

# 비디오에서 동체의 행위인지를 위한 효율적 학습 단위에 관한 연구

권익환<sup>†</sup>, 부베나 하제르<sup>\*\*</sup>, 이도훈<sup>\*\*\*</sup>

## A Study on Efficient Learning Units for Behavior-Recognition of People in Video

Ick-Hwan Kwon<sup>†</sup>, Boubenna Hadjer<sup>\*\*</sup>, Dohoon Lee<sup>\*\*\*</sup>

### ABSTRACT

Behavior of intelligent video surveillance system is recognized by analyzing the pattern of the object of interest by using the frame information of video inputted from the camera and analyzes the behavior. Detection of object's certain behaviors in the crowd has become a critical problem because in the event of terror strikes. Recognition of object's certain behaviors is an important but difficult problem in the area of computer vision. As the realization of big data utilizing machine learning, data mining techniques, the amount of video through the CCTV, Smart-phone and Drone's video has increased dramatically. In this paper, we propose a multiple-sliding window method to recognize the cumulative change as one piece in order to improve the accuracy of the recognition. The experimental results demonstrated the method was robust and efficient learning units in the classification of certain behaviors.

**Key words:** Frame Learning Units, Behavior Recognition, Multiple Support Vector Machine, Multi Sliding Window

### 1. 서 론

최근 민간영역, 공공영역 할 것 없이 감시카메라의 설치가 증가하면서 효율적인 모니터링을 하기 위한 객체의 행동 인지기술에 관심이 높아지고 있다. 범죄가 발생했을 때 동선 확보 등 추적이 쉽기 때문에 개인적, 사회적 대처방안으로 예방 목적으로 설치하는 경우 또한 늘고 있다[1]. 행정자치부조사에 따르면 지자체 감시카메라는 총 12만 5608대이며 이

중 절반이상인 9만 411대가 방범용으로 운용되고 있다[2]. 이처럼 감시카메라의 설치가 증가함에 따라 일반적으로 관리자가 육안으로 모니터링 하는 것에서 효율적인 감시를 위한 지능형 감시시스템의 필요성이 증가하고 있으며 많은 연구가 진행되고 있을 만큼 큰 이슈이다[3,4]. 이 시스템을 구현하기 위해서는 영상처리와 패턴분석 그리고 인공지능 분야의 기계학습을 통한 자동화 기술이 요구된다. 따라서 수집된 많은 양의 영상데이터는 학습을 통해 영상 자체에

\* Corresponding Author : Dohoon Lee, Address: (46241) Research & Lab. Bldg. #1 312 Pusan National University 63-2, Busandeahak St. Geumjeong-Gu, Busan, Korea, TEL : +82-51-510-2491, FAX : +82-51-517-2431, E-mail : dohoon@pusan.ac.kr

Receipt date : Oct. 8, 2016, Revision date : Jan. 18, 2017  
Approval date : Jan. 25, 2017

<sup>†</sup> Electrics and Telecommunications Research Institute  
(E-mail : ickhwan@etri.re.kr)

<sup>\*\*</sup> Dept. of Electrical and Computer Engineering, Graduate School, Pusan National University  
(E-mail : hadjer@pusan.ac.kr)

<sup>\*\*\*</sup> Dept. of Electrical and Computer Science Engineering, Pusan National University

\* This work was supported by a 2-Year Research Grant of Pusan National University(2015.3-2017.2) and partially supported by BK21PLUS, Creative Human Resource.

대한 인식의 범위를 넘어 의미 있는 정보와 내용분석을 통한 더욱 발전된 기능을 제공한다[5,6].

지능형 영상감시시스템에서 행동인식은 카메라로부터 입력되는 영상정보를 이용하여 관심객체의 위치와 패턴을 분석하여 행위에 따른 행동을 분석한다. 이와 같은 행동인식은 크게 세 단계로 구분한다. 입력되는 영상에서 관심 영역의 객체를 감지하고 추적하는 전처리과정과 객체의 행동을 모델링하고 학습에 필요한 특징정보를 얻는 행위모델링, 추출한 특징정보를 이용하여 발생하는 이벤트를 예측하고 최종적으로 행동을 인식하게 된다[7]. 행위를 인지하기 위해 키넥트와 같은 깊이 카메라를 이용하거나 일반 카메라에서 인체를 부위별로 찾아 뼈대를 구성해 인지하는 방법이 있었다. 기존의 뼈대기반의 방법은 신체의 각 부위 움직임으로부터 다양한 행위를 정의할 수 있는 장점을 가지고 있지만 군중이 많은 영상에서 분석이 어렵고 연산 요구량이 증가한다[8]. 따라서 객체의 형태나 위치, 이동궤적 등의 간단한 특징으로 행위를 인지하려는 연구들이 있었다[9,10]. 대부분의 SVM(support vector machine)을 이용한 기계학습 방식의 행위분류는 행위전체를 하나의 단위행위로 보고 학습하는 기법으로 적용되어 사용되고 있다. 이는 객체를 전체적인 시각에서 모델링하므로 간단하고 연산량이 적어 주어진 자원 활용에 효율적이다[11].

본 논문에서는 행위 인지의 정확도를 향상시키기 위해 누적된 변화량을 하나의 정보로 인식하는 멀티슬라이딩 윈도우 방법을 제안한다. 행위를 분류하기 위해 기계학습을 이용하여 비교적 간단한 특징정보를 추출하였다. 행위 시퀀스에 대한 효율적 학습 방법으로 객체의 특정행위에 대응하는 특징정보는 객체 이동 궤적이 존재하는 프레임 시작부터 종료시점까지를 하나의 단위로 구성한다. 프레임율이 높은 경우 불필요한 정보를 계속적으로 획득하게 되며 이는 행위를 인지하는데 있어서 방해요소로 존재한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 연구의 전처리과정과 특징추출, 행위인식 모델에 대해 기존 연구와 비교한 관련연구를 알아본다. 3장에서는 전처리과정을 통해 얻은 행위시퀀스에서 효율적인 학습단위와 오류보정에 대해 설명하고 4장에서는 단위움직임에 따른 행위분류 실험 결과를 나타낸다. 마지막으로 5장에서는 실험결과에 따른 결론 및 향후

진행할 연구 방향을 제시한다.

## 2. 관련연구

움직이는 객체의 정보를 얻기 위해 객체검출 및 추적은 전처리과정으로 기본이 되는 과정이다. 영상에서 움직이는 객체를 검출하기 위해서는 차분영상을 이용하거나 신체정보를 분석하는 방법이 있다. 이때, 영상에서 전경과 배경을 분리하여 객체를 찾아내는 것이 무엇보다 중요하며 크기나 형태의 변화 등을 이용해 특징정보로 사용한다[12]. 관심영역 내에 존재하는 객체는 움직임, 위치, 실루엣, 크기 정보 등을 가지며 모델링 과정에서 객체의 크기나 위치정보는 카메라의 설치 위치나 각도에 따라 다르게 표현되므로 이동하는 객체를 나타내는 크기, 위치정보를 매 프레임마다 지속해서 갱신해야한다.

기존 연구에서 S.Velastin과 N.Bird[13,14]는 화소 단위의 처리와 움직이는 객체의 정보를 결합한 방법으로 배경 영상에서 객체를 찾아낼 때 가장 많이 쓰이는 가우시안 혼합모델(Gaussian Mixture Model)을 사용하여 배경영상과 분리된 객체에 대한 움직임 화소 정보를 이용하였다. 하지만, 이동하는 객체와 배경을 분리하는 단계에서 객체의 수가 많은 경우 겹침이나 잡음 등으로 움직임 위치가 변화되어 다소 부정확한 객체 검출 결과를 갖게 되는 단점을 가지고 있다.

P.Spagnolo[15]는 사라진 물체를 검출하기 위해서 연속되는 프레임의 평균을 이용하여 배경을 생성하였다. 하지만 배경을 생성한 이후 지속해서 갱신을 하지 않아 물체 또는 객체를 정확하게 검출하는 것이 어렵고 예지결과에만 의존하여 복잡한 환경에서 부정확한 결과를 가진다.

E.Yu와 A.Singh[16,17,18]은 사람의 신체 부위를 찾기 위해 실루엣으로부터 뼈대에 기반을 둔 행위로 객체의 행위인식 방법을 제안하였다. 미리 정의된 행위 및 행동의 집합을 이용하여 손동작이나 몸짓을 이용한 여러 가지 행동을 인식할 수 있는 장점이 있지만 객체의 수가 많은 복잡한 환경에서는 연산량이 많아 적용하기 어려운 단점이 있다.

영상에서 객체의 행위는 크게 정상적인 행위와 비정상적인 행위로 구분된다. 정상행위는 일반적인 보행자를 의미하고 비정상 행위는 배회, 실신, 월담 등이 있으며 개인의 시각에 따라 차이가 있다. 비정상

행위를 분석하기 위해 특징정보는 객체의 무게중심, 이동궤적, 실루엣, 크기 정보 등으로 표현된다. 각각 정의된 데이터는 영상에서 객체를 판단하는 기준이 되고 미리 정해진 패턴으로 구분한다. 대부분의 비정상적인 행위의 판단은 움직이는 객체 검출 이후 지속적인 갱신이 이루어져야 원하는 대상으로 분석할 수 있다. Table 1은 객체의 특정 특징정보를 이용하여 비정상적인 행위를 검출하는 알고리즘의 사례를 보여주고 있다.

주변을 목적이 없이 돌아다니는 행위감지는 행위 시퀀스에서 자주 나타나는 패턴이 있기 때문에 이러한 행위를 사전에 감지하고 대응할 수 있다. Table 2는 특정행위를 검출하기 위해서 특정영역에 장시간 반복적으로 머무르는 객체의 배회행위를 검출하기 위한 연구를 보여준다. Lee[19]는 영상에서 객체가 동일한 영역에 궤적이 반복적으로 발생할 때 배회행위의 특성으로 간주하여 검출하는 알고리즘을 제안하였다. 영상을 특정크기의 블록으로 나누어 이동 궤적을 관찰하는 블록기반의 분석 방법을 제안하였다. Thi Thi Zin[20]은 전경과 배경을 분리하여 얻어진 객체의 궤적정보를 토대로 외형 모델 특징과 움직임 모델 특징을 추출하여 배회행위를 검출하였다. Huang [21]은 베이지안(Bayesian) 추적기를 이용하여 움직이는 객체를 추적한 후 영상에 등장하는 객체들의 방문기록을 테이블로 구성한 LV(List of Visitors)을 통하여 배회행위를 검출하는 알고리즘을 제안하였다. 이러한 방법은 영상에서 2차원 궤적으로 분석하

는 방법을 사용한 반면 Park[22]은 움직이는 객체의 실루엣을 이용하여 그림자를 제거하고 카메라 캘리브레이션(calibration)을 이용하여 3차원 궤적분석을 통해 배회행위를 검출하는 방법을 사용하였다.

감시영상에서 불특정한 규칙으로 이동하는 객체를 검출하는 것은 다른 이상행위와 다르게 사고를 미연에 방지할 수 있어 중요한 이슈이다. 일상생활에서 일어나는 불특정행위에서 배회의 사전적 의미는 '어슬렁거리며 정처 없이 이리저리 돌아다니다.'의 의미를 가지고 있다. 명확한 패턴의 규칙과 규정을 내릴 수 없어 알 수 없는 객체의 행위를 판단하는 연구는 범죄 예방에 중요하다고 볼 수 있다. 따라서 본 논문에서는 다양하게 나타나는 배회행위를 인지하기 위해 불특정 패턴 규칙을 가진 영상을 학습 및 실험데이터로 사용하였다.

### 3. 영상프레임 특징정보추출

#### 3.1 제안한 알고리즘의 개요

본 논문에서는 단위 움직임 정보를 정의하고 학습하기 위해서 다중으로 구성된 SVM 기반의 행위인지 알고리즘을 제안한다. Fig. 1은 행위인지 시스템의 전체적인 흐름도를 나타내고 있다.

입력 영상에서 움직이는 객체를 검출하기 위해서 배경 모델링 기법을 이용하여 배경과 전경을 분리한 객체를 검출한다. 검출된 객체의 이동궤적에서 특징정보를 추출하고 추출된 움직임 정보는 이전프레임

Table 1. Various cases of behavior recognition algorithms

Algorithms	Used feature information	Targeted behavior
S.Valastin[13]	Position, Moving, Edge, Silhouette of object	Loitering, Abandoned object
N.Bird[14]	Size, Silhouette, Color of clothes	Loitering
P.Spagnolo[15]	Moving, Haar like feature	Lost and Abandoned object
E.Yu[16]	Silhouette	Jump over wall
A.Singh[17][18]	Position, Silhouette, Background	Abandoned object

Table 2. Various cases of loitering behavior recognition algorithms

Algorithms	Used feature information	Targeted behavior
Lee[19]	Block-based trajectory of object	Loitering
Thi Thi Zin[20]	Markov random walk model	Loitering
Huang[21]	Bayesian tracking, List table of visitors	Loitering
Park[22]	Remove Silhouette, Camera calibration	Loitering

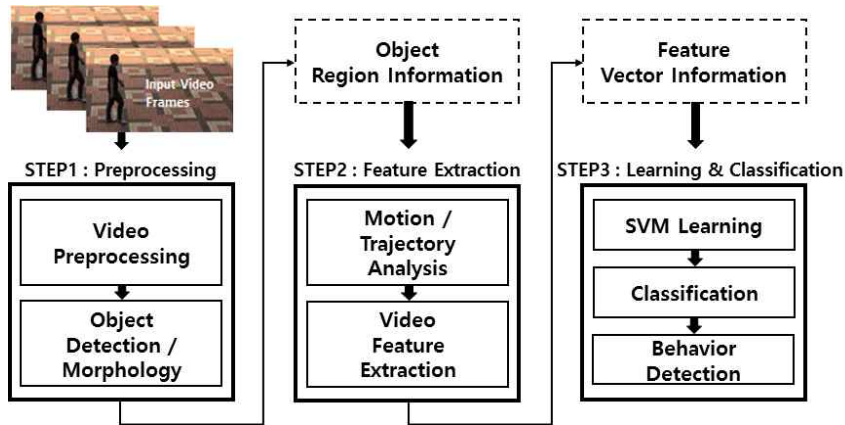


Fig. 1. System for behavior recognition algorithm.

과 현재프레임 사이의 이동변화량인 객체의 중심점  $x, y$  변화량을 이용한다. 두 특징정보를 생성한 후 프레임에 따른 특징벡터를 생성하였다. 움직이는 객체의 이동궤적을 단위 움직임으로 보고 대표적인 기계 학습 알고리즘인 SVM을 이용하여 임의의 특정패턴 행위를 학습하여 분류하였다.

### 3.2 객체 검출을 위한 전처리와 특징정보

객체를 검출하기 위해서 배경과 움직이는 객체의 검출은 Bobick과 Davis[23]가 제안한 모델링된 배경 프레임과 영상에서 현재프레임의 차분 방법을 참고하고 평균 배경 모델을 적용하였다. 대부분의 배경과 전경 분리 알고리즘으로 객체를 검출하기 위해 많이 사용하고 있다.

차분 영상은 배경과 객체로 분리된 영상이며 입력 영상에서 전체 화소를 계산하여 R,G,B 세 결과가 특정 임계값 사이에 있다면 화소를 0으로 그렇지 않으면 255로 한다. 하지만 객체를 추출한 영상에서 미세영역을 객체로 인식하거나 소금과 후추 잡음이 생기는 경우가 있다. 따라서 모폴로지 필터를 이용하여 미세영역과 잡음을 제거한 후 배경과 객체를 분리한 영상  $Background(i,j)$ 를 얻었다.

$$Background(i,j) = \begin{cases} 0, & Threshold_l < B(i,j) < Threshold_h \\ 255, & otherwise \end{cases} \quad (1)$$

움직이는 객체의 추적을 위해서 객체의 윤곽선정보를 이용하였다. 이를 이용하여 객체의 이동정보, 크기비율, 중심점의 변화량 등의 정보를 획득하였다. 본 연구에서는 객체의 단위움직임 정보로 중심점을

이용하였다. 객체가 움직일 수 있는 다양한 방향을 특징으로 하기 위해서 Fig. 2와 같이 객체가 움직이는 변화량을 이용한다. 객체의 중심점  $G_x$ 와  $G_y$ 는 다음 식 (2)와 같이 구하며  $n$ 은 객체의 화소(pixel),  $x, y$ 는  $i$ 번째 화소의 위치를 나타낸다.

$$G_x = \sum_{i=1}^n \frac{x_i}{n}, G_y = \sum_{i=1}^n \frac{y_i}{n} \quad (2)$$

$x, y$ 위치의 중심점의 변화량은 다음과 같이 현재 프레임과 이전 프레임의 차이 값,  $CondG_x$ 와  $CondG_y$ 를 특징정보로 이용하여 다음 식(3)과 같이 구한다.

$$\begin{aligned} CondG_x &= CurrentG_x - PreviousG_x, \\ CondG_y &= CurrentG_y - PreviousG_y \end{aligned} \quad (3)$$

### 3.3 슬라이딩윈도우에 따른 프레임분할

입력영상에서 중심점의 이동변화를 특징으로 추출하기 위해서 1초에 동일한 크기의 프레임으로 추출하는 것이 핵심이다. 실제 영상에서는 객체의 움직

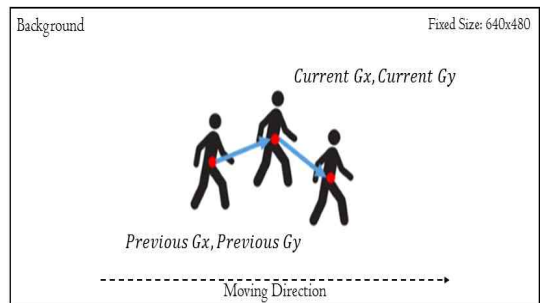


Fig. 2. Feature information between previous frame and current frame.

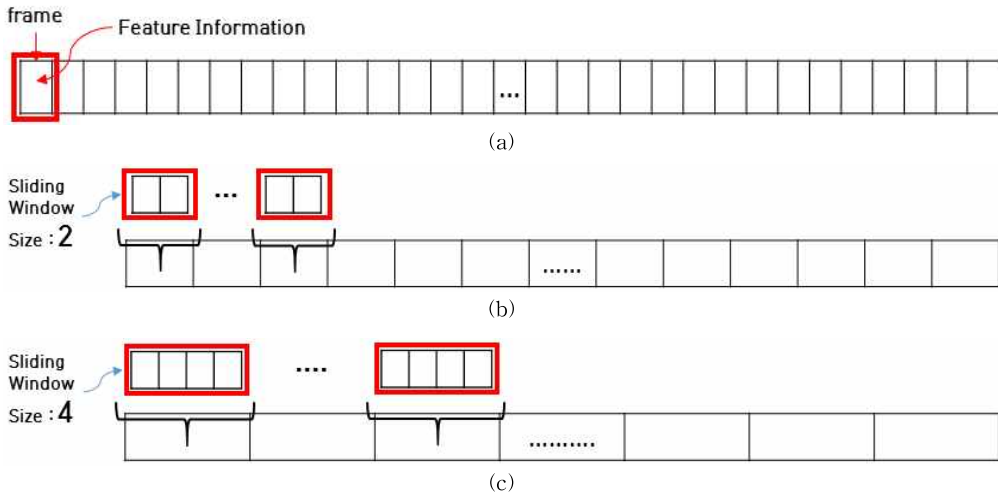


Fig. 3. The proposed method using multi sliding window. (a) Frame sequence, (b) Ex. When the window size is 2, (c) Ex. When the window size is 4.

임에 따른 중심점의 이동 변화량이 많아 방해요소가 포함된 데이터의 수가 많아지기 때문에 프레임 시퀀스에서 멀티슬라이딩 윈도우 방식을 적용하였다. 누적된 변화량 특징정보를 각각 여러 개의 프레임의 특징정보를 하나의 프레임으로 보고 누적합의 평균으로 나타내었다. Fig. 3은 행위패턴을 가진 일련의 행위시퀀스에서 분할된 윈도우의 크기에 따라 여러 개의 프레임을 하나로 샘플링 하는 것을 나타낸다.

여기서  $|w|$ 는 윈도우의 크기를 나타내며 누적평균 값  $\Delta x_k, \Delta y_k$ 은 다음 식 (4)와 같다.

$$\Delta x_k = \frac{\sum_i^{|w|} \Delta x_i}{|w|}, \Delta y_k = \frac{\sum_i^{|w|} \Delta y_i}{|w|} \quad (4)$$

멀티 슬라이딩 윈도우는 윈도우의 크기만큼 이동하며 누적된 특징정보의 평균으로 하나의 프레임으로 본다. 영상에서 30fps 행동프레임은 특징에서 객체가 30프레임(1초)마다 방향을 바꾸는 것은 비현실적이며 같은 값이 연속되거나 잡음이 생기기므로 불필요하다. 이는 분류를 위한 학습시간과 인식률에 영향을 미칠 수 있다. 따라서 하나의 행위시퀀스를 분할하고 샘플링한 특징데이터를 기반으로 학습에 사용하고 가장 적합하고 효율적인 윈도우 크기를 찾는다.

#### 4. 실험 및 결과

본 실험은 Intel(R) Core(TM) i5-4330 CPU @3.00

GHz의 PC와 Windows 10 Home의 64비트 운영체제에서 수행했으며 사용된 프로그래밍언어는 Visual Studio2013 C++(MFC)을 기반으로 OpenCV2.4.11버전과 LibSVM의 일부를 사용하였다.

실험에 사용된 영상은 현실적으로 데이터의 양이 충분하지 않은 점을 감안하여 다중 SVM 분류기의 학습에 따른 성능을 평가하기 위해 10 fold-cross validation의 방법을 채택하였다. 샘플데이터를 10개의 부분집합으로 학습시키고 나머지 한 개의 부분집합으로 학습된 분류기의 성능을 측정하였다. 한편, 실험을 위해 움직이는 객체의 중심점 이동 변화량을 포함하는 일련의 행위패턴을 갖는 시퀀스를 윈도우 크기를 조금씩 늘려가면서 실험하였다.

##### 4.1 실험에 사용된 단위 움직임 정의

본 실험에서는 행위를 인지하기 위해 단위 움직임을 학습하고 정의한다. 실험에 설정된 단위 움직임은 배회행위로 미리 가정하였다. 배회행위는 Table 3과 같은 형태를 가지는 영상으로 설정하였다. B<sub>1</sub>, B<sub>2</sub>, B<sub>3</sub> 각각의 행위영상은 300개로 총 900개의 영상을 실험에 사용하였다.

##### 4.2 슬라이딩 윈도우 크기 변화에 따른 실험결과

실험을 위해 슬라이딩 윈도우의 크기를 3, 5, 7, 9, 10, 11, 13, 15, 20, 30으로 분할하여 샘플링 하였다.

Table 3. Experimental set of loitering behavior

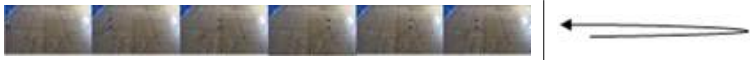


Behavior	Pattern of Actions	The number of input video
B <sub>1</sub>		300
B <sub>2</sub>		300
B <sub>3</sub>		300

Table 4, 5, 6는 각각 미리 정의한 행위에 따른 인식률을 나타내고 있다. 슬라이딩 윈도우를 적용한 결과 기존 행위에 대한 인식률 보다 약 5%이상 정확도를

높이는 결과를 보였다. 이는 객체가 움직이는 속도에 따라 일반적인 걸음걸이의 프레임에서 모든 프레임의 정보를 사용하는 것은 비효율적임을 알 수 있었다.

Table 4. Result of loitering B<sub>1</sub> according to sliding window size

Behavior	Size of Sliding Window( w )										
	data	w =3	w =5	w =7	w =9	w =10	w =11	w =13	w =15	w =20	w =30
Loitering : B <sub>1</sub>	1	86.67	86.67	90.00	90.00	90.00	90.00	90.00	86.67	86.67	80.00
	2	86.67	86.67	90.00	90.00	90.00	90.00	90.00	86.67	86.67	83.33
	3	86.67	83.33	86.67	90.00	90.00	90.00	90.00	86.67	83.33	76.67
	4	83.33	86.67	90.00	90.00	90.00	90.00	90.00	90.00	83.33	63.33
	5	83.33	86.67	86.67	86.67	86.67	86.67	86.67	86.67	86.67	73.33
	6	83.33	83.33	86.67	86.67	86.67	86.67	86.67	86.67	90.00	73.33
	7	83.33	83.33	86.67	86.67	86.67	86.67	86.67	90.00	86.67	70.00
	8	90.00	83.33	90.00	90.00	90.00	90.00	90.00	90.00	90.00	80.00
	9	80.00	86.67	90.00	90.00	90.00	90.00	90.00	90.00	90.00	70.00
	10	83.33	86.67	90.00	90.00	90.00	90.00	90.00	90.00	86.67	76.67
Average(%)		84.67	85.33	88.67	89.00	90.00	90.00	89.00	88.33	87.00	74.67

Table 5. Result of loitering B<sub>2</sub> according to sliding window size

Behavior	Size of Sliding Window( w )										
	data	w =3	w =5	w =7	w =9	w =10	w =11	w =13	w =15	w =20	w =30
Loitering : B <sub>2</sub>	1	83.33	83.33	90.00	90.00	90.00	90.00	90.00	90.00	86.67	80.00
	2	83.33	83.33	86.67	86.67	86.67	86.67	86.67	86.67	83.33	80.00
	3	86.67	86.67	90.00	90.00	90.00	90.00	86.67	86.67	83.33	70.00
	4	86.67	83.33	90.00	90.00	90.00	90.00	90.00	90.00	86.67	76.67
	5	86.67	86.67	86.67	86.67	86.67	86.67	90.00	90.00	83.33	73.33
	6	86.67	86.67	90.00	90.00	90.00	90.00	90.00	90.00	86.67	70.00
	7	83.33	83.33	90.00	90.00	90.00	90.00	90.00	90.00	86.67	70.00
	8	83.33	86.67	86.67	86.67	86.67	86.67	86.67	86.67	86.67	80.00
	9	83.33	83.33	86.67	86.67	90.00	90.00	90.00	90.00	83.33	76.67
	10	83.33	86.67	86.67	86.67	86.67	86.67	86.67	86.67	86.67	73.33
Average(%)		84.67	85.00	88.34	88.34	88.67	88.67	88.67	88.67	85.33	75.00

Table 6. Result of loitering B<sub>3</sub> according to sliding window size

Behavior	Size of Sliding Window( w )										
	data	w =3	w =5	w =7	w =9	w =10	w =11	w =13	w =15	w =20	w =30
Loitering : B <sub>3</sub>	1	83.33	83.33	86.67	88.67	88.67	88.67	88.67	86.67	83.33	83.33
	2	80.00	86.67	90.00	90.00	90.00	90.00	88.67	86.67	83.33	80.00
	3	83.33	86.67	86.67	88.67	86.67	88.67	88.67	83.33	86.67	70.00
	4	83.33	86.67	80.00	80.00	80.00	80.00	83.33	86.67	80.00	70.00
	5	80.00	83.33	86.67	88.67	86.67	88.67	88.67	86.67	83.33	70.00
	6	83.33	83.33	90.00	90.00	90.00	90.00	83.33	83.33	86.67	66.67
	7	80.00	86.67	83.33	83.33	83.33	83.33	88.67	86.67	83.33	80.00
	8	80.00	83.33	86.67	88.67	86.67	88.67	88.67	86.67	83.33	70.00
	9	83.33	83.33	86.67	88.67	86.67	88.67	83.33	83.33	80.00	70.00
	10	83.33	83.33	90.00	90.00	90.00	90.00	83.33	83.33	86.67	76.67
Average(%)	82.00	84.67	88.67	88.67	88.67	88.67	88.67	86.53	85.33	83.67	73.67

본 실험에서는 움직이는 객체의 행위를 인식하기 위해서 가장 효과적인 프레임 분할 윈도우 크기를 알아보는 것으로, 영상에서 객체의 행위에 따른 특징 정보를 추출하기 위한 가장 효율적인 프레임 분할의 크기를 실험적으로 추정하였다. 실험결과 슬라이딩 윈도우의 크기 10에서 11사이에서 88.67% 이상의 인식률로 가장 높은 인식률을 보였으며 7에서 15프레임 사이의 프레임 분할이 영상에서 행위를 인지할 때 가장 효과적임을 확인할 수 있었다. 따라서 영상에서 모든 프레임에서 얻을 수 있는 정보를 사용하는 것은 다소 불필요하고 효과적이지 않음을 알 수 있다. 또한 슬라이딩 윈도우의 크기가 20이상일 경우에는 너무 큰 윈도우의 크기를 하나의 프레임으로 보기 때문에 이 또한 인식률에 영향을 주는 것을 알 수 있다.

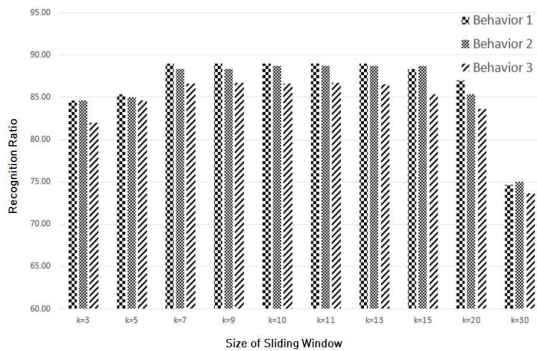


Fig. 4. Average recognition ratio according to sliding window size.

### 5. 토의 및 결론

본 논문에서는 효과적인 행위인식을 위하여 특정 정보를 획득의 크기를 알아보고자 하였다. 이를 위하여 객체의 중심점을 이용하였고 움직이는 객체의 행위를 시작부터 종료시점까지 하나의 행위로 보았다. 단위 움직임 정보를 프리미티브로 정의하고 학습하도록 하였고 멀티 슬라이딩 윈도우의 방법을 이용하여 효과적인 인식률을 실험하였다. 영상에서 얻을 수 있는 정보를 모두 사용하는 것은 불필요하거나 중복, 방해요소가 되는 경향이 있기 때문에 모든 특징을 이용하는 것은 효과적이지 않다. 또한 슬라이딩 윈도우의 크기가 20이상에서는 너무 큰 윈도우의 크기를 하나의 프레임으로 보기 때문에 정보 손실로 인한 인식률에 영향을 주었다. 슬라이딩 윈도우의 크기가 9에서 11사이에서 88.67%의 인식률로 가장 높은 인식률을 보였다. 실험에서 설정한 행위에 대한 평균 인식률은 시공간에서 추출된 프레임 시퀀스가 윈도우 크기 7에서 15일 때 가장 효율적임을 알 수 있었다. 따라서 단위 움직임 정보를 학습에 사용하고 멀티 슬라이딩 윈도우 크기를 적용하여 행위에 적합한 효율적인 특징정보를 추출함으로써 전체 행위인지의 인식률을 제고할 수 있었다.

추후, 객체의 움직임 속도에 연동한 적응형(adaptive) 윈도우 크기의 적용연구가 필요할 것으로 사료되며 대규모 영상데이터들이 기하급수적으로 늘어남에 따라 이를 지원하기 위한 좀 더 일반화된 학습 모델의 개발 및 대용량 처리를 위한 하드웨어 및 플랫폼 개발 지원이 요구된다. 본 결과를 통해 지능형



ITS시스템 분야에서 방향 탐지를 통한 보행자감지, 불법도주차량의 추적이나 IoT홈 시큐리티와 같은 수상한 물체의 행위에 따른 위험방지, 실내에서 독거노인의 움직임 모니터링 등 다양한 타 산업분야에서 활용이 가능할 것으로 예상되며 영상에서 객체의 행동에 패턴분류로 더욱 다양한 행위의 인식과 사전 의도파악 연구 등에 활용하고자 한다.

#### REFERENCE

- [1] N.H. Cho, S.Z. Byun, M.K. Lee, and Y.H. Kim, "Implementation of Pedestrian Crosswalk Monitoring System for Accident Prevention," *Journal of Kalinga Institute of Industrial Technology*, Vol. 14, No. 4, pp. 45-50, 2016.
- [2] Ministry of the Interior, *CCTV Installation and Operation Status, Survey on Actual Situation and Comprehensive Support System for Personal Information Protection*, 2016.
- [3] S.J. Lee, H.G. Kim, J.W. Sa, and Y.W. Chung, "Real-time Moving Detection with Adaptive Generalized Method of Moments in Heterogeneous Computing Environments," *Proceedings of Symposium of the Korean Institute of Communications and Information Sciences*, pp. 1002-1003, 2016.
- [4] T.K. Kim and J.K. Park, "Object Detection and Analysis Technology for Intelligent Video Surveillance," *The Magazine of the IEEK*, Vol. 39, No. 2, pp. 55-65, 2012.
- [5] J.U. Lee, J.Y. Sun, and M.C. Won, "Real-Time Algorithm for Relative Position Estimation between Person and Robot Using a Monocular Camera," *Transaction of the Korean Society of Mechanical Engineers*, Vol. A37, No. 12, pp. 1445-1452, 2013.
- [6] J.G. Ko, Y.S. Bae, J.Y. Park, and K. Park, "Technologies Trends in Image Big Data Analysis," *Electronics and Telecommunications Trends*, 2014.
- [7] Y.Y. Nam, Y.J. Choi, and W.D. Cho, "Object Tracking: A Survey," *Journal of Korean Society for Internet Information*, Vol. 11, No. 1, 129-141, 2010.
- [8] H.S. Chen, H.T. Chen, Y.W. Chen, and S.Y. Lee, "Human Action Recognition Using Star Skeleton," *Proceeding of Association for Computing Machinery International Workshop on Video Surveillance and Sensor Networks*, pp. 171-178, 2006.
- [9] N.J. Kwak and T.S. Song, "Object-Action and Risk-Situation Recognition Using Moment Change and Object Size's Ratio," *Journal of Korea Multimedia Society*, Vol. 17, No. 5, pp. 556-565, 2014.
- [10] Lian, H. Cheng, and B.L. Lu. "Multi-view Gender Classification Using Local Binary Patterns and Support Vector Machines," *Proceeding of International Symposium on Neural Networks*, pp.202-209, 2006.
- [11] B. Yogameena, S. Veera Lakshmi, M. Archana, and S.R. Abhaikumar, "Human Behavior Classification Using Multi-Class Relevance Vector Machine," *Journal of Computer Science*, pp. 1021-1026, 2010.
- [12] Stauffer, Chris, W. Eric.L. Grimson, "Learning Pattern of Activity Using Real-Time Tracking," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, Issue. 8, pp. 747-757, 2000.
- [13] S. Velastin, B. Boghossian, B. Lo, J. Sun, and M. Vicencio-Silva, "PRISMATICA: Toward Ambient Intelligence in Public Transport Environments," *IEEE Transaction Systems, Man and Cybernetics*, Vol. 35, No. 1, pp. 164-182, 2005.
- [14] D.B. Nathaniel, M. Osama, P. Nikolaos, and I. Aaron, "Detection of Loitering Individuals in Public Transportation Areas," *IEEE Transaction Intelligent Transportation Systems*, Vol. 6, No. 2, pp. 167-177, 2005.
- [15] S. Paolo, C. Andrea, L. Marco, and M. Tommaso "An Abandoned/Removed Objects Detection Algorithm and Its Evaluation on



PETS Datasets,” *Proceeding of IEEE International Conference on Video and Signal Based Surveillance*, pp. 17, 2006.

[16] Y. Elden and J.K. Aggarwal, “Human Action Recognition with Extremities as Semantic Posture Representation,” *Proceeding of International Conference Computer Vision, Pattern Recognition*, pp. 1-8, 2009.

[17] J. Aggarwal and Q. Cai, “Human Motion Analysis: A Review,” *Computer Vision and Image Understanding*, pp. 428-440, 1999.

[18] A. Singh, S. Sawan, M. Hanmandlu, V. K. Madasu, B. C. Lovell, “An Abandoned Object Detection System Based on Dual Background Segmentation,” *Proceeding of IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 352-357, 2009.

[19] S.W. Lee, T.Y. Kim, J.H. Yoo, and J.K. Park, “Abnormal Behavior Detection Based on Adaptive Background Generation for Intelligent Video Analysis,” *The Institute of Electronics Engineers of Korea-Signal Processing*, Vol. 48, pp. 111-121, 2011.

[20] T.T. Zin, T. Pyke, and T. Takashi, “A Markov Random Walk Model for Loitering People Detection,” *Proceeding of International Conference on Intelligent Information Hiding and MultiMedia Signal Processing*, pp. 680-683, 2010.

[21] C.H. Huang, M.Y. Shih, Y.T. Wu, and J.H. Kao, “Loitering Detection Using Bayesian Appearance Tracker and List of Visitors,” *Advances in Multimedia Information Processing*, pp. 906-910, 2008.

[22] E.S. Park, H.G. Lee, M.K. Yun, M.Y. Kim, J.H. Kwak, and H.K. Kim, “Loitering Behavior Detection Using Shadow Removal and Chromaticity Histogram Matching,” *Korea Institute of Information Security and Cryptology*, Vol. 21, No. 6, pp. 159-169, 2011.

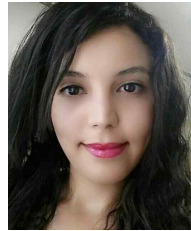
[23] A.F. Bobick and J.W. Davis, “Action Recognition using Temporal Templates,” *Computational Imaging and Vision Series*, Vol. 9, pp. 125-146, 1997.



권 익 환

2014년 부산대학교 정보컴퓨터공학부 학사  
 2016년 부산대학교 전기전자컴퓨터공학과 석사  
 2016년 현재 한국전자통신연구원 (ETRI) 연구원

관심분야: 영상처리, HCI, 딥러닝, 인지기계학습, 데이터 마이닝



부베나 하제르

2009년 University of Constantine 학사  
 2011년 University of Constantine 석사  
 2014년 ~ 현재 부산대 전기전자컴퓨터공학과 박사과정

관심분야: 딥러닝, 자연어 언어 처리



이 도 훈

1986년 부산대학교 계산통계학과 학사  
 1992년 부산대학교 전자계산학과 석사  
 1998년 부산대학교 전자계산학과 박사

1995년 ~ 2005년 밀양대학교 컴퓨터공학과 교수  
 2004년 ~ 2005년 Indiana University 방문연구원  
 2006년 ~ 현재 부산대학교 정보컴퓨터공학부 교수  
 관심분야: Visual Computing, Biometric computing, CNN, BCI