

논문 2021-16-42

다중 도메인 데이터 기반 구별적 모델 예측 트래커를 위한 동적 탐색 영역 특징 강화 기법

(Reinforced Feature of Dynamic Search Area for the Discriminative Model Prediction Tracker based on Multi-domain Dataset)

이 준 하, 원 홍 인, 김 병 학*
(Jun Ha Lee, Hong-In Won, Byeong Hak Kim)

Abstract : Visual object tracking is a challenging area of study in the field of computer vision due to many difficult problems, including a fast variation of target shape, occlusion, and arbitrary ground truth object designation. In this paper, we focus on the reinforced feature of the dynamic search area to get better performance than conventional discriminative model prediction trackers on the condition when the accuracy deteriorates since low feature discrimination. We propose a reinforced input feature method shown like the spotlight effect on the dynamic search area of the target tracking. This method can be used to improve performances for deep learning based discriminative model prediction tracker, also various types of trackers which are used to infer the center of the target based on the visual object tracking. The proposed method shows the improved tracking performance than the baseline trackers, achieving a relative gain of 38% quantitative improvement from 0.433 to 0.601 F-score at the visual object tracking evaluation.

Keywords : Visual object tracking, Deep learning, Reinforced feature, Discriminative model prediction

1. 서 론

시각적 정보의 디지털화를 위하여 다양한 비전센서가 활용되며 일반적인 RGB 카메라센서 이외에도 다중 카메라와 레이저 등을 활용한 다중 도메인의 정보가 응용되고 있다. 최근에는 다중 도메인의 비전 데이터와 인공지능 기술을 응용하여 탐지 (detection), 영상분할 (segmentation), 객체추적 (object tracking)을 비롯한 다양한 기법들이 연구되고 있으며, 객체추적 기술은 이전 영상과 추적 대상의 좌표와 현재 영상의 정보로 사전 정의된 레이블의 정보 없이 현재 추적 대상의 위치와 크기를 실시간 학습하여 연속적으로 추론하는 방법으로 일반적인 사전학습 탐지알고리즘의 연속적 동작과는 구별된다. 객체추적의 성능은 추적 영상에서 실시간으로 추적되는 대상의 위치정보를 의미하는 경계상자 (bounding box)의 네 꼭지점 좌표와 추적 좌표 정답 값의 교차영역 (IoU, Intersection over Union)을 계산하여 평가된다. 기존의 객체추적 기술로 SVM (Support Vector Machine), DPM (Deformable Part Model)을 활용한 성능이

우수한 모델이 존재하나 임베디드 시스템에 적합한 실시간 성능에 한계사항이 있음에 따라, 연산속도가 빠른 SAD (Sum of Absolute Difference), NCC (Normalized Cross Correlation)와 같은 고전적 영상 상관도 판별기반 고속 객체 추적 기술이 산업분야에서는 아직도 많이 활용되고 있다 [1, 2]. 또한, 머신러닝 기술의 급속한 발전 및 확산에 따라 객체 추적 기술 또한 성능향상과 고속화를 위한 연구가 활발히 진행되고 있다. 특히 영상 객체 추적 (VOT, Visual Object Tracking) 경쟁 개발 그룹에서는 구별적 주파수 도메인 상관 필터링 (DCF, Discriminative Correlation Filter), Siamese, ATOM (Accurate Tracking by Overlap Maximization), 및 DiMP (Discriminative Model Prediction) 모델을 기반으로 파생된 고성능의 추적 모델이 새롭게 개발되고 있다 [3-7]. 객체 추적기술의 최신 발전 동향으로 기존의 딥러닝 네트워크 모델의 기술이 융합된 ATOM, DiMP와 같이 사전 학습된 합성곱 계층 모델을 이용하여 목표 객체를 분류하여 실시간 학습 성능을 향상하는 방법이 많이 개발되고 있다. 또한 경계 상자 좌표를 추적하는 방식과 구분되는 세그먼테이션 모델의 접목으로 객체를 구성하고 있는 픽셀영역을 추적하는 모델에 대한 연구로 발전하고 있다.

고성능의 객체 추적 기술의 발전에도 불구하고, 일반적인 상태에서는 원활한 추적이 이루어지지만 객체가 배경에 가려진 후 다시 나타나거나, 크기나 모양이 급격하게 변하는 객체에 대한 연속추적 성능유지에는 여전히 한계사항이 존재

*Corresponding Author (bhkim81@kitech.re.kr)

Received: Jul. 22, 2021, Revised: Jul. 28, 2021, Accepted: Aug. 14, 2021.

J.H. Lee: Korea Institute of Industrial Technology, Kyungpook National University. (Research student)

H.I. Won, B.H. Kim*: Korea Institute of Industrial Technology. (Senior researcher)

※ 본 논문은 한국생산기술연구원 기관주요사업 "멀티 도메인 비전 시스템 AI 인식 기술 연구 (kitech EI-21-0019)"의 지원으로 수행한 연구입니다.

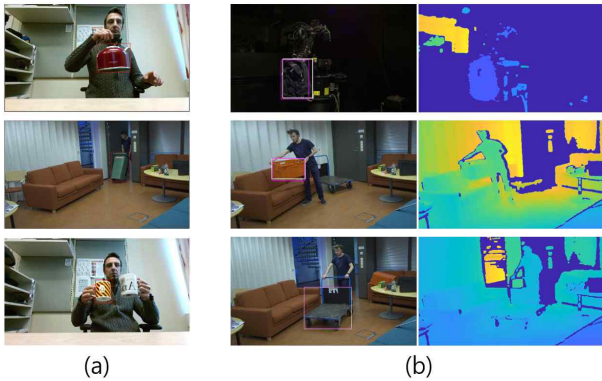


그림 1. 2차원 (RGB) 영상 및 RGB+3차원 데이터를 활용하는 객체 추적 예시. (a) 2차원 (RGB) 객체 추적, (b) RGB+3차원 데이터 기반 객체 추적

Fig. 1. Example of the 2D (RGB), RGB+3D dataset based object tracking. (a) 2D (RGB) tracking, (b) RGB+3D tracking

재한다. 이러한 한계점을 해결하기 위한 수단으로 그림 1과 같이, 기존의 2차원 (RGB) 영상에 3차원 영상 또는 IR (Infra-red) 영상을 융합하는 다중 도메인 데이터 활용 추적 방식이 연구되고 있다 [8-10]. 그 예로, ATCAIS (Accurate Tracking by Category Agnostic Instance Segmentation)는 ATOM에 HTC (Hybrid Task Cascade) 인스턴스 세그멘테이션 기법과 융합하여 동작하며, DDiMP (Depth enhanced DiMP) 모델의 경우 3차원 영상을 함께 이용하여 목표 객체가 배경에 의해 가려진 경우 또는 영상 외로 사라지고, 다시 찾는 경우에 사용된다 [11]. 그러나 여러 모델을 병렬적으로 사용하여 고성능의 객체추적 성능을 구현한 모델 또한 입력 데이터의 객체관련 정보의 손실 및 부족의 상황에서는 성능 저하의 단점을 보여줌에 따라, 본 연구에서는 실시간 객체 추적의 관심 영역에 대한 특징 강화 기법을 통해 추적 성능을 최적화하는 방법을 제안하고자 한다. 제안한 방법의 실험 및 기준모델 (baseline)은 ECCV2020에서 새로 도입된 VOT2020 Challenge의 VOT Toolkit (Trax protocol, Python 3.x)환경에서 수행되었으며, 실험 분석을 위한 데이터세트는 RGBD (RGB+3차원, 80 sequences)를 채택하여 연구를 진행하였다 [12, 13].

II. 관련 연구

추적에 관한 연구는 2차원 영상 데이터 기반으로 상관도 분석 기반의 추적 모델에서부터, 주파수 도메인 상관도 분석, 특징점 추출 기법 순으로 발전해 왔으며 딥러닝 네트워크 기반의 추적 모델의 적용 연구결과들이 현재까지 가장 우수한 성능을 보여주고 있다. 최근에는 다차원 데이터 획득을 위한 센서 기술의 발달로 3차원 정보를 추가적으로 획득하여 추적 성능을 향상시키기 위한 연구가 수행되고 있다.

또한 다차원 데이터의 적용과 함께 복잡한 딥러닝 네트워크 모델의 활용은 추적 성능의 획기적 향상 결과를 보여주었으나 여전히 입력 데이터의 객체 특징 정보 손실에 따른 추적 성능 저하의 문제점 극복을 위한 연구 필요성이 강조된다.

1. 추적을 위한 특징 강화 연구

HoG (Histogram of Oriented Gaussian) 특징 추출 방법과 추적영역의 동적 추정 학습을 적용하여 획기적인 성능향상 결과를 보여준 DSST (Discriminative Scale Scalable Tracking) 방법의 개발 사례로 부터 객체 추적 방법의 최적화를 위한 영상 특징 강화 기법 활용의 중요성을 확인할 수 있다 [14-17]. 영상의 특징 강화 기법으로는 HoG 이외에도 SIFT, SURF, FAST 등의 고전적 방법이 있다 [18-21]. 이러한 영상처리 기반의 특징 추출 방법은 객체의 특징을 추출하기 위해 단일 프레임 영상정보를 활용한 컨벌루션 필터 연산으로 다양한 조건의 프레임에 포함되는 객체의 특징 도출에 한계사항이 있다. 이러한 한계사항을 극복하기 위해 최근에는 RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), MLP (Multi-Layer Perceptron), CNN (Convolutional Neural Network) 등 신경망 구조를 활용한 특징 추출 기법이 활용된다 [22-25]. 그러나 방대한 데이터를 학습한 신경망 구조가 학습되지 않은 데이터에 대한 객체 탐지에 용이한 특징을 추출이 가능하다는 장점과 동시에 동작 속도 저하의 문제로 인해 임베디드 시스템에 활용되기 어렵다는 단점이 존재한다. 노이즈의 영향에 따라 객체의 특징정보 손실의 특성을 극복하기 위한 목적으로 Residual 신경망 네트워크 응용한 MWCNN (Multi-level Wavelet-CNN), RFSUNet (Residual Forward Subtracted U-shaped Network) 방법이 있으며, 노이즈가 제거된 선명한 화질향상 결과를 제공해 주는 동시에 원본의 객체의 섬세한 특징을 왜곡시키는 한계사항을 보여주었다 [26, 27]. 또한 객체 특징 정보의 향상을 위한 영상 화질 향상기법으로 CLAHE (Contrast limited adaptive histogram equalization), AGCWD (Adaptive Gamma Correction With Weighting Distribution), DSINMHE (Dualistic Sub-Image & Non-parametric Modified Histogram Equalization)와 같은 적응적 영역분할 처리 방식은 낮은 분해능의 영상에서 왜곡 없이 객체의 섬세한 특징을 강화할 수 있는 강점을 보여주지만 전체 영상 범위의 연산이 필요함에 따라 비효율적이라는 단점이 있다 [28, 29]. 본 연구에서는 이와 같은 기존 연구의 한계사항 극복을 위하여 신경망 구조의 고성능 특징검출 방법과 효율적 특징 추출 방법에 대한 연구를 진행하였다.

2. 다중 도메인 영상 (RGBD) 객체추적

최근 딥러닝 기술의 급진적인 발전으로 사전학습 기반의 탐지 네트워크 모델을 객체 추적기로 응용한 모델들은 우수

표 1. 사전학습용 데이터세트
Table 1. Pre-training Datasets

Dataset	Number of Classes	Number of Images (Data storage)
LASOT	70	3,500K (41GB)
GOT10k	563	1,500K (65.8GB)
TrackingNet	27	14,000K (1TB)
MSCOCO	91	2,500K (18GB)

한 성능을 보여주었지만 사전에 학습되지 않은 클래스의 객체 정보 입력 및 연속적 추적 상황에서의 객체의 변화가 발생할 경우 추적 성능이 급격히 감소되게 된다. 이러한 한계점을 해결하기 위해 사전학습과 실시간 학습을 병행하여 동작하는 방식의 Siamese 모델과 최근에는 ATOM, DiMP와 같은 융합형 객체 추적 모델들이 개발되었다 [3-11].

추적기술 활용의 실제 사례를 통하여, 복잡한 배경 및 전경의 가림 등의 현상으로 인하여 기존의 2차원 영상 기반의 객체추적 방식은 연속추적 상황에서 빈번히 객체를 놓치는 한계사항을 보여주었으며, 그 한계사항 극복을 위한 대책 연구의 필요성이 강조되었다. 영상 객체 추적 기술의 연구와 관련하여 유럽을 중심으로 국제적인 협력연구가 활발히 수행중인 VOT Challenge는 최근 RGBD 영상 데이터를 추가로 공개하였으며 다중 도메인 기반의 객체추적 알고리즘 경쟁 개발을 주도하고 있다 [12]. 다중 도메인 영상 기반 객체 추적 방식에서 3차원 (depth) 영상은 2차원 (RGB) 영상을 중심으로 동작되는 추적기의 동작을 선택적 보조적으로 개입하여 최적화하는 방식으로 동작한다. 본 연구의 실험과정에서 활용된 RGBD 객체추적기의 대표적 방법인 DDiMP는 2차원 영상 데이터세트를 기반으로 사전 학습을 수행하는 ResNet50 백본 구조의 특징 추출기를 내장하고 있으며, 두 개의 서브 추적기 모델이 융합되어 동작하는 방식이다.

III. 동적 탐색영역 특징강화 추적기

본 연구에서 제안하는 기법은 복합적 영상 객체 추적 네트워크의 성능 최적화를 위한 기법으로 구별적 모델 예측 특징추출 네트워크와 추적 좌표의 정밀 보정 네트워크 구조에 동적 탐색영역 특징강화 피드백 (RFDSA, Reinforced Feature with Dynamic Search Area) 구조를 추가하여 성능을 최적화하는 방법에 관한 것이다. 제안하는 방법은 DiMP와 IoU-Net 등 두 개의 복합적 서브 네트워크 모델인 DDiMP의 성능뿐만 아니라 단일 구조의 네트워크에 실시간 학습 기반 추적영역 분류기를 활용하는 ATOM 모델과 같이 단일 도메인 데이터 (2차원)를 사용하는 방법에도 동일하게 적용이 가능하다.

1. 구별적 모델 예측 특징추출

동적 탐색영역 특징강화 추적기 (RFDSA)에서 활용되는 DiMP 추적기는 ResNet 기반의 구별적 모델 예측 특징 추출 모델을 가지고 있으며, 본 연구에서 활용하는 DiMP 네트워크는 ResNet50 기반의 사전학습 특징 추출기를 활용하였으며 다양한 경우의 객체 탐지 상황을 고려하기 위하여, 대량의 학습 데이터세트를 활용한다. 사용된 학습데이터는 표 1과 같이 다중 도메인 영상 (RGBD) 기반 객체추적 요구사항의 추적 목표 정보인 차량, 사람, 사물, 동물 등의 샘플을 포함하는 4종의 데이터로 LASOT, GOT10k, TrackingNet, MSCOCO 데이터세트로 학습되었다 [30-33].

2. 추적 좌표 정밀 보정 (IoU-Net)

특징 추출기로 계산한 추적 모델의 반응도 맵의 최댓값을 추론하여 반응도 맵의 최댓값 좌표에 따라 객체 추적 좌표를 갱신하는 것이 IoU-Net 보정기법의 기본적인 동작 개념

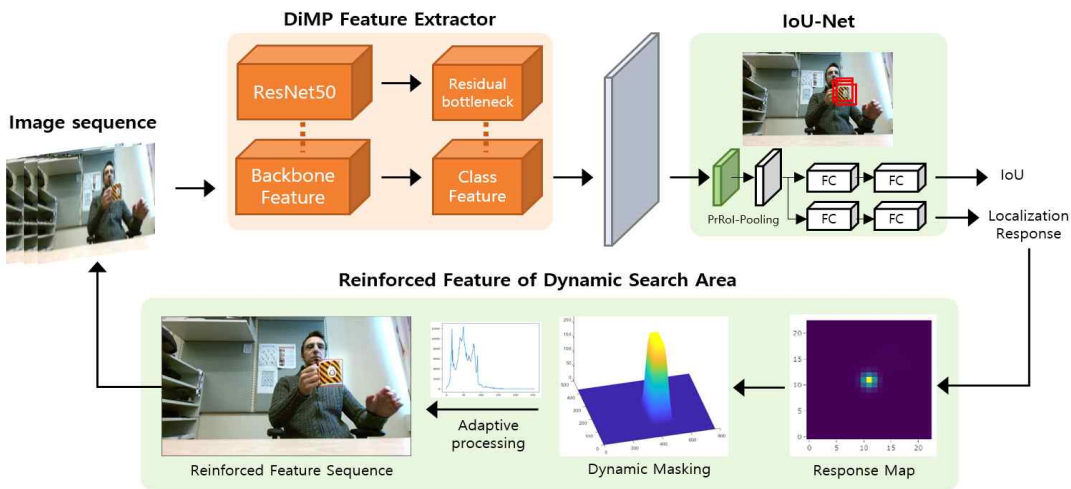


그림 2. 동적 탐색영역 특징강화 기법이 적용된 구별적 모델 예측 추적기 동작 개념도
Fig. 2. Conceptual diagram of reinforced feature for the discriminative model prediction tracker

넘이다. 이러한 추적기의 동작은 입력되는 영상 조건에 따라 추적의 성능이 저하될 수 있다. 예를 들어 영상에서 추적 중인 객체의 특징적 요소가 감소되거나 특징 정보가 부족할 경우, 반응도 맵의 최대 값의 좌표가 다수의 국소 좌표로 분산되거나, 최대값과 최소값의 차이가 줄어들어 현상과 함께, 신뢰도 값이 줄어들게 된다. 이러한 문제점으로 추적기의 연속적 객체추적 동작에서 성능의 오차가 증가되게 된다. 이러한 한계사항을 극복하기 위해 특징 추출기로 계산한 신뢰도 맵은 IoU-Net의 입력으로 전달되며, 추적된 경계 상자와 정답값 간의 IoU를 예측하기 위해 학습된 네트워크로, 추적 좌표의 정밀 보정을 수행하여 보다 정확한 경계 상자를 추정한다.

3. 동적 탐색영역 특징강화

제안한 방법은 기본적 추적기 모델의 반응도 출력좌표로부터 관심 탐색영역 (ROI, Region of Interest)을 지정하고 객체의 존재 확률이 높은 탐색영역에 대해 적응적으로 영상의 특징을 강화하는 처리를 수행한다. 이 경우 전체 영상과 탐색영역의 영상 화질 차이로 인하여 객체의 형상정보가 왜곡될 수 있으므로, 동적 가우시안 마스크 처리를 통하여 객체의 특징 강화 영역이 자연스럽게 처리되도록 보장하였다. 전체적인 동적 탐색영역 특징강화 추적기의 동작 개념은 그림 2와 같다. 첫 프레임에서 영상과 입력되는 초기 추적 좌표를 포함하여 매 프레임의 추적 결과에 대해 앞서 설명한 추적 프로세스를 수행한다. 제안한 방법이 기존의 DiMP 모델과 차별화된 점은, 추적 좌표에 따른 관심 탐색영역 좌표를 생성하고, 영상 특징 강화기법으로 적응형 감마 보정 기법을 활용하여 가중치 분포에 따라 감마 보정 그래프를 능동적으로 변화시키는 형태로 영상의 객체 정보에 따라 적응적으로 특징을 강화 할 수 있게 된다. 그림 3의 (c)에서 보여주는 예시와 같이, 동적 ROI 영역은 추적기 네트워크 출력 반응맵으로부터 생성된 마스크에 가우시안 필터링을 적용하여 ROI 마스크로 설정한다. 해당 마스크는 표 2의 알고리즘 요약 코드에서 $G(x,y)$ 로 표현되었다. G 마스크를 가중치로 주어 원본 이미지와의 융합된 영상을 생성한 결과는 그림 3의 (b)와 같이 객체에 스포트라이트 효과와 같이 ROI 내부의 영역은 특징이 강화되었고 배경의 영역과 불연속 특징이 없도록 융합된 영상이 추적기 네트워크의 입력으로 사용된다. 해당 모델은 ResNet 기반의 특징 추출 학습모델을 가지고 있으며 추적하고 있는 객체 이미지를 이용하여 실시간으로 학습되는 형태로 최종적으로 ResNet50 버전의 특징 추출 네트워크 구조를 사용한다. 실시간 추적하고 있는 대상에 대해, 모델을 업데이트하는 방식은 사전에 학습되지 않은 클래스에 대한 신속한 수렴을 위한 최적화 방법으로 가파른 경사 하강 방법을 사용한다. 따라서 사전에 학습된 클래스임에도 입력된 영상이 객체 특징 정보가 부족하여 클래스 분

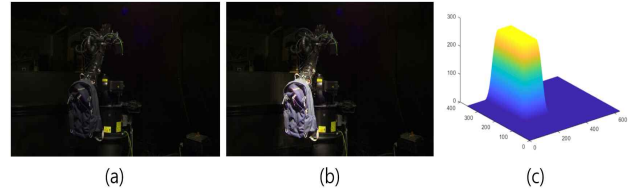


그림 3. 동적 특징강화 처리 결과 (a) 입력 추적 영상 (b) 표적 특징강화 영상 (c) 동적 탐색영역 추론 결과

Fig. 3. Result of reinforced feature (a) Input sequence (b) reinforced feature result (c) dynamic ROI inference result

표 2. 제안한 방법 (RFDSA)의 알고리즘 코드

Table 2. Pseudo code for proposed RFDSA method

Algorithm: $rfDSA(I(x,y), \sigma = 0.1)$
$\in put : I(x,y)$ # color image
$I_{HSV} = BGR2HSV(I)$ # convert BGR to HSV
$H = calcHist(I_{HSV})$ # I_{HSV} : value of HSV
$PDF_w = \max(H) * (\frac{H - \min(H)}{\max(H) - \min(H)})^\sigma$
$CDF_w(i) = \frac{\sum_{k=0}^i PDF_w(k)}{\sum PDF_w}$
$L(i) = 255 * (\frac{i}{255})^{1 - CDF_w(i)}$
For $i = 0, \dots, height$ do
For $j = 0, \dots, width$ do
Intensity = $I(i, j)$
$J(i, j) = L(Intensity)$ # J : value matrix
end
end
$I_{HSV} = J(i, j)$
$E = HSV2BGR(I_{HSV})$ # E : enhanced Image
$S(x,y) = (1 - G(x,y))I(x,y) + G(x,y)E(x,y)$
G : gaussian blurred weight
Return: S

류가 어려운 경우, 제안된 영상 특징강화 기법으로 성능이 향상된다.

IV. 실험 결과 및 분석

1. Ablation Study (최적화 분석)

동적 가중치 분포 영역기반 특징 강화 기법의 파라미터인 σ 값의 변화에 따른 최적화 및 가장 높은 추적 정확도 구현 조건을 확인하기 위하여 ablation study를 수행하였다. σ 파라미터의 최적화 과정은 기존의 성능최적화가 수행된 rFDDiMP 모델을 기반으로 분석하였다. 표 3의 비교 실험 결과와 같이 $\sigma=0.1$ 일 때 가장 높은 성능을 나타내는 것으로 확인되었으며, 과도한 특징 강화 기법의 적용의 경우로 $\sigma=0.5$ 의 경우는 전체적으로 출력 영상의 포화가 시작되며, 객

표 3. RFDSA 기법의 최적화 σ -파라미터 분석 표
Table 3. Ablation study for optimized σ -parameter RFDSA

	$\sigma=0.05$	$\sigma=0.1$	$\sigma=0.5$
Image Quality			
Histogram			
Precision	0.696	0.702	0.698
Recall	0.693	0.696	0.694
F-score	0.694	0.699	0.696

체의 특징이 감소와 동시에 객체 추적의 성능 개선 효과가 감소되는 특성을 보여주었다.

2. 실험 결과

본 연구에서 제안된 기법을 평가하기 위해 2021 VOT-RGBD 데이터셋을 활용하였다 [12]. 데이터 세트는 80개의 시퀀스로 구성되어 있으며 VOT Challenge에서 사용하는 3개의 평가 항목인 정밀도, 재현율, F-score로 평가하였다. 정밀도와 재현율은 예측 확실성 임계값인 τ_θ 에 따라 측정된다. 프레임이 많은 시퀀스의 결과에 편향되지 않도록, 시퀀스마다 측정된 정밀도 (P)와 재현율 (Re)의 평균값으로 결과를 분석한다. F-score는 수식 (1)의 방식으로 계산된다.

$$F(\tau_\theta) = 2P(\tau_\theta)Re(\tau_\theta)/(P(\tau_\theta) + Re(\tau_\theta)). \quad (1)$$

앞서 언급한 DiMP, ATOM, DDiMP와 비교하여 제안한 기법이 추가 적용된 추적기 모델(rfATOM, rfDiMP, rfDDiMP)들의 결과로 그림 4 (b)의 가로축의 값은 정규화된 τ_θ 로, 추적 모델의 추적 신뢰도가 임계값 보다 낮은 경우 출력이 무시되어 세로축의 값인 F-score에 반영된다. 이러한 측정 방법은 장기적인 추적 성능을 평가하기 위한 방법으로, 추적 모델이 추적 대상을 잃었을 경우 재탐색 능력을 간접적으로 평가하기 위한 지표로도 사용된다. 계산된 F-score의 최댓값을 정량적인 평가 지표로 사용한다. 표 4는 정밀도, 재현율, F-score의 정량적인 수치를 보여준다. 제안한 방법의 적용에 따라 모두 성능 향상 결과를 보여주었다. 실험은 PyTorch(Ver. 1.2.0)환경에서 구축되었으며, Intel i9-10900K (64GB RAM), NVIDIA GeForce RTX 2080 Ti GPU 실험 시스템 환경에서 실행되었다. 그림 4의 (a) 정밀도-재현율 곡선에서 rfDDiMP의 결과가 가장 넓은 곡선 아래 면적 (AUC, Area Under the Curve)의 결과로 가장 우수한 성능을 보여준다. rfDDiMP의 경우 F-score 기준으로 기존의 DDiMP에서 0.696수준으로 측정 되었으나 0.699의 수준으로 향상했다. 최대 개선 결과를 보여준 rfDiMP는 그림 5와 같이 F-score 기준으로 기존의 DiMP에서 0.433수준으로 측정 되었으나 0.601의 수준으로 38.8% 향상 효과를 보여주었다.

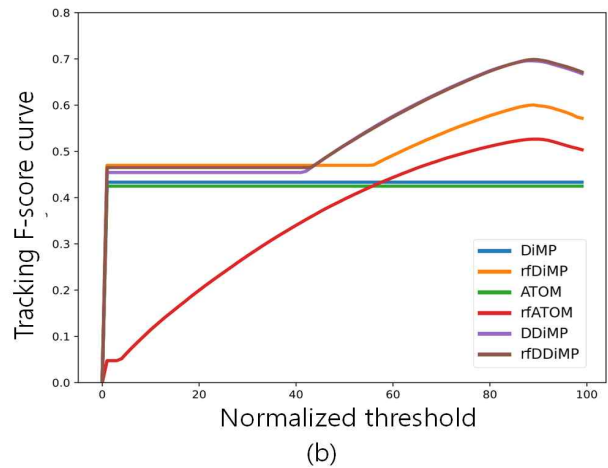
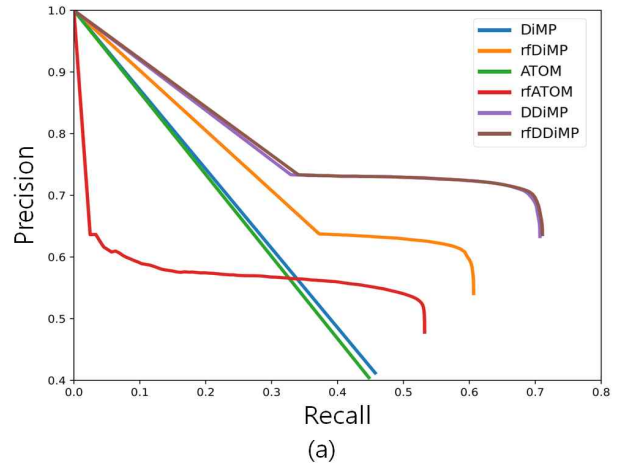


그림 4. 추적 성능 비교 실험 결과 (a) 추적 모델들의 정밀도-재현율 곡선 (b) 추적 모델의 F-score-임계값 곡선
Fig. 4. Comparative result of the performance of the trackers (a) Precision-Recall curves of the trackers (b) Tracking F-score curves of the trackers

표 4. 정밀도, 재현율, f-score으로 평가한 VOT RGBD 데이터 셋에 대한 추적 모델의 비교

Table 4. Trackers comparison on the VOT2021 RGBD dataset in terms of expected precision, recall, f-score

Trackers	Precision	Recall	F-score
ATOM	0.404	0.448	0.425
DiMP	0.412	0.457	0.433
DDiMP	0.704	0.688	0.696
rfATOM	0.529	0.524	0.526
rfDiMP	0.608	0.593	0.601
rfDDiMP	0.702	0.696	0.699

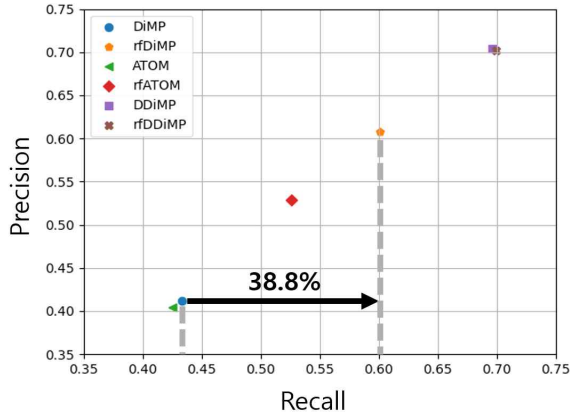


그림 5. 추적기 성능 실험 결과 분석 (F-score rank)

Fig. 5. Analysis of tracker performance (F-score rank)

V. 결론

본 논문에서는 시각적 객체 추적에서의 영상 개선 기법으로 실시간 추적 결과에 따른 동적 탐색영역 특징강화 기법을 제안하였으며, 다양한 추적기법에 제한한 방법을 적용하여 추적 성능 향상 여부를 보여주었다. 제안한 방법은 다차원의 데이터기반 추적기인 DDiIMP 및 단일 데이터 기반의 ATOM 추적기 모델에 적용하여 일관된 성능 향상 결과를 보여주었다. 최적화 파라미터 비교분석 과정을 통하여 제안한 기법의 성능 최적화를 위한 σ 파라미터를 도출하였으며, 제안방법이 적용된 모든 추적기에서 성능 향상효과가 확인되었으며 그 가운데 rDiIMP에서 추적 성능의 가장 큰 개선 효과가 확인되었으며, rDDiIMP의 경우 가장 높은 추적 성능을 보여주었다. 향후 계획으로는 다중도메인 데이터로 활용되는 3차원 영상 데이터의 낮은 해상도와 객체의 색상 정보를 획득할 수 없는 한계사항을 극복하기 위한 연구를 수행하고자 한다. 3차원 데이터는 2차원 영상 데이터와 비교하여 낮은 해상도와 색상 특징의 정보가 부족하지만 공간 및 형상 정보의 탐지 분류 학습모델을 활용하여 객체 추적 알고리즘의 재탐색 단계 탐지 식별력을 향상하는 접근방법으로 네트워크 개발 연구를 수행할 계획이다.

References

[1] L. Hou, W. Wan, K. H. Lee, J. N. Hwang, G. Okopal, J. Pitton, "Robust Human Tracking Based on DPM Constrained Multiple-kernel from a Moving Camera," *Journal of Signal Processing Systems*, Vol. 86, No. 1, pp. 27-39, 2017.

[2] S. J. Lee, B. H. Kim, M. Y. Kim, "Multi-Saliency Map and Machine Learning Based Human Detection for the Embedded Top-View Imaging System," *IEEE Access*,

Vol. 9, pp. 70671-70682, 2021.

[3] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, P. H. S. Torr, "Fully-convolutional Siamese Networks for Object Tracking," *European Conference on Computer Vision*. Springer, Cham, pp. 850-865, 2016.

[4] G. Bhat, M. Danelljan, L. V. Gool, R. Timofte, "Learning Discriminative Model Prediction for Tracking," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6182-6191, 2019.

[5] M. Danelljan, G. Bhat, F. S. Khan, M. Felsberg, "Atom: Accurate Tracking by Overlap Maximization," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4660-4669, 2019.

[6] B. Li, J. Yan, W. Wu, Z. Zhu, X. Hu., "High Performance Visual Tracking with Siamese Region Proposal Network," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8971-8980, 2018.

[7] M. Danelljan, L. V. Gool, R. Timofte, "Probabilistic Regression for Visual Tracking," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7183-7192, 2020.

[8] B. H. Kim, A. Lukezic, J. H. Lee, H. M. Jung, M. Y. Kim, "Global Motion-aware Robust Vobject Tracking for Electro Optical Targeting Systems," *Sensors*, Vol. 20, No. 2, pp. 566, 2020.

[9] B. H. Kim, D. Khan, C. Bohak, W. Choi, H. J. Lee, M. Y. Kim, "V-RBNN Based Small Drone Detection in Augmented Datasets for 3D LADAR System," *Sensors*, Vol. 18, No. 11, pp. 3825, 2018.

[10] B. H. Kim, D. Khan, W. Choi, M. Y. Kim, "Real-time Counter-UAV System for Long Distance Small Drones Using Double Pan-tilt Scan Laser Radar," *Laser Radar Technology and Applications XXIV. International Society for Optics and Photonics*, pp. 11005, 2019.

[11] K. Chen, J. Pang, J. Wang, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Shi, W. Ouyang, C. C. Loy, D. Lin, "Hybrid Task Cascade for Instance Segmentation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4974-4983, 2019.

[12] M. Kristan, A. Leonardis, J. Matas, M. Felsberg, "The Eighth Visual Object Tracking VOT2020 Challenge Results," *European Conference on Computer Vision*. Springer, Cham, pp. 547-601, 2020.

[13] L. Čehovin, "TraX: The Visual Tracking Exchange Protocol and Library," *Neurocomputing*, Vol. 260, pp. 5-8, 2017.

[14] B. H. Kim, M. Y. Kim, Y. S. Chae, "Background Registration-based Adaptive Noise Filtering of LWIR/MWIR Imaging Sensors for UAV Applications," *Sensors*, Vol. 18, No. 1, pp. 60, 2018.

[15] M. Ž. Mekuč, C. Bohak, S. Hudoklin, B. H. Kim, R.

- Romih, M. Y. Kim, M. Marolt, "Automatic Segmentation of Mitochondria and Endolysosomes in Volumetric Electron Microscopy Data," *Computers in Biology and Medicine*, Vol. 119, pp. 103693, 2020.
- [16] B. H. Kim, C. Bohak, K. H. Kwon, M. Y. Kim, "Cross Fusion-based low Dynamic and Simage Enhancement for Infrared Search and Tracking Systems," *IEEE Access*, Vol. 8, pp. 15347-15359, 2020.
- [17] H. M. Jung, B. H. Kim, M. Y. Kim, "Residual Forward-Subtracted U-Shaped Network for Dynamic and Static Image Restoration," *IEEE Access*, Vol. 8, pp. 145401-145412, 2020.
- [18] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Ieee, Vol. 1, pp. 886-893, 2005.
- [19] T. Lindeberg, "Scale Invariant Feature Transform," pp. 10491, 2012.
- [20] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, "Surf: Speeded up Robust Features." European conference on computer vision, Springer, Berlin, Heidelberg, Vol. 110, No. 3, pp. 346-359, 2006.
- [21] E. Rosten, T. Drummond, "Machine Learning for High-speed Corner Detection," European conference on computer vision. Springer, Berlin, Heidelberg, pp. 430-443, 2006.
- [22] S. Kamthania, "A Novel Deep Learning rbm Based Algorithm for Securing Containers," in 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 1-7, 2019.
- [23] W. Li, Z. Dou, L. Qi, "Communication Protocol Classification Based on lstm and dbn," *IEEE Access*, Vol. 8, pp. 91818-91828, 2020.
- [24] I. El-Feghi, A. Tahar, M. Ahmadi, "Efficient Features Extraction for Fingerprint Classification with Multi Layer Perceptron Neural Network," in Eighth International Multi-Conference on Systems, Signals Devices, pp. 1-4, 2011.
- [25] H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. M. Summers, "Deep Convolutional Neural Networks for Computer-aided Detection: Cnn Architectures, Dataset Characteristics and Transfer Learning," *IEEE*, Vol. 35. No. 5, pp. 1285-1298, 2016.
- [26] P. Liu, H. Zhang, K. Zhang, L. Lin, W. Zuo, "Multi-level Wavelet-CNN for Image Restoration", *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, pp. 773-782, 2018.
- [27] H. M. Jung, B. H. Kim, M. Y. Kim, "Residual Forward-Subtracted U-Shaped Network for Dynamic and Static Image Restoration," *IEEE Access*, Vol. 8, pp. 145401-145412, 2020.
- [28] S. C. Huang, C. Fan-Chieh, C. Yi-Sheng, "Efficient Contrast Enhancement Using Adaptive Gamma Correction with Weighting Distribution," *IEEE Transactions on Image Processing*, Vol. 22, No. 3, pp. 1032-1041, 2012.
- [29] Y. Wang, Q. Chen, B. Zhang, "Image Enhancement Based on Equal Area Dualistic Sub-image Histogram Equalization Method," *IEEE transactions on Consumer Electronics*, Vol. 45, No. 1, pp. 68-75, 1999.
- [30] H. Fan, L. Lin, F. Yang, P. Chu, G. Deng, "Lasot: A high-quality benchmark for large-scale single object tracking," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5374-5383, 2019.
- [31] L. Huang, X. Zhao, K. Huang, "Got-10k: A Large High-diversity Benchmark for Generic Object Tracking in the Wild," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [32] M. Muller, A. Bibi, S. Giancola, S. Alsubaihi, B. Ghanem, "Trackingnet: A Large-scale Dataset and Benchmark for Object Tracking in the Wild," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 300-317, 2018.
- [33] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, "Microsoft coco: Common Objects in Context," *European Conference on Computer Vision*. Springer, Cham, pp. 740-755, 2014.

Jun Ha Lee (이 준 하)



2020 School of Electronic Engineering at Kyungpook National University, Republic of Korea (B.S.)

2021~School of Electronic Engineering at Kyungpook National University, Republic of Korea (M.S.)

Career:

2021~2021 Korea Institute of Industrial Technology (Researcher)

Field of Interests: Visual Object Tracking, Visual Object Detection, Multi-domain Object Detection, Visual Reinforcement

Email: wnsqk986@kitech.re.kr

Hong-In Won (원 흥 인)

2012 Mechanical Engineering from Hanyang University (B.S.)

2017 Mechanical Design Engineering from Hanyang University (Ph.D.)

2019~Korea Institute of Industrial Technology (Senior Researcher)

Career:

2017~2019 Postdoctoral Researcher, TU Dresden

Field of Interests: Digital Twin, Augmented Reality, Industry Automation, Machine Learning, Embedded System, Haptic Engineering, Noise and Vibration

Email: luvhayym@kitech.re.kr

Byeong Hak Kim (김 병 학)

2015 School of Electronic Engineering at Kyungpook National University, Republic of Korea (M.S.)

2020 School of Electronic Engineering at Kyungpook National University, Republic of Korea (Ph.D.)

2020~Korea Institute of Industrial Technology (Senior Researcher)

Career:

2007~2015 Engineer at the SAMSUNG THALES, Korea.

2016~2019 Engineer at the Hanwha Systems, Korea.

Field of Interests: IR and 3D Imaging Systems, Visual Object Tracking, ML/DL Object Detection, 3D Laser Radar, Counter Drone Systems, and Gesture Recognition.

Email: bhkim81@kitech.re.kr