

도심 영상에서의 비음수행렬분해를 이용한 차량 인식

반재민*, 이병래*, 강현철°

Vehicle Recognition using NMF in Urban Scene

Jae Min Ban*, Byeong-rae Lee*, Hyunchul Kang°

요약

차량인식은 차량 후보영역 검출단계와 검출된 후보 영역에서 특징을 기반으로 차량을 검증하는 차량 검증단계로 나누어진다. 선형 변환 방식의 특징은 차원 감소 효과와 통계적인 특징을 지니게 되어, 이동이나 회전에 강한 특성을 갖는다. 선형 변환 방식 중 비음수행렬분해(Non-negative Matrix Factorization, NMF)는 부분 기반 표현 방식으로 차량의 국소적인 특징을 기저벡터로 사용하여 희소성을 갖는 특징을 추출할 수 있기 때문에 도심 영상에서 발생하는 차폐 영역에 따른 인식을 저하를 방지할 수 있다. 본 논문에서는 차량 인식에 적합한 NMF 특징 추출 방법을 제안하고, 인식률을 검증하였다. 또한 희소성 제약 조건을 이용하여 기저 벡터에 희소성을 가지는 SNMF(Sparse NMF)와 LVQ2(Learning Vector Quantization) 신경 회로망을 결합하여 차폐 영역에 대한 차량 인식 효율을 기존의 NMF를 이용한 방법과 비교하였다. NMF를 이용하는 특징은 차량이 혼재되어 차폐 영역이 빈번히 발생하는 도심에서도 강건한 특징임을 보였다.

Key Words : vehicle detection, stereo vision, NMF, part based representation, LVQ2

ABSTRACT

The vehicle recognition consists of two steps; the vehicle region detection step and the vehicle identification step based on the feature extracted from the detected region. Features using linear transformations have the effect of dimension reduction as well as represent statistical characteristics, and show the robustness in translation and rotation of objects. Among the linear transformations, the NMF(Non-negative Matrix Factorization) is one of part-based representation. Therefore, we can extract NMF features with sparsity and improve the vehicle recognition rate by the representation of local features of a car as a basis vector. In this paper, we propose a feature extraction using NMF suitable for the vehicle recognition, and verify the recognition rate with it. Also, we compared the vehicle recognition rate for the occluded area using the SNMF(sparse NMF) which has basis vectors with constraint and LVQ2 neural network. We showed that the feature through the proposed NMF is robust in the urban scene where occlusions are frequently occur.

I. 서론

운전자 보조 시스템(driver assistance system)은 지능형 자동차(intelligent vehicle)의 구현에 필수적인 기술로 기계, 전자, 통신 제어 기술을 융합하여

안전성과 편의성을 향상 시킬 수 있다. 그 중 차량 인식 시스템은 충돌 방지, 교통량 수집, 교통 단속 등 다양한 분야에서 사용 가능하고, 결과적으로 운전 중에 발생할 수 있는 위험요소를 방지하거나 많은 사회적 문제를 효율적이고 경제적으로 관리 할

* 주저자 : 인천대학교 정보통신공학과 영상 연구실, gosemf85@nate.com, 준회원

° 교신저자 : 인천대학교 정보통신공학과, hckang@incheon.ac.kr, 정회원

* 한국방송통신대학교 컴퓨터학과, brlee@knou.ac.kr, 정회원

논문번호 : KICS2012-04-229, 접수일자 : 2012년 4월 30일, 최종논문접수일자 : 2012년 7월 2일

수 있게 된다.

차량 인식은 여전히 객체 인식 분야에서 중요한 부분을 차지하고 있다. 특히 영상처리를 이용한 많은 연구가 활발하게 진행되고 있으며 Z. Sun^[2]은 영상을 이용한 차량 인식 시스템은 기본적으로 후보영역 검출단계와 후보영역에서의 특징 추출 후에 인식 성능을 높이기 위한 검증 단계로 구성된다고 설명한다. 후보 영역 검출을 위한 이전 연구에는 차량 전면의 밝기 대칭성^[3]을 이용한 방법, 컬러 정보^[4]나 밝기 정보의 임계값 이상의 히스토그램을 이용하여 차량과 배경을 분리하여 검출하는 방법, 코너점^[6]이나 수평, 수직 에지성분^[7]을 이용한 방법, 배경과 구분되는 차량의 움직임^[8]을 이용한 방법, 거리정보를 이용하는 스테레오 비전^[9,10] 등이 있다. 검증 단계의 연구로는 크게 이미지의 상관관계를 이용하여 검증을 수행하는 템플릿 기반 방법^[11]과 신경망^[12]이나 SVM^[13] 등의 분류기를 이용한 방법으로 나뉜다. 그러나 이전의 연구에서는 차량이 적은 도로에서의 적용을 대상으로 하였으며 최근 복잡한 환경의 도심 영상에 대한 차량 인식에 대한 연구가 이루어지고 있다. 또한 차량의 특성상 실제 적용을 위하여 실시간성을 보장하고, 차종이나 이동에 따른 변형이 발생하더라도 변하지 않는 특징이 요구되기 때문에 이러한 경우 주로 선형 변환 방식을 사용한다.

선형 변환 방식은 대부분 고차원의 데이터를 저차원의 데이터로 줄이는 방식을 사용함으로써 실시간성을 보장한다. 대표적인 선형 변환 방식으로는 데이터의 상관행렬에서 변동이 심한 방향을 찾아내는 PCA(Principal Component Analysis)^[14], 학습 단계에서 클래스의 구분에 중점을 두는 LDA(Linear Discriminant Analysis)^[15], 통계적 독립성을 이용하여 독립적인 성분만을 추출하는 ICA(Independent Component Analysis)^[16] 등이 있다. 하지만 기존의 선형 변환 방식들은 영상의 전체적인 특징으로 표현하기 때문에 복잡하게 차량이 존재하는 경우 인식이 크게 떨어진다. 최근 인간의 두뇌가 영상의 부분적인 특징을 인식하는 기능을 응용한 NMF는 입력 데이터가 양의 값만을 가지는 영상을 대상으로 하는 경우에 유용하다고 알려져 있다. 지역적인 특징을 이용하기 때문에 입력 데이터가 다른 객체에 의해 가려지는 현상이 발생하더라도 목표표 하는 표현과 크게 변화하지 않으므로 복잡한 도심에서의 인식에 다른 방법보다 적합한 방법이다.

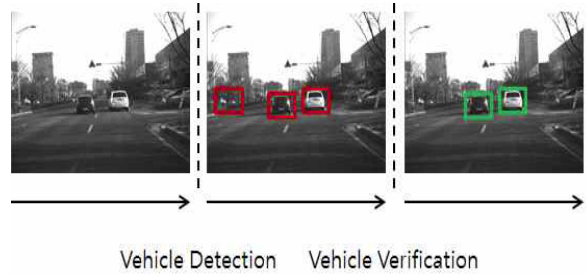


그림 1. 차량 인식 시스템의 단계
Fig. 1. Illustration of the step vehicle recognition system

따라서 본 논문은 NMF를 사용한 차량 인식 방법을 제안한다. 제안하는 방법을 나타낸 그림 1은 차량일 가능성이 있는 후보 영역들을 추출하고 NMF를 이용하여 차량의 공통적인 부분기반 표현들을 자동적으로 추출하여 특징으로 사용한다. 후보영역을 빠르고 간단하게 추출하기 위하여 스테레오 비전을 사용하여 거리에 따른 차량 검출 방법을 사용한다. 본 논문에서 제안한 방법은 후보영역을 찾아내는데 거리 정보만을 사용함으로써 모든 객체 영역에 대한 검출을 수행한다. NMF를 사용한 차량 인식 방법은 차량 표현에 사용되는 특징의 수를 줄여주고 새로운 특징을 찾기 위해 같은 알고리즘을 처리해야 하는 부담을 줄여 실시간성을 보장해 주는 장점이 있다.

II. 차량 인식 시스템

차량의 인식은 차량에 장착된 카메라로부터 영상을 입력 받은 후 먼저 차량의 후보 영역의 추출을 위한 검출 단계와 후보 영역이 차량인지 아닌지를 검증하는 검증단계로 나뉜다. 시스템을 두 단계로 나눔으로써 이전 단계의 후보 검출의 오류를 차량 인식 검증 단계에서 상호 보완하여 인식의 효율을 높이는 효과를 얻을 수 있다.

제안하는 전방 차량 인식 시스템의 전체적인 구성은 그림 2와 같다. 스테레오 카메라로부터 영상을 입력 받아 거리정보를 포함하고 있는 깊이 맵을 생성하여 특정거리에 존재하는 모든 객체를 대상으로 후보영역을 검출한다. 실제 도심 주행 영상은 많은 차폐영역을 가지고 있기 때문에 기존의 후보 영역 검출 알고리즘은 강인하지 못하며 추가적인 처리가 필요하다. 따라서 모든 객체에 대하여 추가적인 처리 없이 후보영역을 검출하기 위하여 거리에 기반을 둔 스테레오 비전을 사용한다.

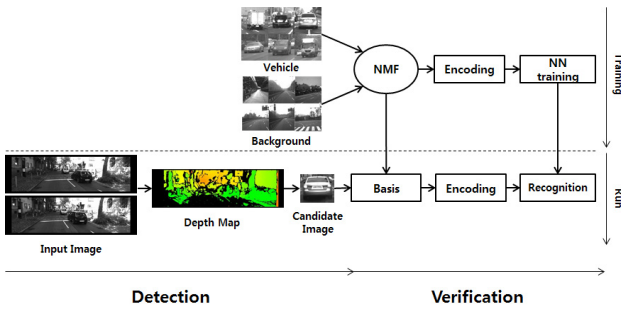


그림 2. 제안된 차량 인식 시스템의 개요
Fig. 2. Overview of proposed vehicle recognition system

검출된 차량의 후보영역에서 학습 단계에서 다수의 차량과 차량이 아닌 학습 이미지들을 열 벡터로 된 집합으로 각각 표현하면 NMF는 학습 이미지 집합과의 차이가 수렴할 때까지 반복하여 계수와 기저로 분해한다. 이때 기저는 임의의 개수를 가지며 분해과정 동안 각 집합의 지역적인 특징들을 학습하고 차량과 배경의 학습 이미지의 집합으로부터 공통적인 부분을 표현한다. 계수는 결과적으로 학습 이미지의 지역적 가중치와 연관된 특징으로 사용되어 차량 집합과 배경 집합이 서로 다른 가중치를 가지게 된다. 그러나 적은 수의 계수를 이용하여 이미지를 표현하기 때문에 차원 감소 효과로 인하여 발생할 수 있는 인식률 저하를 막고 정확하게 분류 경계를 나누기 위하여 신경회로망을 사용하였다. 추출된 계수 벡터 집합을 LVQ2 신경회로망¹⁹⁾에 학습시켜 보다 정확한 분류 경계를 설정한다. 인식단계에서는 입력 영상에서 후보영역을 추출한 후 학습단계에서 구한 기저 벡터를 가지고 새로 입력 받은 이미지의 계수를 구한 후 이전에 학습된 신경회로망에서 분류하여 차량과 배경을 구분하여 인식할 수 있게 된다.

III. 후보 영역 검출

3.1. 스테레오 비전

스테레오 비전 시스템은 다른 위치에 있는 카메라를 이용하여 획득한 2차원의 영상으로부터 3차원 거리 정보를 추출할 수 있는 방법이다. 추출된 정보들은 목표 물체의 검출이나 깊이 정보를 구하는데 이용될 수 있기 때문에 지능형 자동차나 로봇의 자율 주행과 관련된 시스템에 유용하게 이용되어 지고 있다.

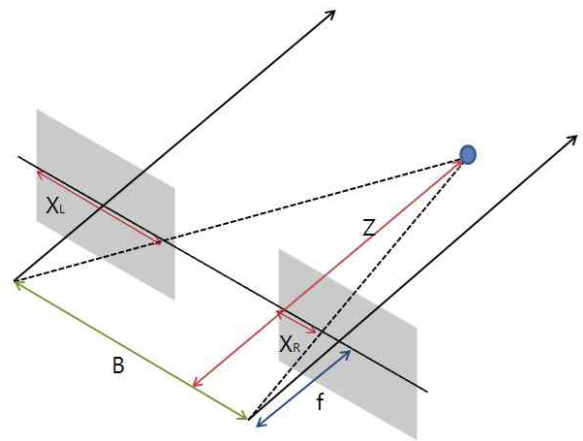


그림 3. 스테레오 비전 시스템의 거리 추정 원리
Fig. 3. The distance estimation principal of the stereo vision system

스테레오 비전의 거리 추정은 삼각법에 의하여 추출되며 좌우 영상의 시차와 카메라의 파라미터들에 의해서 결정된다. 그림 3에서 B와 f는 카메라 파라미터를 구성하는 요소로 B는 카메라간의 거리를 나타내고 f는 카메라의 초점거리를 나타낸다. 또한 X_L 과 X_R 의 차이는 시차를 나타내므로 삼각법에 의한 거리 Z는 식 (1)로 구할 수 있으며 이를 통하여 깊이 맵을 생성한다.

$$Z = \frac{Bf}{X_L - X_R} \quad (1)$$

깊이 맵은 좌우 영상의 대응 화소를 매칭 알고리즘을 사용하여 추출된다. 매칭 알고리즘은 좌 영상을 기준으로 우 영상의 대응화소를 찾는 방법으로 본 논문에서는 실시간으로 동작해야 하기 때문에 블록 기반 매칭 방법을 사용한다. 이 방법은 좌우 영상의 대응점을 찾기 위해 윈도우 내부의 SAD(Sum of Absolute Difference)를 계산하며 이 동하여 영상의 동일한 행에서 가장 잘 매치되는 점을 찾는다.

$$SAD(x, y, d) = \sum_{y=0}^{h-1} \sum_{x=0}^{w-1} |I_L(x, y) - I_R(x + d, y)| \quad (2)$$

식 (2)에서 h와 w는 각각 윈도우의 높이와 너비를 나타내며 d는 시차, I_L 과 I_R 은 좌표에 대응하는 화소값을 나타낸다. 두 영상 사이의 최적의 대응점을 찾기 위하여 d를 변화시켰을 때 최소값을 갖는 d를 최적의 대응점을 나타내는 시차로 선택함으로써

써 깊이 맵을 생성할 수 있다.

그림 4는 좌우 영상을 이용하여 깊이 맵을 생성하였을 때 특정 범위 내의 모든 객체에 대한 후보 영역을 설정하고 설정된 영역을 검출한다.

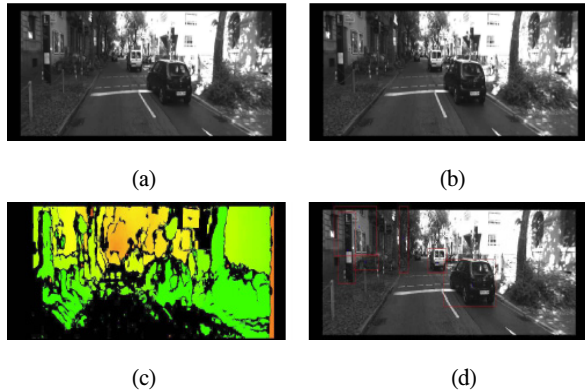


그림 4. 후보 영역 검출 과정
 (a) 좌 영상 (b) 우 영상 (c) 깊이 맵 (d) 후보 영역 검출
 Fig. 4. The candidate region detection process
 (a) left image (b) right image (c) depth map (d) candidate region detection

그러나 (d)에서 검출된 후보 영역들은 차량이 아닌 객체들도 포함되어 있다. 따라서 차량 인식을 통하여 인식률을 높이는 방법이 필요하다.

VI. 후보 영역 검증

4.1. NMF

선형 변환 방식은 컴퓨터 비전이나 기계학습과 같은 고차원 데이터의 분석 문제를 해결하는데 유용하게 사용되어 왔다. 대표적인 선형 변환 방식으로 사용되고 있는 PCA는 주어진 데이터의 상관행렬에서 데이터의 변동이 심한 방향을 나타내는 고유벡터에 대응하는 고유값인 주성분이라 부르는 독립적인 변수를 이용하여 특징을 추출한다. 각 주성분은 데이터 변이의 크기를 기준으로 중요도를 표현하기 때문에 상위 몇 개의 주성분만으로도 원 데이터와의 정보 차이를 최소화하면서 차원을 줄일 수 있다.

PCA이외에 저차원 데이터로의 사상을 위한 변환 방식으로 많이 사용되는 알고리즘으로는 ICA가 있다. ICA는 통계적 특성을 통하여 다차원의 독립적인 성분을 추출해내는 방법이다. PCA가 데이터간의 저차원 성분만을 분해하는 반면 ICA는 고차원의 통계적 특성을 이용하기 때문에 엔트로피와 같은 다양한 특성을 나타낼 수 있다. 그러나 기존의 방식들

은 데이터 전체의 특성을 사용함으로써 데이터 내에 존재하는 지역적인 정보에 대한 표현이 불가능하다. 따라서 차량의 인식에 사용될 경우 배경이나 분류에 영향을 미치게 되고 실제 도로상에 많이 존재하는 가려지거나 겹쳐진 차량에 대해서는 강한한 모습을 보여주지 못하고 인식률을 떨어뜨리는 결과를 만든다¹⁷⁾.

전체적인 특성을 표현하는 기존 선형 변환이 가진 문제점을 해결하기 위하여 부분 기반 선형 변환 방식인 NMF가 연구되었다. NMF를 제안한 Lee와 Seung¹¹⁾에 따르면 안면 인식에 적용하였을 때 얼굴을 눈, 코, 입 등의 사람이 직관적으로 인식 가능한 것들을 기저 벡터로 하여 그것의 선형 조합으로 표현할 수 있다. 그림 5는 하나의 차량 이미지를 NMF의 기저 벡터의 선형 조합으로 나타내었다. 이미지 V 는 차량을 구성하는 프레임, 전조등, 번호판과 같은 성분을 나타내는 기저 벡터 W 와 가중치 계수 H 의 선형 조합으로 이루어진다. 따라서 후보 영역에서 차폐영역이 발생하더라도 특정 기저 벡터에 대응하는 가중치 계수만 작아지기 때문에 전체적인 특징벡터에 많은 영향을 미치지 않는다.

$$\begin{matrix}
 \text{Image} & \approx & \text{Image} & + & \text{Image} & + & \dots & + & \text{Image} & + & \text{Image} & + & \text{Image} & :W \\
 V & & \times & & \times & & & & \times & & \times & & \times & \\
 & & H_{1,m} & & H_{2,m} & & \dots & & H_{r-2,m} & & H_{r-1,m} & & H_{r,m} & :H
 \end{matrix}$$

그림 5. NMF의 선형 조합의 예
 Fig. 5. The example of linear combination of NMF

NMF는 입력 데이터가 양의 값만을 가지고 있는 영상이 대상인 경우 유용하게 사용될 수 있고 데이터의 지역적인 특징을 찾아냄으로써 각 지역별의 미적 특징학습이 가능하다. NMF는 고차원의 데이터 행렬을 저차원의 계수 행렬과 기저 행렬로 분해하고 각 행렬들이 희소특성을 가지기 때문에 부분 기반 표현이 가능하다.

$$V \approx WH \tag{3}$$

식 (3)의 V 는 $n \times m$ 이미지 데이터베이스 행렬이 주어지면 행렬 V 는 $n \times r$ 기저 행렬 W 와 $r \times m$ 계수 행렬 H 로 분해한다. 이때 n 은 한 개 이미지의 픽셀의 개수이고, m 은 이미지 데이터베이스 행렬에 포함된 이미지의 개수이다. r 은 기저벡터의 개수로 rank r 은 일반적으로 $(n+m)r < nm$ 을 만족하여야 한다.

r 이 최대에 가까운 경우 V 와 WH 의 차이를 최소화 할 수 있지만 각 기저 벡터가 데이터베이스의 각 이미지와 같아지므로 부분 기반 표현을 만족하지 못하고, r 이 너무 적은 경우에는 기저 벡터가 충분히 성분들을 표현하지 못하기 때문에 분류의 측면에서 인식을 저하를 가져온다. 따라서 제안된 방법에서는 부분 기반 표현과 분류 능력이 뛰어난 기저벡터를 구성하기 위하여 r 의 가능한 최소와 최대사이의 정수 값 중 최초의 이미지 데이터베이스 행렬 V 와 근사 행렬 WH 의 차이 값이 수렴하는 구간의 r 을 선택하여 이미지 데이터베이스에 적합한 기저 벡터를 학습한다.

$$r = \operatorname{argmin} G(r) \quad (4)$$

$$\text{s.t. } \nabla \cdot G(r) < \epsilon, r = 1, \dots, \left(\frac{n \times m}{n+m}\right) - 1$$

함수 G 를 기저 벡터의 개수 r 에 따른 식 (7)의 목적함수 F 의 벡터로 가정하면 식 (4)는 함수 G 의 기울기가 0에 가까운 임의의 상수 ϵ 보다 작게 되는 제약 조건을 만족시키는 r 을 선택하게 되므로 학습 이미지 데이터베이스에 적응적인 기저 벡터 수를 선택할 수 있도록 한다.

NMF에서는 행렬 V 의 분해를 위하여 W 와 H 를 임의의 양의 값으로 초기화하고 식 (5)와 식(6)과 같은 갱신 규칙을 반복적으로 이용하여 구한다.

$$H_{rm}^{k+1} < -H_{rm}^k \frac{((W^k)^T V)_{rm}}{((W^k)^T W^k H^k)_{rm}} \quad (5)$$

$$W_{nr}^{k+1} < -W_{nr}^k \frac{(V(H^{k+1})^T)_{nr}}{(W^k H^{k+1} (H^{k+1})^T)_{nr}} \quad (6)$$

갱신 규칙은 기울기 하강(Gradient Descent) 방법과 유사하게 동작하며 현재의 W 와 H 에 식 (7)의 목적 함수에서의 양수 부분과 음수 부분사이의 기울기로 구하여진 예측값을 곱하여 갱신을 한다. 이 규칙이 KKT(Karush-Kuhn-Tucker) 최적화 상태를 만족할 경우 목적함수를 최소화하는 값으로 갱신된다. 그리고 갱신된 W 와 H 는 부분적인 편차를 가지게 되고 행렬 상호간, 내부적으로 직교하는 방향으로 변화함으로써 NMF의 특징인 희소 특성을 가지게 된다. 그림 6은 NMF의 기저 벡터를 나타내었다.

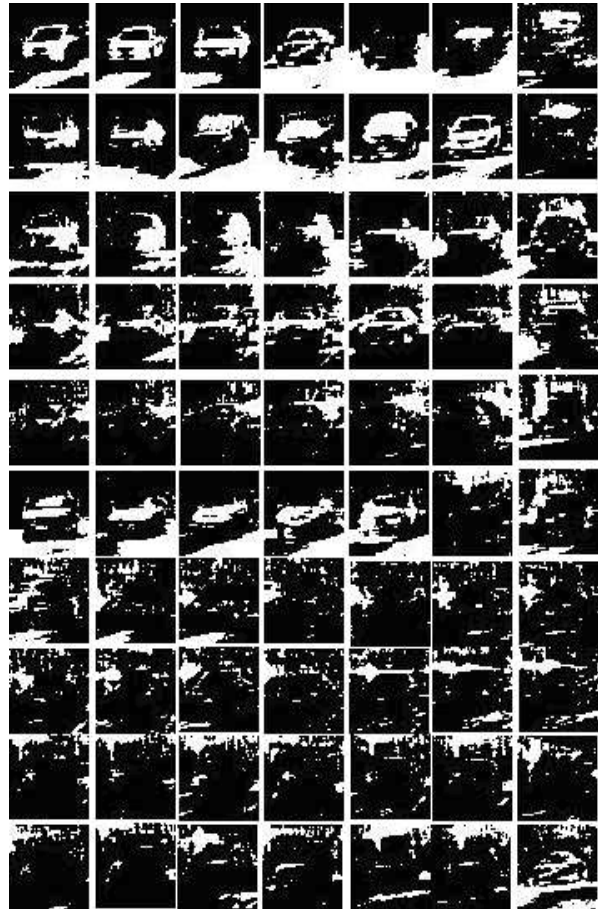


그림 6. NMF의 기저 벡터의 예
Fig. 6. The example of basis vectors of NMF

반복적인 갱신 규칙을 수행하게 되면 V 에 WH 가 근사하게 수렴하고 근사적인 정보만으로도 표현이 가능하기 때문에 목적함수 F 를 정의하여 갱신을 멈춘다.

$$F \equiv \|V - WH\|^2$$

$$\equiv \sum_{i=1}^n \sum_{\mu=1}^m [V_{i\mu} \log(WH)_{i\mu} - (WH)_{i\mu}] \quad (7)$$

식 (7)는 원 데이터와 분해된 근사 행렬 WH 의 차이를 구함으로써 정확도를 반영한다. 목적 함수가 수렴하게 될 때까지 갱신 규칙을 수행하게 되면 각 행렬 내의 표현을 제외한 대부분의 성분들이 0에 가까운 값을 가지게 된다. 따라서 부분 기반 표현에서 중요한 희소성(Sparseness)을 만족하게 된다.

4.2. SNMF

NMF는 때때로 수렴에 실패하거나 결과로 추출된 기저 벡터로 표현된 부분 표현은 여전히 전체를 표현한다. 행렬 W 와 H 의 희소성을 조절할 수 있

다면 NMF는 많은 응용에 적합한 특징 추출이 가능하다. 따라서 희소성에 대한 제약 조건 식 (8)을 추가한 SNMF를 사용한다.

$$S(x) = \frac{\sqrt{n} - (\sum |x_i|) / \sqrt{\sum x_i^2}}{\sqrt{n} - 1} \quad (8)$$

식 (8)의 크기는 n 차원의 벡터가 몇 개의 성분들에 모였는가를 판단한다. 만약 이 함수의 값이 1이라면 오직 한 개의 성분만이 0이 아닌 값을 가지게 되고, 0이라면 모든 성분들이 같은 값을 가지게 된다. 행렬 W 와 H 에 원하는 희소성을 설정하여 식(8)을 각각 적용하면 원하는 희소 특성을 가지는 행렬을 얻을 수 있다. 본 논문에서는 희소 특성을 가지는 기저 행렬과 대응되는 계수 행렬을 특징 벡터로 사용하기 때문에 희소성 제약조건을 식 (9)와 같이 기저 행렬에만 적용하였다. 따라서 희소성을 부가하기 위하여 식 (7)의 목적 함수는 식 (10)로 변형된다.

$$\beta = S(W) \quad (9)$$

$$F = \|V - WH\|^2 + \beta W \quad (10)$$

NMF의 다른 특징으로는 신경 회로망의 구조와 유사하는 것이다. NMF에 의해 분해된 두 행렬은 각 단계에 사용되며 기저 행렬과 계수 행렬이 신경 회로망의 가중치 행렬과 입력 행렬과 유사하게 동작하기 때문에 자체적인 분류 능력이 뛰어나다. 따라서 신경 회로망에 사용될 가중치 행렬을 추가적인 연산 없이 기저 행렬을 사용하여 결정 경계를 정의하더라도 인식률에 큰 영향을 미치지 않는다. 후보 영역 검출 과정으로부터 얻은 후보 영역들은 학습 이미지와 같은 크기로 정규화하고 NMF의 기저 행렬을 이용하여 계수 벡터를 추출한다.

$$W = [W_{vehicle}, W_{BG}] \quad (11)$$

$$H = [H_{vehicle}, H_{BG}] \quad (12)$$

학습 이미지는 추출된 후보 영역들이 배경과 같은 영역을 포함하고 있을 수 있기 때문에 차량 이미지와 배경 이미지를 이미지 데이터베이스에 포함함으로써 추출된 특징이 확실하게 분류 가능한 특징이 되도록 한다. 인식에 적합한 특징 추출을 위하여 식 (11)과 식 (12)에서 기저 행렬과 계수 행렬이 차량과 배경으로 구분되도록 학습한다.

4.3. LVQ 신경회로망

NMF는 국소 영역에 국한된 특징들을 추출함으로써 인식에 바로 적용이 가능하다. 그러나 특징 추출의 영역을 조절하지 못하기 때문에 각 기저 벡터가 여전히 전체적인 특징을 나타내므로 다양한 응용에 적용하기 적합하지 않다. 따라서 희소성에 대한 제약 조건을 조절함으로써 다양한 응용에 적응이 가능한 모습을 보인다. 그러나 SNMF에서 추출된 특징의 경우 희소성을 가지기 때문에 NMF 자체적인 분류능력이 떨어진다. 그러므로 학습 데이터의 특징을 추출하고 인식기를 사용하여 클래스 간의 분류를 돕는다. SNMF에서 분해된 계수 벡터의 특징은 각 기저 벡터에 대한 강도를 나타내며 독립적이거나 희소성을 가지므로 뛰어난 분류능력을 가지고 있다. 따라서 같은 기저 행렬에 의해 분해된 입력 이미지의 계수 벡터를 입력 패턴으로 할 때 학습 데이터의 계수 벡터인 기준 패턴들과의 유사도를 측정해야한다. Kohonen의 LVQ 신경망^[19]은 패턴 분류를 위하여 지도 학습과 비지도 학습이 동시에 가능한 경쟁 학습 규칙을 가지고 있다. 학습 속도가 빠르고, 연산량이 적어 실시간성을 가지는 신경망으로 영상처리와 같은 방대한 데이터를 처리하는데 유용하게 사용된다.

LVQ 신경망의 학습은 훈련 벡터와 출력이 정해지고 훈련 벡터가 입력되면 가중치 벡터와의 유클리디언 거리를 기반으로 출력 뉴런이 경쟁을 한다. 이때 가중치 벡터와 가장 거리가 가까운 한 개의 훈련 벡터만이 승리 벡터가 되고 훈련벡터가 속한 클래스와 가중치 벡터의 클래스가 일치하면 클래스의 가중치 벡터를 훈련 벡터 방향으로 갱신하고, 일치하지 않는 경우에는 훈련 벡터의 반대 방향으로 가중치 벡터를 이동한다.

LVQ는 중복되거나 다차원으로 복잡한 입력 벡터의 경우 신경망의 초기 값을 결정하기 어렵다는 단점이 있다. 초기 값에 따라 경쟁 단계에서 뉴런이 전혀 승리하지 못하거나 데이터를 정확하게 분류하지 못하는 방향으로 가중치 벡터가 이동할 수 있다. 따라서 본 논문에서는 기존의 LVQ를 수정한 LVQ2를 사용한다.

LVQ2의 학습 방법은 한 입력 벡터를 올바르게 분류하는 경우 단 한 개의 가중치 벡터만이 입력 벡터의 방향으로 이동하고, 올바르게 분류하지 못하면 입력 벡터와 가까운 두 개의 가중치 벡터가 동시에 이동한다. 승리 뉴런의 클래스가 다르게 나타난 하나의 가중치 벡터는 입력 벡터와 멀어지는 방향으

로, 같은 경우의 다른 하나의 가중치 벡터는 입력 벡터의 방향으로 이동한다. 입력 벡터 x 가 주어졌을 때 LVQ2의 가중치 벡터의 갱신은 식 (13)와 식 (14)과 같다.

$$W_i(t+i) = W_i(t) + \alpha(t)(x - W_i(t)) \quad (13)$$

$$W_j(t+j) = W_j(t) - \alpha(t)(x - W_j(t)) \quad (14)$$

이때 $W_i(t)$ 는 승리 뉴런 클래스의 가중치 벡터와 같은 가중치 벡터이고 $W_j(t)$ 는 다른 가중치 벡터이다. α 는 학습율로 얼마나 빠르게 최적 결정 경계에 수렴하는가를 결정하게 된다. 이러한 LVQ2 신경망 학습을 통해 최적의 클래스의 중심 벡터인 가중치 벡터를 구하고 인식 단계에서 입력 벡터와 가중치 벡터간의 유클리디언 거리를 비교하여 차량 인식에 적용한다.

V. 실험

본 논문에서는 제안된 NMF와 SNMF에서 추출된 특징의 차량 인식에서의 성능에 대한 유효성을 검증하였다. 실험에 사용된 테스트 영상은 KITTI Vision Benchmark^[21]에서 제공한 스테레오 이미지를 30fps 영상으로 편집하여 사용하였으며 카메라의 보정 파라미터는 각각의 영상에 따라 제공된 파라미터를 사용하였다. 특히 맑은 날 주간을 기준으로 많은 차량이 혼재되어 있거나 눈 덮인 도로와 같은 배경들에 의하여 차폐영역이 빈번히 발생하는 도로의 영상에 제안된 방법을 적용하였다.

학습에 사용된 차량의 학습 이미지는 테스트 영상과 다른 영상에서 스테레오 비전을 이용한 검출 방법을 사용하여 취득하였으며 검출을 수행하는 거리의 최대범위를 약 30m로 설정하였다. 모든 학습 이미지는 48×48 의 크기로 정규화하여 그림 7에 나타내었다. 학습 데이터의 구성은 400개의 차량 이미지와 200개의 차폐영역이 존재하는 차량 이미지, 400개의 배경 이미지로 각 이미지를 1차원 벡터로 변환하여 이미지데이터베이스 행렬 V 에 적용하였으며 따라서 V 는 $2304 \times 1000(n \times m)$ 의 크기를 가지게 되고 기저벡터의 개수를 변화하며 실험하였다.

제안한 방법에서는 각 프레임에 대하여 검출 단계와 검증 단계를 수행하는 동안 추출되는 영역의 수가 감소하기 때문에 실험에 사용된 입력 영상들의 각 프레임에 존재하는 평균적인 전체 차량의 수와 그에 따른 제안된 방법의 평균적인 검출된 영역의 수, 검증된 영역 수를 표 1에서 비교하였다.



(a)



(b)



(c)

그림 7. 학습 이미지의 예 (a) 차량 이미지 (b) 배경 이미지 (c) 차폐영역이 존재하는 이미지

Fig. 7. The example of training image (a) vehicle image (b) background image (c) occluded image

표 1. 한 프레임에서의 객체의 수 비교 Table 1. Comparison of the number of objects in a frame

Ground truth	Detection	Verification
8.7	25.5	9.4

검출 영역은 약 30m내의 모든 객체를 추출하기 때문에 실제로 검출된 영상에서는 창문, 보행자 등의 차량 이외의 객체가 다수 포함되어 있으며 검증 단계에서 이를 보완해주기 때문에 실제 차량의 수와 근사하게 감소하는 것을 볼 수 있다. 그러나 검증 단계에서 추출한 영역은 언제나 차량을 포함하고 있지 않기 때문에 정확한 성능을 계산하기 위하여 인식률과 실패율을 식 (15)와 (16)에 정의하였다. 식 (15), (16)의 $total$ 은 $vehicle_{vehicle}$, $vehicle_{bg}$, BG_{bg} , $BG_{vehicle}$ 의 총합이며 $vehicle_{vehicle}$, BG_{bg} 는 각 클래스가 옳게 구분된 수를 나타내고, $vehicle_{bg}$, $BG_{vehicle}$ 는 반대로 구분된 수를 나타낸다.

$$Recognition\ rate = \frac{vehicle_{vehide} + BG_{bg}}{total} \quad (15)$$

$$Fail\ rate = \frac{vehicle_{bg} + BG_{vehide}}{total} \quad (16)$$

그림 8은 정의된 클래스에 따른 각 클래스의 평균 계수 행렬을 나타낸다. X축은 기저 벡터의 순서를 나타내고, Y축은 계수 벡터의 평균값을 나타낸다. 차량을 나타내는 계수 벡터의 경우 특정 기저 벡터에서 높은 값을 나타내고 있다. 특히 가려진 차량에 대한 계수 벡터는 차량과 배경의 기저 벡터 모두 대응되기 때문에 대부분의 계수 벡터의 값이 차량과 배경의 계수 벡터보다 높게 나타난다. 배경의 계수 벡터의 경우 다양한 기저 벡터와 대응되어 고르게 분포되어 있는 것을 볼 수 있다. 또한 차량의 계수는 세부적으로 차종, 시점, 차폐영역의 정도에 따라 차이를 보이는 계수를 가진다.

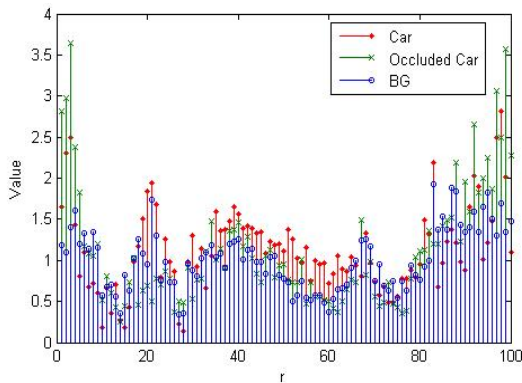


그림 8. 각 클래스에 따른 평균 계수 벡터
Fig. 8. The average encoding vector about each class

그림 9는 기저 벡터의 개수에 따른 초기 목적 함수의 값의 예를 나타내었다. 기저 행렬과 계수 행렬에 임의의 고정된 초기 값으로 초기 목적함수를 계산할 경우 정의된 함수 $G(r)$ 은 n 과 m 의 개수에 대한 특성을 반영한다. 따라서 $G(r)$ 은 부분 기반 표현과 분류 능력을 충분히 가지는 기저 벡터의 개수를 가질 수 있도록 하며 목적함수의 최소값에 근사한 값에서 시작할 수 있도록 한다. 추출된 후보영역으로 학습하였을 때 기저 벡터 개수에 따른 목적함수는 기저벡터의 개수가 약 50개일 때 수렴하기 시작하고 약 150개일 때부터 다시 증가하기 시작한다. 따라서 계산의 복잡도 측면에서 수렴하기 시작하는 최소의 기저벡터의 수를 선택한다.

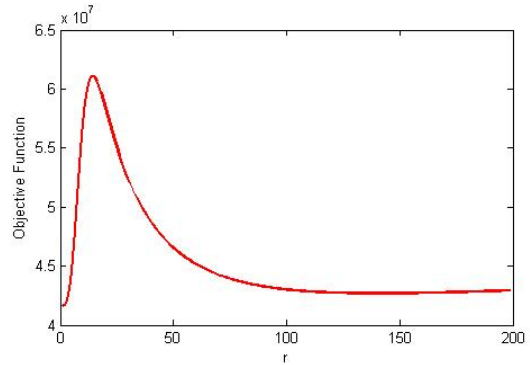


그림 9. 기저 벡터의 개수에 따른 목적 함수
Fig. 9. The objective function value according to the number of basis vectors

선택된 기저 벡터의 수를 검증하기 위하여 인식을 그림 10에 나타내었다. 시작 목적함수의 값이 차이 남에 따라 반복횟수를 다르게 설정하였으며 기저 벡터의 형태 또한 다르게 나타났다. 목적함수가 다시 증가하는 구간인 150개 이상의 기저벡터의 인식률은 수렴구간에서의 인식률에 비해 점차 감소하였으며 50개 이하의 기저 벡터로 설정하는 경우 부분 기반 표현을 만족하는 기저 벡터를 생성할 수 없었다.

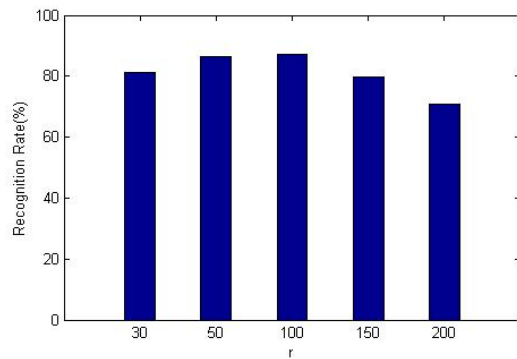


그림 10. 기저 벡터의 개수에 따른 차량 인식률
Fig. 10. The Vehicle recognition rate according to the number of basis vectors

제안된 알고리즘의 성능 평가를 위하여 실시간성을 만족하는 선형 변환 방식 중 가장 대표적인 PCA와 차량 인식에 많이 적용되는 Haar특징^[20]을 식 (15)에서 정의된 인식률을 이용하여 제안한 알고리즘과 비교하였다. 또한 NMF와 SNMF를 비교하기 위하여 같은 기저 벡터의 개수 100개를 가지며 SNMF의 희소성 계수 β 를 0.3으로 설정하였다. 표 2은 차량이 많이 존재하는 영상들에서의 기저 벡터의 개수에 따른 차량의 평균 인식률을 나타내었다.

PCA의 경우 적은 수의 고유값으로 정보의 표현이 가능하기 때문에 상위 100개의 고유값을 선정하여 분류하였고 NMF만을 적용하였을 때는 다른 알고리즘과 동일하게 유클리디언 거리만을 사용하여 분류하였다.

표 2에서 PCA의 경우 검출 이미지의 배경에 대한 영향을 많이 받아 다양한 형태의 차량에 대한 인식률이 좋은 결과를 보이지 못하고 있다. 또한 Haar특징의 경우 전방 진행 차량에 대한 인식률은 높지만 차량의 측면의 경우 인식이 이루어지지 않기 때문에 인식률이 저하되었다. 또한 횡단보도와 창문과 같은 도로 내에 존재하는 다양한 배경 요소를 차량으로 오인식하는 경우가 발생한다. 그에 반해 NMF를 적용하였을 때 차량이 적은 환경에서는 다른 알고리즘과 유사한 인식률을 보였으며 차량과 배경의 패턴 분류에 강인함을 보이기 때문에 차량이 많이 존재하는 환경에서의 인식률이 크게 저하되지 않는다. 또한 약 20fps로 처리하기 때문에 실시간성을 만족한다. 그러나 기저의 개수가 학습데이터의 수에 비하여 너무 적으면 각 기저가 차량을 충분히 표현하지 못하고 최대 가능 개수에 가까워지면 계산량이 급격히 늘어나기 때문에 실시간성을 만족할 수 없고 인식률 역시 저하됨을 알 수 있다. LVQ 신경 회로망을 사용하였을 때 인식률의 큰 변화가 없지만 높은 인식률을 보장함을 알 수 있다.

표 2. 도심 영상에서의 차량 인식률
Table 2. The vehicle recognition rate in the urban scene

Feature	Basis	Recognition Rate(%)
Haar+Adaboost		81.6
PCA	100	74.2
NMF	50	71.8
	100	87.5
	200	69.3
SNMF+LVQ	50	71.4
	100	85.7
	200	86.1

비교한 특징들이 대부분 높은 인식률을 보였으나 실제 응용에서는 조도, 날씨 등의 외부적 요인에 의하여 획득한 영상 내에 차량의 일부가 보이지 않거나 차량 간의 겹침으로 인하여 차폐 영역이 발생한

차량을 많이 접하게 된다. 이는 잠재적 위험을 감지할 수 없다는 것을 의미하며 따라서 실제 적용을 위하여 차폐영역이 발생한 차량에 대한 인식이 필요하다. 그림 11에서 10~50% 가려진 차량 이미지에 대하여 차폐 영역이 발생하는 이미지로 정의하고 차폐영역이 발생하는 영상에 대한 인식 결과를 비교하였다.

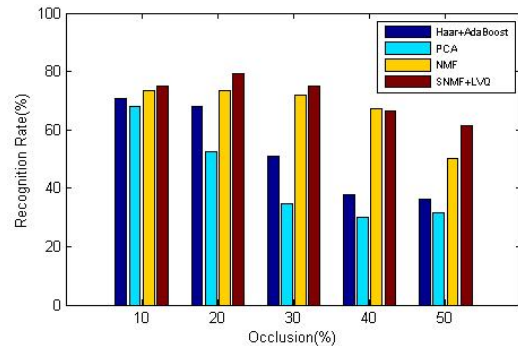


그림 11. 차폐 영역에 대한 차량 인식률
Fig. 11. The vehicle recognition rate about occluded region

PCA를 사용하였을 때 이미지의 전체적인 특징을 사용하기 때문에 작은 차폐영역을 가진 이미지에 대하여 인식률이 민감하게 반응 하지 않았으나 30% 이상을 가린 차량의 경우 인식률이 50%이하로 감소하였고 또한 Haar와 AdaBoost를 사용한 알고리즘은 전면이나 후면 이외의 가려진 차량에 대하여 인식률이 크게 저하되었다. NMF의 경우 학습된 기저 벡터의 형태에 따라 인식률이 변화하였으며 SNMF의 계수 벡터를 입력으로 하는 LVQ 신경 회로망을 사용하였을 때 차폐 영역의 존재에도 불구하고 성능을 유지할 수 있었다. 또한 제안된 알고리즘은 음영 지역에서의 조도 감소로 인한 낮은 대비를 가지는 영상에 대해서 비교한 알고리즘에 비하여 강인함을 보였다.

VI. 결 론

제안된 논문에서는 실제 도로상에서의 차량 인식을 위한 NMF를 통한 특징 추출을 제안하였다. 제안된 차량 인식 방법은 스테레오 비전을 이용하여 특정 범위내의 모든 객체에 대하여 다수의 후보영역을 검출하고 후보 영역으로 추출된 차량 이미지의 지역적인 특징을 차량과 배경에 대한 NMF특징으로 학습된 신경망에 인식하는 것이다. NMF는 차

원 감소의 효과가 있으며 추출된 차량의 후보 영역을 국소적인 부분에 대한 특징의 결합으로 표현함으로써 혼잡한 도로 상에서 가려진 차량의 인식에서 NMF이외의 전체적인 특징을 표현하는 선형 변환 방식과 비교하여 강건함을 갖게 된다. 또한 SNMF는 희소성의 조절을 통하여 특징으로 추출되는 영역을 조절함으로써 다양한 응용에 사용될 수 있다. 또한 NMF에 의하여 분해된 계수 행렬을 보면 차중, 전면, 후면에 따라 각 클래스별 유사한 결과물을 볼 수 있었으며, 이는 추후 과제로 차중 인식이나 중앙선 침범이나 끼어들기 차량 단속과 같은 법률에 관련된 문제 해결에 응용되어 사용될 수 있을 것이다.

References

- [1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol.401, pp.788-791, 1999.
- [2] Z. Sun, "On-Road Vehicle Detection: A Review," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no 9, pp. 694-711, 2006.
- [3] A. Kuehne, "Symmetry-Based Recognition for Vehicle Rears," *Pattern Recognition Letters*, vol. 12, pp. 249-258, 1991.
- [4] S.D. Buluswar and B.A. Draper, "Color Machine Vision for Autonomous Vehicles," *Int'l J. Eng. Applications of Artificial Intelligence*, vol. 1, no. 2, pp. 245-256, 1998.
- [5] C. Tzomakas and W. Seelen, "Vehicle Detection in Traffic Scenes Using Shadows," Technical Report 98-06, Institut fur Neuroinformatik, Ruhr Universitat, Bochum, Germany, 1998.
- [6] M. Bertozzi, A. Broggi, and S. Castelluccio, "A Real-Time Oriented System for Vehicle Detection," *J. Systems Architecture*, pp. 317-325, 1997.
- [7] N. Srinivasa, "A Vision-Based Vehicle Detection and Tracking Method for Forward Collision Warning," *Proc. IEEE Intelligent Vehicle Symp.*, pp. 626-631, 2002.
- [8] A. Giachetti, M. Campani, and V. Torre, "The Use of Optical Flow for Road Navigation," *IEEE Trans. Robotics and Automation*, vol. 14, no. 1, pp. 34-48, 1998.
- [9] R. Mandelbaum, L. McDowell, L. Bogoni, B. Beich, and M.Hansen, "Real-Time Stereo Processing, Obstacle Detection, and Terrain Estimation from Vehicle-Mounted Stereo Cameras," *Proc. IEEE Workshop Applications of Computer Vision*, pp. 288-289, 1998.
- [10] U. Franke and I. Kutzbach, "Fast Stereo Based Object Detection for Stop and Go Traffic," *Intelligent Vehicles*, pp. 339-344, 1996.
- [11] U. Regensburger and V.Graefe, "Visual Recognition of Obstacles on Roads," *Intelligent Robots and Systems*, pp. 73-86, 1995.
- [12] T. Kalinke, C. Tzomakas, and W. von Seelen, "A Texture-Based Object Detection and an Adaptive Model-Based Classification," *Proc. IEEE Int'l Conf. Intelligent Vehicles*, pp. 143-148, 1998.
- [13] Z. Sun, G. Bebis, and R. Miller, "On-Road Vehicle Detection Using Gabor Filters and Support Vector Machines," *Proc. IEEE Int'l Conf' Digital Signal Processing*, July 2002.
- [14] J. Wu and X. Zhang, "A PCA Classifier and Its Application in Vehicle Detection," *Proc. IEEE Int'l Joint Conf. Neural Networks*, 2001.
- [15] S. Mika, "Fisher discriminant analysis with kernels," *IEEE Conference on Neural Networks for Signal Processing IX*, pp. 41-48, 1999.
- [16] A. Hyvarinen, "Fast and Robust Fixed-Point Algorithm for Independent Component Analysis," *IEEE Trans. Neural Networks*, vol.10, no 3, pp. 626-634, 1999.
- [17] N.Naveen, "Parts based representation for pedestrian using NMF with robustness to partial occlusion," *Signal Processing and Communications (SPCOM), 2010 International Conference.*, 2010.
- [18] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol. 5, pp. 1457 - 1469, 2004.
- [19] T. Kohonen, J. Hunninen, J. Kangas, J. Kaaaksonen, and K. Torkkola, "LVQ_Pak : The Learning Vector Quantization Program

Package,” *Technical Report A30*, Helsinki Univ. 1996.

- [20] C. Wu, L. Duan, J. Miao, F. Fang, and X. Wang. “Detection of Front-view Vehicle with Occlusions Using AdaBoost,” *International Conference on Information Engineering and Computer Science*, 2009. ICIECS 2009., 2009.
- [21] KITTI Vision Benchmark, <http://www.cvlibs.net/>, 2010

반재민 (Jae Min Ban)



2011년 인천대학교
정보통신공학과 학사졸업
2011년~현재 인천대학교
정보통신공학과 석사과정
<관심분야> 영상처리, 패턴 인식

이병래 (Byeong-rae Lee)



1985년 연세대학교
전자공학과 학사졸업
1987년 연세대학교 대학원
전자공학과 석사졸업
1992년 연세대학교 대학원
전자공학과 박사졸업
1992년~현재 한국방송통신
대학교 컴퓨터 과학과 교수

<관심분야> 영상처리, 특징 추출

강현철 (Hyunchul Kang)



1982년 연세대학교
전자공학과 학사졸업
1984년 연세대학교 대학원
전자공학과 석사졸업
1989년 연세대학교 대학원
전자공학과 박사졸업
1990년~현재 인천대학교
정보통신공학과 교수

<관심분야> 영상처리, 패턴 인식, 화상통신