Locality-constrained Linear Coding for Vehicle Color Recognition

Rui Zhang[1, a], Chenghui Huang[2, b] and Mei Xie[1, c]

[1]School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 610000, China

[2] School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610000, China

[a]susu_eric@163.com, [b] 1242190863@qq.com, [c] mxie@uestc.edu.cn

Abstract

Automobile information recognition is a key component of intelligent transportation system which gains its popularity on the basis of a rapidly developing society. Color plays an important role in vehicle identification. Given that every vehicle has its own inner structure, and the outer natural environment deforms its main color to some extend in images, the toughest problem of vehicle color recognition is to pick out the region of interest (ROI), which means vehicle body, for identifying its main color. In this paper, we propose a method which combines Bag-of-Word (BoW) and Locality-constrained Linear Coding (LLC) for color recognition whilst preprocessing is introduced to offset the impact made by image degradation. Then, we train the classifier by support vector machine. The experiments are validated on the data set, which are gathering on express ways. The proposed method achieves good effect.

Keywords

Color Recognition, LLC, BoW.

1. Introduction

Automobile information recognition forms a key part of intelligent transportation system (ITS), which witnesses synchronized development with the rapidly developing society, and has received much attention in recent years due to an attached importance to city public security. As one of the conspicuous cues for vehicles, color can provide useful and profuse information in vehicle detection. [1]-[4]However, color recognition is still a strenuous challenge in images or videos according to the following difficulties.

1. Color can be susceptible and sensitive to the variation of a truly natural environment, such as illumination changes, haze, snow, which may cause a significant color deformation.

2. There is a large number of colors of automobile, and meanwhile, different parts of a single vehicle also have different colors. It is necessary to select the appropriate regions for color recognition.

3. The performance of color recognition is also limited by the low quality of images or videos. In a video, it is hard to segment a moving car and find out its main body accurately due to different position of cameras, noise, and overexposure.

Some state-of-art detection methods can provide a precise bounding box for each vehicle. [5]-[8] Thus, we apply the method to vehicle detection. Preprocessing must be introduced after locating vehicles positions, which is an essential step in color recognition, for getting a better quality image.

In image color recognition, the color features are collected from image patches, and then, a classifier is trained for color recognition. Wang et al. [9] proposed a novel way to recognize lip color based on multiple support vector machine to exclude invalid feature. Burghouts et al. [10] introduced an evaluation of local color invariants in the presence of realistic geometric transformations. Park and Kim [11] described an automatic time-saving method for comparing and distinguishing colors with two techniques of dimension reduction. Butzke et al. [12] discussed an approach for automatic detection of vehicles attributes. However, these methods above cannot be applied to vehicle color recognition regarding to the lack of region of interest selection.

Color feature extracting determines the behavior in the general image color recognition. Van de Sande et al. [13] analyzed prevalent color descriptors and their performance to illumination changes. Chen et al. [14] argued a Gaussian mixture model background subtraction algorithm turned to be effective. Li et al. [15] proposed a color normalization operator used in HIS color space. However, what mentioned before all failed to recognize the importance of scene and object which can bring gain in recognition accuracy to a color recognition system respectively.

The spatial information can be described by bag-of Feature (BoF) and bag-of-word (BoW)[16].Sivic and Zisserman originally introduced BoW [17]. The feature is encoded by a codebook, which is learned from a set of features by clustering algorithm. The BoF method disregards the information about the spatial layout of features, hence it is incapable of capturing shapes or locating an object. Comparing to BoF methods, BoW can achieve a high accuracy as seen in [18]-[19]. Inspired by Chen et al. [19], the Bow representation can quantize the color features via a large codebook, and make a classifier get easier access to differentiate them by mapping them into a higher dimensional space. Thus, we choose BoW in our task.

The toughest work for vehicle color recognition is to separate the main body of a car under a natural environment, because every car has its particular appearance and inner structure. What makes it worse is that every part of a single car may be in different colors incipiently, regardless of being transformed or painted when someone owned it, which would make the task more arduous. Fortunately, most cars have a single and dominant color in its body kept unchangeable, so it is a key step to divide the region of interest of a car for color recognition. Since BoW creates a large codebook, method in [19] may have somewhat deficiency in time consumption. Wang et al. [20] surveyed an efficient way of image classification, and propose Locality-constrained Linear Coding (LLC) to extract features. LLC showed its exceptional ability to differentiate images and also expressed that it was a time-saving strategy, even with very large codebooks, LLC can still process multi frames per second. Motivated by Wang, we adopt LLC into our task of vehicle color recognition.

In this paper, we first use BoW to establish a model for vehicle color recognition. BoW representation is an efficient way to discriminate different colors through a way of mapping the initial color features to a higher dimensional space via a large codebook, and the BoW based color feature reaches the better performance in vehicle color recognition. Although BoW has already been applied to ITS in [19], there are still many aspects worth improving. The strategies in ROI selection and feature extracting for object detection and vehicle color recognition are very diverse. It is hard to make a choice due to lacking of theoretical support. The best one is often picked from a huge number of experimental trials. So we introduce LLC, which performs well in image classification. In order to get a satisfactory recognition consequence, we choose ROI as our recognition object using LLC, which is verified that it also has equivalent performance in vehicle color recognition and enhances ability to reduce time consumption. LLC code can achieve an impressive image classification accuracy even with a linear SVM classifier [20]. We do experiments on the data set that is created by Chen et al. [19] for public use. Hence the primary contributions of this paper are:

Combine BoW and LLC which used in image classification, in ITS for vehicle recognition.

Compare some color extracting scheme and then find out a good way of color feature extracting in vehicle color recognition.

The rest of this paper is organized as follows. In Section 2, we discuss the details of our method, which includes the preprocessing and the most important part, image representation. Section 3 shows the experimental results on images and videos. Section 4 is the conclusion of this paper.

2. Approach to vehicle color recognition

As shown in Fig. 1 is the flow chart of our method.

2.1 Preprocessing

The quality of the images or videos taken by the monitors on express ways is usually very poor, which is ascribed to the changeable natural environment. The quality is often susceptible and sensitive to the impact of haze, strong light, and color shift caused by bad weather conditions or inappropriate configuration of cameras. Above all, the impact caused by haze is the most common demerit for getting good quality vehicle images. The poor quality is challengeable for color recognition. We adopt the haze removal method [21] as the preprocessing in our method to eliminate disadvantages caused by haze for getting a better image quality.

The results of the preprocessing are shown in Fig.2 (a) and Fig. (b), the original image which is under thick haze has a very low quality and definition. After the haze removal, the image is much clearer. From Fig.2 (b), it is easier to recognize its main color after preprocessing is gray.

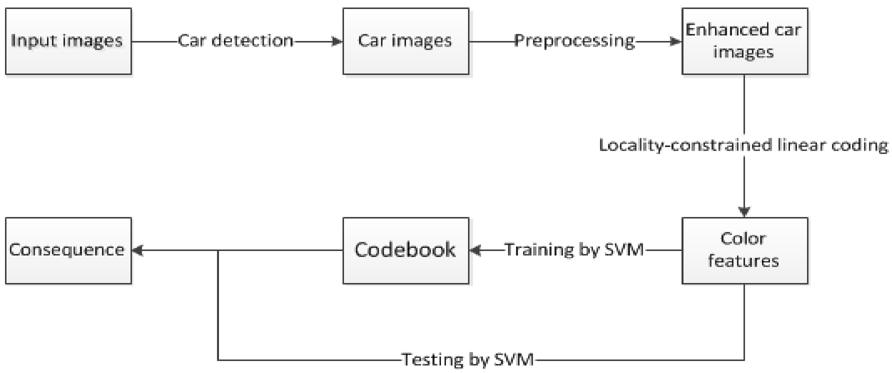

Fig. 1 Flowchart of our method

(a) (b)

Fig. 2 Preprocessing effect (a) Original image (b) after haze removal

2.2 Image color representation by locality-constrained linear coding

After preprocessing, image quality is promoted. The color features of images are extracted from the former images. As we mentioned before, feature extracting determined performance of the whole color recognition system. Kim et al. [22] proposed the optimal dimension of color feature of a vehicle as a color histogram in HSI color space. Chen et al. [19] discussed some useful ways of extracting strategies as well as Van de Sande et al. [13]. There are several histograms successfully used in color recognition, such as RGB histogram, transformed color histogram [13], opponent histogram [13], color moment [13], hue histogram [23], and normalized RGB histogram [24].We introduce LAB histogram [25] in this paper, for its good ability to extract color features. All of the former histograms have their own merits. One of the effective approach is to combine all or most of them, but meanwhile the more histograms use, the more time cost. Considering their own advantages, and the impetus for discovering a time-saving solution on the basis of a satisfactory accuracy, we do a series of trials and find a good way eventually. Thus, we propose a combination with two simple histograms, LAB histogram and normalized RGB histogram, which behave well in experiments.

In our method, the codebook B is learnt from the patch features in training data set by K-means clustering method, according to [26] [27]. We take the representation of LLC [20] to describe the image. Let X be a set of D-dimensional local descriptor extracted from an image, i.e. $X=[X_1, X_2, \ldots\ldots, X_N] \epsilon R^{D\times N}$, Given a codebook with M entries, $B=[b_1, b_2, \ldots\ldots, b_M] \epsilon R^{D\times M}$, different coding schemes convert each descriptor into a M-dimensional code to generate the final image representation.

LLC incorporates locality constraint instead of the sparsity constraint. LLC code uses the following criteria:

$$\min_{C} \sum_{i=1}^{N} \|X_i - Bc_i\|^2 + \lambda \|d_i \otimes c_i\|^2 \quad \text{s.t.} 1^{\mathsf{T}} c_i = 1, \forall i. \tag{1}$$

where $\otimes$ denotes the element-wise multiplication, and $d_i \in R^M$ is the locality adaptor that gives different freedom for each basis vector proportional to its similarity to the input descriptor X_i. Specifically,

$$d_i = \exp\left(\frac{dist(X_i, B)}{\sigma}\right) \tag{2}$$

where $dist(X_i, B) = [\mathrm{dist}(X_i, b_1), \ldots\ldots, \mathrm{dist}(X_i, b_M)]^T$, and $\mathrm{dist}(X_i, b_j)$ is the Euclidean distance between X_i and b_j. σ is used for adjusting the weight decay speed for the locality adaptor. d_i is normalized to be between (0,1] by subtracting $\max(dist(X_i, B))$ from $dist(X_i, B)$. The constraint $1^{\mathsf{T}} c_i=1$ follows the shift-invariant requirements of the LLC code. Note that the LLC code in Eq. (1) is not sparse in the sense of ℓ^0 norm, but is sparse in the sense that the solution only has few significant values, that suggests the LLC can be modified to accelerate its encoding speed. According to [20], we can use the K(K<D<M) nearest neighbors of X_i as the local bases B_i. Specially, the faster approximation of LLC code uses the following criteria:

$$\min_{\tilde{C}} \sum_{i=1}^{N} \|X_i - \tilde{c}_i B_i\|^2 \quad \text{s.t.} 1^{\mathsf{T}} c_i = 1, \forall i. \tag{3}$$

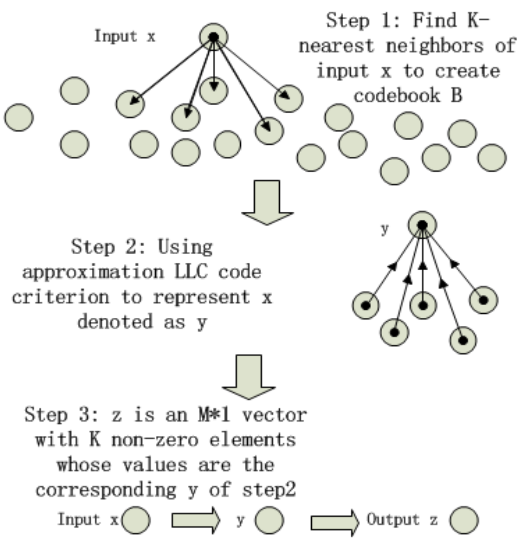

Fig. 3 LLC coding processing

This reduces the computation complexity from $\mathcal{O}(\mathrm{M}^2)to\mathcal{O}(\mathrm{M}+\mathrm{K}^2)$, where K<<M. As K is usually very small, solving Eq. (3) is very fast. For questing for K-nearest neighbors, K-NN search strategy is applied. LLC coding processing is shown in Fig.3.

3. Experiment

The data set contains 15601images of various categories such as truck, limousine, and SUV. The whole data set is classified into 8 classes, including black, blue, cyan, gray, green, red, white, and yellow. All the images in the data set are taken in the frontal view captured by high-definition cameras on express ways. We use SVM for color classification. The number of images with different kind in our data set is listed in Table 1. Some samples in the data set are shown in Fig.4.

In our experiment, the data set is randomly divided into two groups, half for training and the rest for testing. Each experiment is carried out for 5 times, and the final recognition rate is the average of all runs. Specially, we resize all the images into 300*300 for both training and testing. We do experiments on HSV histogram [23], and normalized RGB histogram [24], and LAB histogram [25]. Every histogram is with a dimension of 48 (16 per channel)， and each channel is normalized independently. For every channel, it complies with the distribution where the mean μ=0 and the standard deviation $\sigma = 1$. Local features are computed on an image patch with a grid of size 24*24 and a stride of 8 pixels. The codebook is fixed to 1024 in our experiments.

Table 1 The number of images in the data set

Color	Black	Blue	Cyan	Gray	Green	Red	White	Yellow
Number	3442	1086	281	3046	482	1941	4742	581

Fig. 4 Samples of different colors in the data set

First we use HSV histogram [23], and normalized RGB histogram [24], and LAB histogram [25] respectively. The consequences are shown in Table 2. We find out that both normalized RGB histogram [24] and LAB histogram [25] perform better than HSV histogram, and meanwhile they show a similar performance in color recognition. Considering what we mentioned before, we combine them and experiment again. We unfold that with a combination of normalized RGB histogram [24] and LAB histogram [25], we can get a better effect. The main reason for such promotion is that LAB histogram [25] is able to constrict illumination changes and independent to pigment, which makes up for the demerits of RGB histogram [24]. The result can be seen in the last row of Table 2. In all experiments, we fasten K=5 in LLC coding algorithm, and linear SVM is adopted in all the experiments. What is obvious in Table 2 is that among all the 8 colors, the recognition rates for green and gray are less than the others and the recognition rate for red is always beyond 0.95. Given that red is the most noticeable color in color space, for which red means stop in transportation, we can easily accept the result. AS green and gray are less unified, their features are more diffused and hard to be classified.

Table 2 Average recognition performance of different color descriptors

	Black	Blue	Cyan	Gray	Green	Red	White	yellow
RGB	0.9479	0.8903	0.8862	0.7880	0.7329	0.9799	0.9366	0.9213
HSV	0.9174	0.8743	0.7734	0.6982	0.6136	0.9625	0.8922	0.8677
LAB	0.9203	0.8606	0.8825	0.7442	0.7297	0.9709	0.9129	0.9029
RGB+LAB	0.9538	0.9164	0.9094	0.8066	0.7596	0.9878	0.9389	0.9385

As shown in Table 2, normalized RGB histogram achieves the best performance. There are some extant methods and algorithms which use normalized RGB histogram in vehicle color recognition. Thus we choose this histogram to make comparison experiments to exemplify our method is practical.We do some experiments in our computers and fix the dimension of every histogram into 48(16 per channel) to provide a fair comparison. We list the recognition rates in Table 3.

Table 3 Comparison with other methods using normalized RGB histogram

	Black	Blue	Cyan	Gray	Green	Red	White	Yellow
Original Local Features	0.6446	0.2276	0.7464	0.1216	0.6788	0.8110	0.7819	0.6597
BoW [16]	0.7737	0.7209	0.7623	0.5450	0.6149	0.9371	0.7994	0.8160
BoW +SPM [16]	0.9397	0.8323	0.7971	0.7400	0.5013	0.9638	0.9323	0.8609
BoW+FC [19]	0.9213	0.8414	0.9594	0.7837	0.6473	0.9601	0.9097	0.8646
Our Method(RGB)	0.9479	0.8903	0.8862	0.7880	0.7329	0.9799	0.9366	0.9213
Our Method(RGB+LAB)	0.9538	0.9164	0.9094	0.8066	0.7596	0.9878	0.9389	0.9385

What we propose in this paper gets an average recognition rate of 0.8853 (using RGB only) and 0.9013 (RGB+LAB). It takes 0.376 seconds (RGB+LAB) and 0.251 seconds (RBG only) per image for vehicle detecting and recognition. And the dimension of color features is 96(RGB+LAB) and 48 (RGB only). Table 4 is a conclusion for these two feature extracting methods.

Table 4 Comparison with RGB and RGB+LAB

	Accuracy	Time-consumption	Dimension
RGB	0.8853	0.251	48
RGB+LAB	0.9013	0.376	96

Although our method of RGB+LAB surpasses the way of using RGB only in color recognition rate at a superiority about 0.0260, the latter only consumes nearly 67% time of the former costs,so it is hard to say which one is better. Those two methods all have its own advantages and can be used under different content. The larger the data set is, the more useful our method of RGB is. However, in most cases, method of RGB+LAB is better for its better recognition rate.

4. Conclusion

In this paper, an effective method for vehicle color recognition has been proposed. We combine the BoW representation and Locality-constrained Linear Coding into vehicle color recognition whilst invesgate two good approaches to color feature extracting. The experiment on image data set shows that our method is effective. Our future work might be finding some other features to promote our method further more.

References

[1] T. Xiong and C. Debrunner:Stochastic car tracking with line- and colorbased features, in Proc. IEEE Trans. Intell. Transp. Syst., Oct. 2003,vol. 2, pp. 999–1003.
[2] Q. Li, N. Zheng, and H. Cheng: Springrobot: A prototype autonomous vehicle and its algorithms for lane detection,IEEE Trans. Intell. Transp.Syst., vol. 5, no. 4, pp. 300–308, Dec. 2004.
[3] N. Wu, F. Chu, S. Mammar, and M. Zhou:Petri net modeling of the cooperation behavior of a driver and a copilot in an advanced driving assistance system, IEEE Trans. Intell. Transp. Syst., vol. 12, no. 4, pp. 977–989,Dec. 2011.
[4] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez: Adaptive multicue background subtraction for robust vehicle counting and classification, IEEE Trans. Intell. Transp. Syst., vol. 13,no. 2, pp. 527–540, Jun. 2012.
[5] P. Viola and M. Jones: Rapid object detection using a boosted cascade of simple features, in Proc. IEEE Comput. Soc. Conf. Comput. Vis. PatternRecog., 2001, pp. 511–518.
[6] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan: Object detection with discriminatively trained part-based models, IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 8, pp. 1627–1645,Sep. 2010.
[7] P. F. Felzenszwalb, R. B. Girshick, and D. McAllester: Cascade object detection with deformable part models," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog., San Francisco, CA, USA, Jun. 2010,pp. 2241–2248.
[8] P. Dollár, Z. Tu, P. Perona, and S. Belongie: Integral channel features, in Proc. Conf. Brit. Mach. Vis. Conf., 2009, pp. 91.1–91.11.
[9] J. Wang, X. Li, H. Fan, and F. Li: Classification of lip color based on multiple SVM-RFE, in Proc. IEEE Int. Conf. BIBMW, Nov. 2011,pp. 769–772.
[10] Gertjan J. Burghouts, Jan-Mark Geusebroek: Performance evaluation of local colour invariants in Computer Vision and Image Understanding 113,2009,pp.48-62.
[11] S. M. Park and K. J. Kim: Color recognition with compact color features, in t. J. Commun. Syst., vol. 25, no. 6, pp. 749–762, Jun. 2012.
[12] Miriam Butzke, Alexandre G. Silva, Marcelo da S. Hounsell, Mauricio A. Pillon: Automatic Recognition of Vehicle Attributes-Color Classification and Logo Segmentation in Hifen, Urunguaiana,2008.
[13] K. E. Van De Sande, T. Gevers, and C. G. Snoek: Evaluating color descriptors for object and scene recognition, IEEE Trans. Pattern Anal.Mach. Intell., vol. 32, no. 9, pp. 1582–1596, Sep. 2010.
[14] Zezhi Chen,Nick Pears, Michael Freeman and Jim Austin: Road Vehicle Classification using Support Vector Machine, in IEEE Sep.2009
[15] Xiuzhi Li ,Guangming Zhang, Jing Fang, Jian Wu, Zhiming Cui: Vehicle Color Recognition Using Vector Matching of Template, in IEEE Oct.2010.
[16] S. Lazebnik, C. Schmid, and J. Ponce: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, in Proc.IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog., 2006, vol. 2.pp. 2169–2178.
[17] J. Sivic and A. Zisserman: Video Google: A text retrieval approach to object matching in videos,in Proc. IEEE Int. Conf. Comput. Vis., Nice,France, Oct. 2003, pp. 1470–1477.
[18] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong: Localityconstrained linear coding for image classification, in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog., 2010, pp. 3360–3367.
[19] Pan Chen, Xiang Bai, and Wenyu Liu: Vehicle Color Recognition on Urban Road by Feature Context, IEEE Trans. Intell. Transp. Syst., Feb,2014.
[20] Jinjun Wang, Jianchao Yang, Kai Yu, Feng jun Lv, Thomas Huang, and Yihong Gong: Locality-constrained Linear Coding for Image Classification, Proc. of CVPR 2010.
[21] K. He, J. Sun, and X. Tang: Single image haze removal using dark channel prior, in Proc. IEEE Conf. CVPR, 2009, pp. 1956–1963.

[22] Ku-Jin Kim, Sun-Mi Park, Yoo-Joo Choi: Deciding the Number of Color Histogram Bins for Vehicle Color Recognition, in IEEE Asia-Pacific Services Computing Conference,2008.
[23] J. Van De Weijer, T. Gevers, and A. D. Bagdanov: Boosting color saliency in image feature detection, IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 1, pp. 150–156, Jan. 2006.
[24] T. Gevers, J. Van De Weijer, and H. Stokman: Color feature detection, in Color Image Processing: Methods and Applications, R. Lukac and K. N. Plataniotis, Eds. Boca Raton, FL, USA: CRC Press, 2007,pp. 203–226.
[25] Adrian Ciobanu, Ioan Pavaloi, Mihaela Luca, Elena Musca: Color Feature Vectors Based on Optimal LAB Histogram Bins, in ICDAS,May,2014.
[26] J. Yang, K. Yu, Y. Gong, and T. Huang: Linear spatial pyramid matching using sparse coding for image classification,in Proc. IEEE Comput. Soc.Conf. Comput. Vis. Pattern Recog., 2009, pp. 1794-1801.
[27] K. Yu, T. Zhang, and Y. Gong: Nonlinear learning using local coordinate coding, in Proc. Adv. Neural Inf. Process. Syst., 2009, vol. 22,pp. 2223–2231.