TRAFFIC SIGN RECOGNITION
USING SCALE INVARIANT FEATURE TRANSFORM AND SVM

Xiaoguang HU[a, *], Xinyan ZHU[a] , Deren LI[a], Hui LI [b]

[a] State key Laboratory of Information Engineering in Surveying,Mapping and Remote Sensing,Wuhan ,Hubei
[b] China University of Geosciences, Wuhan,Hubei

KEY WORDS: Traffic sign recognition, SIFT, SVM, Bag-of-words

ABSTRACT:

TSR (Traffic sign recognition) has been studied for realizing drivers assisting system and automated navigation and is an important studied field in ITS (Intelligent traffic system). In this paper, a recognition method of traffic signs separated from real image was studied. Images were divided into several categories according to the actual weather, distance and angle of view etc. SIFT was firstly used to detect keypoints and describe them because the SIFT(Scale Invariant Feature Transform) features were invariant to image scale and rotation and were robust to changes in the viewpoint and illumination. And then the Bag-of-words method was applied to recognition. Finally, the SVM(Support vector machine) approach was used in classification. As shown by the experiment results, the proposed method is effective.

1. INTRODUCTION

With the development of society and economy, the traffic problems become more and more serious, which become one of the bottle-neck of modern city. Traffic accidents are paid much attention by the governments of the world. Under this conditions, ITS (Intelligent Traffic System) is in need. Traffic sign is important information of the road traffic and its detection and recognition has attracted much attention in last decade because of its importance. It involves the technologies of pattern recognition, digital image processing, artificial intelligence, computer vision and so no. However, traffic sign recognition is still an unsolved problem as it can be changed easily depending on its relative location and angle of view against camera and surrounding condition such as weather and daytime. At present, some research has already been carried out:

L. Estevez and N. Kehtarnavaz(Kehtarnavaz 1996)presents the development and realtime implementation of an algorithm capable of recognizing stop, yield, and do-not-enter traffic warning signs. It consists of six modules: color segmentation, edge localization, RGB differencing，edge detection，histogram extraction, and classification，After the histogram of the image has been extracted, this histogram is analyzed to determine which if any sign of interest is present in the image. But this method is very sensitive to the integrity of edge and sufficiency of light. The recognition rate will greatly decreased in low resolution and poor light conditions.A de la Escalera(A de la Escalera 1997) firstly recognized sign based on their color and shape, and used neural network to classify sign.S.H.Hsu,C.L.Hang(S.H.Hsu 2001) apply the MP filters to recognize triangular and circular road sign by scanning every particular MP（matching pursuit）filter across

an image. S. D. Zhu, Y. Zhang, X. F. Lu(S. D. Zhu 2005) firstly convert RGB color space into HIS color space, and extract pixels with red color. Then perform LOG mask operation on the pixels got from step 1, for the detection of edges. By using neural network, we determine the angle pixels, and at the same time, we also determine on which specific angle the pixel is. And finally the author detected the traffic sign by using shape information.Maldonado-Bascon, S.Lafuente-Arroyo, S.etc (Maldonado-Bascon, Lafuente-Arroyo et al. 2007) used Gaussian-Kernel SVMs for the content recognition.Merve Can Kuş, Muhittin Gökmen, Şima Etaner-Uyar(Merve Can Kuş 2008)used a scale invariant feature transform and color classification to recognize the traffic signs.S. Maldonado Bascón, J. Acevedo Rodríguez etc (S. Maldonado Bascón 2009) proposed pre-processing methods and improvements in support vector machines to increase the accuracy.In order to improve the speed of sign recognition, some researchers have begun to try to improve the hardware performance to solve this problem.Matthias Muller, Axel Braun,Joachim Gerlach etc(Matthias Muller 2010) proposed an implementation based on multicore chip.

The remaining parts of this paper are organized as follows: in section 2, the SIFT method and a new traffic signs recognition strategy are described and in section 3 experiment on traffic signs recognition are presented followed by some discussion and conclusion in section 4.

---

* Xiaoguang HU , Ph.D candidate, majors in pattern recognition,GIS and computer vision etc. Michael.hu.07@gmail.com.

A special joint symposium of ISPRS Technical Commission IV & AutoCarto
in conjunction with
ASPRS/CaGIS 2010 Fall Specialty Conference
November 15-19, 2010 Orlando, Florida

2. ALGORITHMS APPLIED

2.1 Scale Invariant Feature Transform

SIFT is invariant to image scale, rotation, affine distortion, changes in 3D viewpoint, noise and changes of illumination, which is excellent for image matching(Lowe 2004). Some recent researches have shown that it is also good for object recognition(Plinio Moreno 2007). So this paper provides an introduction in this section. SIFT is divided into two stages, keypoint detection and keypoint description. Each stage consists of two sub-stages respectively.

1. Scale-space detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.

2. Keypoint localization: At each candidate location, a detailed model is used to determine location and scale. Keypoints are selected based on measures of their stability.

3. Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

4. Keypoint descriptor: The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

SIFT feature is a local feature. Compared with traditional overall feature, it improves the efficiency of the method. Fig.1 shows the SIFT features extracted on sample traffic signs and some corresponding matching points in two sign images.
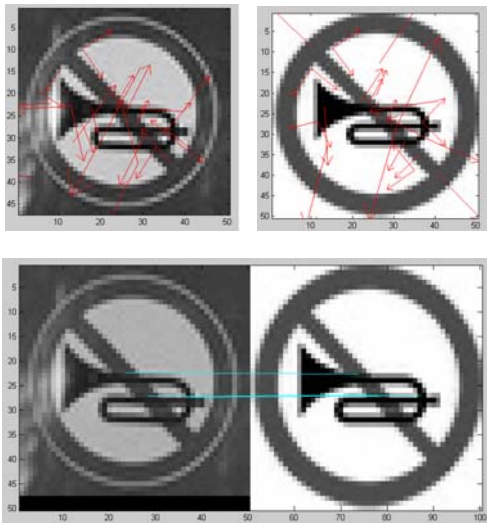

Figure 1. SIFT features on sample images and corresponding matching points

2.2 Bag-of-words

In recent years, " Bag-of-Words " model has been paid more and more attention in the field of object recognition. The image representation method based on "Bag-of-Words"(Fei-Fei and Perona 2005; Sivic, Russell et al. 2005; Gang, Ye et al. 2006) is shown in Fig.2.The basic idea of BoW is to categorize objects in terms of some particular features which called "code-words" and "codebook" consists of all the "code-words".
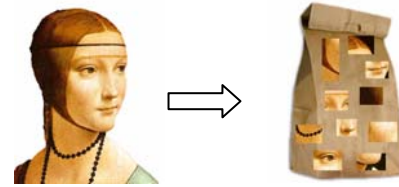
(a) Local features extracted
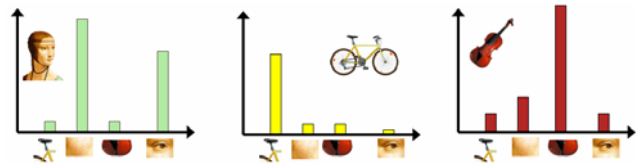
(b) After quantify the local features and construct codebook, an image is converted into a histogram of the number of occurrences of features descriptor in a given image
Figure 2. Bag-of-words model(Fei-Fei 2007)

Bag-of-words is a valid classification from text analysis, two-dimensional image data mapped to a collection of visual keywords, which is not only maintained the local features of the image but also compressed effectively the description of the image. Thus natural language processing techniques and methods can be very effectively applied to the field of object recognition.

2.3 Codebook construction & SVM classifier

Codebook construction process is shown in Fig.3 and as follows:
(1) The SIFT features extracted from certain kind of training traffic sign are clustered by using K-means method and initialize k-means cluster centers with random values. The final cluster center is selected as candidate intrinsic feature.
(2) After calculating the occurrence frequency of cluster center according to the Euclidean distance between each SIFT feature and cluster center. Top n of the occurrence frequency of cluster centers are selected as "code-words" and the size of "codebook" is n.
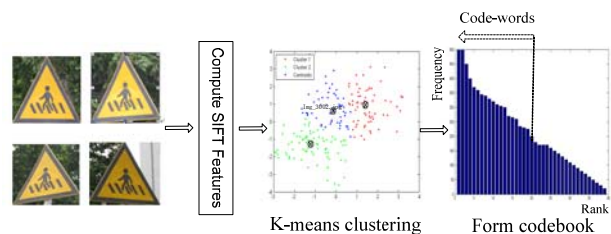

Figure 3.The process of codebook construction

A special joint symposium of ISPRS Technical Commission IV & AutoCarto
in conjunction with
ASPRS/CaGIS 2010 Fall Specialty Conference
November 15-19, 2010 Orlando, Florida

Examples of converted code words by this method are shown in Fig.4 and the size of image is 50*50. The difference of histograms among images is clearly found in the examples.
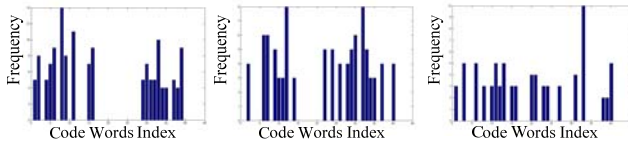

(a) Prohibited from entering   (b) Attention to vehicle

(c) Motor   pedestrian

(d) histogram of (a)   (e) histogram of (b)   (f) histogram of (c)

Figure 4. Converted examples of traffic signs

After extracting code-words from all kinds of traffic signs, all training images are converted into histogram and three groups of SVMs (for Prohibition signs, Warning signs and Indication signs) are trained with these. Process of training is shown as Fig.5.
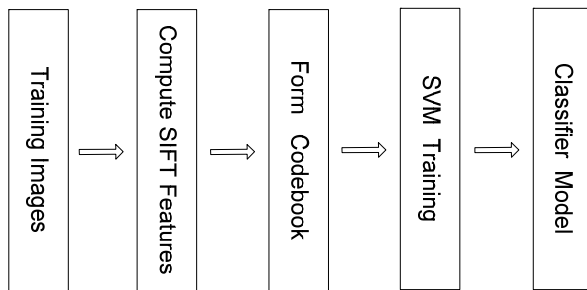

Figure 5. Process of Training

3. EXPERIMENTS

China's current standard is GB5768-1999 after revised in 1999 and mainly consists of warning signs, prohibition signs, indication signs etc(Beijing Traffic Management Bureau 2010). Because of the direct relationship to traffic safety, these three kinds of signs are the main study objects. The purpose of this study is to achieve higher recognition rate and lower false alarm rate for these three objectives in most real scene.
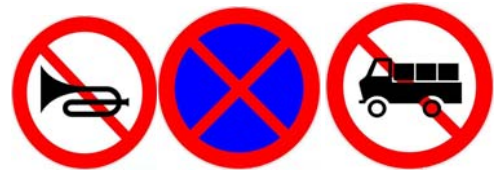
(a) Prohibition signs

(b) Warning signs   (c) Indication signs

Figure 6. Target traffic signs used in experiment

Fig.6 shows the experimental images include Prohibition signs, Warning signs and Indication signs. Total 130 Real images under different weather,distance and angle of shooting are used for evaluating proposed recognition method. The size of images is 50*50, length side is more than 50 pixels are resized to be 50 pixels. Because images in different scale have different entropy, the image must be scaled to the same scale. At the same time, this also suppresses intra-class variation and makes computational cost lower. Partial examples of the real images are shown in Fig.7.

Figure 7. Partial example of the Real Image

This paper uses SVM classifier, which is provided by the library LIBSVM(C. Chang 2001). Images are taken by Canon camera and the software system is implemented with MATLAB R2007b on a Core2 Q6600 processor. The average run time is 0.098 seconds per image.

A special joint symposium of ISPRS Technical Commission IV & AutoCarto
in conjunction with
ASPRS/CaGIS 2010 Fall Specialty Conference
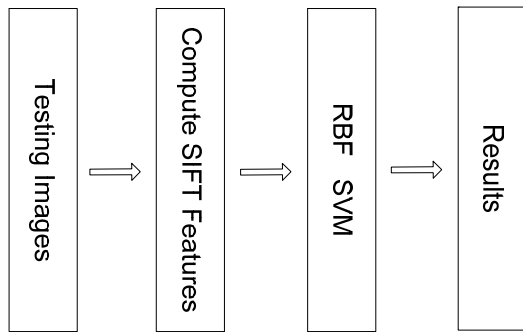November 15-19, 2010 Orlando, Florida

Figure 8. Process of recognition

This paper used radial basis function (RBF) as the kernel function in SVM and process of recognition is shown as Fig.8.While using RBF kernel, two parameters need identify: C and gamma. In the experiment, model selection is performed to search for the best C and gamma by doing grid-search using cross-validation. The best result is observed when C is 2.0 and gamma is 2.0. Tab.1 shows an experimental result.

Table 1.Experimental Result

# of success	121
# of error	9
# of total	130
Recognition rate	93%

4. CONCLUSION & FUTURE WORKS

In this paper, we proposed a TSR method. Experiments considering the difficulties of TSR such as weather, angle of shooting and distance variation and were carried out on training image sets classified according to weather, viewpoint and distance. As discussed in section 2, SIFT method is used to extract all the features of certain types of signs, then formed theirs code book from all the features by using K means and finally classified by the SVM. That means our method recognizes traffic sign by feature distribution. Result shows that the method is robust to weather, distance and angle of shooting, which has higher accuracy rate (93%) and lower calculation cost (0.098s per image).
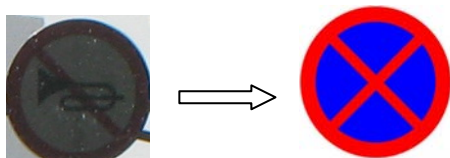

Figure 9. Wrong category

There are two mainly causes of recognition failure. When a traffic sign is seriously polluted, its signal noise ratio is low and the sign cannot be recognized because its SIFT features maybe wrong（Fig.9）.Another cause is the different cluster centers result in different codebook. This would affect the success rate of

identification. Thus, preprocessing image and developing a stable way to obtain codebook are future works.

REFERENCE

A de la Escalera, J. M. A., M.A.Salchs,et al,1997. Road traffic sign detection and classification. *IEEE Transactions on Industrial Electronics*, **44**(6),pp. 848-859.

Beijing Traffic Management Bureau.2010. National standards GB5768-1999 《 Road traffic signs and markings 》. from http://www.bjjtgl.gov.cn/publish/portal0/tab113/.

C. Chang, C. L.,2001. LIBSVM: A Library for Support Vector Machines. from http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Fei-Fei, L.,2007. CVPR2007_tutorial_bag_of_words. from http://vision.cs.princeton.edu/documents/CVPR2007_tutorial_bag_of_words.ppt.

Fei-Fei, L. and P. Perona,2005. A Bayesian hierarchical model for learning natural scene categories. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*.

Gang, W., Z. Ye, et al.,2006. Using Dependent Regions for Object Categorization in a Generative Framework. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*.

Kehtarnavaz, L. E. a. N.,1996. A real-time histographic approach to road sign recognition. Image Analysis and Interpretation, 1996., Proceedings of the IEEE Southwest Symposium on

Lowe, D. G.,2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* **60**(2),pp. 91-110.

Maldonado-Bascon, S., S. Lafuente-Arroyo, et al.,2007. Road-Sign Detection and Recognition Based on Support Vector Machines. *Intelligent Transportation Systems, IEEE Transactions on*, **8**(2),pp. 264-278.

Matthias Muller, A. B., Joachim Gerlach, Wolfgang Rosenstiel,Dennis Nienh¨user, J. Marius Z¨ollner,Oliver Bringmann,2010. Design of an Automotive Traffic Sign Recognition System Targeting a Multi-Core SoC Implementation. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*,pp. 532 - 537

Merve Can Kuş, M. G., Şima Etaner-Uyar,2008. Traffic Sign Recognition using Scale Invariant Feature Transform and Color Classification. *Computer and Information Sciences, 2008. ISCIS '08. 23rd International Symposium on*

Plinio Moreno, M. J. M.-J., Alexandre Bernardino, José Santos-Victor and Nicolás Pérez de la Blanca,2007. A Comparative Study of Local Descriptors for Object Category Recognition: SIFT vs HMAX.[Extended Version]. *Proceedings of Pattern*

A special joint symposium of ISPRS Technical Commission IV & AutoCarto
in conjunction with
ASPRS/CaGIS 2010 Fall Specialty Conference
November 15-19, 2010 Orlando, Florida

Recognition and Image Analysis, Third Iberian Conference, IbPRIA, Girona, Spain.

S. D. Zhu, Y. Z., X. F. Lu,2005. Detection for triangle traffic sign based on neural network. *IEEE International Conference on Vehicular Electronics and Safety*.

S. Maldonado Bascón, J. A. R., S. Lafuente Arroyo, A. Fernndez Caballero,F. López-Ferreras,2009. An optimization on pictogram identification for the road-sign recognition task using SVMs. *Computer Vision and Image Understanding*, **114**(3),pp. 373-383

S.H.Hsu, C. L. H.,2001. Road sign detection and recognition using matching persuit method. *Image and Vision Computing*, **19**(3),pp. 119-129.

Sivic, J., B. C. Russell, et al.,2005. Discovering objects and their location in images. *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*.

A special joint symposium of ISPRS Technical Commission IV & AutoCarto
in conjunction with
ASPRS/CaGIS 2010 Fall Specialty Conference
November 15-19, 2010 Orlando, Florida