

유연성 다중 회귀 모델을 활용한 보행자 이상 행동 예측 모델 연구[†]

정유진¹ · 윤용익²

^{1,2}숙명여자대학교 멀티미디어학과

접수 2015년 8월 6일, 수정 2015년 10월 29일, 게재확정 2016년 1월 14일

요약

최근 강력 범죄 및 우발 범죄가 끊이지 않고 있으면서 사회적 불안감이 고조되고 있다. 이에 따라 방범용 카메라, CCTV (Closed Circuit Television)가 범죄 증거 확보와 치안을 위해 사용되고 있다. 그러나 CCTV는 주로 사후 처리 기능으로 사용하고 있으며 사전에 범죄를 예방하기는 힘들다. 본 연구에서는 CCTV로부터 수집된 보행자 데이터를 이용하여 객체의 행동을 분석하고 위험 행동 여부를 추정하기 위한 유연성 다중 회귀 모델을 제안한다. 유연성 다중 회귀 모델은 필터링, 상황분석, 예측 단계로 구성되어 있다. 먼저 보행자에 대한 환경과 상황에 대해 필터링한 후 상황분석에 대한 정보를 구축하고 관찰 객체에 이상 행동이 결정된다. 마지막으로 연관분석을 통해 객체의 행동이 예측되어 위험 상황을 통지한다. 이를 통해 다중 지역에서 객체의 행동을 추적하여 객체 행동의 위험여부를 알 수 있으며, 행동 예측을 통해 범죄 발생을 예측 가능하다.

주요용어: 상황인지, 연관분석, 이상행동, 행동예측, 확률모형.

1. 머리말

연일 터져 나오는 범죄소식에 사회적 불안감을 느끼고 있다. 사회가 급속도로 발전하면서 5대 범죄를 포함하여, 특히 서로간의 상관관계가 존재하지 않고 불특정 대상에게 행해지는 범죄, 이른바 묻지마 범죄의 발생으로 사람들은 치안 불안감이 고조되고 있다. 이에 따라 보안에 대한 인식 또한 높아지고 있다. 그러나 Bark (2012)에서 묻지마 범죄의 개념은 행위에 대한 특성이 나타나지 않는 모호한 개념이라는 점에 문제가 있다. Kim 등 (2014)연구와 같이 범죄의 특성을 분석하고 관련 기술을 통해 Park과 Jung (2013)과 같이 보행자의 안전을 확보하는 많은 과학적인 방법이 연구되고 있다. 이에 사람들은 범죄 상황이 발생하기 전에 미리 예방할 수 있는 방법에 대해 관심을 두고 있으며, 관심도가 높아짐에 따라 범죄 발생 예측에 대한 필요성이 대두되고 있다.

이러한 범죄 상황 발생 시 보안과 증거확보를 위해 보안 시스템의 대표적인 CCTV시스템의 역할이 커지면서 유동 인구가 많은 지역 사회부터 점차 설치 영역을 넓혀가고 있다. CCTV의 영역이 확장됨에 따라 관련 기술 또한 지속적으로 발전하고 있다. 기존의 감시 시스템은 감시자가 여러 대의 모니터 화면을 통해서 감시하는 시스템이었으나, 오늘날 자동으로 영상을 인식하고 관찰하며 실시간으로 통보하는

[†] 이 논문은 숙명여자대학교 교내연구비지원에 의해 수행되었음 (과제번호 1 - 1503 - 0141).

¹ (140-742) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 멀티미디어학과, 대학원생.

² 교신저자: (140-742) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 멀티미디어학과, 교수.

E-mail: yiyoon@sm.ac.kr

기술이 발전되고 있다. 그러나 우리나라의 개발 및 성장 속도는 빠르게 발전하고 있지만 기술의 완성도는 아직 부족한 상태이다 (Chang 등, 2009).

본 연구의 목적으로는 CCTV로부터 수집된 영상정보의 대표적인 정보인 행동과 환경 정보를 통해 데이터 상관분석 및 유연성 다중 회귀 모델 (FMR: Flexible Multi-level Regression)을 실시하여 환경변수와 행동의 위험 및 이상 행동 여부를 추정할 수 있는 모델을 구축하고자 한다. 이를 통해 객체 행동의 이상 행동 여부와 행동의 위험 여부를 알 수 있으며, 축적된 데이터를 바탕으로 행동 예측을 통해 범죄 발생을 예측하고자 한다.

본 연구의 범위로는 CCTV로부터 객체를 인식하고 파악한 후 객체의 행동과 환경정보가 주어진 상태에서 실시한다. 객체는 2명 이상의 다중객체일 때 분석되며, 한 객체의 행동이 정상적인 움직임으로 보기 어려운 행동이라고 판단되었을 때 주시하게 된다. CCTV의 범위는 다중 지역의 다중 카메라로 설정하였다. 환경변수는 자료를 분류한 후 상관분석을 통한 통계적 분석을 실시하였다. 행동 예측 분석은 R을 사용하고 통계분석은 SPSS18을 사용하였다. 분석된 모델을 통해 4절의 시나리오에 적용할 것이다. 시나리오의 상황은 폭력으로 가정하였다.

2. 이상 행동 정의 및 분류

이상 행동에 대해 명확히 정의된 개념이 없다. 그러나 대부분 Hong (2000)에서 정의한 이상 행동과 같이 이상 행동이란 평균과 같은 정상적 기준에서 벗어나있는 즉, 일정한 기준에서 벗어나 있음을 뜻하는 용어로 일반적이지 않은 움직임이라고 정의할 수 있다.

따라서 본 연구에서 이상 행동이란, 일반적인 상황으로부터 행동에 대한 평가 (assessment)의 기준에서 이탈된 행동을 하거나 범위를 벗어난 경우 이상 행동라고 정의한다.

이상 행동의 분류로는 행동은 현재 연구에서 CCTV로 인식할 수 있는 행동으로 크게 10가지 (Kicking, Punching, Hugging, Pushing, Shaking, Pointing, Running, Walking, Sitting, Standing)로 추출하였으며, 이상 행동의 위험도에 따라 위험 (dangerous), 경고 (caution), 보통 (normal)인 3가지로 분류하였다. 분류된 행동은 Table 2.1과 같다.

Table 2.1 Abnormal behavior clustering

Categorize	Action
Dangerous	Kicking
	Punching
	Hugging
	Pushing
Caution	(head, hand etc.) Shaking
	Pointing
	Running
Normal	Walking
	Sitting
	Standing

수집된 행동을 Table 2.1에 의해 분류한다. 행동의 위험도가 보통 (normal)으로 분석될 경우에는 아무런 조치가 행해지지 않으며 위험하지 않다고 본다. 만약 수집된 행동이 경고 (caution)로 분석될 때에는 객체를 면밀히 관찰하게 된다. 이때 표적 객체를 추적 (target tracking)하며 다음 행동이 위험 정도까지 행해지는지 감시한다. 이때부터 객체의 행동은 이상행동으로 판단된다. 마지막으로 객체의 행동이 위험 (dangerous)단계에 오면 위협을 받는 객체를 포함하여 주변 공공기관에 알람을 주게 된다.

그러나 위험 (dangerous) 단계에서는 이미 범죄가 행한 단계이기 때문에 본 연구에서는 행동의 위험도가 위험 (dangerous) 단계로 가기 전 경고 (caution) 단계에서 이상 행동 및 범죄 징후가 파악되는 것을 목표로 한다.

3. 이상 행동 예측 모델

3.1. 이상 행동 예측 구성도

Figure 3.1은 보행자 이상 행동을 예측하기 위한 전체적인 구성도이다. 본 연구에서는 객체인식과 객체 행동 및 환경에 대한 정보 수집을 마치고 수집된 데이터를 넘겨받은 상태에서의 감시 서비스 미들웨어 (ISM: Inspect Service Middleware) 부분을 다룰 예정이다.

감시 서비스 미들웨어 (ISM)는 크게 2가지인 상황인지 (situation awareness), 예측 (prediction)로 나뉘며 총 4단계인 데이터분류 (data clustering), 행동분석 (behavior analysis), 행동분류 (behavior classification), 행동예측 (behavior prediction)을 거친다. 이번 3절에서는 각각의 단계에 분석 결과 및 알고리즘에 대해 다룰 예정이며, 본 연구에서는 행동분류 (behavior classification)에 대한 내용은 언급하지 않는다.

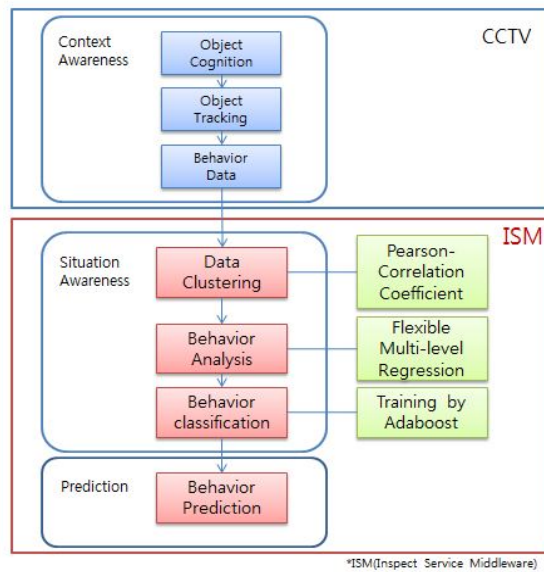


Figure 3.1 Behavior prediction configuration

CCTV의 영상 정보로부터 수집된 데이터로 객체의 움직임을 감지하여 동적 객체를 판별하여 객체를 인식하고 추적한다. 객체를 트래킹하면서 객체에 대한 행동정보를 수집한다. 이때 날씨, 지역, 시간, 거리와 같은 환경정보를 함께 수집한다. 이렇게 전반적인 상황에 대해 인식한다.

감시 서비스 미들웨어 (ISM)에서는 CCTV로 부터 주어진 객체의 행동정보와 환경정보를 분석하여 이상 행동을 잡아내는 단계이다. 환경변수 데이터는 상관분석을 통해 분류·분석되고, 이를 통해 행동과 환경변수간의 행동분석을 실시한다. 유연성 다중 회귀 모델 (FMR)을 통해 수집된 환경변수가 관찰된 객체의 행동에 대해 얼마만큼의 영향을 주는지 분석해 본다.

마지막으로 수집된 행동데이터로부터 행동의 연관 규칙을 통해서 다음 수행할 행동을 예측하게 된다.

3.2. 데이터 정의 및 상관관계 분석

실제 데이터는 카메라가 설치된 3곳의 다중 CCTV를 이용하여 두 명의 객체가 이상 행동을 하는 실험 영상을 촬영하였다. 실험 영상 정보를 수집한 후 xml 데이터를 파싱하여 분석하였다. 먼저 환경변수는 카메라로부터 얻을 수 있는 이상 행동에 영향을 미치는 환경변수 중 가장 중요하다고 생각되는 변수 4가지를 선정하였다. 본 연구에서 고려할 환경변수는 거리, 지역, 시간, 날씨이며 아래와 같이 분류하였다.

• 거리 정보

영상을 통해 인식된 객체와 객체간의 거리를 측정된 값을 말한다. 단, 객체간은 모르는 사이로 가정한다. 객체간의 거리는 Edward (1963)의 공간학 (proxemics)에 의해 4단계로 분류하여 각각에 위험도를 부여한다.

• 지역 정보

CCTV의 위치나 지역 정보 등 위치에 대한 정보를 포함한다. 주로 위험지역을 판별하는 정보로도 사용된다. 지역 정보는 거리의 폭을 말하여 미터 (m)로 측정된다.

• 시간 정보

CCTV를 통해 현재 시간을 파악한다. 시간은 새벽/아침/오전/오후/저녁으로 나뉘며 경찰청의 시간에 따른 범죄발생빈도에 따라 위험도를 나눈다.

• 날씨 정보

운량이나 밝기 등과 같은 주변의 날씨정보를 포함한다. 날씨, 밝기 등 에 의해 위험도 수준이 고려되며 본 연구에서는 운량을 날씨정보로 포함하였다.

이러한 4가지 환경변수를 SPSS를 이용하여 행동과 환경변수간의 상관분석을 실시하였다. 상관분석은 변수들 간에 존재하는 선형관련성의 정도를 측정하여 분석하는 방법이며 이때 선형관련성의 정도를 측정하는 척도로 상관계수를 사용한다. 상관분석 결과는 Table 3.1과 같다.

Table 3.1 Correlation coefficient

	Distance	Area	Time	Weather	Behavior
Distance	1				
Area	.567**	1			
Time	-.861**	-.501**	1		
Weather	.690**	.439**	-.819**	1	
Behavior	-.596**	-.485**	3.97**	-.239**	1

** $p < 0.01$

행동 (behavior)과 거리 (distance), 지역 (area), 시간 (time), 날씨 (weather)를 변수로 하여 상관분석을 실시한 결과이다. 상관분석 결과 행동, 거리, 지역, 시간, 날씨 변수가 유의확률 값이 유의수준 0.05보다 작으므로 상관관계가 존재한다고 할 수 있다. 행동과의 상관분석을 살펴보았을 때 시간 0.397로 이는 두 변인은 통계적으로 유의미한 양의 상관관계에 놓여있으며, 상관의 정도가 높다는 것을 알 수 있다. 나머지 3개의 변수인 거리는 -0.596, 지역은 -0.485, 날씨는 -0.239로 모두 통계적으로 유의미한 음의 상관을 갖으며, 상관의 정도는 비교적 높은 상관임을 알 수 있다.

3.3. Flexible multi-level regression model

3절에서는 앞서 분류한 데이터들의 특성을 모두 고려하여 거리, 지역, 시간, 날씨의 환경변수가 이상 행동에 영향을 끼치는 설명력 정도를 확인하기 위하여, 유연성 다중 회귀 모델 (FMR)을 실시한다. 본 연구에서는 거리, 지역, 시간, 날씨를 고정 변수 (fixed feature), 유연성 변수 (flexible feature)로 나누어 행동의 위험여부를 추정할 수 있는 모형을 구축하고자 하였다.

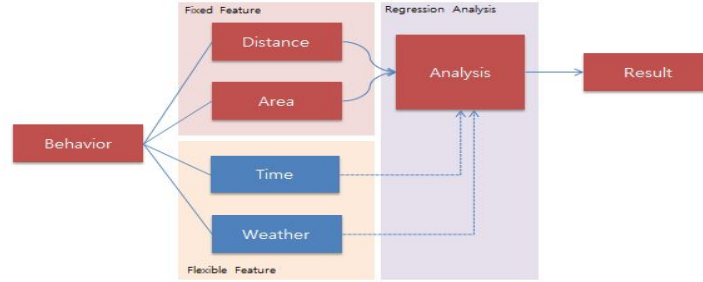


Figure 3.2 Overall flexible multi-level regression analysis model

Figure 3.2을 살펴보면 먼저 행동에 대해 각각의 변수를 2가지로 나누었다. 거리와 지역은 고정된 (fixed)한 변수로 설정하고, 시간과 날씨는 유연한 (flexible)한 변수로 설정하였다. 평소 날씨변수는 변화가 없을 경우 결측 변수로 처리되어 분석되지 않으며, 날씨가 변화할 때 분석되는 변수로 유연한 (flexible)한 변수로 처리된다.

본 연구에서는 행동, 거리, 지역, 시간, 날씨의 변수를 사용하여 적절한 모델을 찾기 위해 식 (3.1)과 같은 통계학적 분석 방법인 회귀분석을 활용하였다. 이모형을 이용하여 독립변수들의 변화에 따른 종속 변수의 변화를 예측할 수 있다. 통계학적 분석 결과 식 (3.2)와 같은 회귀모형을 구축하였다.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \quad (3.1)$$

이 방정식은 독립변수와 종속변수 사이의 관련성을 수학적 모형을 이용하여 측정된 변수들의 자료로부터 추정하고 분석하는데 적합한 방정식이다. 가설검증을 위해 유의확률 (p -value) 값을 확인하여 가설의 채택, 기각 여부를 파악한다. 일반적으로 유의확률의 값은 0.05를 기준으로 한다.

$$\text{behavior} = 5.42 - 0.476 \times \text{distance} - 0.270 \times \text{area} \quad (3.2)$$

식 (3.2)는 유연성 다중 회귀 모델의 고정 변수만을 고려한 방정식이다. 이를 통해 지역변수 -0.476의 회귀계수는 지역변수를 통제하고 거리를 한 단위 증가 시킬 때 행동은 -0.476이 증가되는 것을 알 수 있다. 또한 결정계수 (R^2)는 0.388로 나타났으며, 추정된 회귀선이 총변동의 38.8%를 설명하고 있음을 알 수 있다.

이를 통해 나온 결과는 Table 3.2과 같이 행동에 대한 평가 (assessment)가 이루어진다. 분석된 결과 값을 통해 각각의 범위에 해당되는 위험도가 나오게 되며 행동의 위험여부를 추정할 수 있다.

Range	Degree of Risk
~ 23	Normal
24 ~ 55	Caution
56 ~	Dangerous

결과 값이 23이하인 경우 위험도는 정상으로 행동이 위협하지 않은 상태이며, 24이상 55미만일 경우 위험도는 경고 단계이고, 마지막으로 56 이상의 값이 산출되면 위험도가 매우 위험한 상태라고 판단할 수 있다.

4. 시나리오

Figure 4.1은 이상 행동 예측 모델을 적용한 예상 시나리오이다. 감시 서비스 미들웨어 (ISM) 단계에서는 CCTV로부터 수집된 데이터를 전송 받아 앞서 언급한 상황 인지 (situation awareness), 예측 (prediction) 단계를 거치게 된다.

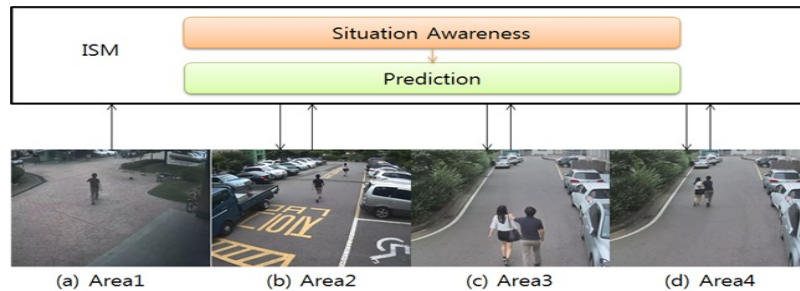


Figure 4.1 Application of scenario

먼저 Area1에서 Camara₁이 다중 객체를 인식하여 인식된 객체를 감시 서비스 미들웨어 (ISM)에 전달한다. Area2서는 인식된 객체들의 행동과 객체간의 거리를 파악한 후 현재 위치에 대한 지역을 파악한다. 계속적으로 Object1의 행동을 추적하여 현재 시간과 날씨의 정보도 함께 수집한다. 유연성 다중 회귀 모델 (FMR)을 통해 관찰된 객체의 행동이 경고행동을 취하게 되면 Camara₂는 Object1의 행동을 면밀하게 감시 관찰하게 된다. 지금까지 수집 및 분석된 객체의 행동을 연관규칙 알고리즘 및 분류 알고리즘을 활용하여 수집된 행동데이터로부터 행동에 대한 패턴을 분류하고 예측하여 이상 행동을 하거나 이상 행동이 일어날 확률이 높을시 관찰 객체와 거리가 가까운 다른 객체에게 알림을 준다.

위에 적용된 모델을 통해 현재 행동을 분석하고 앞으로 일어날 행동에 대해 미리 예측함에 따라 범죄가 일어나기 전 사전 대응이 가능하고 즉각적인 대응이 가능해진다.

5. 결론 및 향후 연구

본 연구에서는 다중 CCTV로부터 범죄 상황 발생 전에 미리 예방할 수 있는 이상 행동 예측 모델을 제시하였다. 또한 사후처리의 용도 및 기존의 모니터링 용도뿐만 아니라 보행자의 이상 행동을 환경변수와 함께 고려하여 분석하였다. 분석된 데이터를 통해서 최종 의사결정에 영향을 미치는 행동 예측 모델이며, 그 중 감시 서비스 미들웨어 (ISM)분석에 대해 중점을 두었다. 유연성 다중 회귀 모델 (FMR)을 통해 행동에 대한 평가 규칙 (assessment rule)을 수립하여 이를 바탕으로 CCTV로부터 객체의 행동을 관찰하여 이상행위를 예측, 검출하였다. 제안된 모델은 자원의 비용을 감축할 수 있으며 범죄가 발생하기 전 객체에 대한 행동을 미리 예측하여 상황판단 및 의사결정을 신속히 할 수 있다. 이를 통해 범죄 사전에 예방을 할 수 있는 적합한 모델이라고 보여 진다. 그러나 환경요소를 활용하여 이상 행동을 분별해 내기 위한 지속적인 행위 분석 모델과 수집 및 분석된 객체의 행동에 대해 분류하는 분류 알고리즘은 추후 연구로 남아있다.

References

- Bark, H. M. (2012). Concept and features of random crime: Crimes against random people. *Journal of the Korean Association of Public Safety and Criminal Justice*, **50**, 226-258.
- Chang, I. S., Cha, H. H., Park, G. M., Lee, K. J., Kim, S. K. and Cha, J. S. (2009). A study of scenario and trends in intelligent surveillance camera. *Journal of the Intelligent Transport System*, **24**, 93-101.
- Cho, K. H. and Park, H. C. (2011). A study insignificant rules discovery in association rule mining. *Journal of the Korean Data & Information Science Society*, **22**, 81-88.
- Edward, T. H. (1963). A system for the notation of proxemic behaviour. *Journal of the American Anthropologist*, **65**, 1003-1026.
- Hahsler, M., Grun, B., Hornik, K. and Buchta, C. (2009). *Introduction to arule- A computational environment for mining association rules and frequent item sets*, The Comprehensive R Archive Network, USA.
- Hong, S. Y. (2000). *Criminal psychology*, Hakjisa, Korea.
- Kim, C. K., Kang, I. J., Park, D. H. and Kim, S. S. (2014). Analysis of the five major crime utilizing the correlation regression analysis with GIS. *Journal of the Korean Society for Geospatial Information System*, **22**, 71-77.
- Lim, J., Lee, K. and Cho, Y. (2010). A study of association rule by considering the frequency. *Journal of the Korean Data & Information Science Society*, **21**, 1061-1069.
- Park, S. H. and Jung, S. H. (2013). *A preliminary research on technology development to ensure the safety of pedestrian*, National Disaster Management Institute, Korea.

Study on abnormal behavior prediction models using flexible multi-level regression[†]

Yu Jin Jung¹ · Yong Ik Yoon²

^{1,2}Department of Multimedia Science, Sookmyung Women's University

Received 6 August 2015, revised 29 October 2015, accepted 14 January 2016

Abstract

In the recently, violent crime and accidental crime has been generated continuously. Consequently, people anxiety has been heightened. The Closed Circuit Television (CCTV) has been used to ensure the security and evidence for the crimes. However, the video captured from CCTV has being used in the post-processing to apply to the evidence. In this paper, we propose a flexible multi-level models for estimating whether dangerous behavior and the environment and context for pedestrians. The situation analysis builds the knowledge for the pedestrians tracking. Finally, the decision step decides and notifies the threat situation when the behavior observed object is determined to abnormal behavior. Thereby, tracking the behavior of objects in a multi-region, it can be seen that the risk of the object behavior. It can be predicted by the behavior prediction of crime.

Keywords: Abnormal behavior, analysis, association, behavior prediction, probability model, situation awareness.

[†] This research is partially supported by a research fund (1-1503 - 0141).

¹ Graduate student, Department of Multimedia Science, Sookmyung Women's University, Seoul 140-742, Korea.

² Corresponding author: Professor, Department of Multimedia Science, Sookmyung Women's University, Seoul 140-742, Korea. E-mail: yiyoonyoon@sm.ac.kr