

가우시안 혼합 모델과 옵티컬 플로우 기법을 이용한 특이행동 인지 기법 연구

박종현*, 임성조*, 강동중*
 *부산대학교 기계공학부
 e-mail : see@pusan.ac.kr

Abnormal behavior detection using Gaussian Mixture Model and Optical Flow

Jong-Hyun Park*, Sung-Jo Lim*, Dong-Joong Kang*
 *School. of Mechanical Engineering, Pusan National University

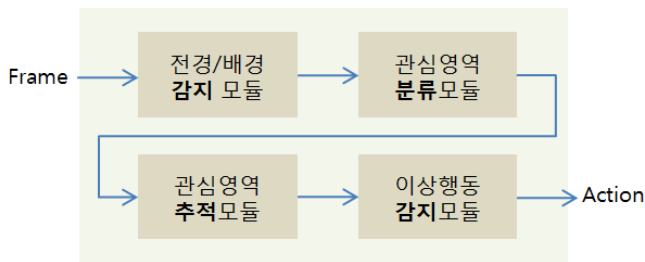
요 약

본 논문에서는 감시시스템이 갖추어진 환경 내에서 발생할 수 있는 특이 행동을 효율적으로 감지하기 위한 기법을 제시한다. 최근 대형 범죄 및 방화 사건 등의 방지목적으로 DVR의 단순 녹화를 벗어나 지능형 감시시스템을 도입하려는 연구가 활발히 진행되고 있다. 그러나 이러한 시스템들은 아직 초기 연구 단계에 있으며 영상내의 관심물체 추출을 위한 전경과 배경의 분리 및 추적 단계에 그치고 있다. 이에 본 논문에서는 가우시안 혼합 모델을 통하여 전경과 배경을 분리하고, 관심영역에 한해서 Optical Flow 기법을 이용하여 폭력상황과 같은 특이 행동의 감지 여부를 판단 할 수 있는 방법에 대해 실험을 통해 평가하였다.

1. 서론

1) 지능형 감시 시스템

최근 대형 범죄 및 방화 사건 등의 빈발에 따라 지능형 감시 시스템(Intelligent Surveillance System)에 관한 연구가 활발히 진행되고 있다[1, 2]. 지능형 감시시스템은 감시 카메라의 영상정보를 디지털화 하여 시스템이 스스로 사물의 위치 및 패턴을 분석하고 위험상황이 발생할 경우에 경보 및 출입 차단 시스템을 작동하거나 보안담당자에게 즉각 현황을 알려줌으로써 사람의 관리로 인해서 발생하는 취약점을 보완할 수 있게 해주는 시스템이다. 이와 같은 시스템을 구축하기 위해서는 일반적으로 (그림 1)과 같은 구조가 필요하다[3, 4].



(그림 1) 비디오 감시 시스템의 구성

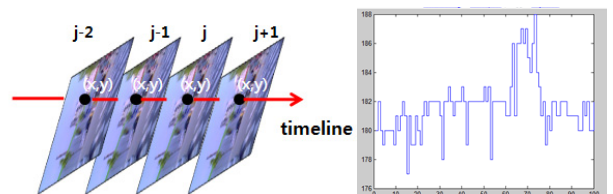
2) 가우시안 혼합모델을 이용한 감시시스템

실시간 영상감지 시스템에 있어서 (그림 1)의 첫 번째 단계인 전경(foreground) 물체와 배경(background)

을 정확하고 효율적으로 분리하는 것은 아주 중요한 문제이다. 이 단계에서 정확하게 물체를 분리해내지 못한다면 다음 단계인 물체 추적(Tracking)과 관심물체(Interesting Object)의 행동분석(behavior analysis)은 매우 어려운 작업이 된다.

효과적인 배경의 제거를 위해서는 적용환경에 대한 고려가 필요하다. 실내에 비해서 조명제어가 어려운 실외에서의 분리는 보다 어렵고 복잡하다[5]. 실외에서는 갑작스런 물체의 등장으로 인한 변화(sudden on-off changes)뿐만 아니라 분수대, 나뭇가지의 흔들림, 광고 모니터의 깜박거림, 호수 물결의 표면 흔들림, 에스컬레이터의 움직임 등을 고려해야 한다. 배경의 제거를 위해서 픽셀 기반의 배경분리 기법인 GMM(Gaussian Mixture Model)이 주로 사용되고 있다 [6, 7].

GMM은 아래의 (그림 2)와 같이 짧은 시간에 대한 frame들이 있을 때, 특정 픽셀은 각각 시공간적으로 독립이며 동일한 확률적 모델을 이룬다고 가정하고 확률적으로 접근한다[8]. 최적의 모델을 찾기 위해서 반복적 절차를 이용하여 학습하며 이를 EM(Expectation Maximum)이라고 한다.



(그림 2) 데이터의 적층을 통한 확률적 모델 획득

시간의 흐름에 따른 각각의 픽셀값을 고려하면 아래의 수식(1)과 같은 누적값의 획득이 가능하다.

$$\aleph = \{x_j = I_j(x, y)\}_{j=1}^n \quad (1)$$

여기서 $I_j(x, y)$ 는 위치(x, y)에서의 픽셀에 대한 시간 j 의 grayscale 값이다. \aleph 에 대한 가우시안 확률 혼합모델은 아래 수식(2)의 PDF(probability density function)을 통해서 얻을 수 있다.

$$f(x|\theta) = \sum_{i=1}^K \omega_i \eta(x|\theta_i) \quad (2)$$

위의 수식(2)에서 ω_i 는 Gaussian component C_i 에 i 번째 정규화된 가중치 값이다. 따라서 $\sum_{i=1}^K \omega_i = 1$ 이다. $\eta(x|\theta_i)$ 는 C_i 를 위한 PDF 이고 이것은 수식(3)과 같이 표현할 수 있다.

$$\eta(x|\theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2\right] \quad (3)$$

관측 데이터의 집합 x 에 대해서 ω_i 번째 성분 파라미터 $\theta = \{\omega_i, \theta_i\}_{i=1}^K$ 로 이루어진 PDF 값을 계산하는 것이며, 이때 확률값을 최대화(Maximum-Likelihood)하는 값을 구하기 위해서 EM 을 반복적으로 학습하고 [9-11] 매 장면마다 평균과 분산에 대한 값을 갱신한다[7]. 본 논문에서 구현된 프로그램 또한 위와 같은 GMM 의 일반적인 방법을 사용하였으며, 위의 수식(2)에서 최근 900frame 에 대해서 각 frame 에 pixel 당 가중치를 0.00125 로 적용하였다. 가우시안 혼합 모델은 일반적으로 3 개에서 5 개를 사용하지만 실험환경에서 얻고자 하는 object 가 사람에 한정이 되어 있으므로 3 개만을 사용하였다.

위와 같이 첫 번째 단계가 이루어 지고 나면, (그림 1)의 두 번째 단계는 첫 번째 단계에서 배경을 제거한 전경물체에 대해서 ROI(region of Interest)를 설정한다. 세 번째 단계에서의 추적은 두번째 단계에서 얻어진 ROI 를 통해서 질량중심을 이용한다. 중심의 궤적과 위치 정보를 통해서 네 번째 과정인 이벤트를 감지한다[12]. 본 논문에서는 위와 같은 과정에 기반한 특이 행동 인지를 목적으로 한다. 이를 위해서 특정환경에서 몇 가지의 행동 패턴을 정의하고 실험을 통하여 패턴인식을 위한 효율적인 방법을 도출하였다.

2. 특이 행동 인식

1)Optical Flow 를 이용한 이상 행동의 인지

Optical Flow 는 실시간 영상에서 특징점이 시간에 따라서 변할 때 이동되는 vector 의 크기와 방향 정보이다. 이러한 데이터는 정보 특징의 적당한 조합을 통해서 여러 가지로 해석이 가능하다[13, 14]. 실제로 일본의 미쯔비시는 엘리베이터에서 일어나는 범죄를 예방하기 위한 방법으로 Optical Flow 를 이용하여 시

스템을 구축하였다[15]. 여기서는 세가지 요소(D: 흐름에 방향, M: 흐름의 크기, Q: vector 의 개수)들의 데이터를 이용하여 아래와 같은 방식을 취하였다.

평균값 : $(\bar{D}, \bar{M}, \bar{Q})$, 표준편차 : $(\sigma_D, \sigma_M, \sigma_Q)$

$$P_d = \frac{D - \bar{D}}{\sigma_D}, P_m = \frac{M - \bar{M}}{\sigma_M}, P_q = \frac{Q - \bar{Q}}{\sigma_Q} \quad (4)$$

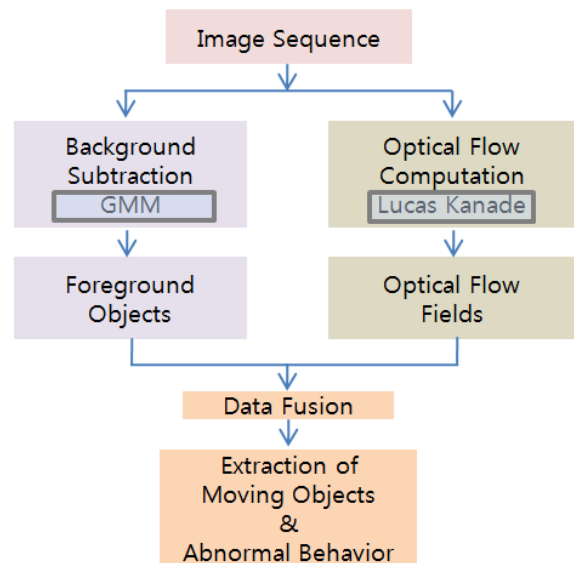
을 활용하여 아래와 같은 수식(5)를 구하여 활용하였다.

$$\text{Estimate Value} = P_d \cdot P_m \cdot P_q \quad (5)$$

수식(5)와 같이 단순히 데이터를 곱하는 형식을 취하였다. 사람의 운동패턴은 평소에는 간단하지만 폭력적인 상황이나 특이한 행동의 경우에는 패턴이 매우 복잡하다는 것을 이용한 것이다. 그러나 엘리베이터의 내부환경(인테리어, 벽면의 유리, 카메라의 위치)을 모두 배경으로 간주하였고 엘리베이터 승차자의 옷의 특징에 따른 변화를 고려하지 않았다. 따라서 특징 값(D, M, Q)을 통해서 이상 행동을 의심할 수는 있지만 명확한 기준을 두고 사용 할 수 없다는 단점이 있다. 본 논문에서는 위와 같은 단점을 극복하기 위해서 vector 의 방향에 대한 평균과 분산, 그리고 vector 의 크기에 대한 평균과 분산 결과를 도출하여 명확한 활용방안을 제시하였다.

2)환경에 따른 패턴 정의와 알고리즘

이상행동을 인지하기 위해서는 적용하고자 하는 환경에 맞게 관찰대상 물체의 패턴을 정의해야 한다. [12]에서는 사무실, 복도, 지하철, 캠퍼스 내에서 일반적인 객체의 움직임을 Foreground pixels 의 추적과 물체간의 휘도(luminance contrast) 차이를 이용하고, blob 의 궤적을 이용하여 이상 행동을 관찰하였다.

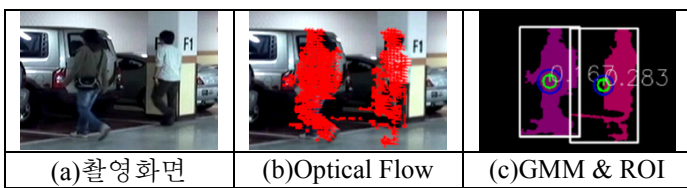


(그림 3)알고리즘 다이어그램

본 논문에서는 위와 같은 이전의 연구 내용을 바탕

으로 하여 주차장이라는 특정환경에서 뛰다, 걷다, 도난, 폭력이라는 4 가지 새로운 행동 패턴을 정의하여 실험을 수행하였다.

위의 (그림 3)은 GMM 모델을 이용하여 추출된 물체와 Optical Flow 를 이용한 데이터를 결합에 의한 제안된 시스템의 다이어그램이다. GMM 은 online-EM 알고리즘[16]을 이용하였으며 Optical Flow 는 [17]의 방법을 이용하였다. 아래의 (그림 4)는 이러한 방법을 이용한 실험 화면이다. (a)는 실제 비디오의 촬영영상이며 (b)는 관심영역에 대한 Optical Flow 영상이다. (c)는 GMM 을 통해서 추출된 Object 에 대해서 데이터를 추출하는 장면이다. 중간의 원은 평균과 분산, 평균 방향을 나타낸다.



(그림 4)GMM 과 Optical Flow 를 물체 및 특징 추출

3)실험

아래의 (그림 5)는 주차장에서 연출된 상황을 통해서 실험을 하는 장면이다. 연출된 상황은 (a)도난, (b)뛰다, (c)폭력, (d)걷다 의 4 가지로 실험을 하였다.

(a)도난 450frame 이상 행동		
(b)뛰다 95frame 이상 행동		
(c)폭력 750frame 이상 행동		
(d)걷다 350frame 정상 행동		

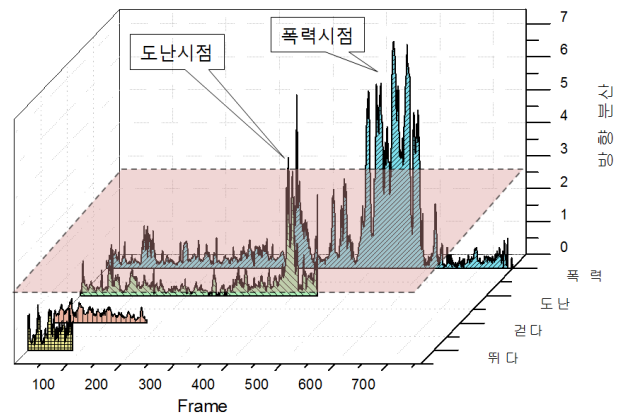
(그림 5)상황에 따른 데이터 추출

3. 결과 및 향후 과제

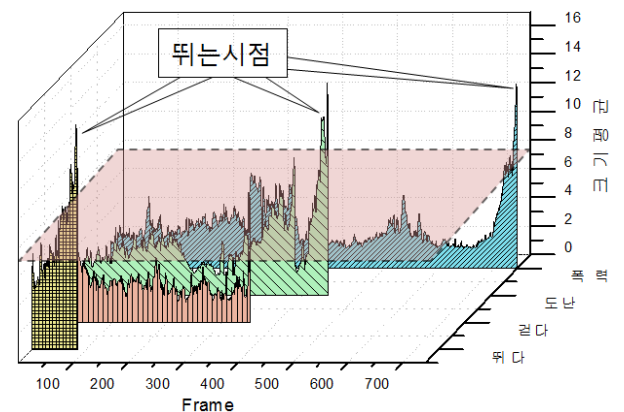
1)실험 결과

(그림 5)의 상황에 따라서 추출된 물체의 Optical Flow 에 대한 ‘방향 분산(그림 6)’, ‘크기의 평균(그림 7)’, ‘크기의 분산(그림 8)’, ‘방향의 평균(그림 9)’을 구하였다. Optical Flow 의 방향은 기준 축을 양의 x 축으로 하는 각도(radian)값이다.

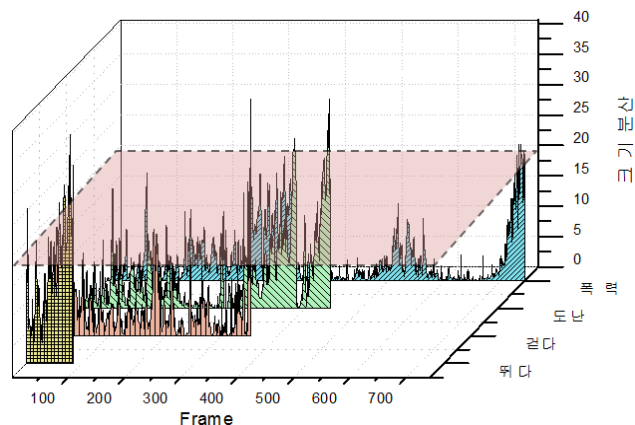
(그림 6)은 vector 의 방향에 대한 분산 값을 살펴보면 도난과 폭력상황에서 peak 지점을 가지는 것을 알 수가 있다. 폭력이나 도난 시에는 vector 의 방향에 대한 변동이 크기 때문이다.



(그림 6)상황에 따른 Optical Flow 방향의 분산



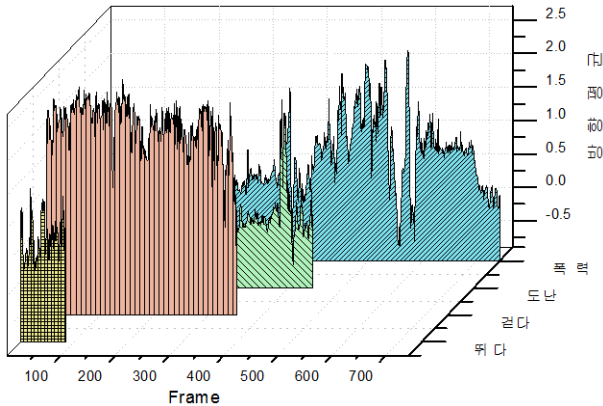
(그림 7)상황에 따른 크기의 평균값



(그림 8)상황에 따른 크기의 분산

(그림 7)은 이상행동 이후에 범의자가 빠른 걸음으로 현장을 뛰어나가는 시점에서 peak 가 발생하는 것을 알 수가 있다. 아래의 (그림 8)도 (그림 7)과 비슷한 그래프의 형태를 가지는 것을 알 수가 있다.

아래의 (그림 9)는 정상적인 행동인 ‘걷다’에서 가장 높은 값을 지속적으로 유지하는 것을 볼 수가 있다. 이것은 주차장이라는 환경에서 차량의 승하차 시에 정상적으로 일어나는 것이며 vector 의 평균방향이 일정하게 유지되기 때문이기도 하다.



(그림 9)상황에 따른 방향의 평균

(그림 6~9)를 통해서 정상적인 행동이 아닌 이상행동 시에는 vector 의 방향과 크기에 대한 평균과 분산이 큰 차이를 보인다는 것을 알 수가 있었다.

본 논문에서는 GMM 과 Optical Flow 의 혼합적인 방법을 이용해서 관심영역에 관한 행동인지에 관련된 데이터를 추출하였다. 이러한 데이터를 활용하여 응용을 위한 다양한 해석이 가능하다. 예를 들어서 (그림 6~9)에서 gray 면으로 표시된 중간의 평면영역은 이상행동을 효율적으로 감지하기 위한 임계치를 설정해 놓은 것이다. 환경에 영향을 받지 않는 적응적인 임계치 결정을 통해 임계치를 넘기면 이상행동으로 간주하는 방법을 통해서 주차장이라는 환경에서의 인지가 효과적으로 이루어 질 수 있다.

2)향후 과제

제안된 알고리즘을 통한 실험에서 이상 행동 인지에 대한 의미 있는 결과를 도출하였으나 실험과정에서 몇 가지 개선 사항이 발생하였다. 첫째 본 논문에서는 추출된 물체에 대해서 그림자를 제거하지 않았다. 그림자의 정보 또한 물체와 관련된 정보 간주했기 때문이다. 그러나 주차장이라는 환경에서 자동차 표면에 반사된 조명 또한 물체로 인식하는 오류가 발생하였다. (그림 6~9)의 그래프를 보면 이상행동을 가지는 그래프가 Multi-modal 형태를 이룬다는 것을 알 수가 있다. 이것은 확실적인 방법을 통해서 보다 유연한 임계치의 설정이 가능 하다는 것을 의미한다. 향후 연구에서는 여러 환경과 보다 많은 상황에서의 실험을 필요로 한다.

참고문헌

[1] 유장희, 문기영, 조현숙, *지능형 영상보안 기술현*

향 및 동향 전자통신동향분석, 2008. 제 23 권(4 호).
 [2] 문현찬, *지능형 CCTV 기술 및 시장 동향*. 주간기술동향, 2008. 제 1361 호.
 [3] T. P. Chen, H.H., A. Bovyrin, R. Belenov, K. and A.K. Rodyushkin, and V. Eruhimov, *Computer Vision Workload Analysis: Case Study of Video Surveillance Systems. Intel Technology Journal*, 2005. 9(12).
 [4] Porikli, F. and O. Tuzel. *Human body tracking by adaptive background models and mean-shift analysis*. 2003.
 [5] Li, L., et al. *Foreground object detection from videos containing complex background*. 2003: ACM New York, NY, USA.
 [6] Elgammal, A., D. Harwood, and L. Davis, *Non-parametric Model for Background Subtraction*. LECTURE NOTES IN COMPUTER SCIENCE, 2000: p. 751-767.
 [7] Stauffer, C. and W. Grimson. *Adaptive background mixture models for real-time tracking*. 1999.
 [8] Stauffer, C. and W. Grimson, *Learning Patterns of Activity Using Real-Time Tracking*. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2000: p. 747-757.
 [9] Friedman, N. and S. Russell, *Image segmentation in video sequences: A probabilistic approach*. University of California, Berkeley, 1997. 94720.
 [10] Dempster, A., N. Laird, and D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society, 1977. 39(1): p. 1-38.
 [11] Neal, R. and G. Hinton, *A view of the EM algorithm that justifies incremental, sparse, and other variants*. Learning in Graphical Models, 1998. 89: p. 355-368.
 [12] Fuentes, L. and S. Velastin, *People tracking in surveillance applications*. Image and Vision Computing, 2006. 24(11): p. 1165-1171.
 [13] Bors, A. and I. Pitas, *Optical flow estimation and moving object segmentation based on median radial basis function network*. Image Processing, IEEE Transactions on, 1998. 7(5): p. 693-702.
 [14] Chunke, Y. and S. Oe, *A new gradient-based optical flow method and its application to motion segmentation*. Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conf. journal of the IEEE, 2000. 2.
 [15] Kage, H., et al. *Pattern recognition for video surveillance and physical security*. 2007.
 [16] KaewTraKulPong, P. and R. Bowden. *An improved adaptive background mixture model for real-time tracking with shadow detection*. 2001.
 [17] Lucas, B. and T. Kanade. *An iterative image registration technique with an application to stereo vision*. 1981.

본 과제는 교육인적자원부, 산업자원부 노동부의 출연금으로 수행한 산학협력중심대학 육성사업의 연구 결과입니다.

본 연구는 산업자원부와 한국산업기술재단의 지역혁신인력양성사업으로 수행된 연구결과임