

옵티컬 플로우 기반 장면 모델링을 통한 교통 영상 내의 이상 상황 인식 시스템

권언혜*, 노승중*, 전문구*
*광주과학기술원 정보통신공학부
e-mail: eonhye@gist.ac.kr
ivyun@gist.ac.kr
mgjeon@gist.ac.kr

Anomaly Detection in Traffic Video Using Optical-Flow Based Scene Modeling

Eonhye Kwon*, SeungJong Noh*, Moongu Jeon*
*School of Information and Communications, GIST

요 약

최근 카메라 센서 및 알고리즘의 발달로 엔터테인먼트 목적의 영상 시스템을 비롯한 공정 기술, 교육 및 의료 등 다양한 목적의 영상 시스템이 개발 되고 있다. 또한 범죄 예방, 사고 상황 인식을 위한 감시 영상 시스템의 연구도 활발히 진행되고 있다. 본 논문에서는 이상 상황 인식을 위한 지능형 교통 시스템에 대해 제안하고자 한다. 제안하는 시스템은 크게 학습 과정과 이상 상황 인식 과정으로 나누어진다. 학습 과정에서는 CCTV와 같은 정적인 카메라에서 촬영된 도로 교통 영상에서 이동 객체의 특징을 추출하고 이를 추적하여 특징 벡터를 구성한다. 구성된 특징 벡터들은 클러스터링 기법을 통해 장면을 모델링하는데 이용되며 최종적으로 이 모델을 이용해 실시간으로 도로 교통 영상에서 이상 상황을 인식할 수 있게 된다. 실험을 통한 성능 평가를 통해 시스템의 우수함을 확인 하였다.

1. 서론

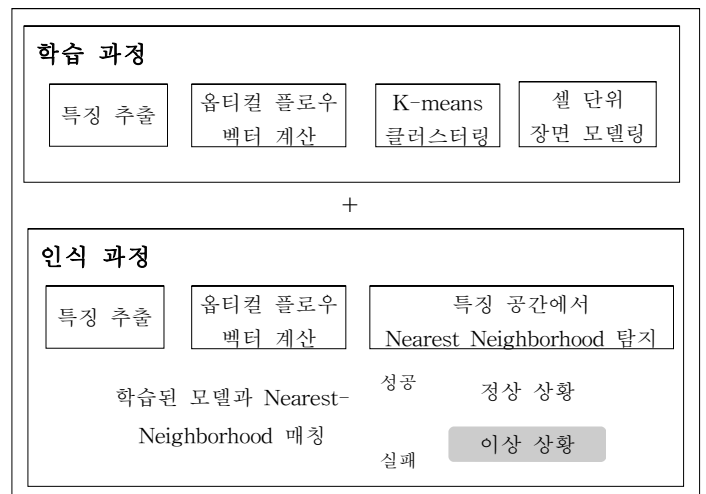
최근 카메라 센서의 발달로 다양한 산업 분야에서 영상 처리 및 패턴 인식 알고리즘을 적용한 어플리케이션이 개발되고 있으며 일부는 이미 실용화 단계에 접어들었다. 스마트폰을 비롯한 스마트 가전제품에서의 엔터테인먼트 목적의 휴먼 컴퓨팅 인터페이스뿐만 아니라 공정 기술, 교육 및 의료 분야 등에서도 영상 처리 시스템이 연구 및 개발되고 있다. 또한 범죄 예방 및 사고 상황 인식을 위한 감시 영상 시스템의 연구도 활발히 진행되고 있다.

본 논문에서는 이상 상황 인식을 위한 지능형 교통 시스템에 대해 제안하고자 한다. 제안하는 시스템의 목적은 CCTV와 같은 저해상도 카메라 센서의 입력 영상을 이용해 무단 횡단, 불법 유턴, 차선 이탈과 같은 교통 이상 상황을 인식함에 있다. 기존 교통 상황 감시 목적의 영상 시스템은 배경 모델링을 이용한 전경 검출 혹은 특정 객체 감지 알고리즘을 통한 객체 검출 이후 이에 대한 추적 결과를 바탕으로 이상 상황을 인식하였다. 하지만 이는 연산량이 많아 실시간 시스템에 적합하지 않을 뿐만 아니라 저해상도 영상에서는 현저하게 성능이 떨어진다는 단점이 있다. 우리는 이를 극복하기 위해 이동 객체가 아닌 특징 기반의 추적 알고리즘을 수행함으로써 화면내의 모션 정보를 획득한다. 먼저, 일정 학습 시간 동안 촬영된 영상에서 추적하기 좋은 특징들을 추출하고 이 특징들에 대하여 희소 옵티컬 플로우를 계산한다. 계산된 옵티컬 플로우 벡

터들이 보인 이동 객체들의 패턴 정보들은 특징 벡터로 변환되어 장면 모델링을 위한 학습 데이터로 이용된다. 장면 모델링은 K-means 클러스터링 기법을 이용하여 이루어진다. 학습 과정이 끝나면 생성된 장면 모델을 이용해 이후 시스템의 입력 영상에서 이동 객체가 보이는 이상 상황을 실시간으로 판별하게 된다.

본 논문은 제안 방법, 실험 결과, 결론의 순서로 구성되어 있다.

2. 제안 방법



(그림 1) 시스템의 프레임워크

지능형 교통 시스템에서 도로의 이상 상황 인식을 위해 본 논문에서 제안하는 방법은 먼저 특징을 추출하고 추출된 특징을 이용해 장면을 모델링하는 학습 과정을 거치게 된다. 이 과정을 거친 후 학습된 장면 모델을 이용해서 실시간으로 도로 내 이동 객체의 이상 상황 여부를 인식할 수 있다. 전체적인 프레임워크는 (그림 1)과 같다.

학습 과정과 이상 상황 인식 과정에 대한 자세한 방법은 이어지는 2.1절과 2.2절로 나누어 설명한다.

2.1. 학습

2.1.1. 특징 추출 및 추적

본 논문은 이동 객체의 궤적을 학습하여 입력 영상의 장면을 모델링하고 이를 이상 상황 인식에 이용한다. 이동 객체의 궤적은 옵티컬 플로우 벡터를 계산 하여 구한다.

감시 영상 시스템은 실시간 처리가 가능한데 큰 의미가 있다. 그러므로 밀집 옵티컬 플로우가 아닌 희소 옵티컬 플로우 계산 방법을 이용하며 이를 위해 추적하기 좋은 특징을 먼저 추출하여야 한다. 본 논문에서는 이 특징으로 Harris가 제안하는 코너 검출 방법보다 성능이 우수한 Shi와 Tomasi가 제안하는 방법으로 코너를 검출하여 이용하였다[1]. 교통 영상에서는 끊임없이 이동 객체들의 패턴이 변하므로 추적할 특징을 시스템에 부하가 없는 정도의 주기로 갱신해주어야 한다. 본 논문에서는 40frame마다 이 특징을 갱신해주었다.

추출된 특징은 모든 프레임에 대하여 Lucas-Kanade가 제안한 방법으로 옵티컬 플로우를 계산하였다[2]. 이 과정을 통해 계산된 옵티컬 플로우 벡터들 중 영상의 비이동 객체에서 발생한 벡터들은 크기가 매우 작으므로 실험적인 수치를 통해 제거해주었다. 의미 있는 장면의 모델링을 위해 최종적으로 이용하는 특징 벡터의 정보는 아래와 같다.

$$feature\ vector = (Pos_x, Pos_y, Degree)$$

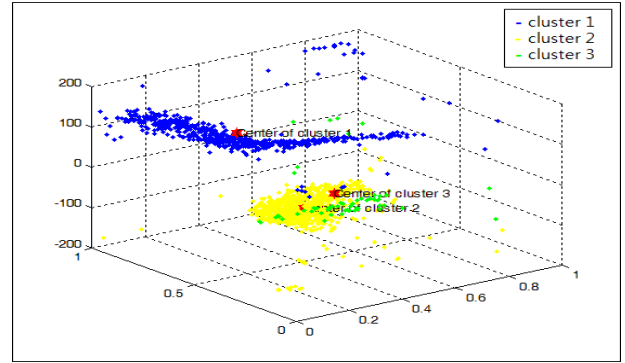
$$Degree = \arctan\left(\frac{Pos_{x_i} - Pos_{x_{i-1}}}{Pos_{y_i} - Pos_{y_{i-1}}}\right)$$

Pos_x, Pos_y 는 이동 객체의 x 축 y 축에서의 위치 정보를 표현하며 $Degree$ 는 원점을 기준으로 이동한 방향의 정보를 표현한다.

2.1.2 장면 모델링

이전 단계에서 학습을 위해 수집된 특징 벡터들은 K-means 알고리즘을 통해 특징 공간에서 클러스터링 된다. 이 때 수집된 특징 벡터들은 일반적으로 이동 객체의 위치에 대한 변화도가 이동 방향에 대한 변화도보다 크므로 이동 객체의 위치에 지배적인 클러스터링이 이루어진다. 하지만 이상 상황을 인식하기 위해서는 객체의 이동 방향에 좀 더 지배적인 클러스터링이 이루어져야 한다. 이

를 위해 Pos_x, Pos_y 의 변화도를 0~1 범위내로 줄여주었다. 이 과정을 거쳐 3차원 특징 공간에서 K-means 클러스터링 과정을 거치게 되면 사용자가 지정한 K 만큼의 클러스터들이 생성된다. (그림 2)는 이 과정을 통해 생성된 클러스터링 결과를 보여준다.



(그림 2) 특징 공간에서 클러스터링 된 결과

지금까지의 과정은 특징 공간에서 이루어진 결과이다. 이를 영상 좌표 공간에서 장면의 모델링을 위해 본 논문은 ‘셀’의 개념을 도입한다. 먼저 입력 영상을 특정 사이즈 ($size_{cell}$)의 격자로 나눈다. 분할된 각각의 격자는 지금부터 ‘셀’이라는 단위로 명명 한다. 최적화된 셀의 크기 ($size_{cell}$)는 입력 영상의 크기 및 도로 교통의 상황마다 다르므로 실험적인 수치로 최적화 해준다. 각각의 셀은 셀 내부에서 발생한 특징들의 클러스터 결과에 대한 히스토그램을 생성한다. 만약 $cell_j$ 에 포함되는 (Pos_x, Pos_y) 에서 발생한 특징점이 클러스터링 결과 $clust_k$ 로 분류 되었다면 히스토그램의 $clust_k$ 의 값을 증가시켜준다. 이 과정을 통해 셀에서 가장 많이 누적된 클러스터 값으로 셀을 레이블링 해준다. 이 과정을 표현한 수식은 아래와 같다,

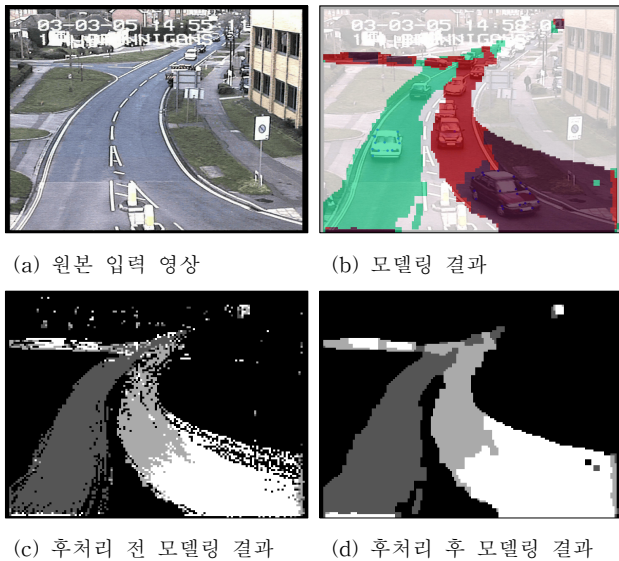
$$Label(cell_j) = \underset{k}{\operatorname{argmax}}(hist(clust_k)_j)$$

$Label(cell_j)$ 는 $cell_j$ 의 레이블링 값을, $hist(clust_k)_j$ 는 $cell_j$ 에서 생성된 히스토그램에서 k 로 분류된 데이터의 수를 의미한다. 이 과정에서 만약 어떤 특징도 나타나지 않은 위치의 셀은 ‘0’의 값으로 레이블링 해준다.

하지만 영상의 잡음으로 인해 간헐적으로 발생한 특징 벡터들도 모두 셀의 히스토그램 생성에 참조되므로 후처리를 통해 이를 제거해주어야 한다. 우리는 이를 위해 이웃하는 모든 셀의 레이블링 값과 다른 값을 가지는 셀에 대해서 이웃하는 셀의 레이블링 값을 따르도록 처리해주었다.

(그림 3)은 장면 모델링 기법의 각 과정이 도출하는 결과를 보여준다. 제안하는 방법을 통해 원본 입력 영상(그림 3-a)은 (그림3-c)와 같이 모델링되며 후처리 과정을 거쳐

최종적으로 (그림3-d)와 같이 모델링 된다. (그림3-b)를 통해 도로 내부의 이동 객체의 패턴에 따라 장면 모델링이 잘 이루어 졌음을 알 수 있다.



(그림 3) 장면 모델링 결과

2.2. 이상 상황 인식

실시간 이상 상황 인식을 위해 학습 단계에서와 마찬가지로 입력 영상에서 특징을 추출하고 추출된 특징에 대해 희소 옵티컬 플로우를 계산하여 이를 3차원의 특징 벡터로 추출한다.

우리는 장면 모델링 단계에서 수행한 K-means 알고리즘을 통해 학습 데이터에 대한 클러스터 결과 및 각 클러스터들의 센트로이드 값을 이미 알고 있다. 그러므로 실시간으로 추출된 이동 객체의 특징 벡터들과 각 클러스터의 센트로이드까지의 유클리디언 거리를 측정할 수 있다.

$$dist_{cent_i} = \sqrt{(pos_x - cent_{x_i})^2 + (pos_y - cent_{y_i})^2 + (deg - cent_{deg_i})^2}$$

이렇게 측정된 거리 중 가장 가까운 센트로이드의 클러스터와 특징 벡터가 발생한 셀의 레이블링 값이 만약 다르면 '이상 상황'으로 인식하게 된다. 그리고 '0'으로 레이블링 된 셀은 학습 당시에 이동객체가 발생하지 않은 지역으로 간주되었으므로 이러한 셀에서 특징벡터가 발견되었을 때도 이는 '이상 상황'으로 인식된다.

이러한 이상 상황 인식 과정은 객체 단위가 아닌 특징 단위에서 이루어진다. 그러므로 영상 내부 잡음이나 간헐적인 오인식이 하나의 특징점에서만 발생해도 입력 영상은 '이상 상황'으로 판단된다. 이런 현상을 방지하기 위해, 본 논문에서는 '이상 상황'으로 판단되는 특징점의 수가 일정 수($t_{anomaly}$) 이상인 프레임이 연속적으로 일정 수(t_{frame}) 이상 들어왔을 때부터 입력 영상을 '이상 상황'으

로 인식 하도록 하였다.

이어지는 3절에서는 구축된 시스템의 성능 평가를 위한 실험 방법과 수치적인 결과에 대해 제시한다.

3. 실험

3.1. 실험 방법

먼저 논문에서 사용하는 데이터 셋은 Home Office에서 제공하는 i-LIDS의 'Parked Vehicle Detection'이다.[3] 이 데이터 셋은 고정된 CCTV에서 촬영된 영상으로 불법 유턴, 차선 이탈, 무단 횡단 등과 같은 이상 상황을 포함하고 있다. 학습 단계에서는 이상 상황을 포함하지 않은 1330프레임의 영상을 이용하였다. 성능 평가를 위한 영상은 이상 상황을 포함한 500 프레임의 10개의 영상을 사용하였다. 영상의 프레임 속도는 25frame/sec다.

본 논문에서 인식하고자 하는 이상 상황은 (표 1)과 같다.

1	무단 횡단
2	불법 유턴
3	차선 이탈

(표 1) 이상 상황 정의

성능 평가를 위해 사용될 지표는 인식율과 false positive 발생 횟수 두 가지다.

$$(1) \text{Detection Rate} = \frac{\text{The Number of Successful Detection}}{\text{The Number of Anomalies}} \times 100$$

$$(2) \text{The Number of False Positive}$$

평가를 위한 영상은 실험에 앞서 발생한 이상 상황의 종류와 프레임 넘버를 통해 Ground Truth를 제작하였고 이에 대한 인식 성공률을 측정하여 성능을 평가 하였다. 실험에 사용된 최적화 수치는 아래 (표 2)와 같다.

특징 갱신 주기	40 frame
셀의 크기 ($size_{cell}$)	5×5 pixel
클러스터의 수 (K)	3
이상 이동 객체로 인식될 특징점 수 ($t_{anomaly}$)	3
이상 상황 인식 위한 연속 프레임 수 (t_{frame})	3

(표 2) 실험에 이용된 최적화 수치

3.2. 실험 결과

10개의 각 평가 영상에서의 실험 결과는 (표 3)와 같다.

이상 상황의 번호는 (표 1)에서 정의한 번호를 의미한다. 인식 여부는 'S'는 성공 'F'는 실패를 의미한다. 평가 영상에서 발생한 이상 상황은 총 16회다.

영상					False Positive 발생 횟수
1	이상 상황	2	3		1
	인식 여부	S	S		
2	이상 상황	1			
	인식 여부	S			
3	이상 상황	1			
	인식 여부	S			
4	이상 상황	3			1
	인식 여부	S			
5	이상 상황	2	1		1
	인식 여부	S	S		
6	이상 상황	1			1
	인식 여부	S			
7	이상 상황	1	2	3	
	인식 여부	S	S	S	
8	이상 상황	1			
	인식 여부	S			
9	이상 상황	1			
	인식 여부	S			
10	이상 상황	3	1	1	
	인식 여부	S	S	S	

(표 3) 실험 결과

최종 *Detection Rate* = 100%

최종 *False Positive* 발생 횟수 = 4회 (하지만 발생한 *False Positive* 현상은 시스템에 큰 영향을 주지 않는 간헐적인 현상임을 확인하였다.)

4. 결론

본 논문은 CCTV와 같은 정적 카메라의 입력 영상을 통한 도로의 교통 이상 상황 인식을 위한 시스템을 제안하였다. 연산량의 개선을 통한 실시간 시스템 구축과 저해상도 영상에서 성능 저하 현상을 개선하기 위해 특징 기반의 유틸리티 플로우를 통해 이동 객체를 추적하였다. 이 과정을 통해 특징 벡터를 구성하고 특징 공간에서 클러스터링 하였으며 이 결과는 다시 영상 좌표 공간에서의 장면을 모델링하는데 이용하였다. 모델링된 장면을 통해 최종적으로 실시간 도로 내 이동 객체의 이상 상황을 인식할 수 있다.

성능 평가를 위한 실험 결과, 논문에서 정의한 이상 상황을 100% 인식하는 높은 성능을 보였으며 *false positive* 는 4회 발생하였다. 하지만 이 또한 장시간 발생한 케이스가 아닌 간헐적인 현상으로 시스템에는 큰 영향을 주지 않음을 확인하였다.

우리는 실험을 통해 제안하는 시스템의 우수함을 확인하였으며 추후 모델링된 장면의 엡지에서 발생하는

False Positive 문제점을 개선하고 이동 객체의 방향뿐만 아닌 속도 패턴에서의 이상 상황 인식에 대하여 연구하고자 한다.

5. 감사의 글

본 연구는 한국 연구 재단 (No. 2012-000-2420)과 광주과학기술원 과학기술응용연구소의 실용화연구개발사업의 연구결과로 수행되었음.

참고문헌

[1] J. Shi and C. Tomasi, "Good features to track," 9th IEEE conference on Computer Vision and Pattern Recognition, June 1994.
 [2] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Proceedings of the 1981 DARPA Imaging Understanding Workshop (pp.121-130), 1981.
 [3] i-LIDS 'Parked Vehicle Detection' data set - <ftp://motinas.elec.qmul.ac.uk/pub/iLids/>