

논문 2011-48SC-5-7

영상감시시스템에서 움직임의 비교사학을 통한 비정상행동탐지

(Unsupervised Motion Learning for Abnormal Behavior Detection in
Visual Surveillance)

정 하 욱*, 장 형 진**, 최 진 영***

(Hawook Jeong, Hyung Jin Chang, and Jin Young Choi)

요 약

본 논문에서는 비교사학을 통해 영상의 방대한 정보를 효율적으로 모델링 하는 방법을 제안하고자 한다. 여기서 이동궤적들은 자연어 처리에 사용되는 알고리즘인 잠재 디리클레 할당 모형(Latent Dirichlet Allocation)에 의해 직진, 좌회전, 우회전 등 각 상황 별로 주제에 따라 그 영역을 효과적으로 분류할 수 있다. LDA를 이용해 주제별로 의미 있는 영역을 분류한 후, 각 주제별로 분류된 궤적을 관측열로 보고 은닉 마르코프 모델(Hidden Markov Model)의 바움-웰치 알고리즘을 사용하여 학습한다. 전향 알고리즘을 사용하여 입력된 행동과 학습된 행동을 비교함으로써 영상내의 행동이 정상인지 비정상인지를 효과적으로 판단할 수 있다. 실험결과 다양한 영상에 대해 의미있는 주제별로 영역이 잘 분류되며 추적에러로 인한 궤적의 노이즈에도 강인하게 물체의 무단횡단, 신호위반과 같은 상황을 효과적으로 탐지하는 것을 확인할 수 있다.

Abstract

In this paper, we propose an unsupervised learning method for modeling motion trajectory patterns effectively. In our approach, observations of an object on a trajectory are treated as words in a document for latent dirichlet allocation algorithm which is used for clustering words on the topic in natural language process. This allows clustering topics (e.g. go straight, turn left, turn right) effectively in complex scenes, such as crossroads. After this procedure, we learn patterns of word sequences in each cluster using Baum-Welch algorithm used to find the unknown parameters in a hidden markov model. Evaluation of abnormality can be done using forward algorithm by comparing learned sequence and input sequence. Results of experiments show that modeling of semantic region is robust against noise in various scene.

Keywords : 비정상행동탐지, 이동궤적 모델링, 잠재 디리클레 할당 모형, 은닉 마르코프 모델, 영상감시

I. 서 론

영상 감시 시스템(Video Surveillance System)은 군사, 보안 및 통행 관리 등의 목적으로 근래에 와서 그

중요성이 부각되고 있으며 그에 따라 CCTV의 사용량도 점차 증가하고 있다. 하지만 이런 영상 감시 시스템의 대부분은 여러 대의 감시 카메라에서 입력된 영상을 여러 대의 모니터에 전시하고 녹화하는 기능만 수행하고, 이렇게 입력된 영상을 통해 이상 여부를 감시하는 것은 사람이 그 역할을 하고 있다. 하지만 사람이 여러 개의 모니터를 동시에 관찰하면서 비정상적인 행동을 탐지하는 것은 매우 힘든 일이며, 모든 상황을 녹화하는 것은 다시 찾아보기도 힘들뿐더러 엄청난 메모리의 낭비이다. 따라서 이를 자동화하기 위한 영상감시 알고

*정회원, **학생회원, ***평생회원, 서울대학교 공과대학 전기·컴퓨터 공학부, ASRI

(Electrical Engineering and Computer Science, Engineering College, Seoul National University)

※ 본 연구는 삼성테크윈 및 지식경제부 BK 21 프로그램의 지원을 받아 수행되었습니다.

접수일자: 2011년4월22일, 수정완료일: 2011년8월10일

리즘의 연구가 활발히 진행되고 있다.

CCTV 영상에서 자동으로 물체의 움직임을 파악하는 것은 지능형 영상 감시시스템에서 매우 중요한 과제이다. 기존의 영상감시시스템^[10-11]은 현재 이동물체의 탐지, 추적에 초점을 두고 있지만 최종 목표는 이런 물체들의 움직임, 행동을 인식하고 해석하는데 있다. 행동인식의 방법 중에서 걷기, 뛰기, 직진, 좌회전 등의 기본적인 패턴(Pattern)을 미리 지정해 주는 방법도 있긴 하지만, 대부분의 경우 감시 영상에서 움직임의 패턴은 미리 지정하기가 어렵다. 따라서 일반적인 상황에서 움직임 패턴들을 자동적으로 학습하는 방법이 다양한 감시 상황에 적용하기에 더 적절하다. 움직임 패턴을 학습하고 모델링 하는 방법은 크게 두 가지로 분류된다. 첫 번째는 물체가 탐지되고, 추적되면서 생긴 궤적(Trajectory)을 모델링하는 방법이고, 두 번째 방법은 탐지추적과 같은 과정을 생략하고 옵티컬 플로우(Optical Flow)와 같은 낮은 수준의 움직임 패턴을 추출하여 이 패턴을 학습하는 방법^[15-16]이다. 하지만 옵티컬 플로우를 활용한 방법들은 학습과정에서 추출된 움직임 벡터들의 시간적 순서는 고려하지 않아 노이즈에 민감하며 비정상 행동을 하는 물체를 지속적으로 탐지하고 관찰하기에 어려움이 있다.

자동으로 궤적의 패턴을 학습하고 물체의 움직임을 분석하는 시스템은 대체로 그림 1과 같은 과정을 따른다^[12]. 움직임 패턴의 모델링과 비정상 움직임의 탐지를 위해서는 기본적으로 이동물체의 탐지와 추적을 통한 궤적이 필요하다. 움직임 궤적을 바탕으로 장면 모델링(Scene Modeling)을 하고 이를 토대로 새로운 영상에 대해서 움직임 패턴의 정상여부를 판단할 수 있다. 즉, 행동인식을 위해 가장 중요한 과정 중 하나가 궤적을 모델링 하는 것이다. 이 모델을 토대로 해서 이상한 궤적을 찾아낼 수도 있고, 감시영상 내에서 입구영역이나 출구영역을 찾아낼 수도 있으며, 이 모델을 토대로 추적알고리즘의 필터로 사용할 수도 있다.

궤적을 모델링하는 기존의 알고리즘의 경우, 대부분 움직임 궤적들을 논문에서 정의한 유사도(Similarity)에

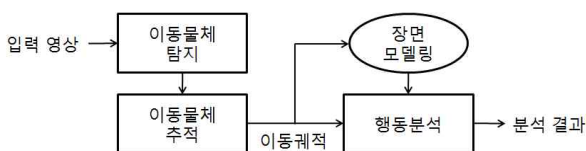


그림 1. 행동분석 블록 다이어그램

Fig. 1. The block diagram for trajectory analysis.

거리를 측정해 거리가 가까운 궤적끼리 클러스터링(Clustering)하는 방법을 취한다. X. Wang^[5]에서는 하우스도르프 거리(Hausdorff Distance) 방법을 개량하여 궤적 사이의 유사도를 측정해 비슷한 궤적끼리 묶는 알고리즘을 제안했다. 하지만 궤적의 공간적인 거리를 측정할 경우 투영 왜곡(Projective Distortion)에 의한 오류가 생길 수 있으며, 수많은 궤적에 대해 모든 궤적 쌍의 거리를 구하는 것은 비효율적인 계산량을 가지게 된다. Weiming Hu^[1]는 fuzzy K-mean 알고리즘을 이용해 움직임 패턴들을 클러스터링하는 알고리즘을 제안했다. Imran Saleemi^[2]와 Arslan Basharat^[4]은 각 픽셀에서 주변픽셀로 움직이는 물체가 움직이기 위한 전이확률(Transition Probability)을 커널 밀도 추정(Kernel Density Estimation)이나 혼합 가우시안 모델(Gaussian Mixture Model)방법으로 확률 공간상에 학습을 시키는 방법을 사용했다. Arslan Basharat^[4]에서는 추가적으로 물체의 크기까지 함께 학습함으로써 정확성을 더 높이려고 노력했다. 하지만 위와 같은 방법들은 감시 영상내에서 모든 픽셀에 확률 모델을 만들기 위해서는 수많은 학습데이터와 학습시간이 필요하다. X.Wang^[3]은 이런 문제들을 극복하기 위해 궤적을 문서(Document), 궤적 중 각 관측값(Observation)을 단어로 생각하고 자연어 처리에 사용되는 HDP(Hierarchical Dirichlet Process)라는 토픽모델(Topic Model)을 이용해 의미영역(Semantic Region)을 구분하는 알고리즘을 제안했다. 이 방법은 적은 데이터에서도 비교적 빠른 속도로 투영왜곡(Projective Distortion)에 강인하게 의미영역을 구분해내는 장점을 가지고 있지만 HDP의 모델의 특성상 각 영역의 시간적 순서를 고려하지 못한다는 단점이 있다.

이러한 문제점들을 해결하기 위하여 본 연구에서는 이동궤적을 X.Wang^[3]에서 제안한 디리클레 할당 모형(Latent Dirichlet Allocation)을 이용해 궤적을 효율적으로 감시영상에서 의미 영역(Semantic Region)을 클러스터링(Clustering)한다. 또한 각 주제별로 분류한 궤적을 관측열(Observation Sequence)로 보고 은닉 마르코프 모델(Hidden Markov Model)^[8]을 학습한다. 학습한 HMM에 대해 새로운 이동물체의 궤적이 나타났을 때 그것이 정상일 확률을 구할 수 있으며, 이 확률이 낮을 경우 비정상 행동으로 판단함으로써 행동인식 성능을 높일 수 있다.

II. 본 론

1. 감시상황에서 이동물체 탐지 및 추적

그림 1의 이상행동 분석 다이어그램과 같이 영상감시에서 행동분석을 위해서는 입력 영상에 대해 이동물체의 탐지 및 추적 결과가 필요하다. 움직이는 물체의 탐지는 이전 프레임의 영상과 현재 프레임의 영상간의 차를 구하는 차영상 방법으로도 가능하지만, 만약 배경에 해당하는 영상을 획득할 수 있다면 현재 프레임의 영상과 배경 영상간의 차를 구하는 배경 차감법이 더 효과적이다. 이 배경 차감법의 경우, 배경이 항상 동일한 환경에서는 정적인 배경 영상으로 충분하지만, 조명이 변하거나 자동차가 주차하는 등 배경이 변하는 경우에는 적절하게 배경을 갱신(Update)해 주어야 한다. 이렇게 지속적인 갱신을 통해 변화하는 배경에 적응할 수 있도록 한 배경 모델을 적응적 배경 모델이라고 하며, C. Stauffer의 가우시안 혼합 모델(Gaussian Mixture Model)^[13]이 대표적이다. C. Stauffer는 변화하는 배경에 대한 강인성을 높이기 위해 가우시안 혼합 모델을 이용하여 배경 모델링하는 방법을 제안하였다. 이는 각 픽셀에 대해 3~5개의 가우시안 분포로 배경을 표현하고, 지속적으로 이 가우시안들을 갱신하는데, 간략화한 EM 과정(Expectation Maximization)을 통해 실시간성을 확보한 것이 특징이다.

또한 움직이는 객체에서 궤적을 추출하기 위해서는 움직임이 탐지되는 물체의 라벨을 지속적으로 유지하는 추적 기술 또한 필요하다. 본 연구에서 사용한 시스템^[14]에서는 기본적으로 이전의 추적 결과와 현재의 탐지 결과를 매칭 테이블에서 상호 비교하여 올바르게 라벨을 유지 한다.

2. 의미영역 분류

자연어 처리에 있어서 LDA^[6]와 같은 토픽모델(Topic Model)은 함께 발생하는 단어들을 모아 하나의 주제(Topic)로 분류해준다. 예를 들어 ‘환율’, ‘주가’와 같은 단어는 항상 같은 문서에서 나타나지는 않겠지만 자주 같은 문서에서 발견될 수 있으므로 ‘경제’라는 주제에서 이 단어들은 다항 확률분포에서 높은 가중치를 가지게 될 것이다. 만약 T개의 주제가 있다고 가정했을때, i번째 단어에 대한 확률은 아래와 같이 주어진다.

$$P(w_i) = \sum_{j=1}^T P(w_i|z_i = j)P(z_i = j). \quad (1)$$

이때 z_i 는 i번째 단어 w_i 에 대응하는 주제를 나타내는 잠재 변수(Latent Variable)이다. $P(w_i|z_i = j)$ 는 j번째 주제에서 단어 w_i 가 생성될 확률이며, $P(z_i = j)$ 는 주어진 문서에서 주제 j가 선택될 확률이다. 하지만 이 확률분포에 대한 직접적인 추론은 불가능하기 때문에 근사적인 MCMC(Markov Chain Monte Carlo)방법^[7]을 이용해 추론한다.

LDA가 궤적을 모델링하는데 적용된다면 주제(Topic)는 비슷한 궤적들이 지나간 자리가 의미하는 셀들의 집합으로 표현될 것이며 이 셀들의 집합을 의미영역(Semantic Region)으로 볼 수 있다. 제안한 알고리즘에서 각 주제는 32×24개의 항을 가지는 다항 확률분포(Multinomial Probability Distribution)를 이루며, 각 문서는 이러한 주제의 혼합으로 만들어졌다고 가정한다. 제안한 알고리즘에서는 1개의 궤적을 문장(Sentence)로 생각하고 수개의 궤적을 모아 하나의 문서(Document)로 만든다. 또한 궤적내의 각 관측값(Observation)을 단어(Word)로 취급한다. 예를 들어 320×240크기의 영상에서 각 셀(Cell)은 10×10의 작은 셀로 나뉘며, 각 셀이 한 단어(Word)를 의미한다. 즉 궤적에서의 각 위치는 32×24개의 단어중 하나로 맵핑(Mapping)되며 어휘목록(Vocabulary)의 크기는 32×24이다.

그림 2에서는 궤적을 단어로 맵핑하는 예시를 보여준다. 왼쪽 그림에서는 이동물체가 지나간 궤적을 보여주며 오른쪽 그림에서는 32×24 크기의 어휘목록(Vocabulary)에서 궤적이 지나간 셀(Cell)을 진한색으로 처리해 보여주고 있다. 이 셀 하나가 토픽모델에서 하나의 단어로 취급되며 하나의 궤적에서 나온 셀들의 집합을 하나의 문장으로 보고, 본 알고리즘에서는 10개의



그림 2. 이동물체의 궤적과 그에 따른 단어 생성
Fig. 2. The trajectory of moving objects and corresponding word generation.

문장을 하나의 문서로 모아서 LDA^[6,7]의 입력으로 사용하였다.

3. 의미영역에서의 시간 모델링

II.2절에서는 LDA를 이용해 감시영상으로부터 의미영역을 분류해냈다. 하지만 LDA는 문서내의 단어들은 얼마든지 순서가 뒤바뀔 수 있다는 가정에서 모델링을 했다는 단점을 가지고 있다. 따라서 앞서 구한 T개의 의미영역별로 은닉 마르코프 모델을 이용해 시간적 순서를 모델링 하고자 한다. 은닉 마르코프 모델(Hidden Markov Model)을 학습하기 위해서는 각 의미영역별로 그에 맞는 관측열(Observation Sequence)의 집합이 필요하다. 따라서 LDA의 결과를 이용해 각 문장을 T개의 주제 중에서 가장 가까운 곳에 할당함으로써 문장을 문서로부터 분리해 T개의 작은 집합으로 나눈다. 예를 들어 n개의 단어로 구성된 어떤 문장

$$w = \{w_1, w_2, w_3, \dots, w_n\}, \quad w_i \in \{1, 2, \dots, W\}, \quad (2)$$

가 있다고 가정했을 때, 각각의 단어에 대응하는 주제의 집합

$$z = \{z_1, z_2, z_3, \dots, z_n\}, \quad z_i \in \{1, 2, \dots, T\}, \quad (3)$$

또한 LDA의 결과로 얻어낼 수 있다. 이때 우리는 집합 z로부터 최빈값(Mode)을 구하고 그 최빈값을 그 문장에서 가장 가까운 주제로 판단해 그 주제의 클러스터에

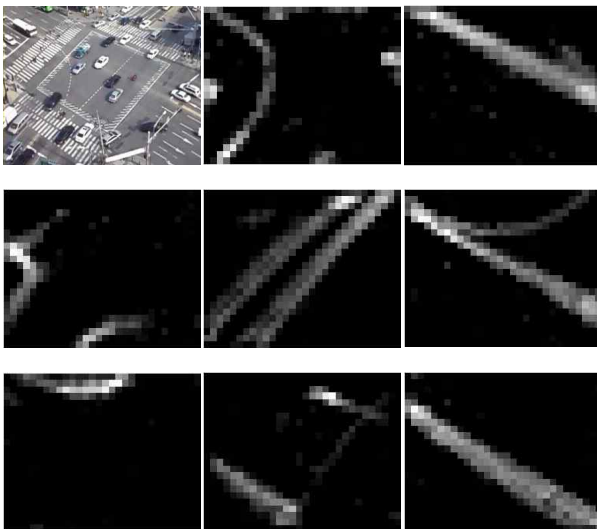


그림 3. T = 8 일 때 각 클러스터에서 구한 단어의 밀도

Fig. 3. Distributions of the 8 semantic regions over space.

w를 추가한다. 이렇게 의미영역에 따라 문장을 분류한 후 각 클러스터에서 구한 단어들의 밀도를 구한 결과는 그림 3에서 확인할 수 있다. 그림 3을 보면 복잡하고 이 동물체의 밀도가 높은 상황에서도 직진, 좌회전, 우회전, 보행신호등의 상황에 따라 각 영역이 잘 나뉘는 것을 확인할 수 있다.

이렇게 구한 각 클러스터의 관측열(Observation Sequences)을 학습 데이터로 이용해서 우도(Likelihood)를 최대화 하는 은닉 마르코프 모델 파라미터 $\lambda = (A, B, \pi)$ 를 바움-웰치 알고리즘으로 추정한다^[9]. 즉 T개의 의미영역에 대응하는 HMM을 결과로 얻을 수 있다. 각각의 HMM은 관측열에서의 단어의 순서에 따라 영역이 나누어져 상태(state)가 모델링될 것을 기대한다. 제안한 알고리즘에서는 10개의 상태(state)를 가진 HMM으로 학습을 했으며 초기 상태전이행렬(State Transition Matrix) A를 각 행(Row)의 합이 1인 임의의 10×10 행렬로 초기화하였다. 관측 심볼 확률 분포(Observation Symbol Probability Distribution) B는 각 클러스터에서 유일한 단어 개수를 심볼의 수로 정했으며 처음에는 동일한 분포를 가진다고 가정하고 초기화 하였다. 학습 결과로부터 물체가 진행하는 방향으로의 상태 천이 확률은 높지만 상태에 따라 나뉜 영역에서 물체의 진행 반대방향으로는 상태 천이 확률이 낮게 모델링 되므로 역주행과 같은 비정상행동도 탐지할 수 있다.

4. 비정상 행동 탐지

영상 감시 시스템에서 비정상 행동을 하는 물체를 찾는 것은 본 알고리즘의 최종목표라고 할 수 있다. 정상 행동은 II.3절에서 만든 T개의 은닉 마르코프 모델(Hidden Markov Model)에서 만들어내는 패턴이라고 정의하고, 테스트 영상에서 물체의 이동 궤적이 T개의 모델 모두에 대해 확률이 낮은 궤적은 비정상으로 판단한다. 테스트 궤적을 HMM의 입력으로 사용하기 위해서는 탐지·추적 알고리즘으로부터 나온 실시간의 궤적 결과를 II.2절과 같은 방법으로 단어의 연속으로 변환하는 작업을 해주어야 한다. 한 프레임의 영상내에서 m개의 물체가 탐지되어 추적되고 있었다고 가정한다면 그 프레임에서 테스트 알고리즘에 입력으로 들어갈 문장은 $\{w_1, w_2, \dots, w_m\}$ 이며 각 문장은

$$\mathbf{w}_i = \{w_1^{(i)}, w_2^{(i)}, w_3^{(i)}, \dots, w_n^{(i)}\}, w_j^{(i)} \in \{1, 2, \dots, W\} \quad (4)$$

로 나타낼 수 있으며 n 은 문장 \mathbf{w}_i 가 이루고 있는 단어의 개수이다. 이때, 학습된 HMM 파라미터 $\{\lambda_1, \lambda_2, \dots, \lambda_T\}$ 에 대하여 관측열이 일어날 확률 $P(\mathbf{w}_i | \lambda_t)$ 를 구하기 위해서는 전향 알고리즘^[9]을 이용한다. 이렇게 T 개의 HMM으로부터 문장 \mathbf{w}_i 가 일어날 확률

$$\{P(\mathbf{w}_i | \lambda_1), P(\mathbf{w}_i | \lambda_2), \dots, P(\mathbf{w}_i | \lambda_T)\}, \quad (5)$$

을 구할 수 있고

$$t^* = \operatorname{argmax}_t \{P(\mathbf{w}_i | \lambda_t)\}, \quad (6)$$

에서

$$\log P(\mathbf{w}_i | \lambda_{t^*}) < \text{threshold}. \quad (7)$$

일 때, \mathbf{w}_i 에 해당하는 궤적을 비정상적으로 판단한다. 다음 장에서는 본문에서 제시했던 알고리즘을 바탕으로 한 실험결과를 분석하도록 하겠다.

III. 실험

실험에서 제안한 알고리즘은 MATLAB으로 구현하였고, 실험에 사용한 영상중 사거리 영상은 실제 교통상황에서 사용되고 있는 CCTV 영상을 인터넷에서 동영상 캡처 프로그램을 이용하거나 디지털카메라로 녹화한 영상이다. 또 다른 영상은 A. Basharat^[4]에서 실험영상으로 사용한 UCF dataset이다. 이 영상의 경우에는 학습데이터는 제공하는 궤적의 텍스트 파일을 이용했고, 테스트의 경우에만 동영상을 이용해서 탐지·추적 알고리즘으로 궤적을 추출하여 비정상행동을 탐지하는데 사용하였다. 연산 속도면에서 살펴봤을때, 사거리 영상의 학습데이터에서 추출된 궤적은 3066개, UCF dataset에서 학습을 위해 사용된 궤적은 2060개였고, 이것을 학습하는데 HMM의 상태개수와 초기 파라미터의 값에 따라 다르지만 대체로 1시간~1시간 30분정도가 소요되었다. 테스트 과정에서는 각 프레임에서 탐지물체의 ID와 위치정보를 미리 받아온다고 가정했을때 비정상 행동탐지에만 걸린 시간은 약 30~50ms로 MATLAB에서도 실시간성을 보장했다.

테스트 영상에서 물체의 이동 궤적을 단어와 문장으

로 맵핑 한 후 그 문장이 학습한 HMM에 대해 일어날 확률을 계산한다. II.4절에서 정의한 방법대로 비정상 행동의 여부를 판단하며 2개의 사거리데이터와 UCF 데이터에 대해서 비정상 행동으로 판단한 경우의 예시는 그림 3과 같다. 사거리데이터의 경우 오랜 시간 직접 CCTV 화면을 보면서 비정상 행위가 일어나는 부분을 녹화해야 하기 때문에 원하는 이상행동을 만들어내기는 힘들었다. 하지만 몇 가지의 비정상 행동이 일어나는 장면을 캡처해 테스트 해본 결과 다양한 상황에 대해서 정상궤적과 비정상 궤적을 잘 분리 해냈으며, 특히 무단횡단이나 차선위반의 상황을 잘 탐지해냈다.

그림 4의 (a)에서는 횡단보도로 길을 건너다가 차도로 빠져나가는 오토바이가 비정상행위로 탐지 되었다. (b)에서도 오토바이가 차도에서 직진을 하다가 횡단보도를 거쳐 인도로 올라가는 것이 비정상행위로 탐지되었다. (c)와 (d)에서는 오토바이와 사람이 무단횡단을 하는 모습이다. 이러한 경로는 평소에 매우 드물게 일어나는 경로이기 때문에 비정상적으로 판단된다. (e),(f)에 나오는 영상의 경우 뒤편에는 차도가 있고 보행자가 다니는 길이 기억자로 나온다. UCF 데이터에서 주어진 궤적에서는 기억자로 나오는 길에만 보행자가 다니도록

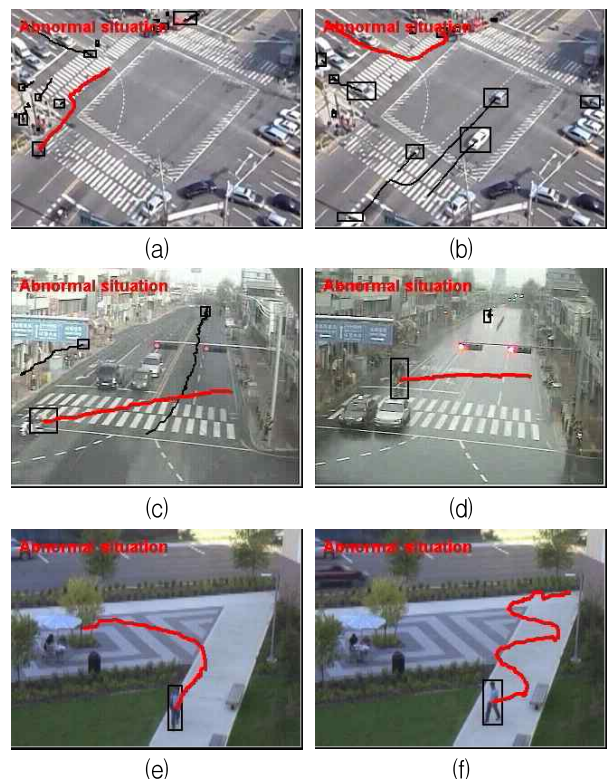


그림 4. 다양한 영상에서 비정상 행동의 탐지 결과
Fig. 4. Various examples of abnormal behaviors.

표 1. 궤적데이터에 대한 정량적 실험 결과
Table 1. The quantitative result of trajectories.

사거리 데이터	이상궤적 탐지율	83.3%
	정상궤적 탐지율	98.2%
UCF 데이터	이상궤적 탐지율	80.0%
	정상궤적 탐지율	92.9%

학습이 되어있는데 (e)는 그 길을 가로질러 감으로써 비정상행위로 판단하게 된 예시를 보여준다. (f)에서는 사람이 의도적으로 지그재그로 걷는 상황을 보여주며 이 행동도 일반적으로 자주 일어나지 않는 상황이므로 비정상으로 판단하게 된다. 결과에서 확인할 수 있듯이, 다양한 영상에서 확률적으로 낮게 일어나는 행동들을 효과적으로 탐지하는 것을 볼 수 있으며, 주변에서 일어나는 정상적인 행동들은 검정색으로 표시되어 정상과 비정상을 잘 구분하고 있음을 보여준다. 표 1에서는 사거리 동영상에서 테스트한 63개의 궤적(정상궤적 57개, 이상 궤적 6개)과 UCF 동영상에서 추출한 52개의 궤적(정상궤적 42개, 이상 궤적 10개)에서 비정상 결과를 탐지한 정량적 결과를 보여준다. 표에서 확인할 수 있듯이 두가지 경우 모두 비정상 궤적은 80%이상, 정상궤적은 90% 이상의 높은 탐지율을 보여주었다.

IV. 결 론

본 논문에서는 비교사적 방법으로 아무런 사전정보 없이 영상감시시스템에서 방대한양의 탐지·추적결과를 효율적으로 모델링하고 그것을 토대로 신뢰할만한 비정상 행동 탐지 결과를 얻었다. 이동궤적 데이터를 양자화(Quantized)해서 문서(Document), 단어(word)화한 후, 단계잠재 디리클레 할당 모형(Latent Dirichlet Allocation)을 이용해 효율적으로 감시영상에서 의미 영역(Semantic Region)을 찾아냈고, HMM을 이용해 의미 영역 내에서 이동 물체가 지나갈 시간적인 순서도 모델링하여 역주행과 같은 이상행동도 효과적으로 탐지할 수 있도록 하였다. 테스트 과정에서는 학습한 HMM에 대해 새로운 이동물체의 궤적이 나타날 확률을 전향 알고리즘(Forward Algorithm)을 사용하여 구했으며, 이 확률이 낮을 경우 비정상 행동으로 판단하였다. 실험 결과에서 확인 하였듯이, 다양한 상황에서 정확하게 알고리즘이 동작하는 것을 확인할 수 있었지만 추적실패에 따른 노이즈와 고정적인 셀 크기와 양자화

(Quantization)의 에러에 따른 문제로 오탐지가 일어나는 경우도 찾아볼 수 있었다.

추후 연구에서는 이러한 단점을 극복할 수 있도록 양자화를 하지 않고 의미영역을 만들어내는 방법에 대해 연구가 수행되어야 할 것이다. 또한 의미영역 사이의 연관관계를 함께 모델링한다면 2개 이상의 의미영역에서 일어나는 움직임들이 동시에 일어날 수 있는지 없는지를 판단함으로써 좀 더 심층적인 장면의 이해가 가능할 것이다. 그리고 고정형 카메라에서 장면을 학습하는 것이 아니라 Pan-Tilt 이동이 카메라에서의 장면을 모델링하고 비정상 행위를 탐지하는 방법에 대해서도 연구를 진행할 필요가 있다.

참 고 문 헌

- [1] W Hu, X Xiao, Z Fu, D Xie, T Tan, and S Maybank, "A system for learning statistical motion patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.28, pp.1450-1464, 2006.
- [2] Imran Saleemi, Khurram Shafique and Mubarak Shah, "Probabilistic Modeling of Scene Dynamics for Applications in Visual Surveillances," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.31, pp.1472-1485, 2009.
- [3] Xiaogang Wang, Keng Teck Ma, Gee-Wah Ng, at al., "Trajectory analysis and semantic region modeling using a nonparametric Bayesian model", *IEEE Conf. Computer Vision and Pattern Recognition*, 2008.
- [4] A. Basharat, A. Gritai, Mubarak Shah, "Learning object motion patterns for anomaly detection and improved object detection," *IEEE Conf. Computer Vision and Pattern Recognition*, 2008.
- [5] X. Wang, K. Tieu and E. Grimson, "Learning Semantic Scene Models by Trajectory Analysis," in *Proceedings of European Conference on Computer Vision (ECCV)*, 2006.
- [6] DM Blei, AY Ng, and MI Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, 3, pp.993-1022, 2003.
- [7] T. L. Griffiths, and M. Steyvers, "Finding scientific topics," *Proceedings of the National Academy of Sciences*, 101 (suppl. 1), pp. 5228-5235, 2004.
- [8] Lawrence R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of IEEE*, Vol.

- 77, NO.2, pp257-286, Feb. 1989.
- [9] 한학용, 패턴인식개론. 서울:한빛미디어, 456-460쪽, 2009.
- [10] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4 : Real-time surveillance of people and their activities," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, pp. 809-830, Aug. 2000.
- [11] R. T. Collins, A. J. Lipton, T. Kanade, et al., "A system for video surveillance and monitoring," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., CMU-RI-TR-00-12, 2000.
- [12] Brendan Tran Morris and Mohan Manubhai Trivedi, "A survey Of vision-based trajectory learning and analysis for surveillance," IEEE Trans. Circuits and Systems for Video Technology, vol. 18, no. 8, Aug., 2008.
- [13] E. Grimson and C. Stauffer, "Adaptive Background Mixture Models for Real Time Tracking," Proc. CVPR Conf., 1999.
- [14] 장형진, 이광무, 김수완, 윤석민, 최진영, 지능형 영상 감시 알고리즘 개발을 위한 통합 시스템, 대한전기학회, CICS 2010 정보 및 제어 학술대회 논문집 2010.10, page(s): 109-110
- [15] T Hospedales, S Gong, and T Xiang, "A Markov Clustering Topic Model for Mining Behaviour in Video," Proc. IEEE Int'l Conf. Computer Vision, 2009.
- [16] Xiaogang Wang, Keng Teck Ma, Gee-Wah Ng, et al., "Trajectory analysis and semantic region modeling using a nonparametric Bayesian model", IEEE Conf. Computer Vision and Pattern Recognition, 2008.

— 저 자 소 개 —



정 하 옥(정회원)
 2009년 서울대학교 전기공학부
 학사.
 2011년 서울대학교 전기·컴퓨터
 공학부 석사.
 <주관심분야 : 컴퓨터 비전, 머신
 러닝>



장 형 진(학생회원)
 2006년 서울대학교 전기공학부
 학사.
 2011년 서울대학교 전기·컴퓨터
 공학부 박사과정.
 <주관심분야 : 컴퓨터 비전, 머신
 러닝>



최 진 영(평생회원)
 1982년 서울대학교
 제어·계측공학과 학사.
 1984년 서울대학교
 제어·계측공학과 석사.
 1993년 서울대학교
 제어·계측공학과 박사.
 1994년~현재 서울대학교 전기공학부 교수.
 <주관심분야 : 영상감시(Visual Surveillance), 신
 경회로망, 적응제어>