The Derived Kernel Based Recognition Method of Vehicle Type

Zhen Chao Zhang[1], Yuan Yan Tang[1], Chang Song Liu[2], and Xiao Qing Ding[2]

[1] Department of Computer and Information Science,
Faculty of Science and Technology, University of Macau, Macau, China
bravobrava@126.com, yytang@umac.mo
[2] Department of Electronic Engineering, Tsinghua University
State Key Laboratory of Intelligent Technology and Systems Beijing, 100084, China
lcs@tsinghua.edu.cn

Abstract. This paper applies the updated derived kernel algorithm into the vehicle type recognition, which is a heated research area based on the method of pattern recognition and digital image processing because of its significant usage on exit and entry administration, traffic and vehicle control and toll collection. The method of two- layer derived kernel on neural response is involved in extracting useful features from vehicle images, for the method itself has better capacity of decreasing the negative influences from different colors and vehicle speeds, background condition interference and blur noises. Some clustering algorithms are employed on the process of templates construction, and the first nearest neighbor algorithm on pattern classification. Since our method can get rid of the disturbances from similar parts of vehicle images and make the most of the features of representative parts, the vehicle type recognition accuracy reaches above 95% as high.

Keywords: vehicle type recognition, derived kernel, templates construction, neural response, classification.

1 Introduction

Recently since the rapid development of vehicle manufacture and its wide usage on many aspects of daily life, there come quite a number of issues concerned on vehicles. As the result, more and more research topics on vehicle area have become significant and applicable, one of which is vehicle type recognition mostly applied on crimes detection, Electronic Toll Collection management, exit and entry administration, and security control of restricted areas.

There have been different methods used on the vehicle type recognition, such as Support Vector Machine, Perimeter-Based Noise Filter, extension-based recognition model [10], and Scale Invariant Feature Transform [11]. Edge detection method is widely and effectively used, in [1], Sobel operator, typically used to find the approximate absolute gradient magnitude at each point in an input grayscale image, together with edge direction and detection method is proposed. An algorithm based on vertical

R. Huang et al. (Eds.): AMT 2012, LNCS 7669, pp. 297–306, 2012.

edge detection and color alternation in grayscale image, is proposed to locate the vehicle in [2]. In [3], canny edge detector is applied for extracting feature information after the process of dilation.

Neural networks can be used in the area of vehicle type identification, because of their ability to derive patterns from complicated or imprecise data. In [4], BP network learning algorithm is improved by using momentum and genetic algorithm after analyzing the defects of the BP network learning algorithm. A new recognition approach based on Radial Basis Function Neural Networks (RBFNN) is mentioned in [5]. Transform methods, represented by Fast Fourier Transform, Discrete Curvelet Transform and Discrete Wavelet Transforms are used in [7] for image feature extraction and vehicle type recognition. In [6], it recognizes vehicle type in different lighting conditions with the application of wavelet and contourlet as tools for feature extraction, which is robust to illumination and scale variation.

As things stand now, some methods mentioned above show the difficulty in dealing with the vehicle type recognition problem with high standard of accuracy. One of the challenges on that area, as we conclude, is how to handle with the similar parts and the representative parts of the vehicle images. According to our experiment and research, the features of similar parts of vehicle images, such as some backgrounds, vehicle wheels and lamps, are easy to make confusion when recognizing; and the information of representative parts, such as vehicle color, landmark and contour is hard to use for discrimination. And the method we propose, with two layer neural response feature extraction and process, which is based on each useful small image patch from more subtle perspective, can get rid of the disturbance from the similar parts of vehicle images, and make the most of the representative parts, so that the features obtained from vehicle images can be better used for enhancing the recognition accuracy.

2 Derived Kernel Algorithm

In this section, the derived kernel algorithm as an updated feature extraction and similarity measurement method is introduced into the recognition of vehicle type. That algorithm is firstly proposed by [12], under development of [13] [14], and it is encouraged to read them for more specific details and other applications.

2.1 Preliminaries

For the whole process of vehicle type recognition, there are four ingredients needed to define for the preparation of derived kernel part listed as followed:

Nested Patches Definition

The initial definition of the derived kernel is made in the case of the nested patches definition on one vehicle image in R^2. By the definition, that vehicle image is proposed to be made up of three different layers of patches: named u, v and Sq. Sq represents the whole vehicle image domain. The left part of Fig. 1 shows the different nested patch domains on one randomly selected vehicle image, where $u \in v$, $v \in Sq$ and $Sq \in R^2$.

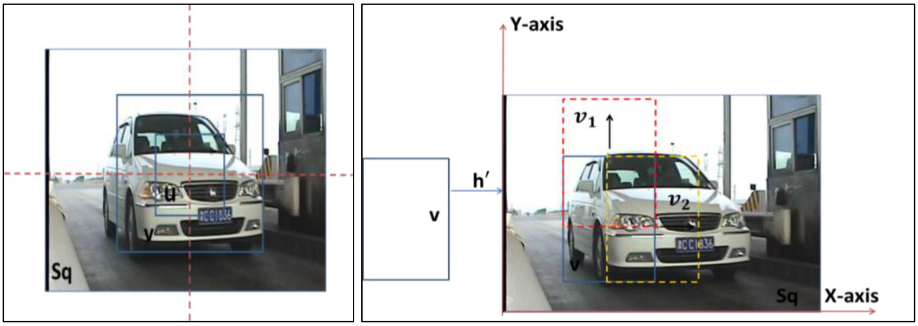

Fig. 1. Nested patches definition & a set of translation v through Sq area

Image Patches Transformation

A set of transformations of a certain vehicle image from a patch to another patch is depicted as the right part of Fig. 1. We define a finite set H_u of transformations that are maps from u to v with h: u→ v, and similarly H_v with h✓: v→ Sq. The transformation can be translation, scaling and rotation. In this paper, considering the actual problems, the sets of transformations are assumed to be finite and limited to only translations for simplicity. As the right part of Fig. 1, we make coordinate patch v transform through the whole image area Sq, and there should come a series of limited coordinate patches named from v_1 to v_n. And on the image we just use v_1 and v_2 as representatives.

Image Function Definition

Assume that we are given an image function on Sq, denoted by I_{Sq}= {f: Sq → [0, 1]} (where f is the image with size Sq), as well as the image function I_u, I_v defined on sub-patches u, v respectively. As Fig. 2, there come two image patches h'_1 and h'_2 through the Sq area.

Image Function Restriction

Besides the ingredients defined previously, there is an axiom which is used to restrict any vehicle images. And f ∘ h: u→ [0, 1] is in I_u if f ∈ I_v and h∈ H_u, similarly f ∘ h′: v→ [0, 1] is in I_v if f ∈ I_{Sq} and h′ ∈ H_v. The mathematical expressions are shown as followed:

$$f \circ h : u \rightarrow [0,1] \in I_u,\ f \in I_v,\ h \in H_u \tag{1}$$

$$f \circ h' : v \rightarrow [0,1] \in I_v,\ f \in I_{Sq},\ h' \in H_v \tag{2}$$

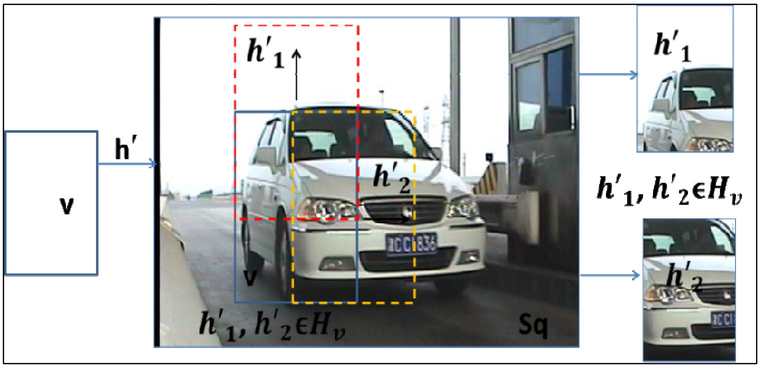

Fig. 2. Image function definition, where h'_1 and h'_2 correspond to two different translation (one is upward movement, another is right side movement)

2.2 Neural Response

In the following part, the neural response is shown based on the above described recursive definition which will be on a hierarchy of local kernel. The bottom-up design style is applied on the construction of derived kernel on neural response.

2.2.1 The derived kernel process begins with the normalized, non-negative, reproducing kernel $I_u \times I_u$, which is denoted by $\widehat{K_u}(f,g)$ in formula (3).The value of $\widehat{K_u}(f,g)$ is the normalized result of $K_u(f,g)$, which is inner product operation of image function $f(x)$ *and* $g(x)$ on patch area u.

$$\widehat{K_u}(f,g) = \frac{K_u(f,g)}{\sqrt{K_u(f,f)K_u(g,g)}} \tag{3}$$

The function f (x) and g (x) on patch area u is belong to I_u, which can be considered as two vehicle images needed to compare for computing the similarity, shown as formula (4). Since the value zero cannot be allowed in the denominator, $K_u(f,f)$=0 is excluded in order to satisfy the normalization.

$$K_u(f,g) = < f(x), g(x) > \tag{4}$$

2.2.2 For the first layer neural response on image function f (x) with one selected template t is denoted as $N_v(f,t)$, which is shown as (5) where f $\in I_v$, h$\in H_u$ and t $\in T_u$.

$$N_v(f,t) = \max_{h \in H} \widehat{K_u}(f \circ h, t) \tag{5}$$

As the left part of Fig. 3, it shows the whole process of one-layer neural response on image function f. For easy understanding, we use t_1' instead of t', h'_1 instead of h'. And $f \circ h'_1$can be considered as a certain patch of the vehicle image of size v, if f $\in I_{Sq}$ is an image patch of size Sq and$h'_1 \in H_v$. $\widehat{K_u}(f \circ h'_1, t'_1)$ is the result of normalized inner product of $f \circ h'_1$and t'_1, and $\widehat{K_u}(f \circ h'_n, t'_n)$ is the result of normalized inner product of $f \circ t'_n$and t'_n. $N_v(f, t'_1)$ is the best match (the maximum of values from $\widehat{K_u}(f \circ h'_1, t'_1)$ to $\widehat{K_u}(f \circ h'_n, t'_1)$) of the template t'_1 in the image patch f.

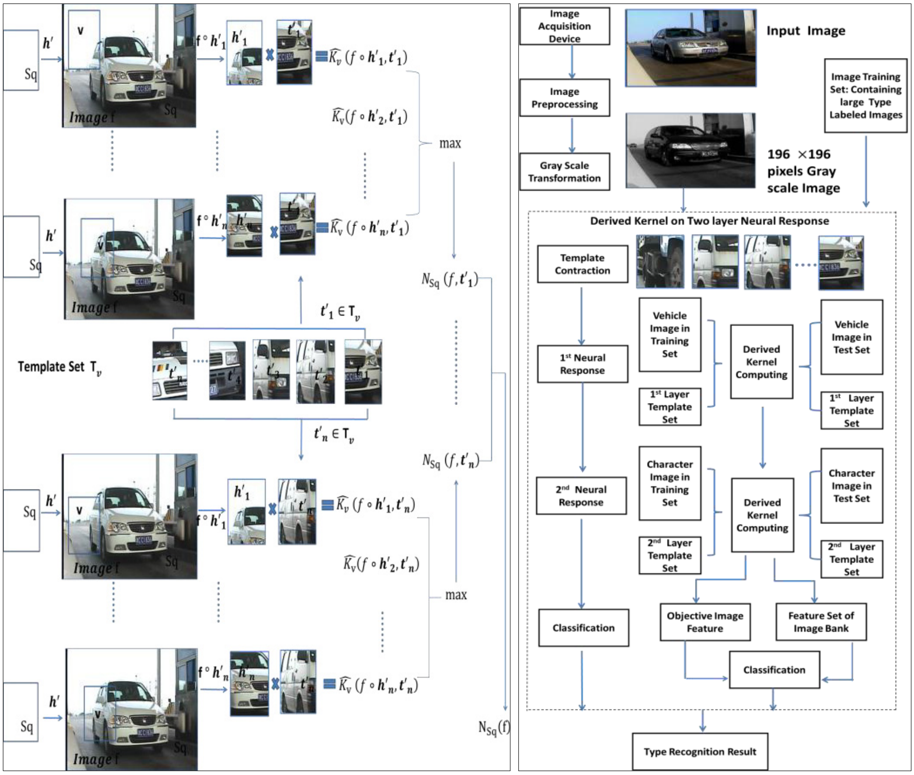

Fig. 3. The process of one layer neural response & the process of vehicle type recognition

The derived kernel of $I_v \times I_v$ is then defined as the formula (7).While f, g $\in I_v$, can be considered as two vehicle images. The value of $\widehat{K_v}(f, g)$ is obtained by normalizing $K_v(f, g)$ according to (4).

$$N_v(f) = \{N_v(f, t_1), N_v(f, t_2) \ldots N_v(f, t_n)\} \tag{6}$$

$$(f, g) =< N_v(f), N_v(g) > \tag{7}$$

For example, for a certain vehicle image, such as image f, after finishing the first neural response, there comes the result set $N_v(f)$, including $N_v(f, t_1), N_v(f, t_2)$ to $N_v(f, t_n)$ shown as (6); so does another vehicle image g, through the first neural response there comes the result set $N_v(g)$.

2.2.3 The process above is repeated by defining the second layer neural response shown as (8)

$$N_{Sq}(f, t') = \max_{h' \in H_v} \widehat{K_v}(f \circ h', t') \tag{8}$$

While in this case f$\in I_{Sq}$, $t' \in T_v$ and $h' \in H_v$. Consequently, the derived kernel on $I_{Sq} \times I_{Sq}$ is given by (10), where f, g$\in I_{Sq}$. Just as the former procedure, the final derived kernel $\widehat{K_{Sq}}$ is obtained by normalizing K_{Sq}.

$$N_{Sq}(f) = \{N_{Sq}(f, t'_1), N_{Sq}(f, t'_2) \ldots N_{Sq}(f, t'_n)\} \quad (9)$$

$$K_{Sq}(f, g) = < N_{Sq}(f), N_{Sq}(g) > \quad (10)$$

3 Vehicle Type Recognition

In the following section, the whole process of vehicle type recognition method is depicted, including image preprocessing, gray scale transformation, derive kernel on neural response, templates construction and classification.

3.1 Image Preprocessing

Since the original images we obtain are randomly sized, and it would have negative effects on image analysis and recognition process. As the result, we make the size of those entire images changed as 196 ×196 pixels.

3.2 Gray Scale Transformation

In this part, we transform images from original RGB type into gray scale type. And the common used formula is shown as (11). Considering 32-bit integer operation and avoiding floating point operation, we do some adjustment and apply (12) for gray scale transformation.

$$\text{Gray} = R \times 0.299 + G \times 0.587 + B \times 0.114 \quad (11)$$

$$\text{Gray} = (R \times 299 + G \times 587 + B \times 114 + 500)/1000 \quad (12)$$

3.3 Derived Kernel on Neural Response

In our research, two layers of neural response are employed. As we can see in Fig. 4, the captured smaller image patches $h_1$and h_2 (which construct one layer patch set named H_u mentioned in section 2) from two different vehicle images have too many similar parts, so that the extracted information would be not enough for type identification. At the same time, with the assistance of the second layer bigger image patches h'_1 and h'_2 (constructing another layer set named H_v), the acquired feature would include more complete and useful information, including smaller parts of vehicles (such as vehicle wheels and lamps) and bigger parts (such as vehicle gate and general contour).

Specifically speaking, in our research, for each image, we would extract six bigger image patches as the first neural response, on each of which the smaller image patches as the second layer neural response would be chosen through the translation, image function and restriction as we mentioned in Section 3.

Fig. 4. The function of first layer and second layer templates comparison

3.4 Template Construction

Template construction counts important on the whole process of vehicle type recognition method based on derived kernel, which connect the mathematical model to a real world setting. In our method, we assume T_u and T_v are discrete and finite template sets as (13) and (14). The template sets $T_u \in I_v$, $T_v \in I_{Sq}$ restrict the templates selection scope.

$$T_U = \{t_1, t_2, t_3, \dots t_n\} \tag{13}$$

$$T_v = \{t'_1, t'_2, t'_3, \dots t'_n\} \tag{14}$$

Templates can be seen as representative image patches frequently encountered in the daily life, and in our research, they can be a certain number of vehicle image patches frequently appeared and they are used for vehicle type matching and recognition. The construction of template sets is considered as important step, and will finally reflects on the accuracy and time efficiency.

We employ some common and standard 196 ×196 pixels vehicle images which we transform into gray-scale ones. And then from those processed images, we extract some templates with specific size. The first step is to select template sets randomly through image translation and transformation. For each vehicle type (there are three types totally, including truck, van and car), 10 smaller image patches with size u are randomly chosen as the first layer templates and 8 larger image patches with size v as the second layer templates. For higher recognition accuracy, the size of image patches should be determined by a series of testing experiments, which will be shown on the following session IV. Since there are three kinds of vehicle types, and for each kind of types, there are four corresponded images to be used for templates selection. So after random selection, the total number of templates is 120 for the first layer and 96 for the second layer. The specific process to choose those templates through translation and transformation way is depicted as Fig. 1, Fig. 2 and their explanation on section 2.

So far, the basic templates construction is finished with random selection method. For gaining the representing capacity of the template set, some clustering algorithms can be applied and the entire templates can be divided into several dozen of classes. The idea is to extract some templates which locate on the central position of a certain class and which have proper ability representing for the other templates on that same class. Specifically speaking, in our method, the K-means clustering and hierarchical clustering algorithm are employed to partitionally and hierarchically select smaller amount of representative templates on two derived kernel layers.

The advantage of applying the clustering algorithms into template construction is as followed:

- It reduces the number of involved templates so that the time consumption on template construction and recognition process is diminished to some extent.
- Since some singularities and their disturbance can be avoided by the clustering algorithm, the representing ability of the selected templates is increased, which helps to enhance the final recognition accuracy.

3.5 First Nearest Neighbor Classification

In our research, the first-nearest neighbor classification method is involved, which requires two sets: training set, and testing set. Together with the template set used for the neural response on derived kernel, there are totally three data sets used.

The images on the training set are selected, preprocessed and transformed from the real-world vehicle images; they will be put into the specific type subset according to their own type. The function of the training set is used to help to recognize the type label of the vehicle images on the testing set by comparing and computing the similarity level of extracted features from images on the training set and testing set. For each vehicle type, we will employ 50 images, and there are three types needed to recognize, including car, van and truck. So for the training subset, there are 150 vehicle images.

The idea of the first-nearest neighbor classification is that an unlabeled testing set is given the label from the closest image in training set. Specifically speaking, the images on the training set are used to compare with the object ones about their feature similarity through hierarchical derived kernel. When a certain image in the training set, as an example, has the largest similarity level with the unlabeled object one (which goes beyond the threshold), the label of that image on the training set would be assigned to the object one as the recognized result. The test set, composed of various object vehicle images is used to test the accuracy of our research.

4 Experiment and Result

The experiment of ensuring the size of each layer templates are depicted as followed. Firstly, we just use one layer derived kernel on neural response for experiment and realize that the template size of 146×146 pixels can lead to the highest recognition accuracy (compared with other template size for one layer). And then we employ two layer derived kernel on neural response and fix the size of the first layer template as 146×146 pixels (the size of vehicle images have already been fixed as 196×196 pixels as we mentioned previously), and as the change of the size of the second layer templates, the recognition accuracy fluctuate. We determine the second layer template size, with which the related experiment result shows the highest accuracy. After the series of experiments, we can conclude that the derived kernel with 146 ×146 pixels first layer template size and 96 × 96 pixels second layer template size neural responses have the greatest recognition accuracy. That's because template with that size include more exclusive and representative feature information with less similarity disturbance.

The experiment of vehicle type recognition on the image bank, or the testing set as we mentioned previously, is presented. All the images employed for training, testing and constructing templates are randomly selected from the internet or some image bank, and they are originally RGB types with different size. Unlike other image bank just depicting the frontal part of vehicles, the images we employ describing the vehicle contour with side view can record more vehicle information (Fig. 5 shows some image examples). On the testing set, we totally employ 150 vehicle images, 50 of which depict cars, 50 vans and 50 trucks.

According to the testing experiment, the recognition accuracy on vehicle images is as high as above 95%. Besides, according to the experiment result the recognition accuracy of truck images is highest compared with the one of the others. The reason is that truck images have more exclusively representative feature, which makes the feature similarity between truck and other vehicles more distances. The following table 1 shows the performance comparison of some related approaches, and some of the statistics on that table are based on [1] [8] [9] [11].

Table 1. Comparision of Different Approaches

Source Paper	**Approach and Accuracy**	
	Approach Accuracy	**Main Proposed Approach**
[8]	87.3%	Different gradient features
[9]	89%	Scale invariant feature transform
[1]	89.22%	Sobel operation and edge detection
Our Paper	Above 95%	Derived kernel based recognition method

5 Conclusion

In this paper, we have described the derived kernel based vehicle type recognition. The vehicle type can be recognized with high level of accuracy with the assistance of first nearest classification and clustering algorithm, since it can avoid the disturbance from similar parts of vehicles and make the most of the information from the representative vehicle parts. Although we have dealt with the template construction and described the clustering algorithm applied together with random selection method on template construction, the templates selection and the number of templates used are still the critical element influencing the time consumption and recognition accuracy, which will be one open topic for future research.

Acknowledgments. The authors wish to thank the State Key Laboratory of Intelligent Technology and Systems of Tsinghua University for the equipment assistance and image acquisition.

References

1. Wang, W.: Reach on Sobel Operator for Vehicle Recognition. In: International Joint Conference on Artificial Intelligence, JCAI 2009, April 25-26, pp. 448–451 (2009)
2. Mello, C.A.B., Costa, D.C.: A Complete System for Vehicle License Plate Recognition. In: 16th International Conference on Systems, Signals and Image Processing, IWSSIP 2009, June 18-20, pp. 1–4 (2009)
3. Munroe, D.T., Madden, M.G.: Multi-class and single class classification approaches to vehicle Model recognition from images. In: Proc. AICS (2005)
4. Wang, Z., Li, S.: Research and implement for vehicle license plate recognition based on improved BP network. In: 2010 International Conference on Computer and Communication Technologies in Agriculture Engineering (CCTAE), June 12-13, vol. 3, pp. 101–104 (2010)
5. Wang, W.: Vehicle Type Recognition Based on Radial Basis Function Neural Networks. In: International Joint Conference on Artificial Intelligence, JCAI 2009, April 25-26, pp. 444–447 (2009)
6. Arzani, M.M., Jamzad, M.: Car type recognition in highways based on wavelet and contourlet feature extraction. In: 2010 International Conference on Signal and Image Processing (ICSIP), December 15-17, pp. 353–356 (2010)
7. Kazemi, F., Samadi, S., Pourreza, H., Akbarzadeh, M.: Vehicle Recognition Based on Fourier, Wavelet and Curvelet Transforms - a Comparative Study. IJCSNS Int. Jou. of Comp. Sci. and Net. Secu. 7(2) (2007)
8. Petrovis, V.S., Cootes, T.F.: Analysis of features for rigid structure vehicle type recognition. In: Proc. of the British Machine Vision Conference (2004)
9. Dlagnekov, L.: Video-based car surveillance: License plate make and model recognition. Master's Thesis, University of California at San Diego (2005)
10. Yigui, Q., Yuanyuan, Y., Ling, Z., Xing, L.: The Extension-Based Recognition Method of Vehicle Type. In: Third International Conference on Convergence and Hybrid Information Technology, ICCIT 2008, November 11-13, vol. 2 (2008)
11. Iqbal, U., Zamir, S.W., Shahid, M.H., Parwaiz, K., Yasin, M., Sarfraz, M.S.: Image based vehicle type identification. In: 2010 International Conference on Information and Emerging Technologies (ICIET), June 14-16, pp. 1–5 (2010)
12. Smale, S., Poggio, T., Caponnetto, A., Bouvrie, J.: Derived distance: towards a mathematical theory of visual cortex. CBCL Paper. MIT, Cambridge (2007)
13. Zhang, Z.C., Tang, Y.Y.: License plate recognition algorithm based on derived kernel. In: 2012 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), July 15-17 (in press, 2012)
14. Smale, S., Rosasco, L., Bouvrie, J., Caponnetto, A., Poggio, T.: Mathematics of the neural response. Found. Comput. Math. 10, 67–91 (2010)