

뉴턴 최적화를 통해 개선된 아다부스트 훈련과 MCT 특징을 이용한 번호판 검출

이영현*, 김대훈**, 고한석**

License Plate Detection with Improved Adaboost Learning based on Newton's Optimization and MCT

Young-Hyun Lee *, Dae-Hun Kim **, Han-Seok Ko **

요약

본 논문에서는 MCT(Modified Census Transform) 특징과 개선된 아다부스트 분류기를 이용한 번호판 검출 알고리즘을 제안한다. MCT 특징은 영상의 국소 지역 패턴을 정수화하여 표현하는 특징으로서 조명 변화에 강인하고 메모리 효율이 높은 장점이 있다. 그러나 패턴을 표현하는 정수형의 MCT 특징값들이 이산적인 특징을 가지기 때문에 아다부스트 훈련 방법을 적용하기 위해서는 룩업테이블(Lookup Table)을 이용하여 분류기를 설계해야 한다. 그동안의 아다부스트 훈련 방법에 대한 최적화 연구는 지수 기준(exponential criterion)을 최소화 하는 방법에 대한 방향으로 연구가 진행되고 있다. 본 논문에서는 MCT 특징을 이용하고 지수 기준의 뉴턴 최적화를 통해 아다부스트 훈련 방법을 개선하여 번호판 검출 성능을 향상 시키는 방법을 제안한다. 번호판 샘플 영상과 필드 테스트 영상에 대한 실험을 통해 제안한 방법의 성능을 고찰하고, 기존의 일반 아다부스트 훈련을 이용한 검출 방법과의 비교 실험을 통해 그 효용성을 입증한다.

▶ Keywords : 번호판 검출, 아다부스트 훈련, 뉴턴 최적화, MCT

Abstract

In this paper, we propose a license plate detection method with improved Adaboost learning and MCT (Modified Census Transform). The MCT represents the local structure patterns as integer numbered feature values which has robustness to illumination change and memory efficiency. However, since these integer values are discrete, a lookup table is needed to design a weak classifier for Adaboost learning. Some previous research efforts have focused on minimization of

• 제1저자 : 이영현 • 교신저자 : 고한석

• 투고일 : 2012. 10. 10, 심사일 : 2012. 10. 15, 게재확정일 : 2012. 10. 23.

* 고려대학교 영상정보처리협동과정(Dept. of Visual Information Processing, Korea University)

* 고려대학교 전기전자전파공학부(School of Electrical Engineering, Korea University)

exponential criterion for Adaboost optimization. In this paper, a method that uses MCT and improved Adaboost learning based on Newton's optimization to exponential criterion is proposed for license plate detection. Experimental results on license patch images and field images demonstrate that the proposed method yields higher performance of detection rates with low false positives than the conventional method using the original Adaboost learning.

▶ Keywords : License plate detection, Adaboost learning, Newton's optimization, MCT

I. 서 론

최근 카메라의 발달과 컴퓨터의 기술 진전으로 다양한 영상처리 서비스가 가능하게 되고, 이를 활용하는 컴퓨터 비전 응용 서비스가 다양화 되고 있다. 그 예로, CCTV를 이용한 감시 보안 장치 (Surveillance & Security system)[1], 로봇 비전을 이용한 휴먼 로봇 인터페이스 장치 (human robot interface system)[2], 뎀스 카메라 (depth camera)를 이용한 게임 엔터테인먼트 장치 (game entertainment system)[3] 등 다양한 분야에서 컴퓨터 비전 기술이 활용되고 있다. 특히, 객체 검출 기술 (object detection)은 컴퓨터 비전 기술 중 가장 중요한 요소 중 하나로, 영상 내에 관심 객체의 위치를 찾는 기술이다. 실제로, 감시 보안 장치에서 움직이는 객체 검출[4] 및 사용자 인식을 위한 얼굴 검출[5], 휴먼 로봇 인터페이스 장치에서 화자 인식을 위한 얼굴 검출[6-7] 및 제스처 인식을 위한 손검출[8], 게임 엔터테인먼트 장치에서 동작 인식을 위한 바디 파트 (body part) 검출[9] 등 객체 검출 기술은 컴퓨터 비전 기술에서 인식을 위한 필수 선행 기술의 역할을 하고 있다.

현대 사회에서 차량의 수가 증가하면서 교통 통제 시스템에 대한 수요가 많아지고 있고, 이를 효율적으로 관리하기 위해 컴퓨터 비전 기술을 적용한 지능형 교통 통제 시스템 (Intelligent Transporting System)가 도입되고 있다. 이 시스템은 CCTV 카메라를 통해 차량 번호판을 인식하여 차량의 신원을 파악함으로써 교통 상황을 제어한다. 그 중 자동차 번호판 검출은 자동차에 대한 정보를 분석하는데 중요한 번호판 인식 기술의 필수 선행 기술이다. 기존의 번호판 검출 기술에는 번호판 배경과 글자 사이에서 나타나는 강한 엣지 (edge) 특징을 이용한 방법[10-11], 번호판 색상정보를 이용하는 방법[12], 사각형 모양을 찾을 수 있는 허프 (Hough) 변환을 통한 번호판의 경계를 찾는 방법[13], 아다부스트, 신경망 알고리즘 등의 훈련 알고리즘을 이용하여 분

류기를 통해 번호판 영역을 찾는 방법[14-15]이 있다. 그 중 훈련 기반의 번호판 검출 방법은 구현이 간단하며, 검출 속도가 빠르고, 주변 조명 변화 등 환경의 영향을 크게 받지 않은 장점이 있기 때문에 이에 대한 연구가 많이 진행되고 있다.

훈련 기반의 번호판 검출에 많이 사용되는 Haar-like 특징은 비교 영역 간의 명도 (intensity) 차이를 이용하여 특징을 표현하는 방법으로서 위치, 모양, 크기에 따라 수많은 형태를 나타낼 수 있기 때문에 아다부스트 학습 과정을 통해 최적의 유효한 특징을 선별함으로써 번호판 검출에 유용하게 사용할 수 있다[14]. 하지만 이 특징은 명도 정보를 그대로 사용하기 때문에 조명 변화와 잡음에 대해 민감한 단점이 있다. 이에 반해 MCT (Modified Census Transform) 특징은 영상의 국소 지역 패턴을 정수화하여 표현하는 방법으로서 조명 변화에 강인하고 메모리 효율이 높은 장점이 있다[16]. 이러한 MCT의 장점을 이용하여 [17]은 번호판 검출 알고리즘과 초해상도 복원 (Super resolution) 알고리즘을 결합하여 번호판 인식을 위한 시스템을 제안하기도 하였다.

아다부스트 알고리즘은 성능이 조금 미흡한 다수 개의 약한 분류기 (weak classifier)를 조합하여 성능이 좋은 하나의 강한 분류기 (strong classifier)를 만드는 훈련 방법이다 [18]. 처음 아다부스트가 소개된 이후 몇 가지 다른 버전의 아다부스트 방법이 제안되었는데, 이 방법들은 지수 기준 최소화 (exponential criterion minimization)를 통해 훈련 방법을 최적화하는 방법으로서 리얼 아다부스트 (Real Adaboost)[19], 젠틀 아다부스트 (Gentle Adaboost)[20] 등이 그 예이다.

일반적으로 아다부스트 훈련 과정은 약한 분류기의 수식적 설계에 따라 구현된다. Haar-like를 비롯한 일반적인 특징값들은 특징값의 크기 변화에 따라 연속적인 특성 변화를 보이기 때문에 문턱치 경계 (thresholding boundary)를 통해 간편히 분류기를 구현할 수 있다. 그러나 MCT 특징은 특징값이 이산적인 특성을 가지고 있기 때문에 문턱치 경계를 사용할 수 없으며, 룩업테이블 (lookup table)을 사용하여 분류기를 설계해야 한다.

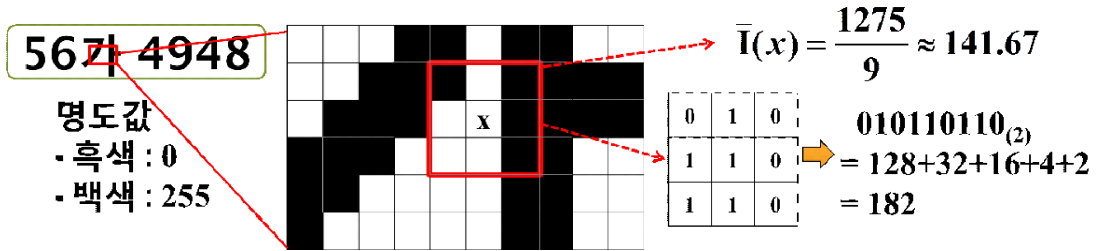


그림 1 . MCT 특징 추출 방법의 예.
Fig. 1. Example of MCT feature extraction.

본 논문에서는 MCT 특징과 뉴턴 최적화를 통해 개선된 아다부스트 훈련을 이용한 번호판 검출 방법을 제안한다. 룩업테이블을 사용하여 MCT 특징에 따른 분류기를 설계하고, 뉴턴 최적화 방법을 통해 아다부스트의 지수 기준을 최소화하는 개선된 아다부스트 훈련 방법을 적용한다.

본 논문의 구성은 다음과 같다. 2장에서는 MCT 특징에 대해서 소개하고 3장에서는 뉴턴 최적화를 통해 개선된 아다부스트 훈련 방법과 룩업테이블을 사용하여 분류기를 설계하는 방법을 소개한다. 4장에서는 설계된 분류기를 이용하여 입력 영상에서 번호판을 검출하는 방법에 대해 설명하고 5장에서는 실험을 통해 제안한 방법의 성능을 고찰한다. 마지막으로 6장에서 결론을 맺는다.

II. 번호판 영상의 특징 추출

MCT 특징은 [21]에서 얼굴 검출을 위해 제안한 특징 추출 방법으로서 특정 위치의 픽셀 x 에 대해 주변 인접 픽셀과의 관계를 패턴으로 표현하는 방법이다. MCT 특징은 명도 값 자체를 그대로 사용하지 않고 주변 픽셀과의 상관 관계를 패턴으로 표현하기 때문에 조명 변화에 강한 장점이 있다. 3×3 커널을 사용하여 MCT 특징을 구하는 방법은 아래 식 (1)과 같다.

$$(\mathbf{x}) = \otimes_{\mathbf{y} \in N' \mathbf{x}} \zeta I(\mathbf{x}), I(\mathbf{y}) \quad (1)$$

여기서 $N' \mathbf{x}$ 는 커널 내의 인접 픽셀들의 집합, $I(\mathbf{x})$ 는 x 의 명도, $I(\mathbf{y})$ 는 커널 내 픽셀의 평균 명도를 나타낸다. $\zeta()$ 는 비교함수로서 $I(\mathbf{y})$ 가 $\bar{I}(\mathbf{x})$ 보다 크면 1, 그렇지 않으면 0을 출력한다. \otimes 는 10진수 변환 연산자로서 $\zeta()$ 의

결과로 나오는 9자리의 2진수 배열을 10진수로 변환한다. 따라서 본 논문에서 사용한 MCT 특징값은 0에서 510까지의 범위를 가진다. 그림 1은 MCT 특징값을 추출하는 과정을 보여주는 예이다. 번호판 영상 내의 어떤 픽셀 위치 x 에서 자기 자신을 포함한 주변 9개 픽셀의 평균 명도를 구하고, 각각의 위치에서 평균 명도와의 크기를 비교한다. 평균 명도보다 큰 픽셀은 1, 그렇지 않으면 0으로 변환하고, 2진수 배열을 10진수 변환하여 픽셀 위치 x 에서의 특징값을 추출한다.

III. 아다부스트 분류기 설계

1. 뉴턴 최적화를 통한 아다부스트 훈련

번호판 영상과 배경 영상의 판별을 위한 분류기 설계는 아다부스트 훈련 방법을 따른다. 아다부스트 훈련 방법은 분류 성능이 떨어지는 다수 개의 약한 분류기들을 조합하여 분류 성능이 좋은 하나의 강한 분류기를 만드는 방법이다[18]. 아다부스트 훈련 과정의 기본적인 순서는 다음과 같다. 먼저 모든 훈련 샘플들에 대한 웨이트 (weight)를 같은 값으로 초기화 한다. 다음 약한 분류기의 후보군들 중에 가장 최소 에러를 가지는 것을 선택한다. 여기서 에러값은 약한 분류기에 의해 잘못 분류되는 샘플들의 웨이트 합으로 구한다. 약한 분류기가 선택되고 나면 선택된 분류기에 대해 잘못 분류된 샘플들은 웨이트를 높이고, 제대로 분류된 샘플들은 웨이트를 낮춰준다. 다음 라운드에서 새롭게 분배된 웨이트를 가지는 샘플들에 대해 최소 에러를 가지는 약한 분류기를 선택하고 샘플들의 웨이트를 그에 따라 업데이트하는 과정을 반복한다. 이 반복 과정은 정해진 라운드의 개수 만큼 약한 분류기가 선택되거나 선택된 약한 분류기들을 조합하여 만든 강한 분류기의 성능이 지정된 에러율보다 낮아지면 멈춘다. 표 1은 아다부스트 훈련 과정을 표현한 것이다. 아다부스트를 이용한 분

표 1. 아다부스트 훈련 과정
Table. Procedure of Adaboost Learning

<p>1. 샘플 웨이트 초기화: $w_i = 1/N, i = 1, \dots, N$</p> <p>2. T 라운드까지 반복 수행 ($t = 1, \dots, T$)</p> <p>(a) 샘플들의 웨이트에 대해 최소 에러를 가지는 약한 분류기 $h_t(x) \in \{1, -1\}$ 선택</p> $= \arg \min_{i=1} \varepsilon_j = w_t(i) [y_i \neq h_j(x_i)]$ $I[y_i \neq h_j(x_i)] = \begin{cases} 1, & y_i \neq h_j(x_i) \\ 0, & \text{otherwise} \end{cases}$ <p>(b) 선택된 약한 분류기의 에러율 계산</p> $\varepsilon_t = \sum_{i=1}^N w_t(i) I[y_i \neq h_t(x_i)]$ <p>(c) 가중치 α_t 계산</p> $\alpha_t = \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right)$ <p>(d) 샘플 웨이트 업데이트 & 노멀라이즈</p> $w_{t+1}(i) = \frac{w_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$ $Z_t = \sum_{i=1}^N w_t(i) \exp[-\alpha_t y_i h_t(x_i)]$ <p>3. 조합된 강한 분류기</p> $H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$
--

류기 설계는 구현이 간편하고 처리 속도가 빠르기 때문에 컴퓨터 비전과 패턴 인식 분야에서 많이 사용되고 있다 [14,16,17,21].

처음 아다부스트 방법이 소개된 이후 보다 개선된 형태의 아다부스트 방법들이 제안되었다. 특히 지수 함수 (exponential function)를 비용 함수 (cost function)로 두고 이를 최소화하는 지수 기준 (exponential criterion)을 사용한 리얼 아다부스트 (Real Adaboost)와 젠틀 아다부스트 (Gentle Adaboost) 방법들에 응용되었다. 개선된 형태의 아다부스트에서 사용되는 지수 비용 함수는 식 (2)와 같다 [20].

$$E[\exp(-yH(x))] \tag{2}$$

여기서 x 는 샘플의 특징값, y 는 샘플 x 의 레이블 (label)이다. $H(x)$ 는 아다부스트 훈련 과정을 통해 설계된 강한 분류기로서 약한 분류기가 순차적으로 조합되는 형태이기 때문에 식 (3)과 같이 가법 모형 (additive model)으로 표현할 수 있다.

$$H(x) = \sum_{t=1} h_t(x) \tag{3}$$

여기서, $h(x)$ 는 약한 분류기를 의미하며 T 는 조합되는 약한 분류기의 개수이다. 이러한 가법 모형에서 이전 라운드까지 조합된 강한 분류기에 순차적으로 조합되는 약한 분류기는 식 (2)의 지수 비용 함수를 최소화하는 방향으로 조합되어야 한다. gentle Adaboost는 이 과정을 뉴턴 최적화 방법 (Newton's optimization)을 통해 접근한다. 비용 함수 $J(H(x) + h(x))$ 를 최소화하는 새로운 강한 분류기 $H(x)$ 를 뉴턴 최적화 방법을 통해 찾는 식은 아래 식(4)와 같이 된다.

$$H(x) \leftarrow H(x) - \begin{pmatrix} \frac{\partial J(H(x) + h(x))}{\partial h(x)} \\ \frac{\partial^2 J(H(x) + h(x))}{\partial h(x)^2} \end{pmatrix} \tag{4}$$

식 (4)의 1차 미분항을 전개하면 식 (5)와 같고, 식 (5)의 결과를 이용하여 2차 미분항을 식 (6)과 같이 구할 수 있다.

$$\begin{aligned} & \frac{\partial (H(x) + h(x))}{\partial h(x)} \Big|_{(x)=0} \\ &= \frac{\partial E \exp(-y(H(x) + h(x)))}{\partial h(x)} \Big|_{h(x)=0} \\ &= -E[y \exp(-yH(x))] \end{aligned} \tag{5}$$

$$\begin{aligned} & \frac{\partial^2}{\partial h(x)} (H(x) + h(x)) \Big|_{h(x)=0} \\ &= - \frac{\partial E [y \exp(-y(H(x) + h(x))) | x]}{\partial h(x)} \Big|_{h(x)=0} \\ &= E [\exp(-yH(x)) | x] \end{aligned} \tag{6}$$

식 (5)와 식 (6)의 결과를 식 (4)에 대입하면 식 (7)과 같이 된다.

$$H(x) \leftarrow -H(x) + \frac{E [y \exp(-yH(x)) | x]}{E [\exp(-yH(x)) | x]} \tag{7}$$

다음, 식 (8)과 같이 샘플의 웨이트 (x, y) 를 정의하고, 식 (9)와 같이 가중치 기댓값 (weighted expectation)을 정의한다.

$$(x, y) = \exp(-y H(x)) \tag{8}$$

$$E [(x, y) | x] \equiv \left(\frac{E [w(x, y)g(x, y) | x]}{E [w(x, y) | x]} \right) \tag{9}$$

식 (8)과 식 (9)의 결과를 식 (7)에 적용하면 최종적으로 식 (10)을 얻는다.

$$H(x) \leftarrow -H(x) + E [y | x] \tag{10}$$

기존의 일반 아다부스트와 개선된 아다부스트와의 큰 차이는 기존의 일반 아다부스트는 약한 분류기 $h(x)$ 가 1 또는 -1 값을 가지지만, 개선된 아다부스트는 $h(x)$ 가 $y | x$ 의 실수값을 가진다는 점이다. 따라서 기존의 일반 아다부스트는 표 1의 2.(d)와 같이 라운드마다 선택된 약한 분류기의 에러율에 따라 샘플 전체가 같은 비율로 웨이트가 업데이트되거나 개선된 아다부스트는 선택된 약한 분류기에 대해 샘플 각각이 가지는 신뢰도에 따라 개별적으로 웨이트가 업데이트된다. 이와 관련된 자세한 내용은 3.2절에서 다룬다.

2. 룩업테이블을 이용한 분류기 설계

명도의 차이, 엣지의 방향, 엣지의 크기 등등 일반적으로 패턴 인식에서 많이 사용되는 특징 추출 방법은 그 특징값이 실수값을 가지는 경우가 많으며 특징값의 크기 변화에 따라 연속적인 특성 변화를 보인다. 즉, 엣지의 크기를 예로 들었을 때, 특징값이 클수록 엣지의 크기가 큰 것이고, 특징값이 작을수록 엣지의 크기가 작다. 따라서 이러한 유형의 특징값에 대해서는 문턱치 경계 (thresholding boundary)를 사용하여 간단하게 분류기를 구현할 수 있다. 그러나 MCT 특징은 특징값이 정수의 형태로 표현되며, 그 값들은 이산적인 특성을 가지고 있다. 즉, 특징값을 표현하는 정수 각각들은 각자의 독립적인 패턴을 표현하고 있으며, 특징값의 크기 변화에 따라 패턴의 연속적인 변화를 표현할 수 없다. 따라서 MCT 특징은 문턱치 경계를 사용할 수 없고, 룩업테이블 (lookup table)을 사용하여 분류기를 설계해야 한다. 룩업테

→ 특징값

0	1	2	...	509	510
1	-1	-1	...	1	1

그림 2. 룩업테이블을 이용한 분류기의 예
Fig. 2. Example of Classifier using Lookup Table

→ 특징값

	0	1	2	...	509	510
1	1	-1	-1	...	1	1
2	-1	-1	1	...	-1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
579	1	1	1	...	-1	1
580	1	-1	-1	...	-1	-1

↓
픽셀
위치

그림 3. 룩업테이블을 이용한 약한 분류기의 예
Fig. 3. Example of Weak Classifier using Lookup Table

→ 특징값

0	1	2	...	509	510
0.02	-0.13	-0.08	...	0.18	0.06

그림 4. 개선된 훈련에서의 약한 분류기 예
Fig. 4. Example of Weak Classifier in Improved Method

이블을 이용한 분류기는 그림 2와 같이 설계할 수 있다. 그림 2의 첫 번째 행의 숫자는 MCT 특징값이고, 두 번째 행의 숫자는 분류 결과이다. 즉, 각 특징값에 대응하는 분류 결과를 찾음으로써 분류기의 역할을 수행한다. 그림 3은 실제 번호판 검출에 사용되는 록업테이블을 이용한 약한 분류기의 예이다. 테이블의 각 열은 특징값을 의미하며, 각 행은 약한 분류기의 후보군을 의미한다. 본 논문에서는 훈련 샘플의 크기를 60x12 로 하였고, 외곽 픽셀은 사용하지 않았기 때문에 총 580개의 픽셀 위치가 약한 분류기의 후보군이 된다.

표 2는 본 논문에서 제안하는 개선된 아다부스트 기반의 MCT 특징을 이용한 분류기의 훈련 방법이다. 먼저 번호판 훈련 샘플과 배경 훈련 샘플 각각의 웨이트를 초기화 한다. 다음 표 2의 2. (a)에 따라 샘플들의 MCT 특징값에 대한 히스토그램을 생성한다. 히스토그램은 모든 샘플에서 같은 위치에 있는 픽셀의 MCT 특징값을 빈(bin)으로 하여 샘플들의 웨이트를 누적하여 생성한다. 예를 들어 번호판 샘플과 배경 샘플이 각각 1000장씩 있고, 샘플의 픽셀 개수가 580개, MCT 특징값의 범위가 0에서 510까지 라면, 1000장의 번호판 샘플의 특징값과 배경 샘플의 특징값에 대해 1번 픽셀 위치부터 580번 픽셀 위치까지 각각 히스토그램을 형성, 511개의 빈 (bin)을 가지는 번호판 샘플에 대한 히스토그램 580개, 배경 샘플에 대한 히스토그램 580개, 총 1160개의 히스토그램을 생성할 수 있다. 생성된 히스토그램은 특징값에 따른 분류기를 설계하는 것에 대한 정보를 제공한다. 예를 들어, 1번 픽셀 위치의 번호판 샘플 히스토그램과 배경 샘플 히스토그램을 비교한다면, 히스토그램의 511개의 빈에 대해 각 빈마다 히스토그램의 값이 번호판 샘플과 배경 샘플 중 큰 쪽에 대해 분류 결과를 내리도록 하는 분류기를 설계하는 것이다. 이 경우 분류되지 않은 쪽의 히스토그램 값은 분류 결과에 대

표 2. 개선된 아다부스트 기반의 분류기 훈련
Table 2. Classifier Learning based on Improved Adaboost

<p>P : 번호판 샘플 수, N_{LP} : 배경 샘플 수 LP : 번호판 특징값, Γ_i^{BG} : 배경 특징값 (i 번째 샘플) $y_i \in \{+1, -1\}$: 레이블 (+1:번호판, -1:배경)</p>
<p>1. 샘플 웨이트 초기화</p> $(i) = \begin{matrix} 1 & 1 \\ 2 & P \end{matrix}, 2N_{BG}$

<p>2. T 라운드까지 반복 수행 ($t=1, \dots, T$) (a) 샘플을 통해 MCT값의 히스토그램 형성</p> $g^{LP}(\cdot, \gamma) = \sum_i w_t(i) I(\Gamma_i^{LP}(\mathbf{x}) = \gamma) I(y_i = 1)$ $g_t^{BG}(\mathbf{x}, \gamma) = \sum_i w_t(i) I(\Gamma_i^{BG}(\mathbf{x}) = \gamma) I(y_i = -1)$ <p>(b) 에러율 계산</p> $\varepsilon(\mathbf{x}) = \sum_{\gamma} \min g_t^{LP}(\mathbf{x}, \gamma), g_t^{BG}(\mathbf{x}, \gamma)$ <p>(c) 최소 에러율의 픽셀 위치 선택</p> $\mathbf{x} = \begin{cases} \operatorname{argmin}_{\mathbf{x}} \varepsilon_t(\mathbf{x}), & \text{if } W_t < n \\ \operatorname{argmin}_{\mathbf{x} \in W} \varepsilon_t(\mathbf{x}), & \text{else} \end{cases}$ <p style="text-align: right;">: 최대 허용 약한 분류기 개수 : t 이전 선택된 약한 분류기 집합, $W_t = W_{t-1} \cup$</p> <p>(d) 약한 분류기를 위한 록업테이블 생성</p> $h_t(\gamma) = g_t^{LP}(\mathbf{x}_t, \gamma) - g_t^{BG}(\mathbf{x}_t, \gamma)$ <p>(e) 샘플 웨이트 업데이트 & 노멀라이즈</p> $w_{t+1}(i) = \frac{w_t(i) \exp(-y_i h_t(\Gamma_i(\mathbf{x}_t)))}{Z_t}$ $Z_t = \sum_i w_t(i) \exp(-y_i h_t(\Gamma_i(\mathbf{x}_t)))$
<p>3. 픽셀 분류기 조합</p> $p_{\mathbf{x}}(\gamma) = \sum_{t=1}^T h_t(\gamma) I(\mathbf{x} = \mathbf{x}_t)$ <p>결과 : 강한 분류기</p> $\Gamma = \bigcap_{\mathbf{x}} p_{\mathbf{x}}(\Gamma(\mathbf{x}))$

한 에러값이 된다. 즉, 특정 픽셀 위치의 어떤 특징값에 대해 번호판 샘플의 히스토그램 값이 0.03이고 배경 샘플의 히스토그램 값이 0.01이라면, 이 픽셀 위치의 어떤 특징값에 대한 분류 결과는 번호판이 되며, 그 분류에 대한 에러값은 0.01이 되는 것이다. 표 2의 2.(b)와 같이 특정 픽셀 위치에 대해 모든 빈의 에러값을 더하면 그 픽셀 위치에 대한 에러율을 구할 수 있다. 다음, 모든 픽셀 위치 중 가장 에러율이 적은 픽셀 위치를 선택하여 그 픽셀 위치에 대한 분류기를 아다부스트

훈련 과정에서 선택된 약한 분류기로 선정한다. 그러나 만약 강한 분류기를 조합하기 위한 약한 분류기의 개수를 n 개 이하로 한정하고자 하고, 이전 라운드에서 이미 n 개의 픽셀 위치가 선택이 되었다면, 이미 선정된 n 개의 픽셀 위치에 대해서만 에러율 비교를 통해 약한 분류기를 선정한다. 이렇게 하면 아무리 아다부스트 훈련 과정의 라운드 수를 증가시켜도 강한 분류기에 조합되는 약한 분류기의 수는 n 개로 고정시킬 수 있다. 선택된 약한 분류기는 표 2의 2.(d)의 식 (11)과 같은 값을 가지는 룩업테이블로 표현된다.

$$(\gamma) = g_t^P(\cdot, \gamma) - g_t^{BG}(\mathbf{x}_t, \gamma) \quad (11)$$

식 (11)을 이해하기 위해서는 먼저 식 (10)의 $|x|$ 을 분석해야 한다. 분류기의 분류 결과는 1 또는 -1의 이진 분류 결과를 가지기 때문에 $E_w[y|x]$ 은 식 (12)와 같이 표현될 수 있다.

$$w[y|x] = P_w(y = 1|x) - P_w(y = -1|x) \quad (12)$$

여기서 x 는 MCT 특징값, $P_w(y|x)$ 는 x 가 주어졌을 때 y 에 대한 확률이다. 즉, $P_w(y = 1|x)$ 는 x 가 특징값 γ 일 때 번호판 샘플의 히스토그램 값인 $g_t^P(\cdot, \gamma)$ 로 대체될 수 있으며, 마찬가지로 $P_w(y = -1|x)$ 또한 $g_t^{BG}(\mathbf{x}_t, \gamma)$ 로 대체될 수 있다. 식 (11)에 의해 표현되는 분류기의 룩업테이블은 그림 2과는 다른 그림 4와 같은 형식으로 표현된다. 그림 2는 룩업테이블 안의 값이 분류 결과인 1 또는 -1만 갖지만, 그림 4는 실수값을 가진다. 이 실수값은 분류 결과에 대한 신뢰도로 볼 수 있다. 값이 양수이면 1의 분류 결과인 번호판으로, 값이 음수이면 -1의 분류 결과인 배경으로 분류하고, 절대값이 클수록 분류에 대한 신뢰도가 높다고 할 수 있다. 약한 분류기가 선택되고 나면 훈련 샘플들의 웨이트를 업데이트 한다. 식 (8)에서 정의된 웨이트에 가법 모델의 업데이트를 적용하면 식 (13)의 업데이트 수식을 구할 수 있다.

$$\begin{aligned} w_{t+1}(x, y) &= \exp[-y\{H(x) + h(x)\}] \\ &= \exp(-yH(x)) \exp(-yh(x)) \\ &= w_t(x, y) \exp(-yh(x)) \end{aligned} \quad (13)$$

업데이트 된 웨이트를 노멀라이즈 하는 과정을 포함하면 표 2의 2.(e)와 같이 표현할 수 있고, 올바르게 분류된 샘플은 웨이트가 감소하고, 잘못 분류된 샘플은 웨이트가 증가하며, 신뢰도가 클수록 증감폭도 커진다. 웨이트 업데이트가 끝나면 다시 표 2의 2.(a)부터 2.(e)까지의 과정을 반복하며, T 라운드까지 종료되면 표 2의 3과 같이 픽셀 분류기를 조합한다. 적은 수의 유한개의 약한 분류기 후보군을 가지는 제한한 방법의 또 하나의 장점은 아다부스트 훈련 과정에서 약한 분류기를 선택하는 라운드의 수가 아무리 증가하여도 강한 분류기에 조합되는 약한 분류기의 수는 한정되어 있기 때문에 연산량이 복잡하지 않아 속도가 빠른 점이다. Haar-like 특징과의 비교를 예로 설명하면, Haar-like 특징을 이용한 방법은 수십만개의 약한 분류기의 후보군 중 T 라운드를 통해 선택된 약한 분류기를 조합하여 강한 분류기를 생성하기 때문에 식 (14)와 같이 강한 분류기의 연산이 표현되며, 라운드 횟수가 수만번 이하인 경우는 수십만개의 후보들 중 같은 인덱스의 약한 분류기가 선택될 가능성이 매우 적기 때문에 강한 분류기 연산을 위해서는 거의 T 번 만큼의 조합 연산량을 필요로 한다.

$$x = h_3^1(x) + h_7^2(x) + h_{108}^3(x) + \dots + h_{83}^T(x) \quad (14)$$

여기서 h 의 위첨자는 약한 분류기의 선택된 순번을 의미하며, 아래첨자는 후보군 중 선택된 약한 분류기의 인덱스를 의미한다. 식 (14)의 연산에 대해서는 위 첨자의 숫자가 중복될 가능성이 적기 때문에 $H(x)$ 의 연산을 위해서는 T 번 만큼의 조합 연산을 수행해야 한다. 그러나 MCT 특징을 이용한 방법은 수백개의 약한 분류기의 후보군에서 약한 분류기를 선택하기 때문에 라운드 횟수가 수천번을 넘어서면 중복 선택되는 약한 분류기가 발생하며, 이는 식 (15)와 같이 정리할 수 있게 된다.

$$\begin{aligned} H(x) &= h_1^1(x) + h_2^2(x) + h_3^3(x) + \dots + h_6^T(x) \\ &= \sum_{i \in G} h_1^i(x) + \sum_{i \in G_2} h_2^i(x) + \dots + \sum_{i \in G_n} h_n^i(x) \\ &= p_1(x) + p_2(x) + \dots + p_n(x) \end{aligned} \quad (15)$$

여기서 n 은 약한 분류기 후보의 개수를 의미하며, G 는 j 인덱스의 약한 분류기가 선택된 순번의 집합을 의미한다.

즉, 훈련 과정에서 같은 인덱스의 선택된 약한 분류기에 대해서는 미리 조합 연산이 가능하기 때문에 x)의 연산을 위해서는 n 번 만큼의 조합 연산만 수행하면 된다. 따라서 제안하는 방법에서는 표 2의 3과 같이 중복 선택된 픽셀 위치의 약한 분류기에 대해서 미리 조합 연산을 수행하여 픽셀 분류기를 만들고, 최종적으로 표 2의 결과인 강한 분류기를 생성한다. 강한 분류기의 값이 양수이면 번호판, 음수이면 배경으로 분류한다.

IV. 분류기를 이용한 번호판 검출

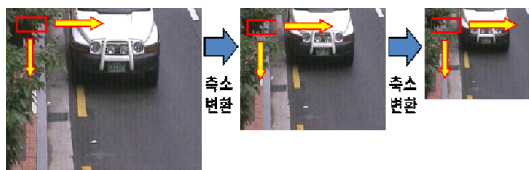


그림 5. 분류기를 이용한 번호판 검출 과정
Fig. 5. License Plate Detection using Trained Classifier

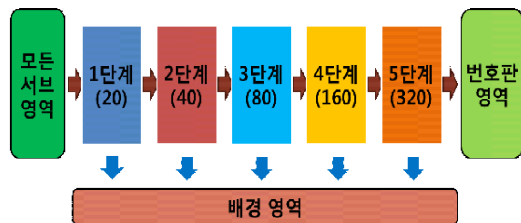


그림 6. 캐스캐이드 검출기 구조
Fig. 6. Detector of Cascade Structure

3장에서 설계된 훈련 샘플 크기의 분류기는 그림 5와 같이 입력된 테스트 영상의 서브 영역에 대해 스캐닝을 수행함으로써 번호판의 위치를 검출한다. 또한 다양한 크기의 번호판 검출을 위해 원본 영상을 축소 변환하면서 스캐닝을 수행한다. 검출 속도 향상을 위해 검출기의 구조는 그림 6과 같이 다섯 단계의 캐스캐이드 (cascade) 구조로 설계하였다. 각 단계의 픽셀 분류기 개수는 20, 40, 80, 160, 320개이다. 각 단계의 분류기는 앞 단계에서 번호판 영역으로 분류된 영역에 대해서만 분류하며, 모든 단계에서 번호판 영역으로 분류된 영역만이 최종 번호판 영역이 된다. 앞 단계의 검출기는 적은 수의 픽셀 분류기 조합만으로 대부분의 영역을 빠른 속도로 분류하며, 뒤 단계의 검출기는 많은 수의 픽셀 분류기 조합으로 보다 정밀하고 정확하게 분류한다. 실제 테스트 결과 약

20만 개 이상의 서브 영역에 대해 1단계에서 80% 이상 배경 영역으로 걸러졌으며, 5단계의 분류기는 50개 이하의 서브 영역에 대해서만 분류를 시행함을 확인하였다.

V. 실험 및 결과

본 논문에서 제안한 번호판 검출기 훈련과 테스트를 위해 도로와 주차장 등에 카메라를 설치하여 정면 각도의 차량 번호판이 담긴 영상을 획득하였다. 번호판 검출기의 각 단계에서 분류기 훈련을 위해 60x12 크기의 영상을 사용하였고, 개수는 번호판 훈련 영상 50,000장과 배경 훈련 영상 100,000장을 사용하였다. 배경 훈련 영상은 번호판을 포함하지 않은 임의의 영상에서 추출하여 사용하였고, 각 단계 마

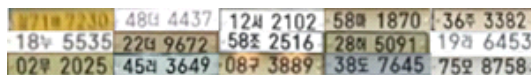


그림 7. 번호판 훈련 영상의 예
Fig. 7. Some Examples of License Plates

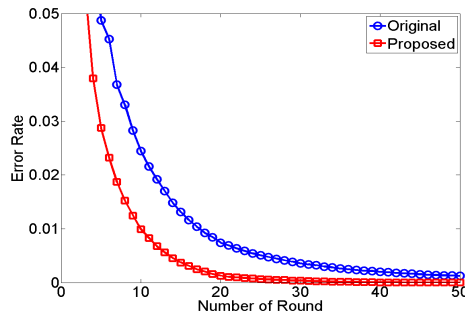


그림 8. 훈련 에러 비교 (20개 픽셀)
Fig. 8. Comparison of Training Error (20 pixels)

다 앞 단계에서 번호판으로 오분류되는 영상을 훈련 샘플로 사용하였다. 따라서 실제로 뒤 단계로 갈수록 훨씬 많은 배경 훈련 영상을 사용하였다고 볼 수 있다. 그림 7은 훈련에 사용된 번호판 영상의 샘플이다. 번호판 영상은 국내에 2006년 11월부터 시행된 일자형 흰색 또는 황색 바탕의 검정글씨 번호판을 대상으로 하였다.

그림 8에서 그림 11까지는 아다부스트 훈련 과정에서 약한 분류기를 선택하는 라운드 횟수에 따른 분류기의 성능을 나타낸 것이다. 분류기는 한 개의 단계로 구성하였으며, 번호판 훈련 샘플 50,000장과 배경 훈련 샘플 100,000장을 사용하였다. 보다 신빙성 있는 평가를 위해 샘플을 번호판 10,000장,

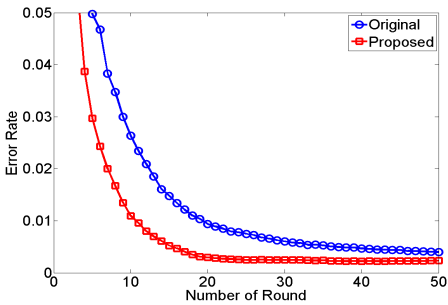


그림 9. 테스트 에러 비교 (20개 픽셀)
Fig. 9. Comparison of Test Error (20 pixels)

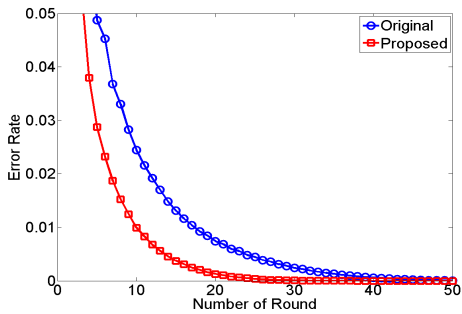


그림 10. 훈련 에러 비교 (320개 픽셀)
Fig. 10. Comparison of Training Error (320 pixels)

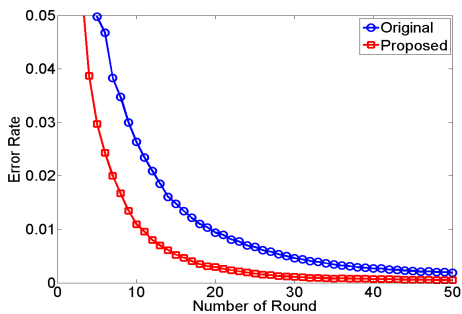


그림 11. 테스트 에러 비교 (320개 픽셀)
Fig. 11. Comparison of Test Error (320 pixels)



그림 12. 번호판 검출 결과 예시
Fig. 12. Examples of Detection Results

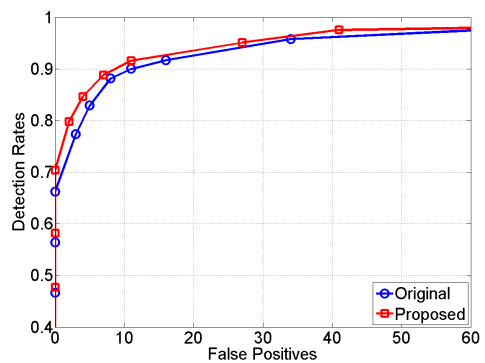


그림 13. 번호판 검출기의 ROC 커브
Fig. 13. ROC curves of Detection Results

배경 20,000장씩 5개의 군으로 나누어 교차검증 (cross validation)을 실시하였다. 즉, 4개 군의 번호판 40,000장과 배경 80,000장을 훈련 샘플로 사용하고 남은 1개 군의 번호판 10,000장과 배경 20,000장을 테스트 샘플로 사용하는 실험을 5번 번갈아 반복하여 평균 성능을 산출하였다.

그림 8은 강한 분류기에 조합되는 픽셀 분류기의 개수가 20개일 때 라운드 횟수에 따른 훈련 샘플에 대한 성능이다. x축은 라운드 횟수, y축은 훈련 에러를 의미한다. 그림을 통해

제한한 훈련 방법이 기존의 훈련 방법[16]에 비해 에러율이 빠르게 감소함을 확인할 수 있다. 기존의 훈련 방법은 230번째 라운드에서 에러율 0에 수렴하였고, 제안한 훈련 방법은 69번째 라운드에서 에러율 0에 수렴하였다. 그림 9는 픽셀 분류기의 개수가 20개일 때 라운드 횟수에 따른 테스트 샘플에 대한 성능이다. y축은 테스트 에러를 의미한다. 두 방법 모두 1000번의 라운드 내에서 에러율 0에 수렴하지는 않았지만 거의 에러율 0에 근접한 결과를 보였으며, 마찬가지로 제안한 훈련 방법이 기존의 훈련 방법 보다 에러율이 빠르게 감소함을 확인할 수 있다.

그림 10과 그림 11은 강한 분류기에 조합되는 픽셀 분류기의 개수가 320개 일 때 라운드 횟수에 따른 성능이다. 픽셀 개수가 20개 일 때와 마찬가지로 제안한 훈련 방법이 기존의 훈련 방법 보다 훈련 에러와 테스트 에러 모두 에러율이 빠르게 감소하였다. 다만, 픽셀 개수가 20개 일 때 보다는 분류기가 정교하기 때문에 훈련 에러가 기존의 훈련 방법은 61번째 라운드에서, 제안한 훈련 방법은 38번째 라운드에서 에러율 0에 수렴하였고, 테스트 에러는 두 방법 모두 1000번의 라운드 내에서 에러율 0에 수렴하지는 않았지만, 픽셀 개수가 20개 일 때 보다는 에러율이 빠르게 감소하였다.

번호판 검출기의 성능을 실제 환경에서 카메라를 통해 획득한 영상에 대해서 평가하기 위해 총 287개의 번호판을 포함하는 640x480 크기의 필드 영상 200장을 필드 테스트 영상으로 사용하였다. 그림 12는 필드 테스트 영상에 대한 번호판 검출 결과의 예이다. 4장에서 소개한 것과 같이 훈련된 분류기는 테스트 영상의 서브 영역을 스캐닝하여 번호판 영역을 찾으며 그림 12와 같은 빨간색 박스로 검출 결과를 나타낸다. 번호판 검출 속도는 2.8GHz CPU의 PC와 Visual C++ 프로그램을 이용하여 수행한 결과 640*480 영상에 대해 초당 15 프레임까지 처리가 가능하였고 320*240 영상에 대해서는 초당 60 프레임까지 처리가 가능함을 확인하였다.

본 실험에서는 제안한 방법의 효용성 증증을 위해 기존의 훈련 방법을 적용한 번호판 검출기와 제안한 훈련 방법을 적용한 번호판 검출기와의 성능 비교 평가를 수행하였다. 그림 13은 제안한 훈련 방법과 기존의 훈련 방법의 필드 테스트 영상에 대한 결과를 ROC 커브로 나타낸 것이다. x축은 오검출(False Positives)로서 200장의 필드 테스트 영상에서 번호판으로 판단되어 발생한 박스 중 번호판이 아닌 것의 개수를 의미하며, y축은 검출률(Detection Rates)로서 총 287개의 번호판 개수 중 올바르게 검출해낸 번호판 개수의 비율을 의미한다. 본 실험에서는 실제 번호판 영역의 80% 이상이 그려진 박스에 의해 포함되고, 그려진 박스 영역을 실제 번호판

영역이 50% 이상 차지할 때 올바르게 번호판을 검출해 냈다고 판단하였다. 그림 13의 실험 결과를 통해 제안한 번호판 검출 방법이 10개 내외의 오검출에 대해 90% 이상의 검출률을 가짐을 알 수 있다. 또한, 기존의 훈련 방법을 이용한 번호판 검출기 보다 제안한 훈련 방법을 이용한 번호판 검출기의 ROC 커브가 좌측 상단에 더 가깝게 그려지고 있기 때문에 제안한 훈련 방법을 이용한 번호판 검출기가 더 좋은 성능을 나타내는 것을 확인할 수 있다.

VI. 결론

본 논문에서는 록업테이블을 이용하여 MCT 특징에 대한 분류기를 설계하고 뉴턴 최적화를 통해 개선된 아다부스트 훈련 방법을 통한 번호판 검출 알고리즘을 제안하고 이를 이용한 번호판 검출기를 구현하였다. MCT 특징과 록업테이블을 활용함으로써 아다부스트 훈련 과정에서 강한 분류기를 조합하는 약한 분류기의 연산량을 감소시킴으로써 검출 속도를 향상시킬 수 있었다. 또한 개선된 아다부스트 훈련 방법을 통해 기존의 방법 보다 빠르게 에러율을 감소시킴으로써 훈련 시간도 단축할 수 있었다. 실험을 통해 제안한 훈련 방법을 이용한 번호판 검출 방법이 기존의 훈련 방법을 이용한 번호판 검출 방법 보다 더 좋은 성능을 보임을 확인하였고, 실시간 처리가 가능함을 확인하였다.

감사의 글

본 연구는 서울시 산학연 협력사업(WR080951)의 연구 결과로 수행되었습니다.

참고문헌

- [1] Younghyun Lee, Taeyup Song, Hanjun Kim, David K. Han and Hanseok Ko, "Hostile Intent and Behaviour Detection in Elevators", 4th International Conference on Imaging for Crime Detection and Prevention (ICDP-11), London, UK, November, 2011.
- [2] Michael A. Goodrich, Alan C. Schultz, "Human-robot interaction: a survey", Journal Foundations and Trends in Human-Computer

- Interaction, Vol.1, No.3, pp. 203-275, February, 2007.
- [3] Ionescu, D., Ionescu, B., Gadea, C., Islam, S., "A Multimodal Interaction Method that Combines Gestures and Physical Game Controllers", Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on , pp.1-6, August, 2011
- [4] Jong Sun Kim, Dong Hae Yeom, Young Hoon Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems", Consumer Electronics, IEEE Transactions on , vol.57, no.3, pp.1165-1170, August, 2011.
- [5] Kyu-Dae Ban, Jaeyeon Lee, DoHyung Kim, Jaehong Kim, and Yun Koo Chung, "Tiny and Blurred Face Alignment for Long Distance Face Recognition", ETRI Journal, vol.33, no.2, pp.251-258, April, 2011.
- [6] Hye-Jin Kim, Ho Sub Yoon, Jae Hong Kim, "User recognition based on continuous monitoring and tracking", Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on, pp.163-164, March, 2011.
- [7] Anwar Saeed, Ayoub Al-Hamadi, Michael Heuer, "Speaker Tracking Using Multi-modal Fusion Framework", Image and Signal Processing Lecture Notes in Computer Science, Volume 7340, pp 539-546, June, 2012.
- [8] Gang-Zeng Mao, Yi-Leh Wu, Maw-Kae Hor, Cheng-Yuan Tang, "Real-Time Hand Detection and Tracking against Complex Background", Intelligent Information Hiding and Multimedia Signal Processing, 5th International Conference on , pp.905-908, September, 2009.
- [9] Feifei Huo, Hendriks, E., Paclik, P., Oomes, A.H.J., , "Markerless human motion capture and pose recognition", Image Analysis for Multimedia Interactive Services, 10th Workshop on, pp.13-16, May, 2009.
- [10] Christos Nikolaos E. Anagnostopoulos, Ioannis E. Anagnostopoulos, Vassili Loumos, Eleftherios Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications", IEEE Transactions on Intelligent Transportation Systems Vol.7 No.3, September, 2006.
- [11] S.Z. Wang, H.M. Lee, "Detection and recognition of license plate characters with different appearances", in Proc. Conf. Intell. Transp. Syst., vol. 2, pp. 979 -984, October, 2003.
- [12] X. Shi, W. Zhao, Y. Shen, "Automatic license plate recognition system based on color image processing", in O. Gervasi et al. (Eds.), Lecture Notes on Computer Science, vol. 3483, Springer, pp. 1159 - 1168, May, 2005.
- [13] T.D. Duan, T.L. Hong Du, T.V. Phuoc, N.V. Hoang, "Building an automatic vehicle license plate recognition system", in Proc. Int. Conf. Comput. Sci. RIVF, pp. 59 -63, February, 2005.
- [14] H. Zhang, W. Jia, X. He and Q. Wu, "Learning-Based License Plate Detection Using Global and Local Features", IEEE International Conference on Pattern Recognition, 2006.
- [15] Gang Li, Ruili Zeng, Ling Lin, "Research on vehicle license plate location based on neural networks", in 1st International Conference on Innovative Computing, Information and Control, September, 2006.
- [16] B. Froba, and A. Ernst, "Face Detection with the Modified Census Transform", In Proceeding of the Sixth IEEE Conference on Automatic Face and Gesture Recognition, pp. 91-96, May, 2004.
- [17] T. Song, Y. Lee, M. Kim, B. Ku, and H. Ko, "Fusion Methods of License Plate Detection and Super Resolution for Improving License Plate Recognition", Journal of The Korea Society of Computer and Information, Vol. 16, No. 4, pp. 53-60, April, 2011.
- [18] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", Journal of Computer and System Sciences,

vol.55, pp.119-139, March, 1997.

[19] R. E. Schapire and Y. Singer, "Improved Boosting Algorithms using Confidence-rated Predictions", Machine Learning, vol.37, pp.297-336, December, 1999.

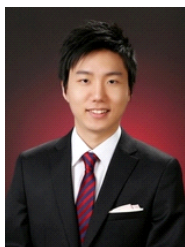
[20] J. Friedman, T. Hastie and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting", Annals of statistics, vol.28, no.2, pp.337-407, August, 2000.

[21] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Conference on Computer Vision and Pattern Recognition, vol.1, pp.511-518, December, 2001.

저 자 소 개



이 영 현
 2007: 고려대학교
 전기전자전파공학부 공학사
 현 재: 고려대학교
 영상정보처리학과 석박사통합과정
 관심분야: 신호처리, 체계검출, 영상인식
 Email : yhlee@ispl.korea.ac.kr



김 대 훈
 2011: 고려대학교
 전기전자전파공학부 공학사
 현 재: 고려대학교
 전기전자전파공학과 석박사통합
 과정
 관심분야: 영상신호처리
 Email : dhkim@ispl.korea.ac.kr



고 한 석
 1982: Carnegie-Mellon University
 전기공학 공학사
 1988: Johns Hopkins University
 전자공학 공학석사
 1992: Catholic Univ. of America
 전자공학 공학박사
 현 재: 고려대학교
 전기전자전파공학부 교수
 관심분야: 영상 및 음성 신호처리,
 패턴 인식
 Email : hsko@korea.ac.kr