

논문 2015-52-5-14

# 비음수 텐서 분해를 이용한 차량 인식

## ( Vehicle Recognition using Non-negative Tensor Factorization )

반재민\*, 강현철\*\*

(Jae Min Ban and Hyunchul Kang<sup>Ⓢ</sup>)

### 요약

차량 인식을 기반으로 하는 능동 제어는 지능형 자동차의 구현에 필요한 핵심 기술이며, 차폐 영역(occlusion)이 빈번하게 발생하는 도심에서 차량을 인식하기 위하여 차량의 부분적인 모습만으로도 차량을 인식할 수 있는 부분 기반 차량 표현이 필요하다. 본 논문에서는 지역적인 특징을 기저벡터로 사용하는 비음수 텐서 분해(non-negative tensor factorization, NTF)를 이용하여 차량을 표현하고, NTF 분해 계수를 특징으로 차량 인식을 검증하였다. 실험 결과, 제안하는 방법이 기존의 비음수 행렬 분해를 사용한 경우에 비하여 보다 직관적인 부분 표현이 가능하며, 도심 영상에서도 보다 강건하게 차량을 인식함을 보여주었다.

### Abstract

The active control of a vehicle based on vehicle recognition is one of key technologies for the intelligent vehicle, and the part-based image representation is necessary to recognize vehicles with only partial shapes of vehicles especially in urban scene where occlusions frequently occur. In this paper, we implemented a part-based image representation scheme using non-negative tensor factorization(NTF) and realized a robust vehicle recognition system using the NTF feature. The result shows that the proposed method gives more intuitive part-based representation and more robust recognition in urban scene.

**Keywords** : 도심 영상, 부분-기반 표현, 비음수 행렬 분해, 비음수 텐서 분해, 차량인식  
non-negative matrix factorization, non-negative tensor factorization, part-based representation,  
vehicle recognition, urban scene

## I. 서론

차선 인식과 차량 및 보행자 등의 도로상의 물체 인식은 지능형 자동차를 구현하기 위하여 필수적인 기능으로 특히 컴퓨터 시각 장치를 활용한 연구가 활발하게

진행되고 있다.<sup>[1~6]</sup>

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 보장할 수 있다. 대표적인 선형 변환 방식으로는 데이터의 상관행렬(covariance matrix)에서 변동이 심한 방향을 찾아내는 주성분 분석(principal component analysis)<sup>[7]</sup>, 학습 단계에서 클래스의 구분에 중점을 두는 선형 판별식 분석(linear discriminant analysis)<sup>[8]</sup>, 통

\* 학생회원, \*\* 정회원, 인천대학교 정보통신공학과  
(Dept. of Information and Telecommunication Eng.,  
Incheon National University)

Ⓢ Corresponding Author(E-mail: hckang@inu.ac.kr)

※ 본 연구는 2012년도 인천대학교 자체연구비 지원에 의하여 연구되었음.

Received ; March 16, 2015      Revised ; April 17, 2015

Accepted ; April 24, 2015

계적 독립성을 이용하여 독립적인 성분만을 추출하는 독립 성분 분석(independent component analysis)<sup>[9]</sup> 등이 있다. 그러나 기존의 대표적인 선형 변환 방식들은 영상의 전체적인 모습을 대상으로 특징을 추출하기 때문에 물체의 모습이 가려지는 경우 인식 성능의 저하를 초래하게 된다. 따라서 차폐 영역이 존재하는 도심 영상에서 차량을 인식하기 위하여서는 부분 기반 영상 표현(part-based image representation)을 이용하여 차량을 표현하여 차량의 부분적인 모습으로도 인식할 수 있어야 한다. 대표적인 부분 기반 표현 기법으로는 비음수 행렬 분해(non-negative matrix factorization, NMF)과 비음수 텐서 분해(non-negative tensor factorization, NTF) 기법이 있다.

인간의 두뇌가 영상의 부분적인 특징을 인식한다는 점을 응용한 NMF는 입력 데이터가 양의 값만을 가지는 영상을 대상으로 하는 경우 유용하다고 알려져 있다.<sup>[10]</sup> NMF는 지역적인 특징을 이용하기 때문에 물체가 다른 객체에 의해 가려지더라도 목표로 하는 표현과 크게 변화하지 않으므로 복잡한 도심 영상에서의 인식에 다른 선형 변환 방식보다 적합한 방법이다. 그러나 NMF는 2차원의 이미지 데이터를 1차원의 벡터로 변환하는 과정에서 공간 구조 정보의 손실과 언제나 유일한 분해를 보장하지 못하는 문제점을 가지고 있다.

텐서는 다차원 배열(multidimensional array)로 이미지와 비디오는 다차원 데이터 배열로 표현될 수 있기 때문에 이미지의 가로, 세로, 색상, 채널, 시간 등을 반영하는 공간으로 표현이 가능하며, NTF는 NMF를 확장한 개념이다<sup>[11]</sup>. NTF는 2차원 데이터 표현을 보존하고 유일한 분해를 보장하며 근사한 재구성 오차(reconstruction error)를 제공하므로 NMF와 비교하여 효율적인 기저 텐서를 생성할 수 있다.

본 논문에서는 비음수 텐서 분해 기법을 이용하여 차량의 부분적인 특징을 표현함으로써 복잡한 도심 영상에서 발생할 수 있는 차폐 영역에 강인하고 기존의 차량 인식 알고리즘 보다 좋은 성능을 갖는 차량 인식 방법을 제안한다. II장에서는 제안하는 비음수 분해 기법에 대하여 소개한다. III장에서는 비음수 분해 기법을 이용한 차량 인식에 대하여 설명하고, 효율적인 기저 텐서를 생성하는 알고리즘과 계산량을 줄이기 위한 방법을 제안한다. IV장에서는 기저 텐서의 차원에 따른 인식을 비교와 제안한 방법과 기존 차량 인식 알고리즘

과의 인식을 비교를 위한 실험을 하고, 마지막으로 V장에서 결론을 맺는다.

## II. 비음수 분해

### 1. 비음수 행렬 분해

NMF는 다변량 분석(multivariate analysis) 기법으로 입력 데이터가 비음수인 경우 유용하게 사용될 수 있으며, 곱셈 갱신 규칙(multiplicative update rule)<sup>[12]</sup>이 제안되면서 얼굴 인식,<sup>[13]</sup> 차량 인식,<sup>[14]</sup> 오디오 인식<sup>[15]</sup> 등의 응용 분야에서 널리 사용되게 되었으며, 희소성 제약을 갖는 희소 NMF(sparse NMF)<sup>[16]</sup>, 직교성 제약을 갖는 직교 NMF(orthogonal NMF)<sup>[17]</sup> 등의 기법이 제안되고 있다. NMF는 부분 기반 표현이 가능함으로써 전체적인 모습을 표현하는 선형 변환 방식이 가진 문제점을 해결하기 위하여 데이터의 지역적인 특징을 찾아 각 지역별 의미적 특징 학습이 가능하다.

NMF를 제안한 Lee와 Seung의 실험<sup>[10]</sup>에서 얼굴 인식에 적용하였을 때 얼굴을 눈, 코, 입 등의 직관적으로 인식 가능한 부분들을 기저 벡터로 하여 기저 벡터의 선형 조합으로 표현할 수 있다는 것을 알 수 있다.

그림 1은 차량 이미지를 NMF의 기저 벡터와 계수 벡터(coefficient vector)의 선형 조합으로 나타내었다. 이미지  $V$ 는 차량을 구성하는 프레임, 전조등, 번호판 등과 같은 성분을 나타내는 기저 행렬  $W$ 와 계수 벡터  $H$ 의 선형 조합으로 이루어진다. 따라서 차량의 후보 영역에서 차폐 영역이 발생하더라도 특정 기저 벡터에 대응하는 계수 벡터의 계수 값만 작아지기 때문에 전체적인 특징 벡터에 큰 영향을 끼치지 않기 때문에 NMF는 차폐 영역에 대하여 강인한 성질을 가질 수 있다.

$V = WH$ 가 되도록 분해하는 것이 이상적이지만 이러한 경우 각 기저 벡터가 이미지의 전체를 나타내므로 인식의 목적에는 적합하지 않기 때문에 인식에 사용되는 기저 행렬과 계수 행렬은 식 (1)과 같이 정의된다.

$$\begin{array}{cccccccc}
 \text{[Image V]} & \approx & \text{[Image H}_{1,m}\text{]} & + & \text{[Image H}_{2,m}\text{]} & + & \dots & + & \text{[Image H}_{r-2,m}\text{]} & + & \text{[Image H}_{r-1,m}\text{]} & + & \text{[Image H}_{r,m}\text{]} & :W \\
 V & & H_{1,m} & & H_{2,m} & & \dots & & H_{r-2,m} & & H_{r-1,m} & & H_{r,m} & :H
 \end{array}$$

그림 1. 'NMF 선형 조합의 예.

Fig. 1. 'An example of linear combination with NMF.

$$V \approx WH \quad (1)$$

이미지 데이터베이스 행렬  $V$ 가  $n \times m$  크기의 행렬로 나타낸다면  $V$ 는  $n \times r$  크기의 기저 행렬  $W$ 와  $r \times m$  크기의 계수 행렬  $H$ 로 분해된다. 이때  $n$ 은 이미지 데이터베이스를 구성하는 이미지 한 장의 픽셀 수를 나타내고,  $m$ 은 이미지 데이터베이스 행렬에 저장된 이미지의 개수를 나타낸다.  $r$ 은 기저 행렬  $W$ 를 구성하는 기저 벡터의 개수나 각 계수 벡터의 차원과 같고 일반적으로  $(n+m)r < nm$ 을 만족해야 하는 제약 조건이 존재한다.<sup>[10]</sup>

$V$ 에 근사하는  $W$ 와  $H$ 를 생성하기 위해  $V$ 와  $WH$ 의 차이를 목적함수  $F$ 로 정의함으로써 반복적인 갱신을 멈추게 된다. 목적함수  $F$ 는 유클리디언 거리 대신 KL(Kullback-Leibler) divergence를 사용하며, 식 (2)의 목적함수  $F$ 가 수렴하기 시작하면  $W$ 와  $H$ 는 부분 기반 표현을 제외한 나머지에서 0에 가까운 값을 가지게 되고 이러한 특성으로 인하여 부분 기반 표현을 만족할 수 있다.

$$F = \sum_{i=1}^n \sum_{\mu=1}^m [V_{i\mu} \log(WH)_{i\mu} - (WH)_{i\mu}] \quad (2)$$

갱신 규칙은 식(3), (4)와 같이 나타낼 수 있다. 갱신 규칙은 곱셈 갱신 규칙(multiplicative update rule)으로, 기울기 하강(gradient descent) 방법과 비슷하게 동작하며, 현재의  $W$ 와  $H$ 의 값에 예측값을 곱하여 목적함수  $F$ 를 최소화하는 방향으로 갱신한다. 갱신된  $W$ 와  $H$ 는 부분적인 편차를 가지게 되고 행렬 상호간 내부적으로 직교하는 방향으로 변화하게 되며, 부분 기반 표현에서 요구하는 희소성(sparseness)을 만족하게 된다.

$$W_{ia} \leftarrow W_{ia} \sum_{\mu} \frac{V_{i\mu}}{(WH)_{i\mu}} H_{a\mu}, \quad W_{ia} \leftarrow \frac{W_{ia}}{\sum_j W_{ja}} \quad (3)$$

$$H_{a\mu} \leftarrow H_{a\mu} \sum_i \frac{V_{i\mu}}{(WH)_{i\mu}} W_{ia} \quad (4)$$

## 2. 비음수 텐서 분해

NMF는 2차원의 이미지를 1차원의 벡터로 변환하는 과정에서 공간 구조 정보의 손실을 초래하게 된다. 이를 극복하기 위하여 2차원 NMF<sup>[18]</sup>와 NTF 등이 제안

되고 있다. 대표적인 텐서 분해 방법으로는 Tucker 분해(Tucker decomposition)모델<sup>[19]</sup>과 병렬 인자 분석(parallel factor analysis, PARAFAC) 모델<sup>[20]</sup>이 있다.

PARAFAC은 다차원 데이터의 분해 방법 중 하나이며, 특이값 분해의 다차원 분해 방법으로 간주된다. 각 차원의 벡터 값들은 텐서의 부여한 특징의 강도로 해석되며,  $n$ 차원의 텐서  $G$ 가 각 차원에 대한 벡터  $u_1, u_2, \dots, u_n$ 가 외적(outer product)의 합으로 표현되는 경우  $n$ 차원의 텐서는 식 (5)와 같이 나타낸다.

$$G = \sum_{j=1}^k u_1^j \otimes u_2^j \otimes \dots \otimes u_n^j \quad (5)$$

$u_1^j, \dots, u_n^j$ 는 랭크가 1인 벡터를 나타내며, 텐서의 랭크는 텐서를 랭크가 1인 벡터의 선형 합으로 표현하는데 필요한 최소한의 개수를 나타낸다. 따라서 텐서  $G$ 의 랭크는  $k$ 가 되며,  $k$ 가 텐서를 구성하는 최소한의 개수일 때 유일한 분해가 존재한다고 말할 수 있다. 행렬의 랭크는 특이값 분해(singular value decomposition)로 얻을 수 있으나 텐서는 그렇지 못하다. Harshman<sup>[20]</sup>은 PARAFAC 모델을 통하여 텐서  $G$ 의 랭크가 최소의 랭크  $k$ 로 근사하도록 만들 수 있다는 것을 소개하였다. PARAFAC 모델은 그림 2와 같으며 식 (6)과 같이 정의될 수 있다.

$$G \approx \sum_{j=1}^k u_1^j \otimes u_2^j \otimes \dots \otimes u_n^j \quad (6)$$

그러나 PARAFAC은 목적함수가 수렴을 보장하지 않고 분해의 유일성을 보장하기 어렵다는 측면에서 Tucker 분해를 사용하는 방법이 연구되기도 하였다.

Tucker 분해는 2차원 주성분 분석(2D-PCA), 다차원 특이값 분해(high-order SVD) 등에서 사용되는 일반적인 모델이다. 그림 3에서  $n$ 차원의 텐서  $G$ 가 주어졌을 때, 텐서  $G$ 는 Tucker 분해를 통하여 코어 텐서  $S$ 와

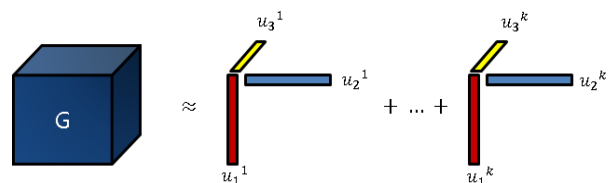


그림 2. PARAFAC 모델의 원리

Fig. 2. The principle of PARAFAC model.

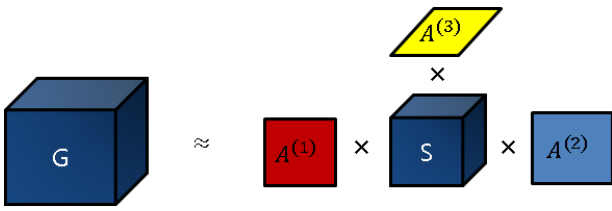


그림 3. Tucker 분해의 원리  
Fig. 3. The principle of Tucker decomposition

계수 행렬  $A^{(1)}, \dots, A^{(n)}$ 로 분해될 수 있으며, Tucker 분해 모델은 식 (7)과 같이 정의할 수 있다.

$$G \approx S \times_1 A^{(1)} \times_2 \dots \times_n A^{(n)} \quad (7)$$

Tucker 분해는 유일한 분해를 갖지만, PARAFAC 보다 속도가 느리다는 단점을 가지고 있기 때문에 인식에 적합하지 않은 경우가 발생한다.

일반적으로 텐서 연산은 곱셈을 행렬처럼 계산하기 위하여 mode-n matricization을 사용한다<sup>[21]</sup>.

NTF의 목적함수는 학습 이미지 텐서  $G$ 와 재구성 텐서  $A^{(1)} \times A^{(2)} \times A^{(3)}$ 의 차이를 최소 제곱 문제 (least-squares problem)로 정의 할 수 있으며, 목적함수는 식 (8)과 같다.

$$F = \| G - (S \times A^{(1)} \times A^{(2)} \times A^{(3)}) \|_F^2 \quad (8)$$

$$\text{subject to : } S, A^{(1)}, A^{(2)}, A^{(3)} \geq 0, S = I$$

$\| A \|_F^2$ 는 Frobenious norm으로 텐서의 모든 요소들의 차의 제곱이다. 분해된 코어 텐서  $S$ 를 단위행렬로 정의하였으며  $A^{(1)}, A^{(2)}, A^{(3)}$ 는 각각 차량 이미지 텐서의 열과 행, 그리고 학습 이미지의 가중치를 나타내며  $G$ 를  $A^{(1)}$ 과  $A^{(2)}$ 의 곱으로 사용되는 기저 텐서와 대응하는 계수 행렬로 정의하는  $A^{(3)}$ 으로 분해한다.

제안하는 NTF의 각 행렬의 분해는 먼저 임의의 값을 mode-n matricization된 각 차원의 행렬에 적용시킨 후 기저 텐서와 계수 행렬에 대하여 NMF의 곱셈 갱신 규칙을 확장시켜 식 (9)와 같이 나타내며 이 갱신 규칙은 텐서  $G$ 를 근사하는 값으로 갱신한다.

$$A^{(i)} \leftarrow A^{(i)} * \frac{G^{(i)} A_{\odot}^{(i)}}{A^{(i)} A_{\odot}^{(i)T} A_{\odot}^{(i)}} \quad (9)$$

$A_{\odot}^{(n)}$ 는 행렬의 텐서의 곱셈 Kronecker product의

한 종류인 Khatri-Rao product로  $A^{(n)} \odot, \dots, \odot A^{(1)}$ 로 나타내며, \* 연산자는 Hadamard product로  $(A * B)_{ij} = a_{ij} b_{ij}$ 와 같이 표현되며, 행렬 간의 각 위치의 성분의 곱을 계산한다. 갱신 규칙은 기울기 하강 방법을 반복하여 분해된  $A^{(i)}$ 를 텐서  $G$ 에 근사하는 방향으로 갱신하며, 목적함수  $F$ 의 값이 수렴 할 때까지 수행한다.

NTF는 NMF와 마찬가지로 비음수 제약 조건을 만족시킬 수 있고, 부분 기반 표현을 만족하는 기저 텐서와 계수 행렬을 생성할 수 있다. NTF는 공간 정보를 포함하고 있다는 점과 차원의 확장이 용이하다는 점을 제외하고는 NMF와 유사하게 동작하며, 신경 회로망의 구조와 유사하다는 성질을 유지하기 때문에 분류 능력이 뛰어난 특성을 가질 수 있다.

### III. 비음수 텐서 분해를 이용한 차량 인식

컴퓨터 시각 장치를 이용한 차량 인식 시스템은 차량 후보 검출 단계와 차량 인식(또는 검증) 단계의 두 단계로 나눌 수 있다.<sup>[1]</sup>

#### 1. 차량 인식 시스템의 개요

본 논문에서는 스테레오 비전(stereo vision)을 이용하여 도로 영상에서 차량 후보를 검출하며, 제안하는 차량 인식 시스템의 전체적인 구성은 그림 4와 같다.

스테레오 카메라로부터 영상으로 부터 거리 정보를 포함하고 있는 깊이 맵(depth map)을 생성한다. 실제 도심 주행 영상에는 많은 차폐 영역이 존재하기 때문에 특정 거리 내에 존재하는 모든 객체를 대상으로 후보 영역을 검출한다.

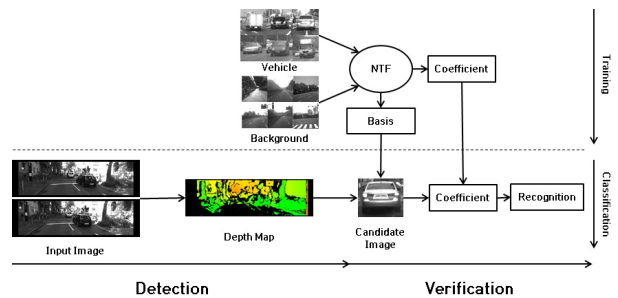


그림 4. 제안하는 차량 인식 시스템의 개요도  
Fig. 4. Overview of proposed vehicle recognition system.

검증 단계에서는 학습 이미지들로 생성한 텐서를 NTF를 이용하여 계수 행렬을 생성한다. 이때 계수 행렬의 차원은 기저 행렬들의 개수와 같으며, 각 집합의 지역적인 특징을 학습하고 차량 이미지의 집합으로부터 차량의 부분 모습을 각각의 기저 행렬들이 표현한다. 계수 행렬은 결과적으로 학습 이미지의 지역적 부분을 나타내는 기저 텐서를 가중치로 사용하게 되고, 새롭게 검출된 후보 영역에서 분해된 계수 벡터와 차이를 구한다. 적은 수의 계수 벡터를 이용하기 때문에 차원 감소 효과로 인하여 빠른 속도로 차량을 인식할 수 있다.

차량 후보 영역 검출은 차량 후보 영역의 위치를 영상 내에서 빠르게 찾고자 하며, 스테레오 비전 기반 방법은 차량의 위치와 거리 정보를 동시에 얻을 수 있는 장점이 있다. 차량 후보 영역을 검출하기 위하여 깊이 맵을 생성하고 생성된 깊이 맵에서 특정 비율과 크기를 만족하는 약 30m 이내의 모든 객체를 추출한다. 깊이 맵은 좌우 영상의 대응 화소를 매칭 알고리즘을 사용하여 추출하기 때문에 두 영상의 대응 위치를 정확하게 검출하는 것이 정확한 깊이 맵을 생성하는데 필수적이며, 실시간성이 요구되기 때문에 매칭 알고리즘의 복잡도가 높지 않아야 한다. 본 논문에서 구현한 차량 후보 영역의 검출은 그림 5에 나타내었다.

두 영상의 밝기 차이의 영향을 줄이기 위하여 전처리 과정이 필요하며, 밝기 정보를 정규화하고 질감을 향상시키는 작업을 수행한다. 스테레오 비전은 질감이 두드러진 점을 매칭할 때 좋은 성능을 보이기 때문에 이를 위하여 Laplacian of Gaussian을 이용하여 전처리 과정을 수행한다. 이는 스무딩 된 영상에서 지역적 최대값을 찾고, 식 (10)의 SAD(Sum of Absolute Difference)를 계산하여 대응점을 검출한다.

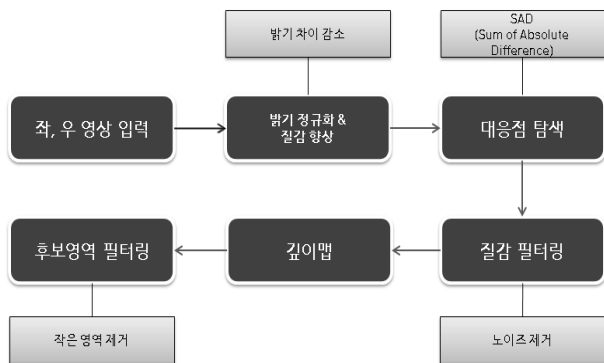


그림 5. 차량 후보 영역 검출의 흐름도  
Fig. 5. The flow of vehicle detection step.

$$SAD(x,y,d) = \sum_{y=0}^{h-1} \sum_{x=0}^{w-1} |I_L(x,y) - I_R(x+d,y)| \quad (10)$$

2. NTF를 이용한 차량 후보 영역 검증

차량 후보 영역 검증 단계는 후보 영역 검출 단계에서 찾아진 차량의 위치에 대한 차량 여부를 판단하는 단계이다.

차량의 일부가 보이지 않거나 겹침으로 인하여 차폐 영역이 존재하는 차량 후보 영역에 대하여 다음과 같은 과정을 통하여 차량을 인식한다.

학습 단계에서는 차량 이미지와 차폐 영역이 존재하는 차량 이미지, 배경 이미지를 차량의 종류, 시점 등의 순서로 이미지 데이터베이스 텐서  $G$ 를 생성한다. 텐서는 3차원 큐브 모양으로 나타낼 수 있으며, 각 차원은 행, 열, 클래스를 나타내는  $W, H, C$  행렬들로 분해되도록 구성하여 그림 6의 학습 과정을 수행한다.

각 차원을 나타내는 분해된 행렬  $W, H, C$ 는 식(9)에 정의된 곱셈 갱신 규칙을 통하여 식 (11), (12), (13)과 같이 갱신되며 가장 앞에 있는 행렬부터 변화하기 시작하며 가장 마지막 행렬이 가장 나중에 변한다. 이는 기저 행렬들이 차량의 부분을 잘 표현하도록 하는데 유용하게 작용한다.

$$W^{(i)} \leftarrow W^{(i)} * \frac{G^{(i)} W_{\odot}^{(i)}}{W^{(i)} W_{\odot}^{(i)T} W^{(i)}} \quad (11)$$

$$H^{(i)} \leftarrow H^{(i)} * \frac{G^{(i)} H_{\odot}^{(i)}}{H^{(i)} H_{\odot}^{(i)T} H^{(i)}} \quad (12)$$

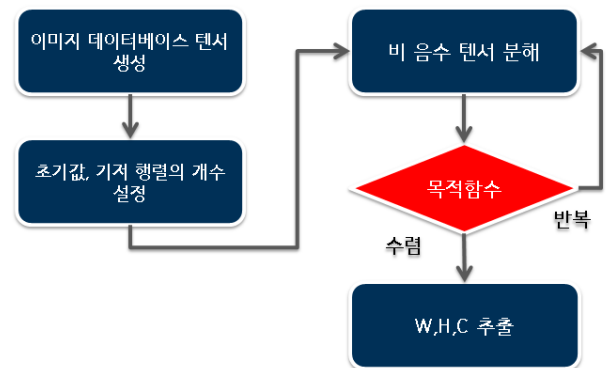


그림 6. 학습 과정  
Fig. 6. Training steps for vehicle verification.



$$C^{(i)} \leftarrow C^{(i)*} \frac{G_{(i)} C_{\odot}^{(i)}}{C^{(i)} C_{\odot}^{(i)T} C_{\odot}^{(i)}} \quad (13)$$

또한 빠른 수렴을 위하여  $W, H, C$  행렬의 초기값을 1로 초기화하고 기저 행렬의 개수를 변경하면서 NTF를 1회 수행하였을 때 식 (14)의 목적함수  $F$ 가 수렴하는 구간의 최소값을 가지는 기저 행렬의 개수  $r$ 를 선택하여 분해한다. 기저 행렬의 개수에 따른 목적 함수  $F$ 의 벡터를  $F(r)$ 라 가정하면,  $F(r)$ 의 기울기가 0에 가까운 임의의 상수  $\epsilon$ 보다 작게 되는 제약 조건을 만족시키는  $r$ 를 선택함으로써 이미지 데이터베이스의 특성에 맞는 기저 행렬의 수를 선택할 수 있다.

$$r = \operatorname{argmin} [F(r)] \quad (14)$$

subject to :  $\nabla \cdot F < \epsilon$

그림 7은 인식 단계의 흐름도를 나타내었다. 차량의 후보 영역이 입력되면 학습되었던 NTF의  $W$ 와  $H$ 를 고정하고 목적 함수가 수렴할 때까지 분해를 반복하여 입력 이미지에 대하여 새로운 계수 벡터  $c$ 를 추출한다. 추출된 계수 벡터  $c$ 는 입력 이미지의 클래스에 따라서 다른 계수 벡터  $c$ 를 생성하기 때문에 학습된 계수 행렬  $C$ 의 각 열과 비교하였을 때 유클리디언 거리를 이용하여 가장 차이가 적은 클래스를 선택하여 구분한다.

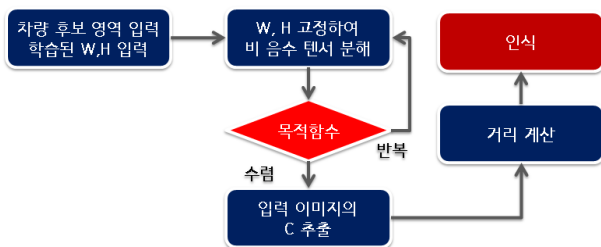


그림 7. 차량 인식 과정  
Fig. 7. The flow of vehicle recognition.

#### IV. 실험 결과 및 분석

##### 1. 실험 환경

제안하는 알고리즘을 성능 평가를 위하여 KITTI Vision Benchmark<sup>[22]</sup>의 1344×391 해상도의 스테레오 이미지를 30 fps 영상으로 편집하여 사용하였으며 카메라 보정 파라미터는 각각의 영상에 따라 제공된 파라미

터를 사용하였다. 실험에 사용된 영상은 맑은 날 주간을 기준으로 전방을 주행하는 차량과 주차되어 있어 겹침이 발생한 차량, 가로수나 표지판 등의 물체에 가려진 차량, 눈이나 그림자에 의해 조도가 다른 차량 등을 다수 포함하고 있다. 실험 환경은 Intel(R) Core(TM) i5-2400 @ 3.16GHz, windows 7 32bit 환경에서 visual studio 2010과 OpenCV 라이브러리를 사용하였다.

학습에 사용된 차량 학습 이미지는 그림 8과 같으며, 검출을 수행하는 거리의 최대 범위를 30m로 설정하였다. 모든 학습 이미지는 48x48의 크기로 정규화 하였으



(a)



(b)



(c)

그림 8. 훈련 이미지의 예. (a) 차량이미지, (b) 배경이미지, (c) 차폐된 이미지

Fig. 8. The example of training images. (a) vehicle images, (b) background images, (c) occluded images

표 1. 각 프레임에 존재하는 검출된 영역 비교  
Table 1. Comparison of the number of objects in a frame.

Ground truth	Detection	Verification
8.7	25.5	9.4

며 정규화된 이미지를 3차원 큐브 형태의 텐서로 구성하여 기저 텐서와 계수 행렬로 분해하였다. 학습 데이터는 총 200개의 이미지를 사용하였으며, 각 클래스 구분은 정상 차량 이미지, 차폐 영역이 존재하는 차량 이미지로 구성되어 있다. 따라서 학습 이미지 텐서는  $2304(48 \times 48) \times 200$ 으로 구성되어 있으며 특징의 개수인 계수 행렬의 수를 증가시키며 기존의 차량 인식 알고리즘과 인식률을 비교하였다.

제안한 방법에서는 각 프레임에 대하여 검출 단계와 검증 단계를 수행하는 동안 후보 영역의 수가 감소하기 때문에 실험에 사용된 입력 영상들의 각 프레임에 존재하는 평균적인 전체 차량 수와 그에 따른 제안된 방법의 평균적인 검출된 영역의 수, 검증된 영역의 수를 표 1에서 비교하였다. 검출 영역은 약 30m 이내의 모든 객체를 추출하기 때문에 실제로 검출된 후보 영역에서는 창문, 보행자 등의 차량 이외의 객체가 포함되어 있으며, 검증 단계에서는 이를 보완하기 때문에 실제 차량 수와 근사하게 감소하는 것을 볼 수 있다.

그러나 차량 인식 시스템에서 추출한 영역은 언제나 차량을 포함하고 있지 않기 때문에 정확한 성능을 계산하기 위하여 인식률과 실패율을 식 (15)와 식 (16)와 같이 정의하였다.

$$Recg\_rate(\%) = \frac{Vehicle_{veh} + BG_{bg}}{total} \times 100 \quad (15)$$

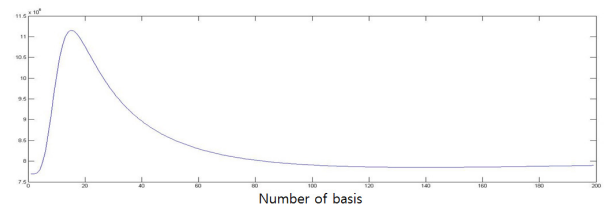
$$Fail\_rate(\%) = \frac{Vehicle_{bg} + BG_{veh}}{total} \times 100 \quad (16)$$

$total$ 은  $Vehicle_{veh}$ ,  $Vehicle_{bg}$ ,  $BG_{veh}$ ,  $BG_{bg}$ 의 합으로 후보 영역의 총 개수를,  $Vehicle_{veh}$ 과  $BG_{bg}$ 는 각 클래스가 바르게 인식된 수를,  $Vehicle_{bg}$ 와  $BG_{veh}$ 는 틀리게 인식된 수를 나타낸다. 차량 인식에 많이 사용하는 Haar 특징과 PCA, NMF, SNMF와 LVQ2 신경회로망<sup>[21]</sup>을 결합한 방법과 제안한 NTF를 비교하여 제안한 차량 인식 시스템에 대한 성능을 평가하였다.

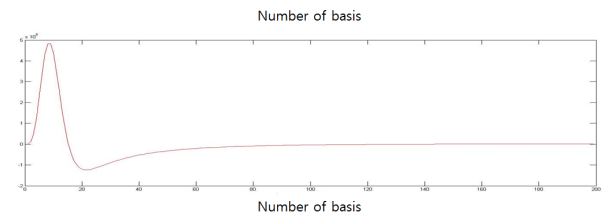
## 2. 실험 및 결과

NTF는 반복하여 기저 행렬들과 계수 행렬로 분해하기 때문에 많은 시간이 요구된다. 따라서 최적의 시작점과 최적의 기저 행렬의 개수의 선택이 필요하다. 최적의 시작점은 기저 행렬의 개수와 관련되어 있기 때문에 기저 행렬의 개수를 변화하면서 목적함수가 수렴하는 구간의 최소값의 기저 행렬의 개수를 선택하여 분해를 시작한다. 그림 9는 NTF의 기저 행렬의 개수에 따른 초기 목적 함수 값의 변화를 나타내었다. NMF의 이미지 데이터베이스 행렬과 이미지 데이터베이스 텐서의 크기는 같은 학습 데이터를 사용하기 때문에 같으며, 기저 행렬과 계수 행렬에 임의의 고정된 초기 값으로 초기 목적함수를 계산할 경우 정의된 목적 함수의 벡터  $F(r)$ 은 행렬과 텐서의 크기에 대한 특성을 반영한다. 따라서  $F(r)$ 은 부분 기반 표현과 분류 능력을 충분히 가지는 기저 행렬의 개수를 얻을 수 있도록 한다.

차량 이미지 데이터베이스 텐서의 기저 행렬에 따른 목적함수의 벡터  $F(r)$ 은 기저 행렬의 개수가 약 40개 일 때 수렴하기 시작하며 약 60개 이상일 때 목적함수의 기울기의 변화가 거의 없어져 인식에 영향을 미치지 않는 중복된 기저를 생성한다. 이는 생성한 텐서의 크기에 대한 요소와 밀접한 관련이 있기 때문에 계산의 복잡도 측면에서 수렴을 시작하는 구간이나 수렴하는 구간의 최소값을 만드는 기저 행렬의 개수를 선택하는



(a) Objective function



(b) Gradient of objective function

그림 9. NTF의 기저 행렬 개수에 따른 목적함수 값  
Fig. 9. The objective function value according to the number of basis vector for NTF.

표 2. 기저 행렬 개수에 대한 차량 후보 영역 인식률  
Table 2. The vehicle recognition rate according to the number of basis matrix.

기저행렬 개수	시작 목적함수	반복 횟수	종료 목적함수	인식률 (%)
20	10768937	148	255099	56
40	8966834	181	99698	85
60	8297491	151	99481	83
80	8024064	146	98865	77
140	7851829	143	101725	79

것이 유리하게 작용함을 알 수 있다.

기저 행렬의 개수  $r$ 에 따른 인식률을 검증하기 위하여 표 2에 기저 행렬의 개수에 대한 차량 후보 영역 100개에 대한 인식률을 나타내었다. 시작 목적 함수의 값이 기저 행렬의 개수  $r$ 에 따라 차이를 보이기 때문에 목적함수  $F$ 가 수렴하기 시작하는 반복 횟수가 다르게 나타나며, 각 기저 행렬의 형태 또한 정확히 차량의 부분을 표현하지 못하는 기저 행렬을 생성하기도 하였다.

목적함수의 벡터  $F(r)$ 이 수렴하기 시작하는 구간인 약 20개 이하의 기저 행렬의 경우 기저 행렬들이 부분 기반 표현을 만족하지 못하고 전체를 표현하게 되는 결과를 보였으며, 다시 증가하는 구간이 시작되는 약 60개 이상의 기저 행렬의 경우 기저 행렬들이 중복되는 경우가 많았으며 차량의 각 부분을 표현하는 기저 행렬을 생성하지 못하였다.

그림 10은 제안된 방법에 따라 기저 행렬의 개수를 설정하고 분해를 하였을 때 생성되는 기저 행렬의 예를 나타내었다. 생성된 기저 행렬들은 차량의 각 부분을 표현하고 있기 때문에 인식 단계에서 유용한 특징으로

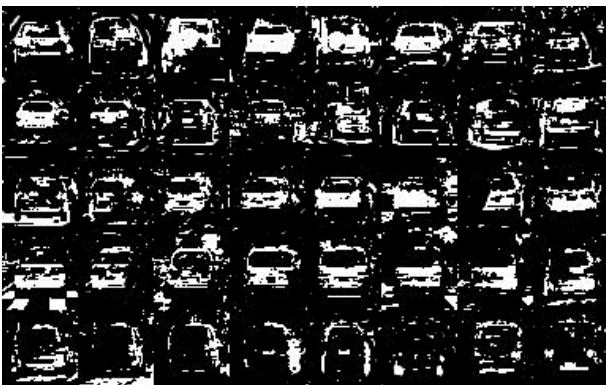


그림 10. NTF 기저 행렬의 예  
Fig. 10. The example of basis matrix of NTF.

사용될 수 있다. 따라서 적절한 기저 행렬의 개수를 선택함으로써 차량의 각 부분을 표현하는 기저 행렬에 나타나는 형태와 형태에 따른 인식률의 변화가 밀접하게 연관되어 있음을 알 수 있다.

기저 행렬의 개수가 선택되면 생성된 기저 행렬과 연관된 계수 행렬의 각 벡터에 따라 클래스가 결정된다. 그림 11은 정의된 클래스에 따른 각 클래스에 대한 NTF의 계수 벡터  $C$ 의 평균값을 나타낸다. X축은 기저 벡터와 행렬의 순서를 나타내고, Y축은 계수 벡터의 평균값을 나타낸다.

차량을 나타내는 계수 벡터의 경우 차량의 부분을 나타내는 기저 벡터와 행렬에 연결된 계수 벡터가 높은 값을 나타내고 있으며, 특히 가려진 차량에 대한 계수 벡터는 특정 기저 벡터와 행렬에 연관된 계수만이 높은 값을 나타내고 있다. 배경의 계수 벡터의 경우 다양한 기저 벡터와 행렬과 대응하기 때문에 계수 벡터의 값이 평균적으로 작으며, 고르게 분포되어 있는 것을 볼 수 있다. 또한 차량의 계수 벡터는 차종, 시점, 차폐 영역의 정도에 따라 차이를 보인다.

### 3. 알고리즘의 성능 비교

NTF의 성능평가를 위하여 차량 인식에 많이 사용하고 있는 Haar 특징과 선형 변환 방식 중 대표적인 PCA, 유사한 부분 기반 표현으로 분해하는 NMF와 SNMF, NTF를 비교하였다.

PCA의 경우 적은 수의 고유값으로 표현이 가능하기 때문에 상위 100개의 고유 값을 선정하여 분류하였고, NMF, SNMF, NTF의 기저 행렬 개수는 목적 함수가 수렴하는 구간의 최소가 되는 개수에 근사화하고 선정된 개수에 대하여 인식률을 실험하였다. 실험된 각 알고리즘에 따른 인식률을 표 3에 나타내었다. 비교한 특징들이 대부분 높은 인식률을 보였으나 실제 응용에서는 조도의 변화, 그림자 등의 외부적인 요인에 의하여 차량의 일부가 보이지 않거나 겹침으로 인한 차폐 영역이 발생하는 경우를 흔히 볼 수 있다. 이는 잠재적인 위험을 감지할 수 없다는 것을 의미하며, 실제 적용을 위하여 차폐 영역이 발생한 차량 후보 영역에 대한 인식률의 비교가 필요하다. 표 4는 차량 후보 영역의 크기에서 약 10~50% 차폐 영역이 발생한 차량 이미지에 대한 인식 결과를 비교하였다.

제안된 NTF를 이용한 차량 인식 방법은 차폐 영역



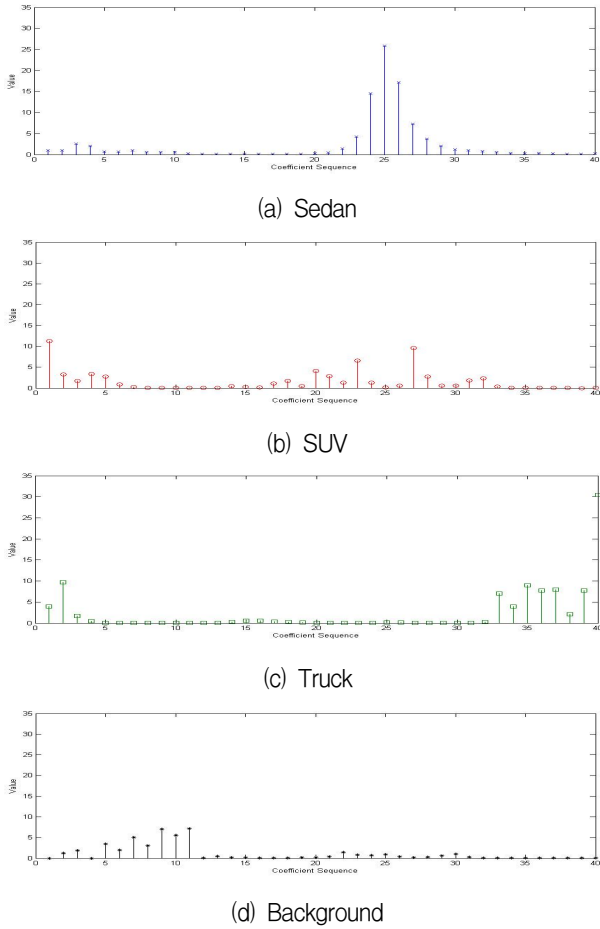


그림 11. 각 클래스에 대한 계수 벡터의 예  
 Fig. 11. The example of coefficient vector for each class.

표 4. 도심 영상에서 알고리즘 별 차량 인식을 비교  
 Table 4. The comparison of the vehicle recognition rate in the urban scene.

Occlusion Method	10	20	30	40	50
Haar+AdaBoost	72.5	70.5	51	39.5	37.5
PCA	71	58.5	37.5	36	38.5
NMF	79	78	78.5	66.5	57.5
SNMF+LVQ	63.5	64	62.5	61.5	60.5
NTF	77.5	80	81.5	75.5	72.5

을 위한 기저 벡터들을 생성함으로써 차폐 영역의 존재에도 불구하고 인식을 유지할 수 있었으며, 음영 지역에서 조도 감소로 인한 낮은 대비를 가지는 후보 영역에 비하여 최대 15%, 평균적으로 5.5%의 인식이 향상됨으로써 강인함을 보였다.

표 3. 도심 영상에서 알고리즘 별 차량 인식을 비교  
 Table 3. The comparison of the vehicle recognition rate in the urban scene.

Feature	Number of basis	Recognition Rate(%)
Haar+Adaboost		81.6
PCA	100	74.2
NMF	100	87.5
SNMF+LVQ	200	86.1
NTF	40	90.4

V. 결론

본 논문은 차폐 영역이 빈번히 발생하는 도심 영상에서 적용 가능한 NTF를 이용한 차량 인식 방법을 제안하였다. 도심 영상은 다른 차량에 의해 가려진 차량, 표지판, 가로수 등에 의해 부분만 보이는 차량 등의 차량 인식에 있어 인식을 저하시키는 요소들이 다수 포함되어 있어 차폐 영역을 포함한 차량 후보에 대하여 인식 단계에서 강인하여야 한다. 따라서 본 논문에서는 이러한 문제점을 해결하기 위해서 이용하여 특정 범위 내의 객체들을 후보 영역으로 추출한 뒤 후보 영역에서 차량의 부분을 표현하는 NTF의 기저 텐서와 계수 행렬을 이용하여 차량 인식을 수행하였다.

많은 차량이 겹쳐져 있거나 시설물, 사람에 의해 가려진 환경의 도심 영상을 대상으로 실험한 결과 기존의 차량 인식 알고리즘보다 제안한 방법이 높은 인식을 보이는 것을 확인 할 수 있었다. 기존의 차량 인식 알고리즘은 차폐 영역에 대하여 비교적 낮은 인식을 보이기 때문에 이는 잠재적 위험을 감지 할 수 없다는 것을 의미하지만, 제안한 방법은 차폐 영역에 대하여 크게 영향을 받지 않는 장점이 있다. 또한 NTF는 NMF와 다르게 다양한 특성을 정의하여 차원 확장이 가능하기 때문에 실제 응용에서 고려되는 조도, 날씨 등의 외부적인 요인에 의하여 인식이 저하되는 점을 방지할 수 있다. 따라서 본 논문에서 제안한 방법이 기존의 차량 인식 알고리즘보다 차량 인식 성능이 향상 되었으며, 특히 차폐 영역이 존재하는 후보 영역에 대한 인식률은 NMF보다 최대 12% 향상됨을 확인할 수 있다.

하지만 곱셈 갱신 규칙은 구현이 간편한 장점을 가지고 있지만 수렴속도가 느리다는 단점을 지니고 있다.

따라서 본 논문에서는 적절한 기저의 개수를 찾는 방법을 제안하였다. 그러나 학습 데이터의 수가 증가하면 많은 계산량으로 인해 기저 텐서와 계수 행렬의 생성에 많은 시간이 요구된다. 따라서 최적화에 기반하여 보다 빠른 갱신 규칙에 대한 연구가 필요하다. 또한 재구성 텐서와 원본 텐서의 차이가 수렴하는 구간에서 갱신 법칙을 중단하기 때문에 정확한 규칙이 존재하지 않으며, 갱신 회수에 따라 생성되는 기저 텐서의 형태도 변화할 수 있다. 이는 차량 학습 데이터가 조금이라도 변경될 경우 다른 분해를 생성하기 때문에 유일성(uniqueness)의 측면에서 고려해야하는 사항이 될 수 있다.

또한 매 프레임마다 차량을 인식하는 방법은 높은 인식률을 얻을 수 있지만 매번 각 후보 영역에 대하여 텐서 분해를 반복하기 때문에 많은 연산을 요구한다. 따라서 인식된 차량에 대한 추적(tracking) 등을 통하여 인식된 차량에 대한 연산을 줄여주는 방법이 필요하다.

## REFERENCES

- [1] Z. Sun, "On-Road Vehicle Detection: A Review," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.28, no.9, pp.694-711, 2006.
- [2] D. Balcones et al., Real-time Vision-Based Vehicle Detection for Rear-End Collision Mitigation Systems," *LNCS 5717*, pp. 320-325, 2009.
- [3] N. Srinivasa, "A Vision-Based Vehicle Detection and Tracking Method for Forward Collision Warning," *IEEE Intelligent Vehicle Symp.*, pp.626-631, 2002.
- [4] Joel C. McCall and Mohan M. Trivedi, "Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation," *IEEE Trans. on Intelligent Transportation Systems*, Vol. 7, No. 1, pp. 20-37, March 2006.
- [5] Wei Liu, XueZhi Wen, Bobo Duan, Huai Yuan and Nan Wang, "Rear Vehicle Detection and Tracking for Lane Change Assist," *2007 IEEE Intelligent Vehicle Symposium*, Istanbul, Turkey, pp. 252-257, June 2007.
- [6] P. Kelly, N. E. O'Connor and A. F. Smeaton, "A Framework for Evaluating Stereo-Based Pedestrian Detection Techniques," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 18, No. 8, pp. 1163-1167, Aug. 2008.
- [7] J. Wu and X. Zhang, "A PCA Classifier and Its Application in Vehicle Detection," *IEEE Int'l Joint Conf. Neural Networks*, 2001.
- [8] S. Mika, "Fisher discriminant analysis with kernels," *IEEE Conference on Neural Networks for Signal Processing IX*, pp.41-48, 1999.
- [9] A. Hyvarinen, "Fast and Robust Fixed-Point Algorithm for Independent Component Analysis," *IEEE Trans. Neural Networks*, vol.10, no.3, pp. 626-634, 1999.
- [10] D. D. Lee and H. S. Seung, "Learning the Parts of Objects by Non-negative Matrix Factorization," *Nature*, vol. 401, pp.788-791, 1999.
- [11] A. Shashua and T. Hazan, "Non-negative Tensor Factorization with Applications to Statistics and Computer Vision," *International Conference on Machine Learning*, Bonn, Germany, 2005.
- [12] D. D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," *Advances in Neural Information Processing Systems*, Vol. 13, pp. 556-562, MIT Press, 2001.
- [13] S. Z. Li, X. W. Hou, H. J. Zhang and Q. S. Cheng, "Learning spatially localized part-based representation," *IEEE Int. Conference on Computer Vision Pattern Recognition*, Kauai, Hawaii, pp. 207-212, 2001.
- [14] M. Hwang and H. Kang, "A Vehicle Recognition Using Part-based Representations," *LNEE 235*, pp. 309-316, 2013. 6.
- [15] Y. C. Cho and S. Choi, "Nonnegative features of spectro-temporal sounds for classification," *Pattern Recognition Letters*, vol. 26, no. 9, pp.1327-1336, 2005.
- [16] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol.5, pp.1457 - 1469, 2004.
- [17] Seungjin Choi, "Algorithms for Orthogonal Nonnegative Matrix Factorization," *IEEE Int. Joint Conference on Neural Networks*, Hong Kong, pp. 1828-1832, 2008. 6.
- [18] D. Zhang, S. Chen and Z. Zhou, "Two-Dimensional Non-negative Matrix Factorization for Face Representation and Recognition," *LNCS 3723*, pp.350-363, 2005.
- [19] L. de Lathauwer, B. de Moor, and J. Vandewalle, "A Multilinear Singular Value Decomposition," *SIAM J. Matrix Anal. appl.*, vol.21, no.4, pp.1253-1278, 2000.
- [20] R. A. Harshman, "Foundations of PARAFAC

- Procedure: Models and Conditions for an Exploratory. Multi-modal Factor Analysis,” *UCLA Working Papers in Phonetics*, 1970.
- [21] M. K. Gullu, E. Yaman, and S. Erturk, “Image sequence stabilization using fuzzy adaptive Kalman filtering,” *Electronic Letters.*, vol.39, no.5, pp. 429-431, Mar. 6, 2003.
- [22] KITTI Vision Benchmark, <http://www.cvlibs.net/>, 2010.
- [23] T. G. Kolda and Brett W. Bader, “Tensor Decompositions and Applications,” *SIAM Review*, Vol. 51, No.3, pp.455-500, 2009.

---

— 저 자 소 개 —

---



반 재 민(학생회원)  
 2011년 인천대학교 정보통신  
 공학과 학사 졸업.  
 2013년 인천대학교 정보통신  
 공학과 석사 졸업.  
 <주관심분야 : 영상처리, 컴퓨터  
 비전, 화상통신>



강 현 철(정회원)  
 1982년 연세대학교 전자공학과  
 학사 졸업.  
 1984년 연세대학교 전자공학과  
 석사 졸업.  
 1989년 연세대학교 전자공학과  
 박사 졸업

1990년~현재 인천대학교 정보통신공학과 교수  
 <주관심분야 : 영상처리, 컴퓨터 비전, 화상통신>