 탐색영역 내부에 위치하도록 하는 탐색영역 예측방법은 다중채널상관필터 추적 시 연산량의 큰 증가 없이 급격히 위치가 변하는 표적을 추적할 수 있도록 하는 방법이다. 이에 본 절에는 다중채널상관필터 추적 전에 수행되는 Lucas-Kanade 기반 탐색영역 예측방법을 소개한다.

Lucas-Kanade 기반 탐색영역 예측방법은 Lucas-Kanade 알고리즘을 이용하여 이 전 프레임의 표적 위치를 기준으로 현재 프레임의 필터 적용 영역인 탐색영역을 예측하는 방법이다. 탐색영역 예측을 통해 급격히 위치가 변하는 표적에 대해서도 탐색영역의 크기를 증가 없이 표적을 지속적으로 추적할 수 있다.

Table 1. Pyramidal Lucas-Kanade Algorithm

$\mathbf{g}^{L_m} = [00]^T$ initialization of guess for $L = L_m$ down to 0 with step of -1 $\bar{\mathbf{v}} = [00]^T$ initialization of Lucas-Kanade for $k = 1$ to K with step of 1 Obtain optical flow $\bar{\eta}^{-k}$ $\bar{\mathbf{v}}^{-k} = \bar{\mathbf{v}}^{-k-1} + \bar{\eta}^{-k}$ guess for optical flow end of for-loop on k $\mathbf{d}^L = \bar{\mathbf{v}}^{-K}$ optical flow at level L $\mathbf{g}^{L-1} = 2(\mathbf{g}^L + \mathbf{d}^L)$ guess for level L-1 end of for-loop on L $\mathbf{d} = \mathbf{g}^0 + \mathbf{d}^0$ final optical flow at level 0

Lucas-Kanade 알고리즘은 식 (5)와 같이 연속된 프레임 I, J 간 $u_x \leq x \leq w_x$ 와 $u_y \leq y \leq w_y$ 범위의 영역에서 명암도의 차이 ϵ 를 최소화하는 광학 흐름인 x 축 변위 d_x 와 y 축 변위 d_y 를 예측하는 알고리즘이다.[10]

$$\epsilon(d_x, d_y) = \sum_{x=u_x}^{w_x} \sum_{y=u_y}^{w_y} (I(x, y) - J(x + d_x, y + d_y))^2. \quad (5)$$

Lucas-Kanade 알고리즘에서 급격하고 큰 움직임에 대한 광학 흐름의 예측 정확도를 높이기 위해 [11]과 같이 영상 프레임의 가우시안 피라미드를 활용한다. 가우시안 피라미드를 활용한 변위예측 방법은 그림 2와 같다. 인접한 프레임 영상 I, J에 대해 각각 $L_m + 1$ 개 레벨의 피라미드 영상을 생성한다. 레벨 0 영상이 프레임 원 영상이고, 레벨 1에서 레벨 m 까지 프레임 영상에서 m 번 가우시안 필터링 후 다운샘플링을 한 영상이다. 레벨 L_m 에서 추정된 거리 \mathbf{g}^{L_m} 에 K 번 반복 연산을 통해 획득한 광학 흐름 \mathbf{d}^{L_m} 을 더하여 레벨 L_m 영상에서의 거리를 추정한다. 레벨 L_m 영상에서 추정된 거리는 레벨 $L_m - 1$ 영상에서 초기 추정 거리로 사용되어 새로운 광학 흐름을 구한다. 이러한 과정을 $L_m + 1$ 번 반복하면 레벨 0의 프레임 원 영상에서의 거리를 예측할 수 있다.

따라서 일반적인 Lucas-Kanade 알고리즘의 목적함수인 식 (5)에 피라미드를 적용한 피라미드 Lucas-Kanade 알고리즘은 식 (6)과 같이 표현할 수 있다.

$$\epsilon^L(d_x^L, d_y^L) = \sum_{x=u_x}^{w_x} \sum_{y=u_y}^{w_y} (I^L(x, y) - J^L(x + d_x^L + g_x^L, y + d_y^L + g_y^L))^2, \quad (6)$$

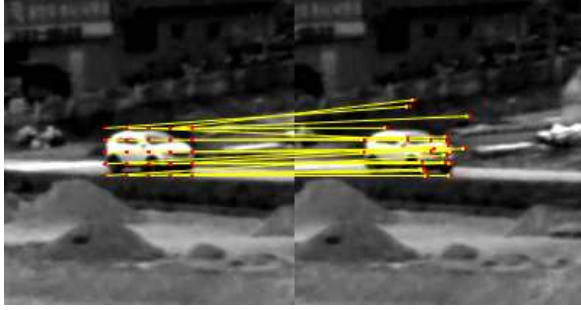


Fig. 3. The result of 25 points optical flow using Pyramidal Lucas-Kanade algorithm

$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	m_3	m_2	m_1
M_0	M_1	M_2	M_3	m_4	X	m_0
$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$	m_5	m_6	m_7
M_4	M_5	M_6	M_7			

Fig. 4. Eight directional Kirsch edge masks and edge response

L 은 가우시안 피라미드 레벨을 의미하고 범위는 $L = \{1, \dots, L_m\}$ 이다. $\mathbf{g}^L = (g_x^L, g_y^L)$ 는 초기 추정 거리를 의미하고 각각 x 와 y 방향의 거리를 의미한다. 지금까지 설명한 알고리즘은 표 1과 같이 정리할 수 있다.

탐색영역을 이 전 프레임의 표적위치를 기준으로 이동변위를 예측하기 위해 그림 3과 같이 추적 게이트 내부에 일정한 간격으로 배열된 25개의 점을 중심으로 하는 패치를 추출해 피라미드 Lucas-Kanade 알고리즘을 적용하면 다음과 같이 총 25개의 변위 $\mathbf{d}_{total} = [\mathbf{d}_i]_{i=1}^{25}$ 를 획득할 수 있다. 획득한 25개 변위 중 식 (7)과 같이 중간 값을 가진 변위를 이 전 프레임의 표적의 중심 위치 기준으로 이동해야 할 최종 변위 \mathbf{d}_{sa} 로 설정한다.

$$\mathbf{d}_{sa} = \mathbf{d}_i \quad i^* = \text{med}[\mathbf{d}_i]_{i=1}^{25}, \quad (7)$$

$\text{med}[\cdot]$ 는 중간 값의 인덱스를 찾는 연산자를 의미한다.

2. 다중채널 지역적 모서리 패턴 특징 추출

다중채널 지역적 모서리 패턴 (LEP)은 영상의 지역적 공간 구조를 표현하는 특징정보이다. LEP는 영상의 각 픽셀 위치에서 총 여덟 방향의 모서리의 상대적 강도를 측정한다. 모서리 강도는 상대적으로 영상 밝기나 노이즈에 덜 민감하다. 따라서 LEP 특징정보는 영상의 지역적 형태를 잘 표현할 수 있을 뿐만 아니라 커브, 코너와 같은 형태의 정보도 담을 수 있다.[8]

일반적으로 LEP 특징정보는 그림 4와 같이 여덟 방향의 모서리 정보를 얻을 수 있는 Kirsch 모서리 마스크 M_l 를 이용한다

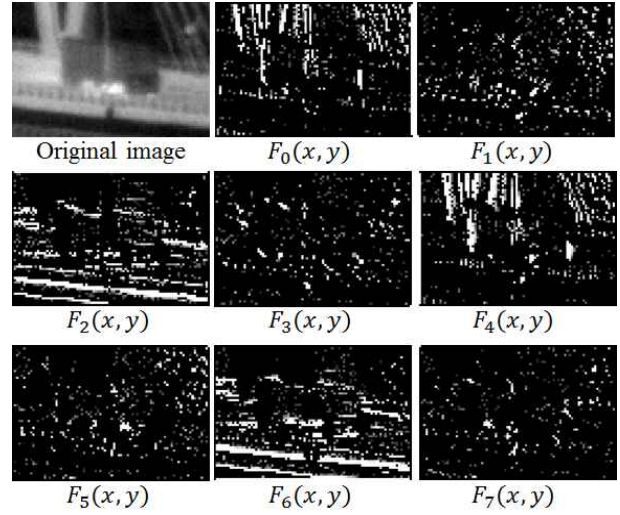


Fig. 5. An example of multi-channel LEP

영상의 각 픽셀 위치에서 여덟 방향의 모서리 강도를 측정하기 위해서 식 8과 같이 각 픽셀 위치에서 마스크를 사용해 각 방향마다 모서리 강도 $m_l(x, y)$ 를 얻는다.

$$m_l(x, y) = \sum_{y'=-1}^1 \sum_{x'=-1}^1 I(x+x', y+y') M_l(x', y'), \quad (8)$$

$$l = \{0, \dots, 7\},$$

$I(x, y)$ 는 표적 패치를 나타내고, l 은 여덟 방향을 의미한다. 그림 4와 같이 중간 픽셀 위치를 기준으로 총 여덟 방향의 모서리 강도 m_l 를 표현할 수 있다.

본 논문에서는 획득한 LEP 특징정보를 다중채널로 확장하여 다중채널상관필터 추적에 적용한다. 다중채널상관필터에 적절한 다중채널 LEP는 식 (9)와 같이 Kirsch 마스크를 사용해 얻은 모서리 강도가 최대인 방향을 선택한다.

$$l^* = \max_l [m_l]_{l=0}^7. \quad (9)$$

각 픽셀 위치에서 모서리 강도가 최대로 선택된 방향에 대해 식 (10)과 같이 모서리 강도 m_l 을 다른 방향에 대해서는 0을 배정한다.

$$F_l(x, y) = \begin{cases} m_l, & \text{if } l = l^* \\ 0, & \text{else } l \neq l^*. \end{cases} \quad (10)$$

$F_l(x, y)$ 는 다중채널 LEP의 l 번째 방향 채널의 특징정보를 의미한다.

임의의 표적 패치에 대해 위에서 설명한 다중채널 LEP 특징정보를 획득한 예시는 그림 5와 같다. 원 표적 패치는 다리를 건너는 버스 표적을 묘사하고 있다. 원 표적에서 표현된 버스 표적과 다리의 형태가 각 채널 별 특징정보에 표현된 것을 확인할 수 있다.

V. Experiments

1. Experimental setup

Table 2. Experimental Parameters

Parameter		Value
Learning rate(λ)		$1e-2$
Padding		1
Gaussian sigma		0.03125
HOG	Orientation	9
	Cell size	1
	Block size	2

Table 3. Experimental dataset list

Dataset name	Target size	# of frame
Tower IR Black Bus1	58×24	400
Tower IR Black Bus2	43×30	550
Hill EO Cart	59×57	1390
Tower IR Black Truck1	27×15	290
Tower IR Black Truck2	22×15	200
Hill IR Cart	52×57	800
Building EO Truck	67×63	166
Building EO Car	127×63	385

실험은 기본적으로 매트랩으로 구현하였고, 일부 함수에 한해 C언어로 구현하였다. 모든 실험은 Intel i7 3.4GHz CPU의 4GB RAM 환경에서 수행하였다. 모든 실험은 동일한 파라미터를 적용해 실험하였고, 파라미터 내용은 표 2와 같다. 본 실험에서 사용한 영상은 그림 6과 같이 총 여덟 개의 영상으로 구성된다. 실험 영상 중 세 개는 가시선 영상(EO)이고, 나머지 다섯 개는 열화상 영상(IR)으로 구성되어 있다. 모든 영상은 ground truth를 포함하고 있고, 영상의 세부적인 내용은 표 3과 같다.

본 논문에서 제안한 방법인 탐색영역 예측방법과 다중채널 LEP 특징 추출방법을 각각 검증한다. 탐색영역 예측방법의 경우 급격히 위치가 변하는 표적을 추적하기 위해 전 프레임 표적의 중심위치를 기준으로 탐색영역을 설정하는 방법과 탐색영역을 예측하는 방법을 비교해 비교실험을 한다. 다중채널 LEP 특징 추출방법은 기존의 명암도 특징과 HOG 특징을 비교해 성능적 우위를 비교한다.

2. Comparison of the search area estimation

탐색영역 예측방법은 먼저 정성적인 방법으로 비교 검증한다. 급격히 위치가 변하는 표적에 대해 전 프레임의 표적위치를 중심으로 탐색영역을 설정하는 기존 방법과 제안하는 탐색영역 예측방법으로 설정하여 표적을 추적하는 방법을 비교하였다. 실험에 사용한 여덟 개의 실험 영상 중에 Building EO Truck과 Building EO Car 영상에서 급격히 위치가 변하는 표적을 포함하고 있다. 표 2에 padding이 1이기 때문에 탐색영역은 추적 게이트 너비의 두 배 길이로 설정한다.

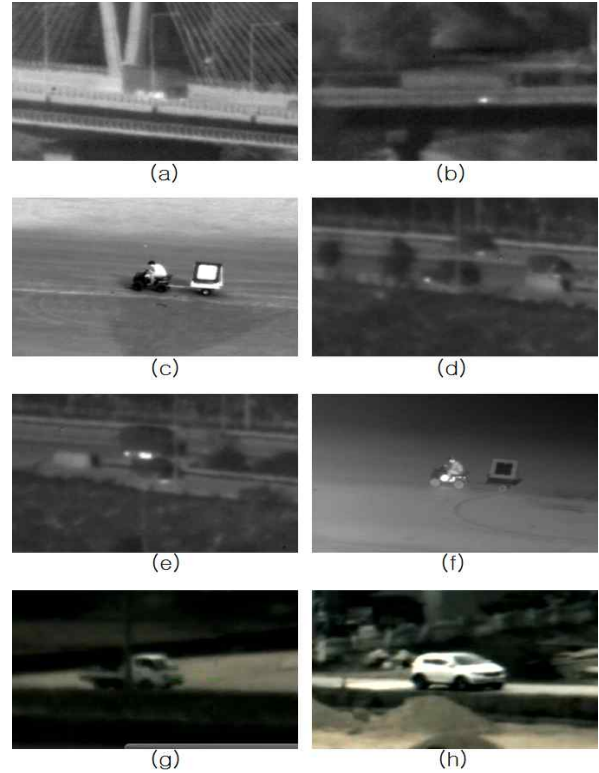


Fig. 6. Experimental dataset (a) Tower IR Black Bus1, (b) Tower IR Black Bus2, (c) Hill EO Cart, (d) Tower IR Black Truck1, (e) Tower IR Black Truck2, (f) Hill IR Cart, (g) Building EO Truck, (h) Building EO Car

그림 7은 비교 실험 결과이다. 먼저 Building EO Truck 영상에서 전 프레임 표적 중심위치 기준으로 탐색영역을 설정한 경우 146번 프레임까지 표적이 탐색영역 내부에 위치하기 때문에 표적을 안정적으로 추적한다. 하지만 147번째 프레임에서 표적의 위치가 급격히 변하면서 초점이 흐려지고, 탐색영역에 표적의 일부만 포함되면서 표적을 놓치게 된다. 제안하는 탐색영역 예측방법을 사용하면 147번째 프레임에서도 표적이 탐색영역 내부에 모두 포함되어 표적을 정확하게 추적한 것을 확인할 수 있다. 이후 프레임에서도 표적이 지속적으로 탐색영역 내부에 위치하고 안정적으로 추적을 유지한다.

Building EO Car영상도 동일한 양상을 보여준다. 전 프레임 표적 중심위치 기준으로 탐색영역 설정 시 279번째 프레임까지 표적의 대부분 영역이 탐색영역에 포함되기 때문에 추적 상태를 유지한다. 하지만 280번째 프레임에서 표적 위치가 급격히 변하면서 표적의 일부만 탐색영역에 포함되고 결국 표적을 놓치게 된다. 반면 제안하는 탐색영역 예측방법을 사용하면 279번째 프레임까지 탐색영역 내부에 표적의 모든 영역을 포함하고 있고, 280번째 프레임에서도 대부분의 표적 영역을 탐색영역 내부에 포함하여 추적 상태를 유지한다.

또한 탐색영역을 확장하는 기존 방법과 예측하는 제안하는 방법의 계산효율성을 비교하기 위해 한 프레임 소요 시간을 측정하였다. 탐색영역 확장의 경우 너비와 높이를 두 배 확장했을 경우에 대해 시간을 측정하였고, 여덟 개의 영상에 대해 한 프

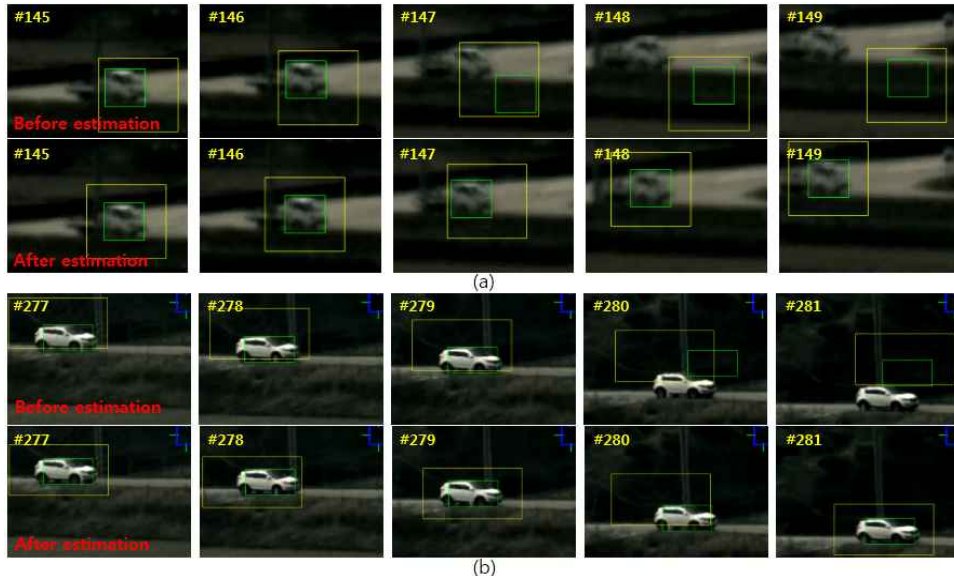


Fig. 7. Comparison of before and after the proposed search area(yellow line) estimation and tracking gate(green line). (a) Building EO Truck, (b) Building EO Car.

레이프 소요시간을 평균하였다. 한 프레임 소요시간 측정결과 탐색영역을 확장한 경우 49.2ms의 시간이 소요되었고, 탐색영역을 예측한 경우 35.1ms의 시간이 소요됨을 확인할 수 있었다.

실험을 통해 급격히 위치가 변하는 표적을 지속적으로 추적하기 위해 제안하는 탐색영역 예측 방법을 검증하였다. 탐색영역 예측을 통해 이동 변위 차이가 큰 표적을 지속적으로 탐색영역 내부에 포함시킬 수 있음을 보였고, 여덟 개의 영상에 대해 탐색영역을 확장하는 기존 방법 대비 14.1ms의 시간을 줄이는 결과를 확인할 수 있다. 따라서 빠르게 이동하는 표적을 추적하기 위해 탐색영역 예측방법이 필요함을 확인할 수 있었다.

3. Comparison of the proposed multi-channel LEP

제안하는 다중채널 LEP 특징정보 추출 방법을 검증하기 위해 기존의 명암도와 HOG 특징정보와 비교 실험 한다. 전반적인 추적성능 검증을 위해 precision plot을 사용하여 성능을 확인한다. Precision plot은 추적 게이트와 ground truth간 거리에 따른 추적 성능의 정확도를 확인할 수 있다. 추적 게이트와 ground truth간 거리가 가까운 상태에서 높은 거리 정확도를 보여줄 때 좋은 성능을 나타낸다.

그림 8은 여덟 개의 실험 영상에 대한 precision plot 결과이다. Tower IR Black Bus1, Tower IR Black Bus2, Hill IR Cart 영상에 대해서 15픽셀 위치 오차에서 명암도, HOG, 그리고 제안하는 다중채널 LEP 모두 표적을 완벽하게 추적하는 것을 볼 수 있다. 반면 급격한 표적의 위치변화가 있는 Building EO Truck, Building EO Car를 포함해 모든 영상에서 제안하는 특징정보가 좋은 분별 성능을 나타내고 있음을 확인할 수 있다.

제안하는 특징정보는 표적의 다양한 방향의 모서리 성분을 표현하고 있다. 이러한 모서리 성분은 영상의 밝기나 노이즈에 강인하고, 표적의 구조를 표현할 수 있다. 이와같은 특징은 명암도나 HOG 특징 대비 영상의 밝기나 노이즈에 덜 민감하기 때문에 배경대비 표적에 대한 분별력을 가질 수 있고, 전반적으로 높은 성능을 얻을 수 있다.

VI. Conclusions

본 논문에서는 다중채널상관필터 추적 방식을 개선하기 위한 방법을 살펴보았다. 이를 위해 급격히 위치가 변하는 표적에 대해서도 지속적인 추적 유지도를 높이기 위해 탐색영역 예측 방법을 제안하였고, 다중채널상관필터의 표적과 배경 간 분별력을 높이기 위한 다중채널 LEP를 제안하였다. 탐색영역 예측 방법은 피라미드 Lucas-Kanade 방법을 이용해 표적의 광학 흐름을 예측하고 이를 탐색영역 변위로 하여 급격히 위치가 변하는 표적에 대해서 탐색영역 크기의 증가 없이 추적할 수 있도록 하였다. 또한 다중채널 LEP는 Kirsch 마스크를 이용해 여덟 방향의 모서리 강도를 측정하고, 모서리 강도가 가장 큰 방향의 값을 특징 값으로 획득하는 방법이다. 본 논문에서는 모서리 방향에 따라 채널을 나눠서 다중채널 LEP를 제안하였고, 다중채널상관필터 추적에 적용하여 추적 성능을 측정하였다.

총 여덟 개의 ground truth를 포함한 영상을 사용해 시험을 수행하여 각 방법의 효율성을 검증하였다. 급격히 변위가 변하는 표적의 이동을 포함한 두 개의 영상에서 제안하는 탐색영역 예측을 했을 때 표적을 지속적으로 추적하고, 반면 예측을 하지 않았을 때 추적을 실패함을 보였다. 또한 탐색영역의 너비와 높이를 두 배 확장 했을 때 대비 예측했을 때, 28.7%의 시간이 줄어들어 본 방법이 시간효율적임을 증명했다. 더불어 모서리

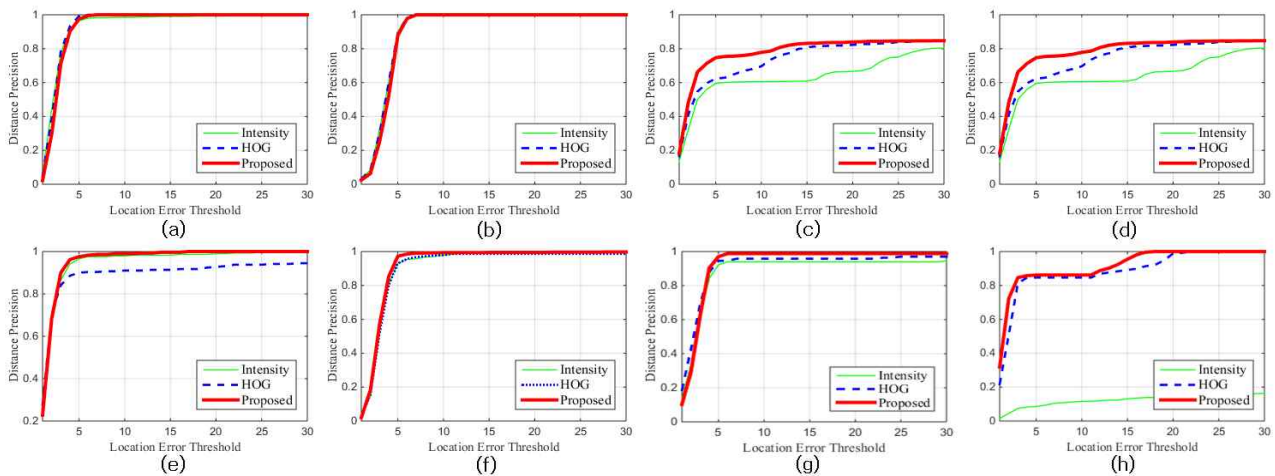


Fig 8. Precision plots of the experimental dataset (a) Tower IR Black Bus1, (b) Tower IR Black Bus2, (c) Hill EO Cart, (d) Tower IR Black Truck1, (e) Tower IR Black Truck2, (f) Hill IR Cart, (g) Building EO Truck, (h) Building EO Car

기반 특징인 제안하는 다중채널 LEP 특징 추출 방법을 사용했을 때, 네 개의 영상에서 기존 명암도나 경사도 기반 특징 대비 거리 오차 5픽셀에서 높은 추적 정확도를 보여 제안하는 다중채널 LEP가 기존 방법 대비 분별력이 있음을 보였다. 앞으로 연구에서는 더욱 더 많은 실험영상을 이용해 제안하는 방법을 실험하고, 특징정보 융합을 통해 성능개선을 확인하겠다.

REFERENCES

- [1] Hare, S., Golodetz, S., Saffari, A., Vineet, V., Cheng, M. M., Hicks, S. L., & Torr, P. H., "Struck: Structured output tracking with kernels." *IEEE transactions on pattern analysis and machine intelligence*, 38(10), pp. 2096–2109, 2016.
- [2] Danelljan, M., Shahbaz Khan, F., Felsberg, M., & Van de Weijer, J. "Adaptive color attributes for real-time visual tracking." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp. 1090–1097, 2014
- [3] DANELLJAN, Martin, et al. "Discriminative scale space tracking." *IEEE transactions on pattern analysis and machine intelligence*, 2016.
- [4] M. Kristan et al., "The visual object tracking VOT2014 challenge results," In *ECCV Visual Object Tracking Challenge Workshop*, 2014.
- [5] Kristan, M., Matas, J., Leonardis, A., Felsberg, M., Cehovin, L., Fernandez, G., ... & Pflugfelder, R. "The visual object tracking vot2015 challenge results." In *Proceedings of the IEEE international conference on computer vision workshops*, pp. 1–23, June 2010.
- [6] Kiani Galoogahi, H., Sim, T., & Lucey, S. "Multi-channel correlation filters." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3072–3079, 2013.
- [7] Danelljan, M., Häger, G., Khan, F., & Felsberg, M. "Accurate scale estimation for robust visual tracking." In *British Machine Vision Conference*, Nottingham, BMVA Press. September 1–5, 2014.
- [8] Jabid, T., Kabir, M. H., & Chae, O. "Local directional pattern (LDP) for face recognition," In *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*, pp. 329–330, January 2015.
- [9] Bolme, D. S., Beveridge, J. R., Draper, B. A., & Lui, Y. M. "Visual object tracking using adaptive correlation filters." In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2544–2550, June 2010.
- [10] Baker, S., & Matthews, I. "Lucas–kanade 20 years on: A unifying framework." *International journal of computer vision*, pp. 221–255. 56(3), 2004.
- [11] Yves Bouguet, Jean. "Pyramidal implementation of the lucas kanade feature tracker," Intel Corporation, Microprocessor Research Labs, 2000.

Authors



Eun-Joon Kim received the B.S.(2011) degrees in Media Communication Engineering from Hanyang University and M.S.(2014) degrees in Electrical Engineering from KAIST, Korea. He is

currently researcher in the Agency for Defense Development(ADD) since 2014. He is interested in visual tracking and object detection.