Proceedings of the 33rd Chinese Control Conference
July 28-30, 2014, Nanjing, China

Traffic Sign Detection Based on Co-training Method

FANG Shengchao, XIN Le*, CHEN Yangzhou

College of Electronics Information and Control Engineering, Beijing University of Technology, Beijing 100124
E-mail: xinle@bjut.edu.cn (the corresponding author)

Abstract: To improve the performance of traffic sign detection and recognition systems in real implementation for the outdoor challenging environment, we propose a robust traffic sign detection algorithm based on co-training learning methods with a small number of manually labeled initial samples (opposite to collect all possible views) in this paper. With consideration on the various appearances of different traffic signs in real environment, two kinds of redundant textual descriptors are extracted for reinforcing the discrimination ability of traffic sign detection classifier from background. First, a novel traffic sign candidate regions extraction method is used based on probability map image built from multiple color-histogram back-projection. Secondly, a small number of labeled samples are used to train two classifiers respectively: one is AdaBoost with MB-LBP (multi-block local binary pattern) features and the other is SVM (support vector machines) with HOG (histograms of oriented gradients) features. Then, on the basis of co-training semi-supervised learning framework, the newly labeled samples with higher confidence from one classifier are used to update the training samples of the other one. Because of the constant increment of each training samples, the performance of traffic sign detection is highly improved which is evaluated intensively in the results of our experiment.

Key Words: Traffic sign detection, Co-training, MB-LBP feature, AdaBoost classifier, HOG feature, SVM classifier

1 Introduction

Visual driver assistance system, without doubt, should be one of the most major topics in the research area of intelligent vehicles, in which traffic sign detection and recognition are two of the most important functions considering the necessary processing modules. Generally speaking, traffic sign detection and recognition are implemented as two separated and closely connected processing modules. The detection of traffic sign accurately is the premise of recognition, and is the focus in our paper.

To address the problem of changing light, bad weather conditions, poor illumination, occlusion, etc., in traffic sign detection, many algorithms have had been proposed by researchers over the years, which have traditionally been divided into two kinds: color-based methods and shape-based methods. In consideration of the sensitivity of RGB space to lighting changes, a different relation between RGB components was used in [1], where the red component was used as a reference. Nevertheless, most researchers Maldonado[2], de la Escalera[3] have used HSI (Hue Saturation Intensity) family spaces. Because of the color information in these spaces are insensitive to lighting changes. But the main weakness of the above approaches lies in the fact that color information about the sign images tends to be unreliable. The illumination of the scene can vary considerably, depending on the time of day, weather conditions, shadows etc. And even though the initial thresholds used for segmentation have been set perfectly, they can't be adaptive for the whole outdoor situation of not long time duration.

Rather than color information, other algorithms used shape information. Xin [4] used two-step randomized Hough transform segmentation method. But many robust shape-based detectors such as Hough circle transform are slow to compute over large sizes of the sign images. Furthermore, parts of the traffic signs can be occluded, making the detection harder.

In recent years, some more complicated methods have been proposed for a high level of robustness of traffic sign detection in the actual implementation. Yuan et al. extract HOG features in gray image and then detect traffic signs with SVM[5], Haar wavelet features were used in [6]. However, these methods require a large number of training samples of traffic signs under various appearances to obtain higher performance. One of the solutions for this problem is to manually collect all possible views. However, traffic signs have various appearances caused by angle change, fading, different lighting conditions in the real environment. It is nearly impossible to manually collect all possible views. Greenhalgh [7] proposed a new solution to solve the problem by applying random distortions to generate large training sets. However, the training sets still can't reflect all kinds of traffic signs in real complicated environment.

To solve the problem above fundamentally, we propose a co-training method to improve the performance and accuracy rate of traffic sign detection efficiently. The proposed method consists of two steps: (a) candidate regions extraction, (b) construction of classifier based on co-training. In step (a), for the large set of manually classified various traffic signs, we grouped them into a series of different subsets with different illumination states under each color of interest (whether red, blue or yellow) in traffic system. When represented as color-histogram separately, the color information with different lighting conditions is considered explicitly. By back-projection with these multiple color-histograms, we can build a probability map insensitive to the interference. In this way, all traffic signs with various conditions can be considered consistently. By thresholding

* This work is supported by National Nature Science Foundation of China (61273006), National High Technology Research and Development Program of China (863Program)(2011AA110301), and Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No.20111103110017, 20111103120015).

the probability map appropriately, candidates of traffic sign regions can be extracted with less miss detection in protean illumination scenario. Next, in step (b), we extract MB-LBP features and HOG features from a small number of manually labeled initial samples (opposite to collect all possible views), two classifiers are trained consequently, which are AdaBoost based on MB-LBP features and the SVM based on HOG features respectively, so that both of them will have some essential identification ability at the beginning. Then, on the basis of co-training semi-supervised learning framework in the training sets, the new samples got from the above algorithms are added to mutual sample sets to increase the number of training samples. Because of the two features have the redundancy, the detected positive and negative samples will contain the images which were missed out or falsely detected mutually. Due to the increment in the number of samples, the robustness of the new re-training classifiers had been greatly improved so that the classifiers can detect the traffic signs more accurate. When for the final traffic sign detection, first, extract candidate regions in the current image based on (a), and then detect traffic signs in the candidate regions using either classifier constructed in (b).

2 Candidate Regions Extraction

2.1 Annotation and Grouping of Traffic Sign Training Samples

First, a large number of traffic signs were extracted manually from images containing traffic signs in real environment. Then, these samples were further clustered into a series of subsets according to the following two rules in turn: (1) the color of traffic signs (In this paper, we consider red, blue and yellow); (2) the illumination conditions in traffic sign samples (There are three conditions considered in this paper: normal light conditions; low light conditions - bad weather, dusk, shade, etc.; strong light conditions). Fig. 1 shows parts of sample images of traffic signs that already classified.

(a) (b) (c)

Fig. 1: Parts of sample images of traffic signs

2.2 Multiple Color-histograms

In accordance with the requirements for the classification of traffic sign samples, we can get nine class sample subsets. For each subset of traffic sign samples, we establish a color histogram to represent color probability distribution of traffic signs in the corresponding light conditions.

Based on the designed visibility of traffic signs with specific information, the Hue and Saturation channels in the specific type of traffic signs with one kind of dominant color have the same distributions, represented as the 2D histogram in our paper, no matter the condition when the photo is captured. The 2D histogram can be computed readily from the image in the training sets. The value in 2D histogram should be normalized to 0~255 according to (1):

$$f(b) = \mathrm{FLOOR}(binsVal(b) * 255 / \max Val) \quad (1)$$

where $f(b)$ is the value of corresponding bins after normalized, $binsVal(b)$ is the value of corresponding bins before normalized, $\max Val$ is the maximum value of all bins, FLOOR is to get the integer value that less than or equal itself.

2.3 Rectification of the Range of Color-histograms

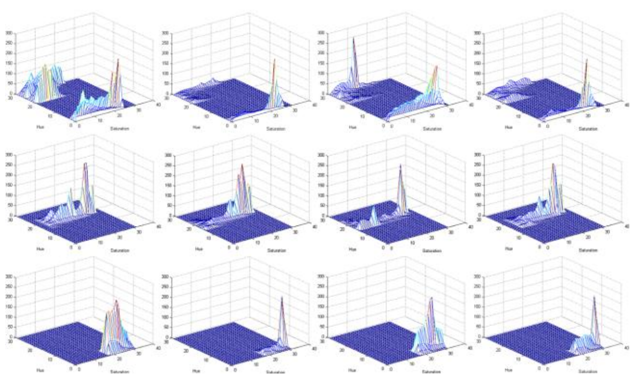
Because the hue and saturation information is unstable when the value of saturation is too small, the color in background will become great noise. So the 2D histogram of hue and saturation will be rectified further as:

$$\mathrm{Red}(i,j) = \begin{cases} \mathrm{True} & (H(i,j) \le 15 \text{ or } H(i,j) \ge 290) \\ & \text{and } S(i,j) \ge 10 \\ \mathrm{False} & \text{else} \end{cases} \quad (2)$$

$$\mathrm{Blue}(i,j) = \begin{cases} \mathrm{True} & 180 \le H(i,j) \le 280 \\ & \text{and } S(i,j) \ge 10 \\ \mathrm{False} & \text{else} \end{cases} \quad (3)$$

$$\mathrm{Yellow}(i,j) = \begin{cases} \mathrm{True} & 20 \le H(i,j) \le 65 \\ & \text{and } S(i,j) \ge 150 \\ \mathrm{False} & \text{else} \end{cases} \quad (4)$$

where $\mathrm{Red}(i,j)$, $\mathrm{Blue}(i,j)$, $\mathrm{Yellow}(i,j)$ represent the color label, red, blue and yellow, of a pixel in (i,j) respectively. $H(i,j)$, $S(i,j)$ represent a pixel's hue and saturation.



(a) (b) (c) (d)

Fig. 2: Histogram of each interested color under different lighting conditions

Fig. 2 shows the final trained color-histogram of each subset after the rectification. As shown in the first three columns (a)(b)(c) (each row correspond to the red, blue and yellow traffic signs in turn), the probabilistic distribution of all sorts of traffic signs under the various conditions of different illumination can be exhibited compactly. In order to visually compare the effects of partition based on lighting conditions, the histogram of all samples in one kind of dominant color (no matter what kind of light conditions) is also shown in (d). Obviously, the color-histograms of one kind of color in various lighting situations have a great difference, so the color-histogram of all samples in this specific color could not represent the divergence of the color probabilistic distribution of traffic signs explicitly in different light conditions.

2.4 Computation of Probability Map Images based on Multiple Color-histograms

For the current image waiting to be processed, the color value in each pixel location will be replaced with the corresponding probability in these 2D histograms, which is called as the image of color probability map. If the 2D histograms constructed from the trained images are really the distribution of the traffic sign, the region having the positive candidate sign will have a high value in the color probability map images.

In order to process the current image without knowing the color and light conditions of the pixel beforehand, by the back-projection with the trained multiple color-histograms, we can get the probability value of one pixel $P(i,j)$ from multiple color-histograms in different light conditions according to (5).

$$P(i,j)=\max(f_{ijk}(b)) \quad k=1\sim 9 \tag{5}$$

where $f_{ijk}(b)$ represent the value of corresponding bins after normalized for the pixel $P(i,j)$ in histogram k . By traversing all the pixels of the entire image, the probability map of the current image will be built.

2.5 Extraction of Sign Candidates as Regions of Interest (ROIs)

By thresholding the probability map appropriately, we can get the threshold image. To get the ROIs, First, use of a 3×3 median filter to remove small noises, and then connect the breakpoint with morphology closing algorithm, finally, eliminate the detected connected component regions which its width or height are less than 15 pixels or its area are less than 200 pixels. In Fig. 3, (a) is the original image, (b) is the probability map of the original image, (c) is the image after thresholding the probability map (the threshold is 55 in this paper), (e) is the image after median filter and morphology closing algorithm in image (c), (f) is the final ROIs, and (d) is an image after thresholding in the normalized RGB color space directly as mentioned in [8] (NRGB). As can be seen clearly from Fig. 3, the boundary of the detected traffic sign is clearer under the illumination changes in our proposed algorithm. However, the results used in [8] can appear discontinuous status, because this method cannot be effective obviously for the extraction of interested area with a serious variant of illumination condition, which will inevitably lead to the miss detection of traffic sign's candidates ultimately. Because of the use of multiple color-histograms trained from different lighting conditions explicitly, our method is more robust to lighting changes. Detailed analysis is described in the experimental result section.

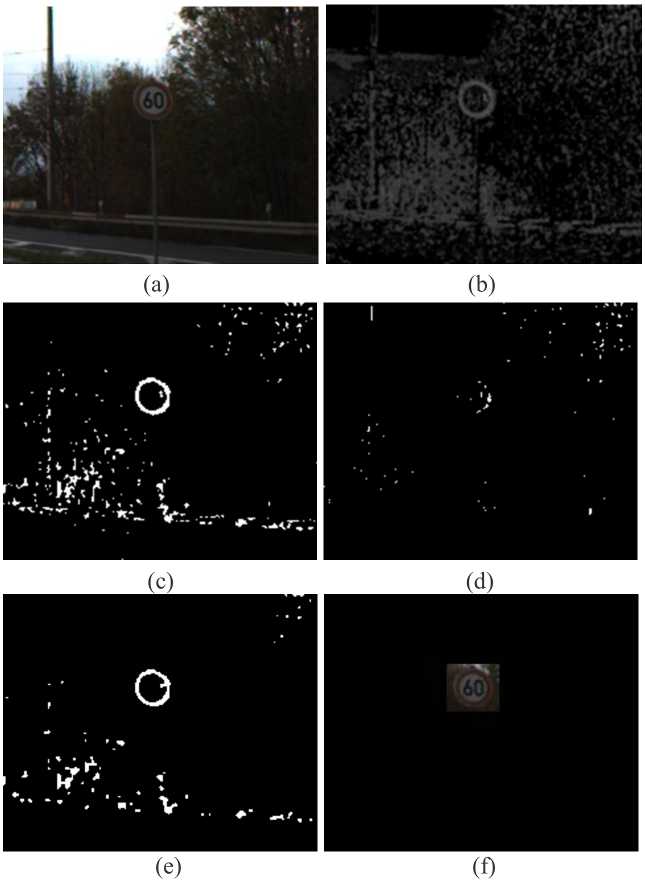


Fig. 3: ROIs extraction for the current sign image as the candidates

3 Construction of Classifiers Based on Co-training

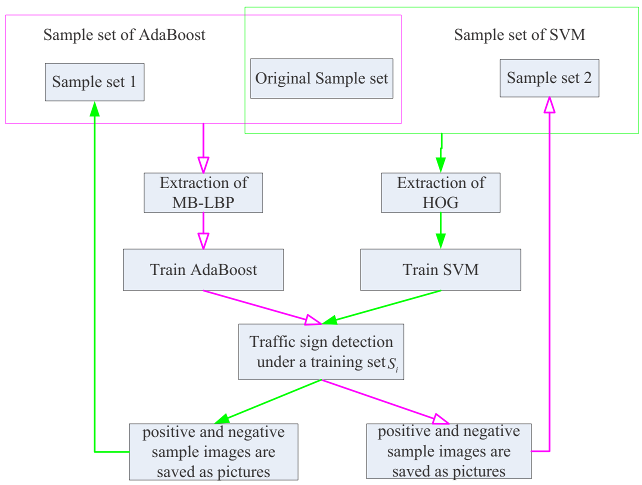


Fig. 4: Flow chart of construction of classifiers based on co-training method

In this paper, two kinds of features MB-LBP[9] and HOG[10] could satisfy the requirement of the redundancy. The characteristics of the redundancy need to satisfy the compatibility (the classifiers are able to identify the samples correctly by the two subsets of features) and the independence (a trait for the classification of the sample correctly or not that has nothing to do with the judgment for another features, and at least one of samples has different classification results). Therefore, MB-LBP features were taken to train AdaBoost classifier and HOG features were taken to train the SVM classifier from a small amount of samples, and then we can use two trained classifiers to supervise each other learning and improve the algorithm accuracy in training set S_i. In our paper, we use Gentle AdaBoost classifier and SVM classifier which have a detailed description in [11] and [12]. The specific process is shown in Fig. 4.

First, two classifiers are trained in each feature subset space according to the original sample set that have been labeled, so that both of them will have some identification ability. In our paper, the AdaBoost classifier is trained based on MB-LBP, and SVM based on HOG. And then, the unlabeled samples are classified with the two classifiers respectively, several samples chosen with higher confidence level will be add to another classifier's training set along with its category label to "teach" the classifier. Each classifier using a different classifier to add samples together with the original samples set to retrain. This process will be repeated until satisfy the end of the cycle number.

The co-training algorithm is described in detail below:

(1) Provide L labeled samples (include positive and negative samples) and some pictures contains unlabeled samples (training set S_i).

(2) Train AdaBoost based on MB-LBP from the L labeled samples. At the same time, train SVM based on HOG from the same L labeled samples.

(3) According to the following co-training learning training in each process cycle:

- Detection of traffic signs under the training set S_i, choose number m positive samples with the maximum confidence level and number n negative samples with the minimum confidence level;
- Add the positive and negative samples classified from unlabeled sample set based on one classifier to another's training sample set;
- Assemble the original samples and the new added samples, retrain AdaBoost and SVM.

(4) Process is repeated until cycle number equals to i.

4 Experimental Results

4.1 Dataset

In this paper, we choose 3553 positive samples and 5116 negative samples to train the initial AdaBoost and SVM. All samples are collected from in-vehicle camera images. The size of in-vehicle camera images are 640×480. The positive samples used to train AdaBoost with MB-LBP have been normalized to the same size 24×24. The positive samples used to train SVM with HOG have been normalized to the same size 40×40.1200 images are prepared for the iterative training, and each iteration use 300 images, the number of iterations is 4, so the training sets S_i are $S_1 \sim S_4$. We also prepared 500 images (containing at least one traffic sign) as test set to test our algorithm.

4.2 Results of Candidate Extraction

To test our algorithm's performance in candidates extraction, we use the test set introduced above. At last, 1215 traffic signs (In our paper, we concern prohibitory, danger and mandatory signs with red, blue and yellow color in China) are included. Through the data in Table 1, we can conclude that our proposed method can reduce the miss detection in candidates extraction of traffic signs, which means our method has a better robustness to the change of illumination.

Table 1: Results of candidates' extraction

Method	Traffic sign	Miss detection	Miss detection/%
Our	1215	49	4.03%
NRGB	1215	100	8.23%

4.3 Training Classifiers based on the Initial Samples and It's Experimental Results

The stages of AdaBoost is 20, max false alarm rate is 0.5, and min hit rate is 0.999. The kernel type is RBF (Radial Basis Function) in SVM. The minimum size of detection window is defined 15×15, and magnification is 1.2. Finally, the detected windows are merged around a traffic sign. Use the two initial classifiers respectively to detect traffic sign in the test set. The results are shown in Fig. 5 and Fig. 6, and top-left image is under strong light, bottom-left image is under low light, bottom-right image is under normal light, and in top-right image the red traffic sign is half occluded by shadow.

Fig. 5: Results of traffic sign detection with AdaBoost

Fig. 6: Results of traffic sign detection with SVM

4.4 Traffic Sign Detection Results based on Co-training

Let the AdaBoost classifier and the SVM classifier run in training sets $S_1 \sim S_4$ in turn based on co-training, positive and negative sample images that satisfy co-training's demand are saved as pictures. Because the candidate regions contain a lot of regions that don't belong to traffic signs, we set the maximum number of negative samples extracted are 3000. And in order to ensure the new labeled samples that added to the other sample set is correct, we choose the positive and negative sample images that have high confidence level, as for AdaBoost we choose the sample image as positive sample when a traffic sign is detected as traffic sign five times in different scales, and if in AdaBoost's first stage a sample is judged as negative sample, we save it as negative sample image; as for SVM if the discriminant function's value is greater than 2 in sample image, save it as positive sample image, if the discriminant function's value is less than -2 in sample image, save it as negative sample image. After each cycle of co-training, there are some new labeled sample images wrong classified, eliminate those wrong classified samples, and continue to run the next cycle.

Fig. 7: Results of traffic sign detection with AdaBoost after co-training

Fig. 8: Results of traffic sign detection with SVM after co-training

After the iterations, there are 1446 positive sample images and 3000 negative sample images extracted with AdaBoost, 1663 positive sample images and 3000 negative sample images extracted with SVM. Detect traffic signs with the retrained classifiers. The results after co-training are shown in Fig. 7 and Fig. 8.

Table 2 shows the Recall and Precision during co-training. Fig. 9，10 show the performance before and after co-training. Through the statistics and experimental effects, we can conclude that: due to the features that we choose have redundancy, after co-training based on MB-LBP and HOG, the performance of AdaBoost and SVM is highly improved. The Recall of AdaBoost is increased from 78.3 to 89.5, and SVM is from 81.1 to 91.3. The Precision of AdaBoost is increased from 85.4 to 90.1, and SVM is increased from 79.2 to 88.4.

Table 2: Results of the proposed method

	Initial/%	S_1/%	S_2/%	S_3/%	S_4/%
Recall(Ada)	78.3	82.5	86.1	88.4	89.5
Precision(Ada)	85.4	87.3	88.3	89.2	90.1
Recall(SVM)	81.1	84.4	88.5	90.0	91.3
Precision(SVM)	79.2	82.3	84.5	87.7	88.4

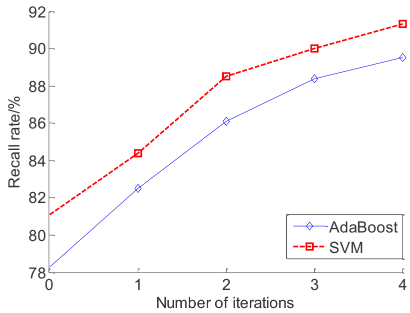


Fig. 9: Recall rate before and after co-training

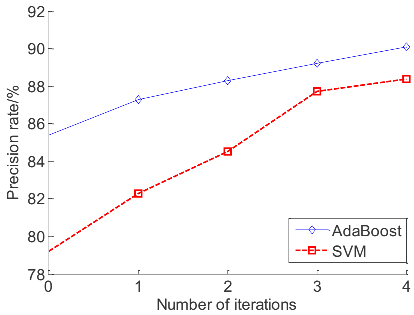


Fig. 10: Precision rate before and after co-training

5 Conclusion

In this paper, we proposed a method to improve the performance of traffic sign detection. In the candidates extraction stage, we proposed a novel method. Through building color-histograms under different lighting conditions, the method is more robust to lighting changes and can reduce the miss detection rate in comparison with the NRGB method. By the choice of two features (MB-LBP and HOG) which have redundancy and strong ability to distinguish traffic sign from background and based on co-training from semi-supervised learning methods with a small number of labeled samples to extract more samples to retain AdaBoost and SVM, experimental results showed that the proposed method improved the performance of traffic sign detection based on classifiers.

References

[1] A. de la Escalera, L.E. Moreno, M.A. Salichs, Road traffic sign detection and classification, *IEEE Transactions. on Industrial Electronics*, 44 (6): 848–859, 1997.

[2] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jiménez, Road-sign detection and recognition based on support vector machines, IEEE Transactions on. Intelligent Transportation Systems, 8(2): 264-278, 2007.

[3] A. de la Escalera, J.M. Armingol, M. Mata, Traffic sign recognition and analysis for intelligent vehicles, Image and vision computing, 21(3): 247-258, 2003.

[4] Z. Xin, Z. Ye, Q. Feng, The Segmentation of Ban Traffic Signs Based on Two-step Randomized Hough Transform, in proceedings of the 31st Chinese Control Conference, 2012:7274-7278.

[5] X. Yuan, L. Liu, C. Li, Unifying visual saliency with HOG feature learning for traffic sign detection, Proceedings of the IEEE Intelligent Vehicles Symposium, 2009: 24-29.

[6] L. Chen, Q. Li, M. LI, Traffic Sign Detection and Recognition for Intelligent Vehicle, 2011 IEEE Intelligent Vehicles Symposium (IV), 2011:908-913.

[7] J. Greenhalgh, M. Mirmehdi, Real-Time Detection and Recognition of Road Traffic Signs, IEEE Transactions on. Intelligent Transportation Systems, 13(4): 1498-1506, 2012.

[8] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jimenez, Goal evaluation of segmentation algorithms for traffic sign recognition, IEEE Transactions on. Intelligent Transportation Systems, 11(4): 917-930, 2010.

[9] L. Zhang, R. Chu, S. Xiang, Face detection based on multi-block LBP representation, Advances in Biometrics, International Conference, 2007:11-18.

[10] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:886-893.

[11] J. Friedman, T. Hastie, R. Tibshirani, Additive logistic regression: A statistical view of boosting (with discussion and a rejoinder by the authors), The Annals of Statistics, 28(2): 337–407, 2000.

[12] B. Boser, I. Guyon, V. Vapnik, A training algorithm for optimal margin classifier, In Proceedings of 5th ACM Workshop on Computational Learning Theory,1992: 144-152.