

일반논문 (Regular Paper)

방송공학회논문지 제21권 제4호, 2016년 7월 (JBE Vol. 21, No. 4, July 2016)

<http://dx.doi.org/10.5909/JBE.2016.21.4.562>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

회전에 강인한 고속 이진패턴을 이용한 실시간 교통 신호 표지판 인식

황민철^{a)}, 고병철^{a)‡}, 남재열^{a)}

Real-time Traffic Sign Recognition using Rotation-invariant Fast Binary Patterns

Min-Chul Hwang^{a)}, Byoung Chul Ko^{a)‡}, and Jae-Yeal Nam^{a)}

요약

본 논문에서는 다양한 교통 표지판 중에서 운전자의 안전운행에 밀접하게 관계가 있는 속도 표지판을 인식하는 연구에 초점을 맞추고 있다. HOG (histogram of gradient)와 LBP (local binary patterns)가 객체 인식을 위한 대표적 특징이지만, 이러한 특징들은 패턴을 생성할 때 목표 객체의 회전을 고려하지 않음으로써 객체의 회전에 약한 특성을 가지고 있다. 따라서 본 논문에서는 회전에 강인한 이진 패턴을 생성하기 위해 FRIBP (fast rotation-invariant binary patterns)를 제안하고 있다. 본 논문에서 제안하는 FRIBP 알고리즘은 히스토그램에서 불필요한 레이어를 삭제하고 비교연산과 시프트 연산을 제거하여 빠르게 원하는 특징을 추출할 수 있도록 설계되었다. 제안된 FRIBP 알고리즘은 GTSRB (German Traffic Sign Recognition Benchmark) 데이터에 적용되어, 다른 비교 알고리즘과 유사한 성능을 보여주었다. 또한, 12,630개의 테스트 데이터에 대해 기존의 방법들보다 약 0.47초가 향상된 인식 속도를 보여주었다.

Abstract

In this paper, we focus on recognition of speed-limit signs among a few types of traffic signs because speed-limit sign is closely related to safe driving of drivers. Although histogram of oriented gradient (HOG) and local binary patterns (LBP) are representative features for object recognition, these features have a weakness with respect to rotation, in that it does not consider the rotation of the target object when generating patterns. Therefore, this paper propose the fast rotation-invariant binary patterns (FRIBP) algorithm to generate a binary pattern that is robust against rotation. The proposed FRIBP algorithm deletes an unused layer of the histogram, and eliminates the shift and comparison operations in order to quickly extract the desired feature. The proposed FRIBP algorithm is successfully applied to German Traffic Sign Recognition Benchmark (GTSRB) datasets, and the results show that the recognition capabilities of the proposed method are similar to those of other methods. Moreover, its recognition speed is considerably enhanced than related works as approximately 0.47second for 12,630 test data.

Keyword : Traffic sign recognition, rotation-invariant binary patterns, random forest, convolutional neural networks

a) 계명대학교 컴퓨터공학부(Dept. of Computer Engineering, Keimyung University)

‡ Corresponding Author : 고병철(Byoung Chul Ko)

E-mail: niceko@kmu.ac.kr

Tel: +82-53-580-5235

ORCID: <http://orcid.org/0000-0002-7284-0768>

※ 이 논문은 연구는 2016년도 계명대학교 계명스칼라연구기금과, 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (NRF-2011-0021780)의 지원을 받아 진행되었습니다.

· Manuscript received April 12, 2016; Revised May 27, 2016; Accepted July 4, 2016.

I. Introduction

With ongoing research in intelligent automobile systems, advanced driver assistance systems (ADASs) which constitute the technology underlying intelligent automobiles, has also attracted considerable attention. Although a driver can receive information regarding traffic signs from a navigation system using a geographic database and global positioning system information, pre-recorded data do not recognize any changes in situation along routes in real time. Hence, in order to control the speed of a smart car, the recognition of a traffic speed sign using camera is essential.

Many camera based traffic sign recognition algorithms using convolutional neural networks (CNNs) have recently been developed [1-5]. A CNN learns the multiple stages of invariant features using a combination of supervised and unsupervised learning. Although CNN-based method showed a higher classification rate based on the German Traffic Sign Recognition Benchmark (GTSRB) [10], a detailed algorithm for detecting traffic signs on a real road and incorporating the detection into the classifier in real time was not proposed [6]. Unlike CNN-based methods, Gim et al. [6] used a two-class boosted random forest with low-dimensional oriented center symmetric-local binary patterns by changing original local binary patterns (LBP).

Histogram of oriented gradient (HOG) and LBP are representative features for shape-based traffic sign recognition because they provide an effective way to capture shape information. For HOG, the image is divided into blocks and the directional histogram of the gradient value is obtained from each block. Although HOG is the most popular feature used for pattern recognition, its intensive computational demand is one of its drawbacks. LBP generates binary patterns by comparing the values of the surrounding pixels to those of the center pixel, and accumulates the data as histogram to extract the desired feature. However, this method compares only adjacent pixels, and has the dis-

advantage of considering information from a very small domain. Further, since it simply compares differences between pixel values, the extracted feature reflects pixel information with noises.

II. Feature Extraction and Traffic Sign Recognition

1. Fast Rotation-invariant Binary Patterns

Multi-scale block local binary patterns (MB-LBP) [7] compares the sum of the values of the center blocks with those of neighboring blocks. Compared to LBP, which compares pixel units, this method has the advantage of being resistant to noise. It can freely decide the block size according to image resolution or application. However, MB-LBP has a weakness with respect to rotation, in that it does not consider the rotation of the target area when generating patterns. To generate a binary pattern that is robust against rotation and, thus, compensates for the drawback of the MB-LBP, the rotation-invariant binary pattern (RIBP) [8] algorithm that is improvement over the MB-LBP, is proposed. As shown in Fig. 1, this algorithm finds the pattern with the minimal value from the numbers resulting from the application of rotation shift to the extracted binary pattern. It then uses this pattern instead of the original data. For example, if the binary pattern is “1000 0000”, this pattern has eight patterns based on rotation shift from left to right by one bit. From these eight patterns, this algorithm selects the minimum value—“0000 0001”—as the pattern. Therefore, unlike MB-LBP, this method is robust against image rotation.

Although the RIBP is invariant to image rotation, it needs more computational time for bit shifting, and some patterns have the same minimum values as when the total number of patterns is 256 shown in Fig. 1. Therefore, it

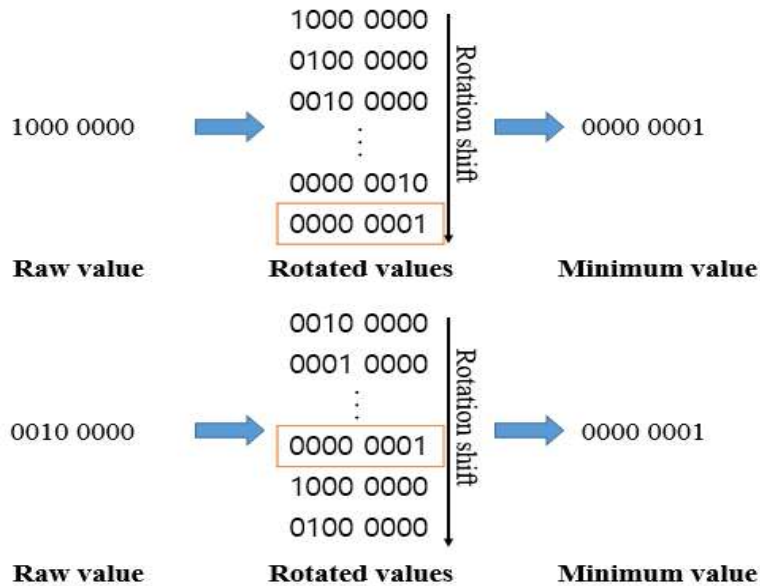


그림 1. 회전이동에 의한 최솟값 추정의 예
 Fig. 1. Example of minimum value estimation by rotation shift

is inefficient to use a shift operation and a comparison operation each time, merely to obtain the minimum value that can be obtained by rotational shift. Therefore, in this study, a fast rotation-invariant binary pattern (FRIBP) algorithm is proposed by deleting unused bins of the histogram, and eliminating the shift and comparison operations in order to

quickly extract the desired feature.

All possible cases of the minimum value are first investigated by the rotation shift of 256 patterns that can be expressed by eight-bit data. Among the 256 patterns, many patterns have the same minimum value after rotation shift as shown in Fig. 1. Therefore, we can express duplicate

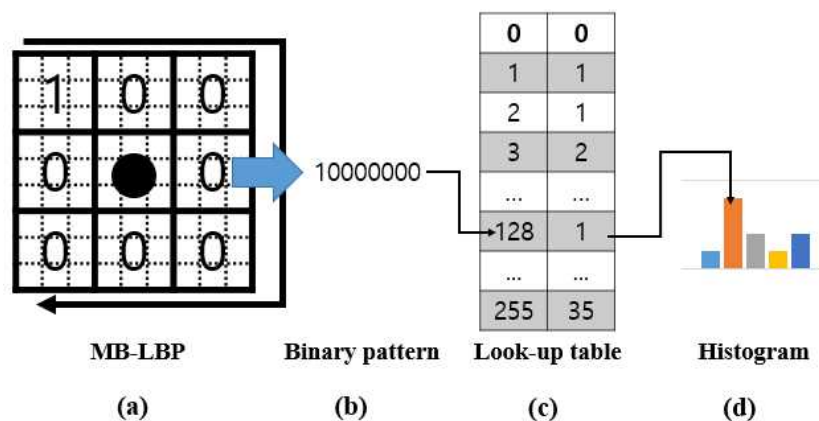


그림 2. 룩업 테이블을 이용한 특징 추출 (a) 비교 연산을 이용하여 계산된 이진 데이터 계산 (b) MB-LBP로 계산된 이진 패턴 (c) 룩업 테이블 (d) 최종 누적 히스토그램

Fig. 2. Feature extraction using the look-up table: (a) Binary data calculated using a comparison operation (b) Binary pattern from MB-LBP (c) Look-up table. (d) Final accumulated histogram

numbers as one number and finally, 256 patterns are reduced by 36 patterns. From the 36 minimum values obtained through this process, we can define a look-up table that maps the binary pattern to a minimum value pattern, as shown in Fig. 2. After the binary pattern is obtained using MB-LBP, the look-up table (see Fig. 2 (b)) is referenced to obtain the corresponding minimum value pattern, instead of trying to find the minimum value by rotation shift and comparison. By using the look-up table, 36-dimensional feature histogram is generated from a local patch region.

2. Traffic Sign Recognition using Random Forest

The CNN and support vector machine (SVM) are the most commonly used classifiers for traffic sign recognition. CNNs and SVM perform well but require many calculations. Thus, they incur the disadvantage of a long time to learn and classify. In particular, because the CNN classifier has the shortcoming of slow speed of classification, it needs to work through many convolutions and down-samplings to extract features. However, for the recognition of traffic signs, real-time recognition is very important, and the CNN classifier is thus not appropriate for the real-time recognition of traffic signs, such as in the case of advanced driver assistant system.

Therefore, we use the random forest [9] classifier that exhibits excellent classification performance and, at the same time, has the advantage of quick learning and classification speeds. Random forest is an ensemble classifier with a binary decision tree that makes it grow on the basis of a subset extracted randomly from the dataset. First, noise is removed from the test image using the preconditioning process to form a uniform image. Then, the feature vector of the processed image is extracted using the proposed FRIBP algorithm. The extracted feature vector is entered into the learned random forest classifier, and follows each

decision binary tree to the end node.

The final classification result F_c can be obtained by using the following equations:

$$P(c_i|L) = \frac{1}{T} \sum_{t=1}^T p(c_i|l_t) \quad (1)$$

$$F_c = \operatorname{argmax}_i P(c_i|L) \quad (2)$$

where, $L = l_1, l_2, \dots, l_T$ denotes the probability distribution obtained from each tree, and T represents the number of trees. From each tree l , the probability distribution $p(c_i|l_t)$ of c_i is obtained, and the highest probability c_i is generated as classification result F_c .

III. Results and Discussion

To assess the performance of the proposed algorithm, the German Traffic Sign Detection Benchmark (GTSRB) [10] dataset was used. The GTSRB contains 40 types of road signs, including eight types of speed signs. These eight types are 1) 20 km, 2) 30 km, 3) 50 km, 4) 60 km, 5) 70 km, 6) 80 km, 7) 100 km, and 8) 120 km. Each class contains 200 - 2,300 images for a total of 2,452 images, including non-speed sign classes in order to distinguish between speed signs and non-speed signs. The signs varied approximately from 20×20 pixels to 100×100 pixels in size. A PC with a GeForce GTX 760 Graphics Card, a 3.6-GHz i7 processor, and 16 GB of memory was used in the experiment. For recognition performance evaluation, we used the average precision rate. Among the benchmark algorithms, following five methods are used:

- Ground-Truth (GT) : classification results by two people through the naked eye
- Com. of CNN [2] : combining a CNN with eight lay-

ers

- Multi-scale CNN [11] : a CNN by feeding first stage and second stage features
- RIBP [5] : RIBP feature with random forest
- Random forest [12] : HOG feature with random forest

Since only precision results have been released for the GTSRB, we first compared performance based on precision, as shown in Fig. 3.

Figure 3. shows the performance evaluation of five methods. The proposed algorithm exhibited performance very close to the classification rate of GT (97.63%) as 97.31%. Although performance of the proposed algorithm was somewhat inferior to that of the top two CNN based algorithms [2,11] by 0.98% - 1.84%, two algorithms took two to three seconds per image in terms of processing time

because they involved a large number of calculations. On the contrary, the proposed algorithm took only 0.0008s per image, and is thus considerably more appropriate for real-time processing. In case of RIBP [5], the average precision is same with the proposed method because RIBP and FRIBP are using same feature histogram. Compared to the random forest [12] method with HOG features, the proposed FRIBP algorithm was less sensitive to the rotation of images and noise. Moreover, it showed impressive performance and exhibited performance improvement of approximately 0.001s because it used the look-up table to reduce the number of shift operations.

To assess the computation speed of the FRIBP feature extraction method, we compared it with the RIBP [8] feature extraction method. In this experiment, the same learning and test data were used for random forest. As shown

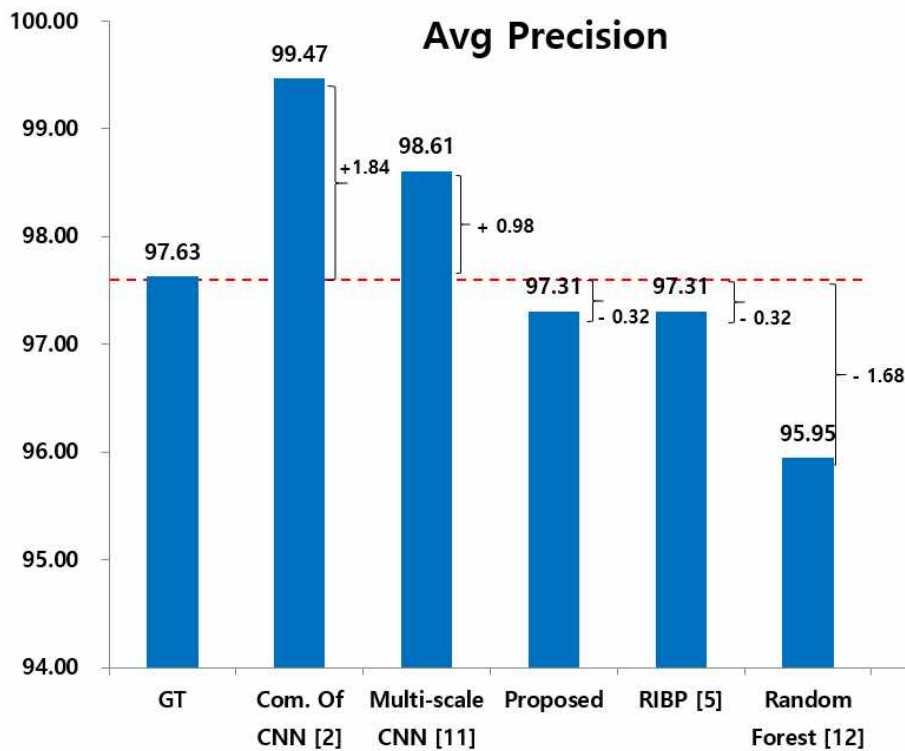


그림 3. 5개의 상위 알고리즘과 제안하는 방법, ground-truth (GT)의 분류 성능 비교

Fig. 3. Comparison of the classification performance of the five top algorithms, the proposed method, and GT

in Fig. 4, the proposed FRIBP showed speed enhancement of approximately 0.47s for 12,630 test images. The speed was enhanced because, as explained above, a shift calculation was not executed during the feature extraction process, but a look-up table was referenced to enter the minimal values right away.

IV. Conclusion

In this paper, we proposed a traffic sign recognition algorithm based on random forests that uses the proposed FRIBP feature. In experiments, the proposed feature exhibited strong classification performance and fast processing speed, superior to those of other algorithms in cases involving complex backgrounds, brightness changes, rotations, and many other variations.

For future works, we plan to apply FRIBP to Boosted random forest and recurrent convolutional neural network to improve the recognition accuracy. Furthermore, if FRIBP is applied to pedestrian detection that has the deformable pose changes, it probably shows better perform-

ance than HOG and LBP in term of accuracy and detection speed.

참 고 문 헌 (References)

- [1] J. Stallkamp, M. Schlipsing, J. Salmen and C. Igel, "The German traffic sign recognition benchmark: a multi-class classification competition," IEEE Int. Conf. Neural Networks, pp. 1453-1460, 2011.
- [2] D. Cireřan, U. Meier, J. Masci and J. Schmidhuber, "A committee of neural networks for traffic sign classification," IEEE Int. Conf. Neural Networks, pp. 1918-1921, Aug. 2011.
- [3] Y. Wu, Y. Liu, J. n Li, H. Liu, and X. Hu, "Traffic sign detection based on convolutional neural networks," IEEE Int. Conf. Neural Networks, pp.1-7, Aug. 2013.
- [4] Y. Zeng, X. Xu, Y. Fang, K. Zhao, "Traffic Sign Recognition Using Deep Convolutional Networks and Extreme Learning Machine," Lecture Notes in Computer Science, vol. 9242, pp. 272-280, 2015.
- [5] Y. Zeng, X. Xu, Y. Fang, and K. Zhao, "Traffic Sign Recognition Using Extreme Learning Classifier with Deep Convolutional Features," The 2015 Int. Conf. Intelligence Science and Big Data Engineering, pp. 1-10, June, 2015.
- [6] J. W. Gim, M. C. Hwang, B.C. Ko and J. Y. Nam, "Real-time Speed-Limit Sign Detection and Recognition using Spatial Pyramid Feature and Boosted Random Forest," 12th Int. Conf. Image Analysis and Recognition, pp.437-445, July, 2015.
- [7] S. Liao, X. Zhu, Z. Lei, L. Zhang and S.Z. Li, "Learning multi-scale block local binary patterns for face recognition," Journal of Biometrics, pp. 828-837, 2007.

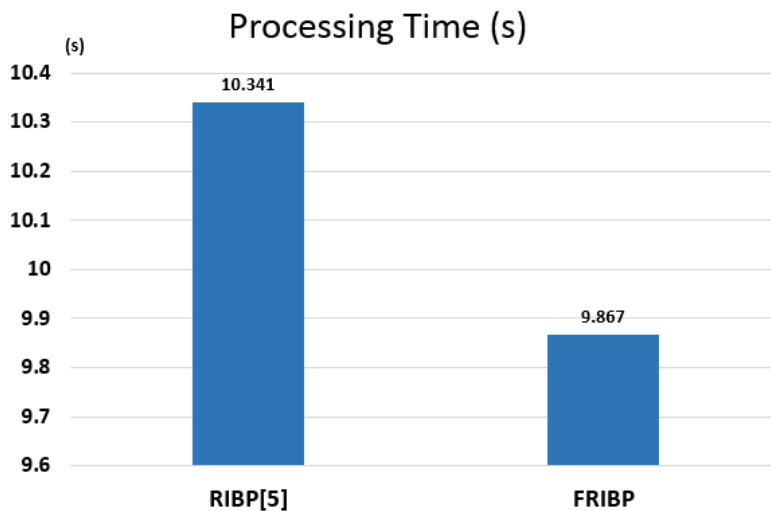


그림 4. 12,630장의 테스트 영상에 대해 기존의 RIBP 알고리즘 [8]과 제안하는 FRIBP 알고리즘간의 수행시간 비교

Fig. 4. Comparison in terms of operation time using 12,630 test images between the previous RIBP algorithm [8] and the proposed FRIBP algorithm

- [8] S. Yin, P. Ouyang, L. Liu, Y. Guo and S. Wei, "Fast traffic sign recognition with a rotation invariant binary pattern based feature," Journal of Sensors, vol. 15, pp. 2161-2180, 2015.
- [9] L. Breiman, "Random Forests," Machine Learning, vol. 45, pp. 5-32, 2001.
- [10] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," IEEE Int. Conf. Neural Networks, pp. 1-8, Aug. 2013.
- [11] P. Sermanet, Y. LeCun, "Traffic sign recognition with multi-scale convolutional networks," IEEE Int. Conf. Neural Networks, pp. 2809-2813, Aug. 2011.
- [12] F. Zaklouta, B. Stanculescu, and O. Hamdoun, "Traffic sign classification using k-d trees and random forests," IEEE Int. Conf. Neural Networks, pp. 2151-2155, Aug. 2011.

저 자 소 개



황 민 철

- 2014년 : 계명대학교 컴퓨터공학과 졸업
- 2014년 : 계명대학교 컴퓨터공학과 석사과정 입학
- 2016년 : 계명대학교 컴퓨터공학과 석사 졸업(공학석사)
- ORCID : <http://orcid.org/0000-0002-8239-8934>
- 주관심분야 : 컴퓨터비전, 패턴인식



고 병 철

- 1998년 : 경기대학교 전자계산학과 졸업(이학사)
- 2000년 : 연세대학교 대학원 컴퓨터과학과 졸업(공학석사)
- 2004년 : 연세대학교 대학원 컴퓨터과학과 졸업(공학박사)
- 2004년 3월 ~ 2005년 8월 : 삼성전자 통신연구소 책임연구원
- 2005년 9월 ~ 현재 : 계명대학교 컴퓨터공학과 교수
- ORCID : <http://orcid.org/0000-0002-7284-0768>
- 주관심분야 : 비전기반 ADAS, 비전기반 화재감지, 딥러닝, 의료영상처리



남 재 열

- 1983년 : 경북대학교 전자공학과 졸업(공학사)
- 1985년 : 경북대학교 대학원 전자공학(공학석사)
- 1991년 : University of Texas at Arlington 전기공학(공학박사)
- 1985년 5월 ~ 1987년 7월 : 한국전자통신연구소 연구원
- 1991년 9월 ~ 1995년 2월 : 한국전자통신연구소 선임연구원
- 1995년 3월 ~ 현재 : 계명대학교 컴퓨터공학과 교수
- ORCID : <http://orcid.org/0000-0001-7288-283X>
- 주관심분야 : 영상압축, 영상통신, 멀티미디어 시스템