

일반논문 (Regular Paper)

방송공학회논문지 제25권 제5호, 2020년 9월 (JBE Vol. 25, No. 5, September 2020)

<https://doi.org/10.5909/JBE.2020.25.5.770>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 360도 영상 공간에서 활성화 뷰포트 기반 이벤트 검출

심 유 정<sup>a)</sup>, 이 명 진<sup>a)†</sup>

### Activated Viewport based Surveillance Event Detection in 360-degree Video

Yoo-jeong Shim<sup>a)</sup> and Myeong-jin Lee<sup>a)†</sup>

#### 요 약

360도 영상은 일반 화각 영상과 다른 프레임 구조를 갖기 때문에 기존 영상 보안 이벤트 검출 알고리즘들을 360도 영상에 바로 적용 시 왜곡에 의한 성능 저하가 발생할 수 있다. 본 논문에서는 360도 영상 공간에서 활성화 뷰포트 기반 이벤트 검출 기법을 제안한다. 다중 객체 영역들을 포함하는 활성화 뷰포트들을 생성하고, 뷰포트 내 객체 검출과 360도 영상 공간에서 객체 추적을 수행하며, 영역 기반의 보안 이벤트를 검출한다. 제안 방법은 360도 ERP 시퀀스들에 대해 성능이 평가되었고, 활성화 뷰포트를 사용하지 않은 이벤트 검출 방법보다 30% 이상의 recall, 30% 이상의 false negative rate 성능 향상을 보였다.

#### Abstract

Since 360-degree ERP frame structure has location-dependent distortion, existing video surveillance algorithms cannot be applied to 360-degree video. In this paper, an activated viewport based event detection method is proposed for 360-degree video. After extracting activated viewports enclosing object candidates, objects are finally detected in the viewports. These objects are tracked in 360-degree video space for region-based event detection. The proposed method is shown to improve the recall and the false negative rate more than 30% compared to the conventional method without activated viewports.

Keywords : 360-degree video, activated viewport, surveillance event detection, object tracking

a) 한국항공대학교 항공전자정보공학과(Dept. of Electronics and Information Engineering, Korea Aerospace University)

† Corresponding Author : 이명진(Myong-jin Lee)

E-mail: artistic@kau.ac.kr

Tel: +82-2-300-0421

ORCID: <https://orcid.org/0000-0002-3136-2819>

※ 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 기초연구사업(2018R1D1A1B07050603)과 2020년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원(N0002431, 2020년 산업전문인력역량강화사업)을 받아 수행된 연구임. (This work was partly supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No. NRF-2018R1D1A1B07050603), and by Korea Institute for Advancemnet of Technology(KAIT) grant funded by the Korea Government(MOTIE)(N0002431, The Competency Development Program for Industry Sprcialist)).

· Manuscript received July 1, 2020; Revised August 28, 2020; Accepted August 28, 2020.

Copyright © 2020 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

## I. 서론

지능형 영상 분석은 영상보안 분야에 적용되어 침입, 배회, 유기, 폭력 등의 이벤트 검출에 사용된다. 영상보안 이벤트 검출 시스템은 객체 검출, 추적, 이벤트 검사 단계를 포함한다. 영상 내 객체 검출은 Convolutional Neural Network (CNN)<sup>[1-6]</sup>, 배경 차분<sup>[7,8]</sup> 방식들이 있다. 영상 내 객체 추적을 위해 프레임별 검출 객체를 시간 방향으로 연결하는 ‘tracking by detection’ 구조가 사용된다. 이 구조로 설계된 지능형 이벤트 검출 시스템에서 객체 검출과 추적 성능은 이벤트 검출 성능에 영향을 미친다<sup>[5-7]</sup>.

기존 일반 화각 영상에 대한 영상보안 이벤트 검출 연구는 활발히 이루어져 왔지만, 최근 활용도가 증가하고 있는 360도 영상 내 보안 이벤트 검출 연구는 거의 진행되지 않고 있다. 360도 영상은 기존 일반 화각 영상과 달리 전 방위를 시청 가능하므로 원하는 시점을 선택하며 시청 가능하며, 시청자의 관심에 따라 한 영상에 여러 뷰포트가 존재한다. 기존 일반 화각 영상에 대한 보안 이벤트 검출 기법들은 360도 영상 포맷의 구조적 왜곡을 고려하지 않고 있으며, 이에 따른 객체 오검출, 추적 성능 저하, 이벤트 검출 성능 저하가 발생할 수 있다<sup>[3-5]</sup>. 360도 영상에서의 객체 검출 방법에는 target plane을 생성하여 객체를 검출하는 방법<sup>[9]</sup>, 영상에 CNN을 적용하기 위해 kernel을 생성하는 방법<sup>[3,4]</sup>이 존재한다. 구면 다각형에 360도 영상을 투영 후 객체 검출을 위한 kernel을 설계하는 방식은 투영 모델의 특성에 따른 객체 특징의 손실 문제가 존재한다<sup>[3]</sup>. 360도 영상과 투영 영상의 kernel간 관계 학습을 통해 CNN 모델을 적용시키는 방식은 작은 객체 검출 문제가 존재할 수 있다<sup>[4]</sup>.

본 논문에서는 360도 영상공간 내 보안 이벤트 검출을

위해 target plane인 객체 후보 영역 중심의 활성화 뷰포트 (activated Viewport; aVP) 생성, 활성화 뷰포트 내 객체 검출, 360도 영상공간 내 객체 추적, 보안 이벤트 검출 방법을 제안한다. 활성화 뷰포트 사용을 통해 시청자가 보는 왜곡이 없는 뷰포트 영상에서의 객체 검출뿐만 아닌 다른 일반 화각 영상 처리 알고리즘 적용에도 사용할 수 있다<sup>[9]</sup>.

## II. 본론

본 절에서 360도 영상에서 침입과 배회 이벤트 검출 방법을 제안한다. 제안 방법은 그림 1과 같이 활성화 뷰포트 생성, 객체 검출 및 추적, 이벤트 검출 단계로 구성된다.

### 1. 360도 영상에서 활성화 뷰포트를 이용한 객체 검출

360도 ERP (equi-rectangular projection) 영상에서 배경 차분과 YOLO 검출기를 이용하여 객체 후보 영역을 검출한다<sup>[1]</sup>. YOLO 검출기는 360도 ERP 프레임에 적용 시 영상 내 왜곡과 고해상도 프레임의 정규화 과정 때문에 성능이 저하될 수 있다. 이를 보완하기 위해 배경차분으로 탐지된 변화 영역을 YOLO 검출 객체 영역과 함께 객체 후보 영역으로 설정한다.

그림 2는 여러 객체를 포함하는 활성화 뷰포트를 검출하는 과정이고, 그림 3은 이 과정의 pseudocode이다. 배경차분으로 라벨링 되거나 YOLO 검출기로 검출한 객체 후보 영역  $l_n$ 의 무게중심에 기설정 화각의 후보 뷰포트  $VP(l_n)$ 를 생성한다. 각 후보 뷰포트는 다른 후보 뷰포트들과 일정 비율 이상 겹치지 않으면 활성화 뷰포트로 확정된다(lines 9-11).

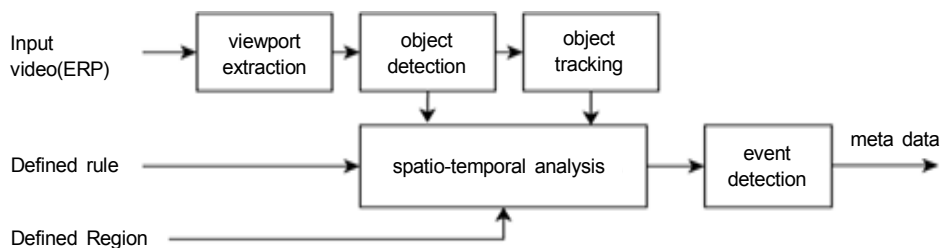


그림 1. 제안하는 360도 영상에서 보안 이벤트 검출 방법  
 Fig. 1. Proposed surveillance event detection in 360-degree video

각 후보 뷰포트가 다른 후보 뷰포트들과 일정 비율 이상 겹치면 이들을 포함하는 **OverlapS** 집합을 구성한다 (line 4-8). 이 집합의 후보 뷰포트의 중심은 집합의 무게중심에 가장 가까운 후보 뷰포트의 무게중심으로 설정한다 (line 12). 이 과정은 프레임 내 좌측에서 우측으로 후보 뷰포트를 스캔하면서 반복해서 수행되며, 현 단계의 후보 뷰포트는 이전 단계의 후보 뷰포트와 일정 비율 이상 겹치지 않았을 때 활성 뷰포트로 확정한다 (line 13).

그림 2(a)에서 임의의 두 후보 뷰포트 사이의 겹침 비율 (**Intersection over Union; IoU**)이 문턱치보다 적어서 각 후보 뷰포트는 활성 뷰포트로 확정된다. 그림 2(b)에서  $VP(l_1)$ 과  $VP(l_2)$ 의 IoU와  $VP(l_1)$ 과  $VP(l_3)$ 의 IoU가 문턱치보다 커서 이들 후보 영역들을 모두 포함하는 하나의

활성 뷰포트를 생성한다.  $VP(l_4)$ 는 다른 후보 뷰포트와 겹치지 않으므로 활성 뷰포트로 확정된다 (lines 8-10). 그림 2(c)에서  $VP(l_1)$  기준으로  $VP(l_2)$ 와  $VP(l_3)$  모두 일정 비율 이상 겹치지만, 기 설정된 활성 뷰포트의 크기가  $l_1, l_2, l_3$ 를 모두 포함할 수 없어서, 좌측부터 스캔하면서  $l_1$ 과  $l_2$ 만을 포함한 활성 뷰포트를 생성한다.  $l_3$ 는 나머지 후보 영역들과 겹침 비율을 따져서 다른 활성 뷰포트에 포함된다.

### 2. 360도 영상 공간에서 객체 추적

추출된 활성 뷰포트 별로 YOLO 검출기를 이용하여 사람 객체를 검출한다. 검출된 객체 영역의 CNN 특징 벡터를 계산하고, 객체의 뷰포트 내 좌표는 360도 ERP 공간 좌표

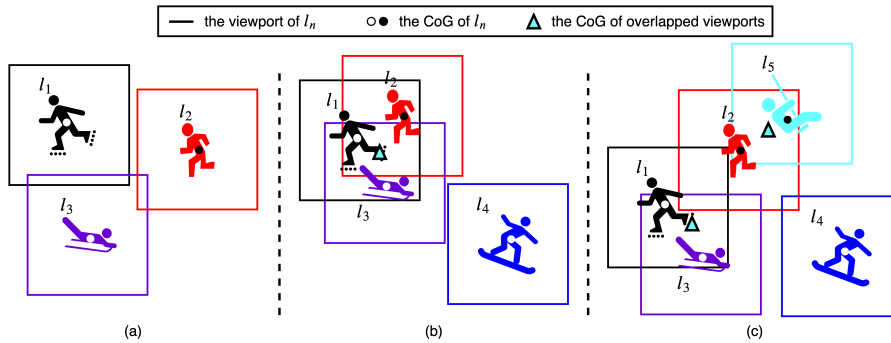


그림 2. 제안하는 활성 뷰포트 생성 방법  
Fig 2. Proposed activated viewport generation method

```

1   $V = \phi$ 
2  for  $i = 1$  to  $|L|$  do
3     $OverlapS = \phi$ 
4    for  $j = i+1$  to  $|L|$  do
5      if  $VP\_IoU(l_i, l_j) > 0$  then
6        if  $l_i \notin OverlapS$  then  $OverlapS = OverlapS \cup \{l_i\}$ 
7         $OverlapS = OverlapS \cup \{l_j\}$ 
8      end
9    if  $OverlapS == \phi$  then
10      $V = V \cup VP(l_i)$ 
11   else
12      $l_{cog} = FindRepresentative(OverlapS)$ 
13     if  $VP(l_{cog}) \notin V$  or  $IoU(VP(l_{cog}), V_{|V|-1}) == 0$ 
14       then  $V = V \cup VP(l_{cog})$ 
15   end
16    $OverlapS = \phi$ 
17 end
18 return  $V$ 

```

그림 3. 제안하는 활성 뷰포트 추출 알고리즘  
Fig. 3. Proposed activated viewport extraction algorithm

로 변환된다. 이들 정보를 이용하여 360도 ERP 공간에서 CNN 특징 벡터와 data association 기법에 기반한 객체 추적을 수행한다<sup>[2]</sup>.

### 3. 360도 영상에서의 이벤트 검출 방식

본 절에서는 침입과 배회 이벤트 검출 방법을 제안한다. 표 1은 이벤트별 감지 영역과 규칙을 정의한 것이다. 영역 기반의 이벤트 검출을 위하여 각 이벤트를 감지하기 위한 영역을 선정한다. 그림 4는 360도 ERP 프레임에 침입과 배회 이벤트 감지 영역을 표시한 영상이다.



그림 4. 영상 내 감지 영역 설정 예시  
 Fig. 4. The example of setting up an event detection zone

제안방법은 추적된 객체의 바운딩 박스 기준으로 객체의 감지 영역 포함 비율과 포함 여부를 다음과 같이 판단한다.

$$R = \frac{S(A \cap B)}{S(B)} \quad (1)$$

$$I(O_i) = \begin{cases} True, & R \geq 1.0 \\ False, & R < 1.0 \end{cases} \quad (2)$$

여기에서 A는 감지 영역, B는 객체 영역,  $S(\cdot)$ 는 영역의 넓이를 나타낸다.

침입은 1명 이상의 사람 객체가 감지 영역에 진입한 경우이다. 침입 이벤트는 식(2)의 영역 포함 여부만으로 검출된다. 배회는 1명 이상의 사람 객체가 감지 영역에서 10초 이상 방향성 없이 이동하는 경우이다. 배회 감지 영역에서 객체가 일정 시간 이상 정지하면 배회로 판단한다. 또한, 감지 영역 내에서 한 방향으로 일관된 이동이 아닌 여러 방향으로 이동하는 객체 존재 시 배회로 판단한다<sup>[9]</sup>.

일관된 이동의 방향성을 판단하기 위해 감지 영역 진입 시와 현재 위치 사이의 magnitude of displacement (MoD)를 다음과 같이 정의한다.

$$MoD = \sqrt{L_s^2 - L_c^2} \quad (3)$$

객체  $O_i$ 의 배회 여부는 다음과 같이 판단한다.

$$LT(O_i) = \begin{cases} True, & t(O_i) > th_t \text{ and } N(MoD) > th_d \\ False, & otherwise \end{cases} \quad (4)$$

여기에서  $N(MoD)$ 은 일정 크기 이상으로  $MoD$  감소하는 경우의 횟수를 나타내며,  $t(\cdot)$ 는 객체의 감지 영역 체류시간이다.  $L_s$ 와  $L_c$ 는 객체의 감지 영역 진입 위치와 감지 영역 내 현재 위치를 나타낸다.  $th_t$ 는 객체의 감지 영역 내 배회 시간 임계 값이며,  $th_d$ 는 객체의 방향 전환 여부에 따른 객체의 방향 기준 배회 판단 임계 값이다.

그림 5는 객체의 이동 궤적과 변위를 나타낸 그림이다. 그림 5(a)의 객체는 직선으로 이동하지는 않으나, 특정 방향으로 방향성을 가지고 지속적으로 이동하며, 객체의  $MoD$ 는 계속 증가한다. 따라서  $N(MoD)$ 는 변화하지 않는다. 그러나 배회의 정의로부터 감지 영역 내에서 방향성 없이 이동하는 경우는 그림 5(b)와 유사한 이동 궤적을 보인다. 그림 5(b)의 객체는 시작점에서 멀어졌다가 다시 돌아오는 경우이며,  $MoD$ 는 증가했다 감소하므로  $N(MoD)$ 는 증가한다.

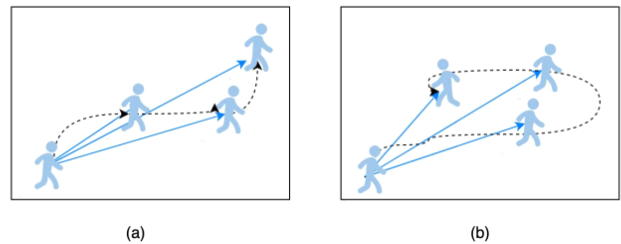


그림 5. 객체의 이동궤적과 변위  
 Fig. 5. The moving trajectories and displacement of the object

### III. 실험 및 실험 결과

제안 방법의 성능 평가를 위해 3840x1920 해상도의 360

도 ERP 시퀀스 11개에 대해 실험을 진행하였다. ERP 프레임에서 5x15 이상 크기의 사람 객체 GT를 생성하였다.

표 1은 활성 뷰포트를 사용하여 360도 영상에서의 객체 추적을 진행하는 경우의 성능을 활성 뷰포트를 사용하지 않는 경우와 비교한 것이다. 활성 뷰포트의 사용으로 인해 영상에서의 객체 재현율이 높아지는 것을 볼 수 있으며, 전반적인 객체 검출 성능이 증가했음을 보여준다.

표 2는 전체 시퀀스들에 대한 활성 뷰포트를 사용하는 제안 방식과 사용하지 않는 기존 방식의 성능을 비교한 것이다. 이벤트 검출 성능은 객체 검출 성능에 의해 크게 영향을 받는다. 기존 방식은 제안 방식보다 전체적으로 낮은 recall rate를 보이는데, 이는 360도 영상 내 객체의 크기가 작고 객체 위치에 따라 다른 왜곡이 존재하기 때문에 객체 검출이 제대로 이루어지지 않았기 때문이다. 제안 방식은 이와 같은 왜곡을 피하기 위해 객체 후보 영역들로부터 활성 뷰포트를 검출하여 객체를 검출하였기 때문에 높은 이벤트 검출 성능을 보인다.

표 1. 활성 뷰포트 사용 여부에 따른 객체 검출 성능 비교

Table 1. Object tracking performance comparison based on usage of aVP

Tracker	recall	prcn	MOTA <sup>1)</sup>
without aVP	52.0	86.8	43.7
with aVP(proposed)	80.8	83.4	63.5

표 2. 활성 뷰포트 사용 여부에 따른 성능 비교

Table 2. Performance comparison based on usage of aVP

		GT	FP	TP	FN	prcn <sup>2)</sup>	recall <sup>3)</sup>	FNR <sup>4)</sup>
without aVP	Loitering	34	0	14	20	1.00	0.41	0.59
	Intrusion	32	0	20	10	1.00	0.67	0.33
with aVP (proposed)	Loitering	34	2	32	1	0.94	0.97	0.03
	Intrusion	32	0	31	1	1.00	0.97	0.03

그림 6은 뷰포트 사용 여부에 따른 객체 검출 결과 이미지를 일부 확대한 모습이다. 그림 6(a)는 기존 방식의 결과

이며, 그림 6(b)는 활성 뷰포트를 사용한 제안 방식의 결과이다. 우측 상단 객체는 시간에 따라 계단을 오르는데, 이는 360도 ERP 프레임에서 적도에서 극 지역으로 객체가 이동하는 경우로서 객체의 왜곡이 증가한다. 검출 객체는 하얀 바운딩 박스로, 추적 중 객체는 파란색 바운딩 박스로 표시된다. 그림 6(a)의 좌측 하단 객체는 검출되었지만, 이전 검출 객체와 연결되지 않았으며, 우측 상단의 객체는 검출과 추적 모두 이루어지지 않았다. 그림 6(b)에서 좌-우 객체 모두 추적 중이며, 우측 상단의 객체가 계단을 오르는 동안 왜곡의 영향 없이 계속 추적되었다.



(a) (b)

그림 6. 활성 뷰포트 사용 여부에 따른 결과 화면

Fig. 6. The result images based on usage of aVP

#### IV. 결론

360도 영상 공간에서 활성 뷰포트 기반 보안 이벤트 검출 기법을 제안하였다. 360도 영상 내 위치에 따른 검출 성능 저하를 막기 위해 객체들을 포함하는 활성 뷰포트들을 생성하고, 뷰포트 내 객체 검출과 360도 영상 공간에서 객체 추적을 수행하였다. 실험을 통해 제안 방식이 360도 영상 내 보안 이벤트 검출에 높은 성능을 보임을 확인하였다.

#### 참고 문헌 (References)

[1] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement,"

1) Multiple Object Tracking Accuracy(GT 대비 추적 결과의 FP, missed target, ID switch의 비율)[10]

2) TP/(TP+FP) 모델이 True 라고 분류한 것 중에 실제 True인 것의 비율

3) TP/(TP+FN) 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율

4) FN/(TP+FN) 실제 True인 것을 모델이 True가 아니라고 분류한 것의 비율

- arXiv preprint arXiv:1804.02767, 2018.
- [2] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," *IEEE International Conference on Image Processing*, pp. 3464-3468, Sep 2017.
- [3] Lee, Yeonkun, Jaeseok Jeong, Jongseob Yun, Wonjune Cho, and Kuk-Jin Yoon, "SpherePHD: Applying CNNs on a Spherical PolyHeDron Representation of 360° Images," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9181-9189, 2019.
- [4] Su, Yu-Chuan, and Kristen Grauman. "Kernel transformer networks for compact spherical convolution," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9442-9451, 2019.
- [5] Yoo-jeong Shim, Min-seok Lee, Seok-ho Baek, Myeong-jin Lee. "Multi-object tracking method and system for 360-degree video," *Summer Annual conference of IEIE*, Jun:700-701, 2019.
- [6] Nayak, Rashmiranjan, Mohini Mohan Behera, Vaisyaraju Girish, Umesh Chandra Pati, and Santos Kumar Das, "Deep Learning Based Loitering Detection System Using Multi-Camera Video Surveillance Network," *IEEE International Symposium on Smart Electronic Systems*, pp. 215-220, 2019.
- [7] Park, Keon-woo, Doo-sik Kang, Jun-sik Kim, and Myeong-jin Lee, "Loitering Detection based on Regional Histogram and Object Velocity," *Summer Annual conference of KICS*, pp.1304-1305, 2016.
- [8] Kim, Seajin, Byung-jin Lee, Jae-won Jeong, and Myeong-jin Lee. "Multi-object tracking coprocessor for multi-channel embedded DVR systems," *IEEE Transactions on Consumer Electronics*, pp.1366-1374, 2012.
- [9] Korea Internet&Security Agency, "Intelligent CCTV Solution Performance Test and Certification Guide," pp. 30, 2018.
- [10] Bernardin, Keni, and Rainer Stiefelagen. "Evaluating multiple object tracking performance: the CLEAR MOT metrics," *EURASIP Journal on Image and Video Processing*, pp. 1-10, Dec 2008.

---

저 자 소 개



심 유 정

- 2019년 2월 : 한국항공대학교 항공전자정보공학부 학사
- 2019년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사과정
- ORCID : <https://orcid.org/0000-0003-0385-8101>
- 주관심분야 : 영상처리, 임베디드시스템



이 명 진

- 2001년 8월 : KAIST 전자전신학부 박사
- 2001년 3월 ~ 2004년 2월 : 삼성전자 System LSI 사업부 책임
- 2004년 3월 ~ 2007년 2월 : 경상대학교 전기전자공학전공 조교수
- 2007년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
- ORCID : <https://orcid.org/0000-0002-3136-2819>
- 주관심분야 : 영상통신, 영상처리, 임베디드시스템