

Support Vector Machines를 이용한 효율적인 차량 인식 알고리즘

황원준, 송명철, 고한석
고려대학교 전자공학과

The Efficient Vehicle Recognition Algorithm using Support Vector Machines

Wonjun Hwang, Myungchul Song, Hanseok Ko
Dept. of Elcectronics Engineering, Korea University
E-mail : {wjhwang, mcsong, hsko}@ispl.korea.ac.kr

Abstract

In this paper, we describe an intelligent method to detect types of vehicles using Support Vector Machines focused to the Intelligent Transportation System (ITS) applications such as in the CCD based Electronic Toll Collection System (ETCS). This algorithm can be used the various fields of ITS applications. Support Vector Machines employed in this paper has been recently proposed as a very effective method for 3D image recognition. And our proposed feature extraction method using the singular values that directly come from pixels at input images. Consequently, The low calculation load and the high recognition rate in spite of image rotation and various noises are one of merits of proposed method.

I. 서론

차량의 수가 기하 급수적으로 증가함에 따라서 그에 대한 효율적인 관리 및 통제에 대한 요구 또한 증가하게 되었다. 이러한 차량에 대한 정보를 처리하는 분야로 최근 ITS에 대한 연구가 활발히 진행되고 있다. 본 논문에서는 영상 센서를 활용하여서 차량의 종류를 분류하는 영상기반 ETCS을 개발하기 위한 알고리즘을 제안하고 있다. 특히, 이러한 영상 기반 시스템의 장점

은 인간의 시각과 같은 영상을 처리하게 됨으로써, 좀 더 인간 중심적인 Human Computer Interface를 제공할 수 있다는 장점이 있지만 상대적으로 환경 변화에 민감하여 높은 성능을 보여주기 어렵다는 단점을 내포하고 있다. 또한, 영상 기반의 센서만을 이용하여 차종을 인식할 경우, 외형이 서로 비슷한 모습을 보여주고 있기 때문에 좋은 성능을 보여주지 못하고 있는 실정이다. 이러한 차량 인식에 대한 연구는 최근 들어서 일본 및 국내에서 가장 주목받고 있는 분야 중에 하나이다. 현재 일본에서는 이와 관련된 다양한 논문들이 나와 있다. Kagesawa[1]의 경우, 적외선 센서만을 이용해서 낮과 밤에 상관없이 차량을 인식하는 하드웨어를 개발했고, Kimachi[2]의 논문에서는 Stereo Vision을 이용하여서 Overlapping이 일어날 경우에도 차량에 대한 인식을 가능하게 하였다. 그러나 차량에 대한 3D 인식이나 그 정확성에 대해서는 많은 연구가 이루어지지 않은 실정이다.

영상 인식의 경우는 많은 특징점들을 가지고 있다. 이러한 많은 특징점들로 인해서 계산속도가 떨어지게 된다. 이러한 것을 극복하기 위해서 본 논문에서는 입력 영상을 같은 크기로 조정한 후 Singular Value Decomposition을 이용하여서 얻어낸 Singular Value들만을 특징점으로 사용을 하고 있다. 또한 인식 알고리즘으로는 SVM(Support Vector Machines) 방법을 채택하고 있다. 이러한 SVM은 3차원 외형 기반 인식방

법에 활용하기 좋은 특성을 보여주고, 훈련데이터가 확실할 경우 매우 정확한 인식률을 보여준다는 장점이 있다.

II. Support Vector Machines

SVM에 대한 연구는 70년대부터 소개가 되었지만, 최근 들어 Vapnik[3]에 의해서 이론이 정리가 되었고 주목을 받기 시작했다. 이러한 SVM은 Pontil[4]에 의해서 영상 응용의 발판이 마련되었다.

SVM은 클래스수가 2가지일 경우, 그 클래스간의 거리가 최대화 하면서 같은 클래스에 속하는 데이터들은 같은 쪽에 위치하게 만드는 Hyperplane을 찾게된다. 이러한 Hyperplane을 Optimal Separating Hyperplane (OSH)라 한다.

선형적으로 분류가 가능한 경우 기본적인 Hyperplane 수식은 다음과 같다.

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \quad i=1,2,\dots,N \quad (1)$$

$$\begin{cases} \mathbf{x}_i \in \mathbf{R}^n & : \quad n\text{차원 입력 값} \\ y_i \in \{-1, 1\} & : \quad \text{속하는 클래스에 따른 라벨값} \\ \mathbf{w}, b & : \quad \text{각각 Weight 행렬, Bias 값} \end{cases}$$

Margin은 가장 근접한 점들 사이의 거리에 2배인 $2/\|\mathbf{w}\|$ 값이 되고, 이 값을 최대화시키는 Hyperplane을 구해내는 것이 기본적인 목적이 된다. 이러한 Margin값을 최대화하는 값을 구하기 위해서 Lagrange Multipliers를 이용하여서 각각의 특징점에 해당하는 벡터 $\mathbf{a}^0 = (a_1^0, \dots, a_N^0)$ 값을 구해낸다. 이렇게 구해진 \mathbf{a}^0 값을 이용하면 식(2)과 같이 \mathbf{w}_0 값을 구할 수 있고, b 값을 구해진 \mathbf{a}^0 값등을 통해서 얻을 수 있다.

$$\mathbf{w}_0 = \sum_{i=1}^N a_i^0 y_i \mathbf{x}_i \quad (2)$$

$$f(\mathbf{x}) = \sum_{i=1}^N (a_i^0 y_i \mathbf{x}_i \cdot \mathbf{x} + b_0) \quad (3)$$

이러한 훈련을 통해서 구해진 OSH는 식(3)은 수식이 되고 후에 입력 데이터를 넣을 경우 출력값의 부호를 보고 2개의 클래스중에 어느곳에 속한 것인지 판별을 하게 된다.

보통 영상에서는 SVM은 선형적인 OSH를 추출하기 위해서 특징점의 차원을 크게 만든다. 그래서 대부분의 논문[4]들은 선형적인 OSH만을 추출하고 있다. 그러나 이렇게 할 경우 특징점의 차원이 커지게 되고 입력 및 훈련 데이터의 양이 많아지게 되어서 시스템의 전체적인 속도가 떨어지게 된다. 그래서 본 논문에서는 훈련 데이터를 커널을 통해서 매핑을 시킨 후 비선형적인 OSH를 만들어 냈으므로써, 특징점 차원을 많이 줄일 수

있었다. 이러한 커널의 종류는 Spline, Polynomial, Radial Basis Function(RBF)등이 있다.

$$f(\mathbf{x}) = \sum_{i=1}^N (a_i^0 y_i K(\mathbf{x}_i \cdot \mathbf{x}) + b_0) \quad (4)$$

$K(\mathbf{x}_i, \mathbf{x})$ 는 선택된 커널을 뜻하고 결국 선형적인 경우의 최적 OSH인 식(3)에 입력값과 훈련값을 커널을 통과시켜주기만 하면 된다.

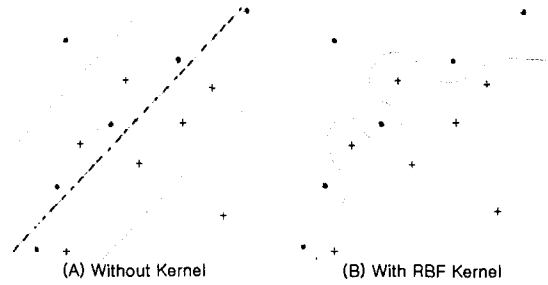


그림 1. 커널 유무에 따른 OSH 모습

그림 1 (A)와 같이 선형적으로 두 개의 클래스가 분류가 되지 못할 경우 커널을 사용할 경우 그림 1 (B)와 같이 OSH를 알맞게 만들어 낼 수 있다.

III. Aspect Based Recognition Algorithm

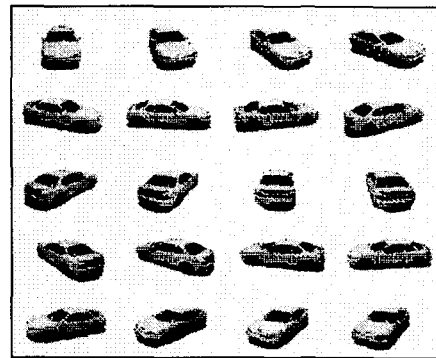


그림 2. 다양한 각도에서 촬영한 영상들

외관이나 외형을 기준으로 삼아서 3차원 물체 인식을 하는 알고리즘은 최근 95년 Murase[5]에 의해서 제안 되었다. 이러한 인식 방법은 360도 모든 방향에서 인식이 가능하기 때문에 많은 응용 분야에 적용이 되고 있는 실정이다. 이와 함께 외형적인 특징뿐만 아니라 색깔까지 특징으로 사용하여 인식률의 향상[6]을 꾀하기도 한다. 그러나 본 논문의 경우 인식을 하는 대상이 차량이기 때문에 색깔 정보는 특징점으로 사용하지 않고 있다.

그림 2와 같이 외형을 기준으로 3D 인식을 할 경우 여러 각도의 입력 영상을 가지고 훈련을 시킨 후, 이와 다른 각도의 영상을 입력 데이터로 사용함으로써 그 성능을 평가하고 있다. 본 논문의 경우 한 차종 당 약 3도씩 돌아가면서 찍은 104장의 영상을 가지고 실험을 하였다.

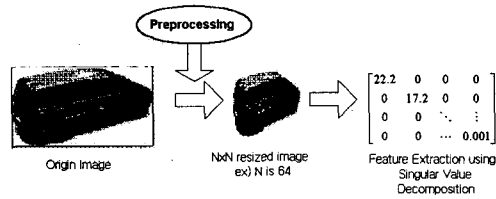


그림 3. 특징점 추출 과정

본 논문에서 제안한 특징점 추출 방법은 그림 3과 같은 형태로 이루어진다. 우선 흑백인 입력 영상에서 차량만을 추출해 낸다. 이후 추출된 차량을 일정한 크기의 정사각형으로 영상 크기를 다시 맞추어 주고, 이렇게 정사각형 영상을 직접적으로 Singular Value Decomposition (SVD)의 입력 값으로 넣음으로써 Singular Value값을 추출하고 있다. 이렇게 얻어진 Singular Value값은 언제나 양수이고, 나오는 값의 수는 일정한 크기로 맞추어진 정사각형의 크기와 같다.

이러한 Singular Value값에는 영상의 Pixel값의 변화에 대한 정보 및 전체 영상에 분포하고 있는 에너지 정보를 가지고 있게 된다.

IV. 시물레이션 결과

본 논문에서 사용하는 모든 차량 영상은 실물과 비슷한 모형 차량을 사용하였고, LPC-U20을 사용하여 영상을 촬영하였다. 또한 시물레이션틀로는 Matlab 5.3을 이용하였다.

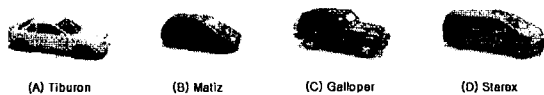


그림 4. 논문에서 사용된 차량의 종류

본 논문에서 사용한 차종은 총 4가지로 그림 4과 같다. 각 차종마다 3도씩 돌아가면서 찍은 104장씩 총 416장의 영상을 가지고 있고 각각의 차종에 대한 훈련의 경우 180도를 돌아가면서 찍은 사진 중 6도씩 돌아가는 26장을 사용하였다. 360도를 돌아가면서 찍은 52장을 훈련 데이터로 사용하지 않은 이유는 차량의 경우 좌우가 대칭이 되기 때문에 180도를 넘어서서 찍은 영상의 경우 거울에 비친 영상과 같기 때문이다. 이러한

거울 영상(Mirror Image)일 경우 똑같은 결과를 보여 주기 때문에 180도 이전의 26장을 훈련 데이터로 사용하고 있다. 그리고 시스템의 성능을 평가하는 입력 영상의 경우는 훈련데이터 보다 3도씩 360도를 다 돌아간 52장을 사용하고 있다.

전처리 과정이 없을 경우 다양한 실험 결과는 표 1, 표 2, 표 3에 잘 나타나 있다. 차량의 종류를 인식할 때 커널의 종류에 따른 인식률은 표 1과 같이 Linear 보다 Exponential RBF가 더 좋은 성능을 보여주고 있다. Exponential RBF는 다음과 같은 수식을 가진다.

$$K(x, y) = \exp\left(-\frac{|x-y|}{2\sigma^2}\right) \quad (5)$$

차량 영상을 일정한 크기의 정사각형 크기로 만드는 Window 크기에 따른 인식률을 보면 특징점이 더 많이 나올 수 있는 128×128일때가 64×64, 32×32일 때 보다 각각 약 12%, 22% 정도 성능 개선을 볼 수 있었다. 입력 영상의 회전의 경우 90도 단위의 회전에 경우 90%이상의 좋은 인식률을 보여주었다.

표 1. 커널 종류에 따른 인식률

	[Linear]		[Exponential RBF]	
	32x32	Execution Time	32x32	Execution Time
Tiburon	96.15%	100.1 Sec	88.46%	102.1 Sec
Matiz	53.85%	106.0 Sec	73.08%	98.8 Sec
Galloper	69.23%	105.0 Sec	73.08%	119.7 Sec
Starax	65.38%	104.1 Sec	65.38%	109.0 Sec
Total	71.15%	103.8 Sec	75%	107.4 Sec

표 2. Resize된 영상 크기에 따른 인식률

	Exponential RBF			
	64x64	Execution Time	128x128	Execution Time
Tiburon	100%	102.1 Sec	92.31%	140.2 Sec
Matiz	88.46%	123.9 Sec	100%	148.1 Sec
Galloper	73.08%	127.0 Sec	100%	159.3 Sec
Starax	80.77%	128.9 Sec	96.15%	143.3 Sec
Total	85.58%	120.48 Sec	97.12%	145.48 Sec

표 3. 입력 영상의 회전에 따른 인식률

	Exponential RBF & 128x128			
	90°	Execution Time	180°	Execution Time
Tiburon	92.31%	136.5 Sec	96.15%	136.3 Sec
Matiz	96.15%	150.1 Sec	100%	144.3 Sec
Galloper	88.46%	138.5 Sec	100%	138.8 Sec
Starax	96.15%	139.2 Sec	96.15%	138.9 Sec
Total	93.26%	141.08 Sec	98.08%	139.58 Sec

제안된 알고리즘의 경우 잡음이 시스템의 성능에 막대한 영향을 미치지 않았다. 그래서 영상들을 바로 사용하지 않고 위너 필터와 같은 적응형 필터를 전처리 과정으로 사용하였다. 이와 같은 전처리를 하고 나서 인식률인 표 4를 보면 전처리 과정이 없을 때와 같거나 더 좋은 성능을 보여주고 있음을 알 수 있다. 이와 같이 전처리 과정을 포함했을 경우 시스템의 성능에는 별다른 영향을 미치지 않고, 잡음이 첨가되었을 경우에도 나쁘지 않은 성능을 보여 줄 수 있다.

실험에서 잡음은 입력 영상에 가우시안 노이즈와 스파

클 노이즈를 각각 주었다. 가우시안 노이즈의 경우 평균은 0 분산은 10^{-3} 이고 스파클 노이즈의 경우 분산이 10^{-2} 이었다. 전처리 과정이 없을 경우 잡음이 첨가되었을 때는 80% 이하의 낮은 인식률을 보여주었지만, 전처리 과정이 있을 경우는 평균적으로 80%이상의 인식 결과를 보여주고 있고 윈도우 크기가 클 경우는 좀더 좋은 인식률을 보여줌을 표 5, 6을 통해 알 수 있다.

표 4. 무잡음에 전처리 과정이 있을 경우 인식률

	Exponential RBF			
	[64x64]	Execution Time	[128x128]	Execution Time
Tiburon	92.31%	133.2 Sec	96.15%	156.3 Sec
Matiz	92.31%	114.3 Sec	100%	159.8 Sec
Galloper	88.46%	114.3 Sec	100%	157.5 Sec
Starex	69.23%	116.8 Sec	96.15%	156.3 Sec
Total	85.58%	119.75 Sec	98.08%	157.48 Sec

표 5. 가우시안 잡음 첨가 되었을 때 인식률

	Exponential RBF			
	[64x64]	Execution Time	[128x128]	Execution Time
Tiburon	76.92%	118.0 Sec	92.31%	147.1 Sec
Matiz	53.85%	95.3 Sec	100%	138.9 Sec
Galloper	84.62%	154.3 Sec	100%	143.9 Sec
Starex	80.77%	121.4 Sec	96.15%	130.1 Sec
Total	81.54%	122.25 Sec	97.12%	140 Sec

표 6. 스파클 잡음이 첨가되었을 때 인식률

	Exponential RBF & [128x128]	
	Recognition	Execution Time
Tiburon	80.77%	126.8 Sec
Matiz	76.92%	123.6 Sec
Galloper	88.46%	127.6 Sec
Starex	100%	131.3 Sec
Total	86.53%	127.33 Sec

V. 결론

본 논문에서는 영상기반의 ETCS제작을 위한 차량의 종류를 인식하는 알고리즘을 제안하고 있다. 3D 영상의 인식의 경우 SVM을 이용할 경우 좋은 성능을 얻을 수 있었다. 특히, 실시간 시스템 구현을 위해서 특징점의 수를 줄이면서도 좋은 성능을 보여주도록 시스템을 구성하였다. 또한 잡음이 첨가될 경우 이를 극복하기 위해서 적용적인 필터를 전처리과정에 도입함으로써 어느 정도의 잡음에 민감하지 않은 결과를 얻을 수 있었다.

VI. 향후 연구 계획

향후 개선해야 할 사항으로는 다음과 같은 사항이 있다. 실제 차량을 사용한 것이 아닌 차량 모델을 사용하여 실험을 했기 때문에 실제 차량에 바로 적용을 시키기에는 미흡한 점이 많다. 이와 함께 Window기법 및 Partition 기법 사용하여 잡음이 첨가된 환경에 좀더 강한 알고리즘 개발함과 동시에 Data Fusion 기법을 이용하여서 전체적인 시스템의 성능을 향상시키는 연구

가 진행중이다.

Acknowledgement

본 연구는 한국과학재단의 특정기초 연구 (과제번호: 1999-1-303-005-3)의 지원으로 수행되었습니다.

참고 논문

- [1] M. Kagesawa, S. Ueno, K. Ikeuchi, H. Kashiwagi, "Local Feature Based Vehicle Recognition In Infra-Red Images Using Parallel Vision Board," Proceedings of the ICIRS, Kyongju, Korea, pp. 1828-1833, Oct, 1999.
- [2] M. Kimachi, Y. Wu, S. Ogata, "A vehicle Recognition Method Robust against Vehicles' Overlapping Based on Stereo Vision," Proceedings of the IEEE ICI, Tokyo, Japan, pp. 865-869, Oct, 1999.
- [3] V.N. Vapnik, "The Nature of Statistical Learning Theory," New York, 1995
- [4] M. Pontil, A. Verri, "Support Vector Machines for 3D Object Recognition," IEEE Trans. PAMI, Vol. 20, No. 6, Jun, 1998.
- [5] H. Murase, S.K. Nayar, "Visual Learning and Recognition of 3-D Object From Appearance," Internation Journal of Computer Vision, Vol. 14, pp. 5-24, 1995.
- [6] D. Roobaert, M. Van Hulle, "View-Based 3D Object Recognition with Support Vector Machines," Proceedings of the NNSP, pp. 77-84, Aug, 1999.