Journal of Korea Robotics Society (2021) 16(4):336-344
https://doi.org/10.7746/jkros.2021.16.4.336 ISSN: 1975-6291 / eISSN: 2287-3961

혼잡 환경에서 강인한 딥러닝 기반 인간 추적 프레임워크

A Robust Deep Learning based Human Tracking Framework in Crowded Environments

오 경 석[1]·김 성 현[1]·김 진 섭[1]·이 승 환†
Kyungseok Oh[1], Sunghyun Kim[1], Jinseop Kim[1], Seunghwan Lee†

Abstract: This paper presents a robust deep learning-based human tracking framework in crowded environments. For practical human tracking applications, a target must be robustly tracked even in undetected or overcrowded situations. The proposed framework consists of two parts: robust deep learning-based human detection and tracking while recognizing the aforementioned situations. In the former part, target candidates are detected using Detectron2, which is one of the powerful deep learning tools, and their weights are computed and assigned. Subsequently, a candidate with the highest weight is extracted and is utilized to track the target human using a Kalman filter. If the bounding boxes of the extracted candidate and another candidate are overlapped, it is regarded as a crowded situation. In this situation, the center information of the extracted candidate is compensated using the state estimated prior to the crowded situation. When candidates are not detected from Detectron2, it means that the target is completely occluded and the next state of the target is estimated using the Kalman prediction step only. In two experiments, people wearing the same color clothes and having a similar height roam around the given place by overlapping one another. The average error of the proposed framework was measured and compared with one of the conventional approaches. In the error result, the proposed framework showed its robustness in the crowded environments.

Keywords: Detectron2, Human Tracking, Crowded Environment, Kalman Filter

1. 서 론

최근 무인 감시 및 로봇 서비스 분야의 관심이 증가하며 관련 로봇 기술에 대한 연구가 활발히 이루어지고 있다. 특히, 4차 산업혁명으로 인한 딥러닝 기술의 발전은 인간과 같은 물체를 탐지하는 기술 발전에 큰 도움을 주고 있다[1]. 하지만, 혼잡한 환경에서 인간 탐지 및 추적 기술은 인간과 인간의 겹침(overlapped) 또는 가려짐(occlusion) 상황에 의해 탐지에 실패하고, 잘못된 대상을 추적하는 결과를 초래할 수 있다. 이를 극복하기 위해 해당 상황을 인식하고, 인간의 키와 같은 다양한 정보를 활용하는 방안이 필요하다[2,3]. 본 논문에서는 위의 상황에서도 강인하게 목표 대상을 추적하기 위해, 딥러닝 및 칼만 필터를 이용하여 가중치 기반 인간 추적 프레임워크를 제안한다.

2. 관련 연구

이 분야의 초기 기술 중 하나는 카메라를 사용하여 인간을 감지하고 추적하였으며[4], 윤곽 기반 접근 방식과 색상 기반 접근 방식을 결합하여 활용하였다. 로봇은 비전 기반의 방법으로 옷의 가장자리 색상 및 질감을 사용하여 개인을 식별하여 대상을 추적하였다[5]. 추적 과정에서 대상자와의 거리를 추정하기 위해 다리를 감지하는 레이저 거리 데이터와 인간의 얼굴을 감지하는 시각 정보를 사용하는 융합 기술을 채택하였다[6]. Choi 등[7]은 인간 추적 분야에서 실제 어린이의 검출 과정 중 거리와 각도를 활용했다. 추적 대상은 칼만 필터를 통해

Received : Sep. 15. 2021; Revised : Nov. 5. 2021; Accepted : Nov. 9. 2021
※ This research was supported by Kumoh National Institute of Technology (2019-104-036)
1. BS Student, School of Electronic Engineering, Kumoh National Institute of Technology, Gumi, Korea (lshh0000@kumoh.ac.kr, tjdgus556@naver.com, zhr287@naver.com)
† Associate Professor, Corresponding author: School of Electronic Engineering, Kumoh National Institute of Technology, Gumi, Korea (leesh@kumoh.ac.kr)

추적 대상의 움직임을 예측하여 ROI (Region of Interest)를 선정 후 HOG (Histogram of Gradient)를 이용해 선정되거나[8], 현재의 위치를 기준으로 다음 프레임에서 기준거리 이내의 후보군을 탐색 후 신체 색상 평균값의 차이가 작은 인간이 선정된다[9]. 만약, 추적 대상이 FoV를 벗어났다면 등록된 추적 대상과의 유사도를 측정해 추적 대상을 복원하였다[10]. 그러나, 이러한 연구들은 환경에 따라 대상에 따라 추적의 정확도가 낮아지는 단점이 있다. 이러한 문제를 해결하기 위해 최근 인간 탐지 및 추적분야에서 딥러닝 기술이 활용되고 있다[11]. 특히, 학습이 비교적 쉽고 빠른 객체 탐지 속도를 가진 Detectron2를 이용한 객체 추적연구가 활발히 진행되고 있다[12]. 추적 정확도 향상을 위해 추적 대상의 특징으로 키[3]와 폭이 활용되었다[2].

본 논문은 혼잡 환경에서 강인한 딥러닝 기반 인간 추적 프레임워크를 제시하였다. 인간 추적 기술이 다양한 서비스 분야에서 사용되기 위해 인간이 감지되지 않거나, 혼잡 상황에서도 강인하게 추적되어야 한다. 이를 위해 3장에서 미검출 및 혼잡 상황 인식, 딥러닝 기반의 인간 탐지 및 가중치 기반의 추적 기술로 구성된 제안된 프레임워크를 상세하게 기술한다. 4장에서는 제안된 방법의 성능을 평가하기 위해 같은 색상의 옷을 입고 비슷한 키의 사람이 추적 대상과 겹치는 두 가지 실험을 진행하였다. 5장에서는 실험 결과에 대한 분석을 진행하여 제안된 방법이 기존 방법들 보다 강인한 추적 성능을 가짐을 보이고자 한다.

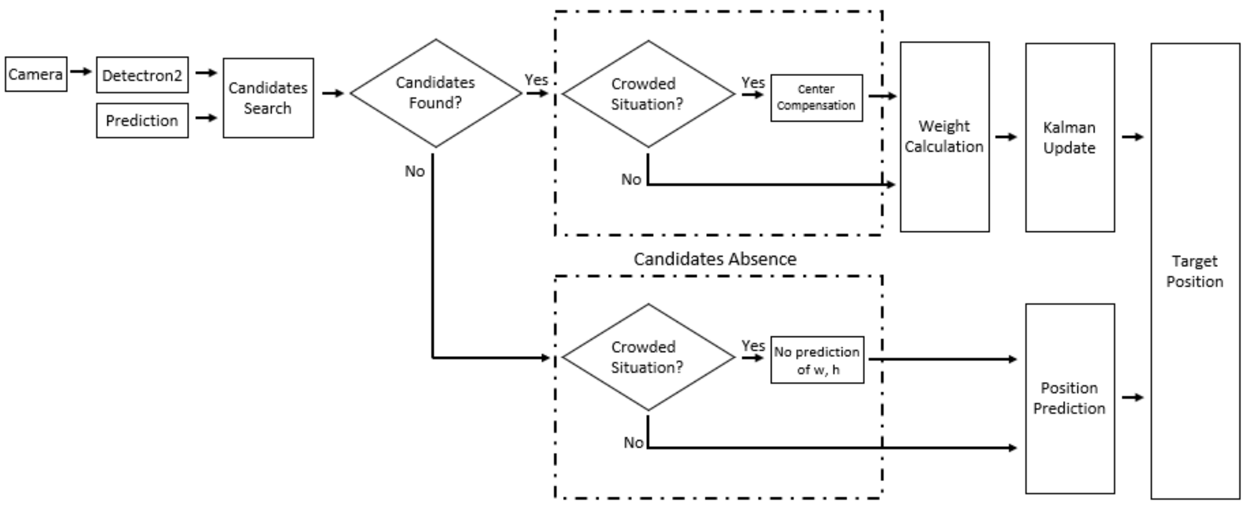
3. 강인한 인간 추적 방법

[Fig. 1]은 제안된 프레임워크의 전체 구조를 보여준다. 칼만 필터의 초기값은 영상 초기 추적 대상만 단독으로 있는 상황에서 Detectron2의 경계 상자(Bounding Box)를 통해 구해진다. 기준 인간 이미지는 칼만 필터의 초기값 선정이 완료된 후 ORB 특징점 매칭을 위해 획득한다.

추적 대상 선정을 위해 ORB 특징점, 경계 상자의 폭, 높이 그리고 중심 좌표의 가중치를 계산하고 활용한다. 다음 세부 문단에서 상세히 설명한다.

3.1 용어(Terminology)

본 논문에서 제안하는 알고리즘을 설명하기 위해 [Table 1]은 관련 용어가 기술되어있다. $x^i_{k,\,candi}$, $y^i_{k,\,candi}$, $w^i_{k,\,candi}$, $h^i_{k,\,candi}$은 Detectron2를 통해 검출된 후보군의 중심 좌표 및 경계 상자의 폭과 높이를 나타낸다. $Z_{k,\,x}$, $Z_{k,\,y}$, $Z_{k,\,w}$, $Z_{k,\,h}$은 제안된 알고리즘을 통해 선정된 추적 대상의 중심 좌표 및 경계 상자의 폭과 높이를 나타낸다. $\hat{x}_k$, $\hat{y}_k$, $\hat{w}_k$, $\hat{h}_k$은 칼만 필터를 통해 추정된 추적 대상의 중심 좌표 및 경계 상자의 폭과 높이를 나타낸다. $\hat{x}^-_k$, $\hat{y}^-_k$, $\hat{w}^-_k$, $\hat{h}^-_k$은 추적 대상의 예측치를 나타낸다. $\hat{x}_b$, $\hat{y}_b$, $\hat{w}_b$, $\hat{h}_b$은 혼잡 상황 직전의 추적 대상의 중심 좌표 및 경계 상자의 폭과 높이를 나타낸다. $x_{k,\,c}$와 $y_{k,\,c}$은 본 논문에서 제안하는 방법을 통해 보상된 추적 대상의 중심 좌표를 나타낸다. $\triangle x_k$와 $\triangle y_k$은 추적 대상 중심 좌표의 변화량을 나타낸다. $\triangle w_k$와 $\triangle h_k$은 추적 대상의 폭과 높이의 변화량을 나타낸다. $N^i_{k,\,candi}$은 기준 인간 이미지와 후보군의 매칭된 ORB 특징점 수를 나타낸다. $W^i_{k,\,candi}$은 추적 대상 선정을 위한 후보군의 가중치를 나타낸다. x_k, y_k, w_k, h_k는 k시간에서 경계 상자의 중심 좌표, 폭 그리고 높이를 의미한다. $\triangle x_k$, $\triangle y_k$, $\triangle w_k$, $\triangle h_k$는 다음과 같이 정의된다.



[Fig. 1] The proposed framework

[Table 1] A list of symbols and their definition

Symbol	Definition
$x^i_{k,candi}$, $y^j_{k,candi}$, $w^j_{k,candi}$, $h^i_{k,candi}$	The state of the i-th candidate for center, width and height at k
$Z_{k,x}$, $Z_{k,y}$, $Z_{k,w}$, $Z_{k,h}$	Measurement of target at k
$\hat{x}_k$, $\hat{y}_k$, $\hat{w}_k$, $\hat{h}_k$	Estimation state at k
$\hat{x}^-_k$, $\hat{y}^-_k$, $\hat{w}^-_k$, $\hat{h}^-_k$	Prediction state at k
$\hat{x}_b$, $\hat{y}_b$, $\hat{w}_b$, $\hat{h}_b$	Estimated state prior to crowded situation
$x_{k,c}$, $y_{k,c}$	Compensated measurement for target center at k
$\triangle x_k$, $\triangle y_k$	Variation of target center at k
$\triangle w_k$, $\triangle h_k$	Variation of width and height at k
$N^i_{k,candi}$	The number of matched orb-features for the i-th candidate at k
$W^i_{k,candi}$	The weight of the i-th candidate at k

$$\triangle x_k = \frac{x_{k+1} - x_k}{\triangle t}, \tag{1}$$

$$\triangle y_k = \frac{y_{k+1} - y_k}{\triangle t}, \tag{2}$$

$$\triangle w_k = \frac{w_{k+1} - w_k}{\triangle t}, \tag{3}$$

$$\triangle h_k = \frac{h_{k+1} - h_k}{\triangle t}, \tag{4}$$

본 논문에서 '혼잡 상황'은 추적 대상의 경계 상자가 추적 대상이 아닌 다른 대상의 경계 상자에 포함되는 것으로 정의한다. 식 (5), (7)을 모두 만족한다면 추적 대상이 왼쪽에서 오른쪽으로 겹쳐지는 혼잡 상황을 나타낸다. 식 (6), (7)을 모두 만족한다면 추적 대상은 오른쪽에서 왼쪽으로 겹쳐지는 혼잡 상황을 나타낸다.

$$x^i_{k,candi} - w^i_{k,candi}/2 < Z_{k,x} + Z_{k,w}/2 < x^i_{k,candi} + w^i_{k,candi}/2, \tag{5}$$

$$x^i_{k,candi} - w^i_{k,candi}/2 < Z_{k,x} - Z_{k,w}/2 < x^i_{k,candi} + w^i_{k,candi}/2, \tag{6}$$

$$y^i_{k,candi} - h^i_{k,candi}/2 < Z_{k,y} < y^i_{k,candi} + h^i_{k,candi}/2. \tag{7}$$

제안된 알고리즘에서는 추적 상태를 혼잡 상황(Crowded situation)과 온전한 상황(Non-crowded situation)으로 분류하였다. 혼잡 상황은 앞서 언급한 바와 같이 추적 대상이 다른 대상과 일부 겹쳐진 상황 또는 완전히 가려진 상황으로 정의하였다. 온전한 상황은 추적 대상이 다른 대상과 겹쳐지지 않은 상황을 의미한다.

3.2 인간 탐지 기술

Detectron2는 Facebook AI Research (FAIR)에서 Mask R-CNN을 벤치마킹하여 속도 및 확장성을 개선한 딥러닝 API이며, 경계 상자의 좌표, 분류명 및 정확도를 얻을 수 있다. 본 연구에서는 사전학습 된 Faster R-CNN ResNet101 FPN 모델[13]을 이용하여 연구를 진행하였다. Detectron2는 제안된 알고리즘에서 초기 추적 대상 선정, 대상의 지속적 추적을 위한 후보군 선정에 활용되었다.

Detectron2에서 검출된 후보군의 $w^i_{k,candi}$, $h^i_{k,candi}$와 추적 대상 예측치 $\hat{w}^-_k$, $\hat{h}^-_k$의 비율이 식 (8), (9)를 만족한다면 검출된 후보군에 추적 대상이 존재한다고 판단한다.

$$1 - \alpha < \frac{w^i_{k,candi}}{\hat{w}^-_k} < 1 + \alpha, \tag{8}$$

$$1 - \alpha < \frac{h^i_{k,candi}}{\hat{h}^-_k} < 1 + \alpha, \tag{9}$$

3.3 인간 추적 기술

3.3.1 추적 대상 추정

본 논문에서 추적 대상의 정보는 기본적으로 칼만 필터를 이용하여 추정되었다. 제안된 인간 추적 알고리즘을 위한 칼만 필터는 식 (10), (11)을 통해 모델링 되었다.

$$x_{k+1} = Ax_k + w_k, \tag{10}$$

$$z_k = Hx_k + v_k, \tag{11}$$

위 수식에서 x_k와 z_k는 8×1 벡터로 앞장에서 설명한 바와 같이 경계 상자의 중심 좌표, 폭 그리고 높이에 대한 정보를 포함하고 있다. $H = I$이며 w_k와 v_k는 각각 공분산에 대한 부분으로 $w_k \sim N(0,Q)$와 $v_k \sim N(0,R)$로 정의된다. A는 전이행렬로 다음과 같이 정의된다.

$$A = \begin{pmatrix} 1 & 0 & \triangle t & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \triangle t & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \triangle t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \triangle t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{12}$$

3.3.2 온전한 상황에서의 추적 과정

칼만 필터의 초기값 선정이 완료된 후, 다음 프레임부터 추적

과정이 시작된다. Detectron2에서 검출된 후보가 1명이고, 그 후보가 식 (8), (9)를 만족한다면 추적 대상으로 선정된다.

[Table 2]는 식 (8), (9)를 만족하는 후보가 2명 이상 존재할 때 추적 대상 선정을 위한 알고리즘을 나타낸다. 알고리즘의 입력은 식 (13)에서 후보군과 추적 대상 예측치와의 중심 좌표 추적 오차 $d^i_{k,dist}$, 식 (14)에서 후보군과 추적 대상 예측치와의 폭 추적 오차 $d^i_{k,width}$, 식 (15)에서 후보군과 추적 대상 예측치와의 높이 추적 오차 $d^i_{k,height}$ 이다. 추적 대상 선정을 위해 $d^i_{k,dist}$, $d^i_{k,width}$, $d^i_{k,height}$가 가장 작은 후보의 가중치를 각각 증가시키고, $N^i_{k,candi}$ 가 가장 큰 후보의 가중치를 증가시켜 각 후보의 가중치 $W^i_{k,candi}$ 를 얻는다.

출력 Target_ID는 어떤 후보의 $W^i_{k,candi}$ 가 가장 높은지를 나타낸다. Target_ID를 통해 Detectron2에서 검출된 후보군 중 추적 대상이 선정된다. 만약, 동일한 가중치로 인해 후보를 선정할 수 없다면, 추적 대상은 식 (16)에서 후보군과 추적 대상 예측치와의 면적 추적 오차 $d^i_{k,area}$가 작은 후보에게 가중치를 추가로 부여하여 선정된다.

칼만 필터는 선정된 추적 대상의 $Z_{k,x}$, $Z_{k,y}$, $\triangle x_k$, $\triangle y_k$, $Z_{k,w}$, $Z_{k,h}$, $\triangle w_k$, $\triangle h_k$를 칼만 필터의 관측치로 사용하여 상태가 갱신된다. 칼만 필터의 추정치 $\hat{x}_k$, $\hat{y}_k$는 추적 대상의 최종적인 위치가 된다.

[Table 2] Algorithm for target detection

Input: $d^i_{k,dist}$, $d^i_{k,width}$, $d^i_{k,height}$, $N^i_{k,candi}$, $W^i_{k,candi}$
Output: Target_ID
If Eq. (8-9) is satified, **then**
for i←0 to N **do**
$W^i_{k,candi} = 0$
$d^i_{k,dist}$, $d^i_{k,width}$, $d^i_{k,height}$ computed from Eq. (13-15)
end for
m←$\arg\min_i(d^i_{k,dist})$
$W^m_{k,candi} \leftarrow W^m_{k,candi} + 1$
n←$\arg\min_i(d^i_{k,width})$
$W^n_{k,candi} \leftarrow W^n_{candi} + 1$
o←$\arg\min_i(d^i_{k,height})$
$W^o_{k,candi} \leftarrow W^o_{k,candi} + 1$
p←$\arg\max_i(N^i_{k,candi})$
$W^p_{k,candi} \leftarrow W^p_{k,candi} + 1$
Target_ID ← $\arg\max_i(W^i_{k,candi})$
If Target_ID is not determined, **then**
$d^i_{k,area}$ computed from Eq. (16)
q←$\arg\min_i(d^i_{k,area})$
$W^q_{k,candi} \leftarrow W^q_{k,candi} + 1$
Target_ID ← $\arg\max_i(W^i_{k,candi})$
end if
end if

$$d^i_{k,dist} = \sqrt{(\hat{x}^-_k - x^i_{k,candi})^2 + (\hat{y}^-_k - y^i_{k,candi})^2}, \tag{13}$$

$$d^i_{k,width} = \left| 1 - \frac{w^i_{k,candi}}{\hat{w}^-_k} \right|, \tag{14}$$

$$d^i_{k,height} = \left| 1 - \frac{h^i_{k,candi}}{\hat{h}^-_k} \right|, \tag{15}$$

$$d^i_{k,area} = \left| (\hat{w}^-_k \times \hat{h}^-_k) - (w^i_{k,candi} \times h^i_{k,candi}) \right|. \tag{16}$$

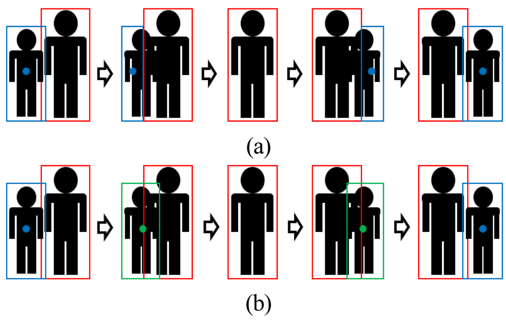
3.3.3 혼잡 상황에서의 추적 과정

[Fig. 2]는 추적 대상이 온전한 상황에서 혼잡 상황으로의 과정을 나타낸다. [Fig. 2(a)]는 중심 좌표를 보정하기 전 추적 과정이고 [Fig. 2(b)]는 제안된 알고리즘을 통해 중심 좌표를 보정한 추적 과정을 나타낸다. 파란색 상자는 추적 대상을 나타내고 빨간색 상자는 추적 대상과 겹쳐지는 대상이다. 파란색 원은 추적 대상의 중심 좌표 $Z_{k,x}$, $Z_{k,y}$를 나타낸다. 초록색 원은 보정된 추적 대상의 중심 좌표 $x_{k,c}$, $y_{k,c}$를 의미한다. 겹쳐지는 대상은 선정된 추적 대상과 거리가 가장 가까운 후보가 선정된다.

추적 과정은 경계 상자만 겹쳐지는 상황, 추적 대상의 일부가 가려져서 경계 상자가 축소되는 상황, 추적 대상이 완전히 가려지는 상황으로 나눠진다.

경계 상자만 겹쳐지는 상황은 영상의 두 프레임으로 판별된다. 추적 대상은 경계 상자의 중심 깊이(Depth)를 활용하여 겹치는 대상보다 뒤에 존재하는지 판단된다. 추적 대상의 경계 상자가 식 (5), (7)을 만족하고 연속된 두 프레임에서 경계 상자들의 겹쳐진 폭이 증가하고 있다면, 해당 상황은 혼잡 상황으로 판단되고 겹쳐지기 직전의 $\hat{x}_k$, $\hat{y}_k$, $\hat{w}_k$, $\hat{h}_k$값을 $\hat{x}_b$, $\hat{y}_b$, $\hat{w}_b$, $\hat{h}_b$에 할당한다.

추적 대상의 일부가 가려져서 경계 상자가 축소되는 상황에서는 추적 대상의 중심 좌표가 실제 움직임을 따라가지 못



[Fig. 2] The compensation effect in a crowded situation. (a) shows the estimated center x-y coordinate of a target without the compensation process. (b) shows the estimated center x-y coordinate of the target that is more accurately tracked by the compensation process

한다. 따라서, 추적 대상이 겹쳐지는 대상보다 좌측에 있다면 식 (17), (19)를 통해 중심 좌표를 보정하고 우측에 있다면 식 (18), (19)를 통해 보정한다.

$$x_{k,c} = Z_{k,x} - \frac{Z_{k,w}}{2} + \frac{\hat{w}_b}{2}, \tag{17}$$

$$x_{k,c} = Z_{k,x} + \frac{Z_{k,w}}{2} - \frac{\hat{w}_b}{2}, \tag{18}$$

$$y_{k,c} = Z_{k,y} - \frac{Z_{k,h}}{2} + \frac{\hat{h}_b}{2}. \tag{19}$$

추적 대상이 좌측에서 진입하여 혼잡 상황이 시작되었다면, 추적 대상은 우측으로 나오거나 좌측으로 나올 수 있다. 우측으로 나온 상황을 판별하기 위해 영상의 두 프레임을 활용한다. 추적 대상이 우측에서 관측되고, 두 경계 상자의 겹쳐진 폭이 연속적으로 감소하고 있다면 추적 대상은 우측으로 빠져나와 혼잡 상황에서 벗어났다고 판단한다. 만약, 추적 대상이 좌측에서 관측되고, $x_{k,c} < \hat{x}_b$를 만족한다면 추적 대상은 좌측으로 빠져나와 혼잡 상황에서 벗어났다고 판단한다.

추적 대상을 관측가능한 혼잡 상황에서 칼만 필터는 선정된 추적 대상의 $x_{k,c}$, $y_{k,c}$, $\triangle x_k$, $\triangle y_k$, $Z_{k,w}$, $Z_{k,h}$를 칼만 필터의 관측치로 사용하여 상태가 갱신 된다. 칼만 필터의 추정치 $\hat{x}_k$와 $\hat{y}_k$는 추적 대상의 최종적인 위치가 된다.

3.3.4 추적 실패 대응 방법

추적을 진행하는 과정 중 Detectron2에서 인간을 검출하지 못했거나, 추적 대상 선별을 위해 Detectron2에서 검출된 경계 상자의 $w^i_{k,candi}$, $h^i_{k,candi}$와 칼만 필터의 예측치 $\hat{w}^-_k$, $\hat{h}^-_k$의 비율이 식 (8), (9)를 만족하지 못한다면 추적 대상 선정에 실패한 것으로 판단한다. 추적에 실패했을 때, 온전한 상황과 혼잡 상황 모두 추적 대상은 예측치만 이용하여 추적된다. 예측된 정보는 다음 프레임에서 Detectron2로부터 검출된 후보군 중 $w^i_{k,candi}$, $h^i_{k,candi}$와 $\hat{w}^-_k$, $\hat{h}^-_k$의 비율이 식 (8), (9)를 만족하는 후보가 존재한다면, 해당 후보가 추적 대상으로 재선정된다. 만약, 비율을 만족하는 후보가 2명 이상일 경우 [Table 2]를 통해 추적 대상이 재선정된다.

4. 실 험

이 장에서는 인간이 혼잡 상황 속에서 제안된 방법과 기존 방법[3,14]의 추적 성능을 검증하였다. 실험에 사용된 파라미터 α는 추적 대상을 판별하는 기준으로 사용되었다. 다음 5장에서 파라미터 선정과 기존 방법의 실험결과를 분석한다.

4.1 성능 평가 방법

제안한 방법의 성능은 식 (20)의 평균 오차를 계산하여 검증하였다. $x^i_{k,true}$, $y^i_{k,true}$, $x^i_{k,est}$, $y^i_{k,est}$는 실제 추적 대상 상반신의 중심 좌표와 각 방법을 통해 추정된 추적 대상의 중심 좌표를 의미한다. $x^i_{k,true}$, $y^i_{k,true}$는 촬영된 영상 속에서 수작업을 통해 추적 대상 상반신의 중심 좌표를 얻었다. n은 촬영된 전체 이미지의 수를 의미한다.

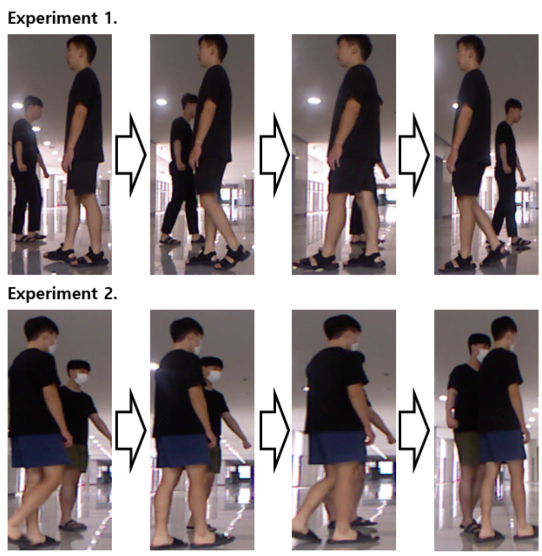
$$E_{tracking} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(x^i_{k,true} - x^i_{k,est})^2 + (y^i_{k,true} - y^i_{k,est})^2}. \tag{20}$$

4.2 실험 환경

실험은 마이크로소프트 社의 RGB-D 센서인 Kinect Xbox360를 통해 촬영된 2 가지의 시나리오로 진행되었다.

[Fig. 3]은 각 시나리오에 따라 진행된 실험 1과 실험 2에서의 혼잡 상황을 나타낸다. 실험 1에서는 추적 대상이 다른 대상과 나란히 걷고 있는 상황에서 추적 대상이 가려지는 상황과 추적 대상은 움직이지 않고 다른 대상의 움직임에 의해 가려지는 상황을 포함한다. 실험 2에서는 추적 대상과 다른 대상이 대각선으로 움직이는 상황에서 추적 대상이 가려지는 상황과 두 대상이 나란히 걷고 있는 상황에서 추적 대상이 가려지는 상황을 포함한다. 실험에 참여한 인간은 3명이고 모두 동일한 색상의 상의를 입고 있다. 영상 초기에는 칼만 필터의 초기값 선정을 위해 추적 대상만 단독으로 촬영하였다. 이후, 나머지 2명이 영상에 등장하고 임의의 방향으로 움직였다.

촬영된 영상에서 인간 검출을 위해 Detectron2에서 제공하



[Fig. 3] The snapshots of the crowded situation in each experiment

는 사전학습 된 Faster R-CNN ResNet101 FPN 모델을 사용하였다. 기준 인간 이미지는 영상 초기에 추적 대상만 단독으로 나오는 상황에서 Detectron2의 경계 상자를 통해 얻어진다.

실험 참여자들은 동일한 색상의 상의를 입고있지만, 하의의 색상은 모두 다르다. 실제 환경에서는 같은 색상의 옷을 입은 사람들이 많이 존재한다. 그렇기에, 같은 색상의 옷을 입고 비슷한 키의 사람이 존재하는 혼잡 상황에서 강인한 추적 성능을 보이고자 색상이 같은 상반신 이미지만 이용해 추적 성능을 검증하였다.

4.3 ORB 특징점 매칭을 통한 추적 실험

기존 방법[14]을 참고하여 Detectron2를 통해 검출된 후보군과 기준 인간 이미지와의 ORB 특징점 매칭을 통해 매칭된 특징점의 개수가 가장 많은 후보를 추적 대상으로 선별한다. 특징점 매칭을 위해 이진 기술자들의 한 종류인 BRIEF에 의해 추출된 기술자들을 해밍 거리를 통해 특징점들을 매칭한다. [Fig. 4]는 기준 인간 이미지와 검출된 후보의 ORB 특징점 매칭 실험 결과 중 일부이다. 빨간색 상자는 실제 추적 대상을 나타내고 파란색 상자는 기존 방법을 통해 선정된 추적 대상을 의미한다. 실험 결과로부터 혼잡 상황이 아님에도 추적에 실패한 것을 알 수 있다.

4.4 색상을 통한 추적실험 및 결과

기존 방법[3]을 참고하여 기준 인간 이미지와 Detectron2를 통해 검출된 후보군의 HS histogram의 상관관계를 식 (21), (22)를 통해 계산한다. H_1, H_2는 기준 인간 이미지와 검출된 후보군의 HS histogram을 나타내고 d는 두 histogram의 상관관계를 나타낸다. d가 1에 가장 가까운 후보가 추적 대상으로 선별된다.

$$d(H_1,H_2) = \frac{\Sigma_I (H_1(I)-\overline{H}_1)(H_2(I)-\overline{H}_2)}{\sqrt{\Sigma_I (H_1(I)-\overline{H}_1)^2 \Sigma_I (H_2(I)-\overline{H}_2)^2}}, \tag{21}$$

(a) (b)

[Fig. 4] Experimental results using ORB feature matching: In most cases, a target is well estimated, as shown in (a). But this method sometimes fails to estimate the target, as shown in (b)

$$\overline{H}_k = \frac{1}{N}\sum_J H_k(J). \tag{22}$$

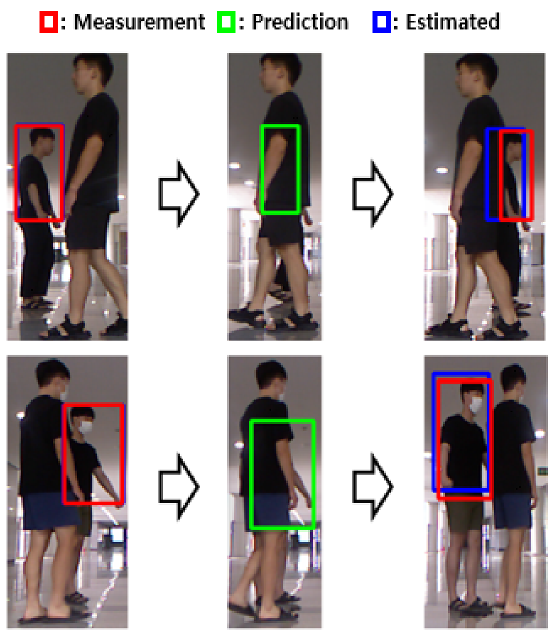
[Fig. 5]는 색상을 통한 추적실험 결과이다. 빨간색 상자는 실제 추적 대상을 나타내고 파란색 상자는 기존 방법[3]을 통해 선정된 추적 대상을 의미한다. ORB 특징점 매칭 실험과 유사하게 혼잡 상황이 아님에도 추적에 실패한 것을 알 수 있다.

4.5 제안된 방법을 통한 추적실험

[Fig. 6]은 본 논문에서 제안된 방법을 통해 혼잡 상황에서의 추적 과정을 나타낸다. 추적 대상이 다른 대상과 겹쳐져 완전히 가려지는 상황에서 추적 대상의 예측치를 통해 추적 대상의 움직임을 따라가는 것을 볼 수 있다. 빨간색 상자는 Detectron2에서 관측된 경계 상자를 나타내고 초록색 상자는

(a) (b)

[Fig. 5] Experimental results using the color feature: In most cases, a target is well estimated, as shown in (a). But this method sometimes fails to estimate the target, as shown in (b)



[Fig. 6] Human tracking results using the proposed method in crowded situations

[Table 3] Performance comparison between the proposed method and the conventional approaches: results of $E_{tracking}$ (unit: pixel)

	Conventional Method[14]	Conventional Method[3]	Proposed Method
Experiment 1	86.7	125.3	28.9
Experiment 2	76.0	119.2	12.6

식 (8), (9)를 만족하지 못해 추적 대상을 놓친 상황에서 추적 대상의 예측된 경계 상자를 나타낸다. 파란색 상자는 갱신된 칼만 필터 추정치의 경계 상자를 나타낸다.

[Table 3]은 본 논문에서 제안된 방법과 기존 방법을 참고하여 구현된 기존 방법들의 성능 평가를 보여준다. 추적 대상과 비슷한 색상의 상의를 입고 있거나 혼잡 상황 속에서 기존 방법들의 추적 정확도는 높지 않았다. 하지만 제안된 방법과 같이 칼만 필터, 경계 상자의 폭, 높이 그리고 중심 좌표 기반의 가중치와 ORB 특징점 기반의 가중치를 함께 이용할 경우 추적 성능이 향상되었다.

5. 분 석

이 장에서는 파라미터 α의 선정 이유와 그 원인을 분석하고, 기존 방법들의 실험결과를 분석하고자 한다.

5.1 추적 대상 판별 기준값

추적 대상 판별 기준값 α는 식 (8), (9)에서 Detectron2를 통해 검출된 후보군의 경계 상자와 칼만 필터에서 예측된 높이와 폭의 비율 경계값을 의미한다. [Table 4]는 추적 대상 판별 기준 값 α에 따른 추적 성능 평가를 보여준다. 실험 1에서 α가 5% 일 때 평균 오차가 크게 발생했다. 실험참가자들은 같은 색상의 상의를 입었고 비슷한 키를 가졌다. 이러한 상황에서 식 (8), (9)를 통해 추적 대상을 선정하는데 α가 5%에서는 추적 대상의 비율 경계값 범위가 좁아서 실제 추적 대상이 선정되지 못했다. 추적 대상의 예측치를 통해 진행하다가 추적 대상이 잘못된 대상으로 복원되었다. 그 후, 칼만 필터는 잘못된 추적 대상의 관측치를 통해 갱신 되었다. 다음 프레임에서 식 (8), (9)를 만족하는 후보가 여러명 존재하더라도 [Table 2]에서 잘못된 추적 대상의 선정되어 5%에서 $E_{tracking}$가 크게 발생했다.

[Table 4] Peformance evaluation of human tracking according to parameter α: results of $E_{tracking}$(unit: pixel)

	$\alpha=5\%$	$\alpha=10\%$	$\alpha=20\%$
Experiment 1	47.7	28.9	30.8
Experiment 2	13.1	12.6	25.8

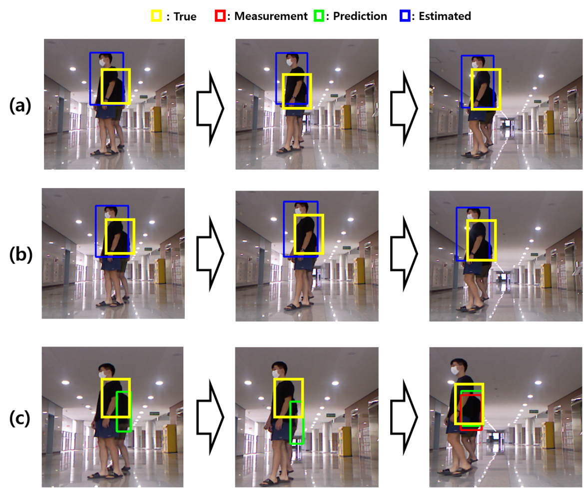
1번 실험과 2번 실험에서 α를 다르게 설정하여 사전 실험을 진행하였고, 두 실험 모두 α가 10%일 때 가장 높은 추적 성능을 보여 제안된 알고리즘을 이용한 실험에서는 α를 10%로 선정 및 적용하였다.

5.2 혼잡 상황에서 기존 방법 및 제안한 방법 성능 비교

ORB 특징점 기반의 방법은 FAST 알고리즘을 통해 영상에서 특징점을 검출한다. FAST 알고리즘은 임의의 픽셀과 주변 16개 픽셀의 밝기를 비교해 연속적으로 임의의 픽셀보다 밝거나 어두우면 특징점으로 검출한다. 기존 방법에서 HS histogram은 색상과 채도를 통해 히스토그램으로 표현한다. 기존 방법들은 빛에 민감한 특징을 가지고 있으며, 실험참가자들의 옷의 색상이 동일하기에 잘못된 추적 대상을 선정할 수 있다.

[Fig. 7]는 추적 대상이 다른 대상과 겹쳐지는 상황에서 기존 방법을 통한 추적 과정을 보인다. 노란색 상자는 실제 추적 대상을 나타낸다. 빨간색 상자는 검출된 후보의 측정치를 나타낸다. 초록색 상자는 추적 대상이 다른 대상에 의해 완전히 가려졌을때 예측치를 나타낸다. 파란색 상자는 기존 방법을 통해 선정된 추적 대상을 나타낸다. [Fig. 7(a)], [Fig. 7(b)]에서 추적 대상이 다른 대상과 완전히 겹쳐져 영상에서 사라졌을 경우 기존 방법에서는 추적 대상과 비슷한 특징을 가진 잘못된 추적 대상을 선정한다.

본 논문에서 제안하는 방법은 칼만 필터, 경계 상자의 폭, 높



[Fig. 7] Experimental results of conventional approaches in the crowded situation: Results using the color feature in (a). Results using the ORB feature in (b). Results using the proposed method in (c)

이 그리고 중심 좌표 기반의 가중치와 ORB 특징점 기반의 가중치를 이용하여 혼잡 상황에서도 대상을 강인하게 추적하였다. 특히, [Fig. 7(c)]에서 비슷한 색상의 대상이 추적 대상과 겹치는 상황에서도 추적에 성공하였다. 2번의 실험을 통해 기존 방법에 비해 추적 오차가 크게 감소함을 알 수 있었다.

6. 결 론

본 논문에서는 혼잡 환경에서 강인한 딥러닝 기반 인간 추적 프레임워크를 제안하였다. 제안된 프레임워크는 인간 미검출과 혼잡 상황 문제를 해결하기위해 각 상황을 인식하고 인간 탐지 및 추적을 위한 기술로 구성된다. Detectron2는 인간 후보군을 추출하기 위한 용도로 활용되었다. 특히, 혼잡 상황에서는 Detectron2의 결과와 추적 대상의 예측된 위치 및 경계 상자 정보를 함께 사용하였다. 대상 추적은 칼만 필터를 사용하여 수행되었다. 실험은 같은 색상의 상의를 입고 비슷한 키의 사람들이 추적 대상과 겹치는 두 가지 형태로 진행되었다. 제안된 방법의 강인한 추적 성능은 실험 결과 및 분석을 통해 검증되었다.

References

[1] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 53, pp. 1-74, 2021, DOI: 10.1186/s40537-021-00444-8.

[2] K. Y. Oh, S. H. Kim, J. S. Kim, Y. Kwon, and S. H. Lee, "An Improved Deep Learning based Human Tracking Framework using Kalman Filter," *KIIT Conference*, Jeju, Korea, pp. 520-522, 2021, [Online], https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE10569022.

[3] R. Algabri and M. Choi, "Deep-Learning-Based Indoor Human Following of Mobile Robot Using Color Feature," *Sensors*, vol. 20, no. 9, pp. 1-19, May, 2020, DOI: 10.3390/s20092699.

[4] C. Schlegel, J. Illmann, H. Jaberg, M. Schuster, and R. Worz, "Vision Based Person Tracking with a Mobile Robot," *British Machine Vision Conference*, Southampton, UK, pp. 418-427, Jan., 1998, DOI: 10.5244/C.12.42.

[5] N. An, S. Y. Sun, X. G. Zhao, and Z. G. Hou, "Online context-based person re-identification and biometric-based action recognition for service robots," *2017 29th Chiness Control and Decision Conference*, Istanbul, Turkey, pp. 3369-3374, Oct., 2017, DOI: 10.1109/CCDC.2017.7979088.

[6] N. Bellotto and H. Hu, "Multisensor-Based Human Detection and Tracking for Mobile Service Robots," *IEEE Transactions on Systems Man, and Cybernetics*, vol. 39, pp. 167-181, Feb., 2009, DOI: 10.1109/TSMCB.2008.2004050.

[7] M. T. Choi, J. S. Yeom, Y. H. Shin, and I. J. Park, "Robot-Assisted ADHD Screening in Diagnostic Process," *Journal of Inteelligent and Robotic Systems*, vol. 95, pp. 351-363, Jun., 2018, DOI: 10.1007/s10846-018-0890-9.

[8] C. J. Seo, "A Study on Multi Target Tracking using HOG and Kalman Filter," *Korean Institute of Electrical Engineers*, vol. 64, no. 3, pp. 187-192, Sep., 2015, DOI: 10.5370/KIEEP.2015.64.3.187.

[9] S. B. Han, S. Y. Na, S. Y. Kim, and D. H. Kim, "Real-time Human Tracking based on Distance and Color," *Korean Institute of Information Scientists and Engineer*s, Jeju, Korea, pp. 1964-1966, 2018, [Online], https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07503558.

[10] S. Y. Park and D. H. Lee, "A Development of Visual Human Tracking System for Path Guidance Robot," *Information and Control Symposium*, Gunsan, Korea, pp. 39-40, 2019, [Online], https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09262782.

[11] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger", *2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 6517-6525, 2016, DOI: 10.1109/CVPR.2017.690.

[12] S. Noor, M. Waqas, M. I. Saleem, and H. N. Minhas, "Automatic Object Tracking and Segmentation Using Unsupervised SiamMask," *IEEE Access*, vol. 9, pp. 106550-106559, July, 2021, DOI: 10.1109/ACCESS.2021.3101054.

[13] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," *2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 2117-2125, 2017, DOI: 10.1109/CVPR.2017.106.

[14] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski "ORB: an efficient alternative to SIFT or SURF," *IEEE International Conference on Computer Vision*, Barcelona, Spain, pp. 2564-2571, 2011, DOI: 10.1109/ICCV.2011.6126544.

오 경 석

2016~현재 금오공과대학교 전자공학부

관심분야: Robot Vision, Image Processing, Deep Learning

김 진 섭

2020~현재 금오공과대학교 전자공학부

관심분야: Image Processing, Deep Learning, Robotics

김 성 현

2016~현재 금오공과대학교 전자공학부

관심분야: Image Processing, Deep Learning, Robotics

이 승 환

2008 경북대학교 전자전기컴퓨터학부(공학사)
2010 서울대학교 전기컴퓨터학부(공학석사)
2015 서울대학교 전기컴퓨터학부(공학박사)
2018~현재 금오공과대학교 전자공학부 조교수

관심분야: Multi-robot SLAM, Multi-robot Coverage Path Planning