UNIVERSITY OF CALIFORNIA
RIVERSIDE

Vehicle Logo Detection and Recognition Through The Traffic Surveillance Camera

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Electrical Engineering

by

Siyu Li

December 2015

Thesis Committee:

Professor Bir Bhanu, Chairperson
Professor Yingbo Hua
Professor Matthew Barth

The Thesis of Siyu Li is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor, without whose help, I would not have been here.

To my parents for all the support.

ABSTRACT OF THE THESIS

Vehicle Logo Detection and Recognition Through The Traffic Surveillance Camera

by

Siyu Li

Master of Science, Graduate Program in Electrical Engineering
University of California, Riverside, December 2015
Professor Bir Bhanu, Chairperson

In order to deal with recognition problem of the low-resolution and poor quality images captured from traffic surveillance camera in intelligent transportation system, a vehicle logo super-resolution method using two-dimensional canonical correlation analysis(2D CCA) is presented. While numerous approaches have been proposed to do the recognition, they can only achieve good performance on the images with high resolution. In this thesis, a novel combination of morphological operation and sliding window technique is adopted to detect the vehicle logo, and then using 2D CCA to do super-resolution reconstruction which improves performance significantly compared to traditional methods. It analyses the performance of proposed syste on a large vehicle image dataset.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Intelligent transportation systems (ITS) have attracted more and more attention in recent years. As a part of ITS, vehicle recognition systems are being developed rapidly due to its commercial value in many surveillance applications. [22, 15, 25, 20, 8, 25, 26] Vehicle logo recognition (VLR) is an area with limited research compared to the mature License plate recognition (LPR) systems, which also play an important role in road surveillance applications. However, the image captured by the surveillance camera usually has low resolution which is insufficient for direct detection and recognition of the logos. Even though many methods exist that solve logo recognition problem when we can get high resolution picture of the vehicle. But in the real world, we can only acquire low resolution images due to limitations of the camera which makes the most of methods useless. Therefore, developing the VLR method which works with low resolution images is important for ITS.

In this thesis, a novel logo detection and recognition method is proposed. First, a fusion method that combines morphological operations and a multi-scale sliding window

technique to detect the vehicle logo is presented. Later a method using 2D CCA to reconstruct the low resolution vehicle logo image is proposed. The vehicle images captured by the outdoor surveillance cameras inevitably suffer from noise, illumination changes and other variations which result in low resolution and inferior quality images. Traditional techniques achieve poor performance under such conditions, so we develop a learning based super-resolution(SR) approach for vehicle logos which are highly structured. An SR approach by An *et al*. [1] applies one-dimensional canonical correlation analysis(1D CCA) to the principal component analysis(PCA) coefficients of high-resolution(HR) and low-resolution(LR) logo images to enhance the coherence of their neighbourhood structure. However, the 1D CCA was not designed specifically for the image data, the image has to be first converted into a one dimensional(1D) vector to fit the image data into 1D CCA formulation. On the other hand, an image can be inherently represented in a 2D matrix. The appearance information is lost when an image is reshaped in to a vector. To address this problem, 2D CCA which is specifically suitable for image analysis has been proposed [9]. Finally, we extract the histogram of oriented gradients(HOG) features for the super-resolved images and feed them into an support vector machine(SVM) classifier. The proposed method is assessed on a database that includes 1070 images belonging to 16 vehicle manufacturers under different types of conditions, it achieves a very good performance when compared with traditional methods.

Chapter 2

Related Work

Automatic license plate recognition (ALPR) techniques are well developed and mature which are routinely used in various ITSs. However, Vehicle Logo Recognition (VLR) area has seen relatively limited research activity. The vehicle images are obtained in the outdoor environment which suffers from uncontrolled lighting and various other environmental factors. Additionally, the area of the logo is very small when compared to the whole image, making VLR more difficult than ALPR. Using the rear-view vehicle images, Dlagnekov and Belongie [5] take advantage of Scale Invariant Feature Transform (SIFT) [11] to find the logo and then identify the manufacturer. However, the computational cost of this method is expensive which makes it unsuitable for real-time applications. Most VLR methods employ detection methods such as template matching and various histogram-based methods, and then use a coarse to fine approach to crop the ROI. In [24], the authors use template matching and edge-orientation histograms to address the logo recognition problem. In order to solve the problem of high computational cost of SIFT descriptor, Psyllos *et al.* [17]

used a sift-based enhanced matching scheme which yields a better performance and reduces the computational cost. They also proposed a system using SIFT descriptor with probabilistic neural network [16]. However, the performance is not robust when the illumination or viewpoint changes. Llorca *et al.* [10] presented HOG and SVM framework for vehicle logo recognition which uses the images captured by a traffic surveillance camera. A sliding window technique combined with a majority voting scheme is used in the approach. In order to detect vehicle logo in low resolution images reliably and fast, Boguslaw and Michal [3] employ the structural tensor to detect logo areas which are then classified to the car brands with help of a classifier operating in the multi-dimensional tensor spaces. Peng *et al.* [14] proposed a novel approach based on statistical random sparse distribution feature and multi-scale scanning for VLR.

Table 2.1: Related Work Summary

Publication	Principle of the Method	Vehicle Data
Dlagnekov et al.(2005)	LPR, MMRSIFT AdaBoost	Rear-view, street
Wang et al.(2007)	template matching	Front-view, parking
Psyllos et al.(2010)	Enhanced matching, SIFT-based	Front-view, internet
Minich et al.(2011)	SIFT, Interior Fourier descriptor	Front-view, head-on
Psyllos et al.(2011)	VMMR, SIFT, Neural network	Front-view, 110 images, 11 class
Pearce et al.(2011)	MMR, LNHS, KNN	Front-view, 262 images, 10 class
Sam et al.(2012)	Harr-like features, Adaboost	Front-view, 200 images, 10 class
Llorca et al.(2013)	Sliding window, HOG, SVM	Front-view, 3579 images, 27 class
Cyganek et al.(2014)	Tensor classifiers, HOSVD	Front and rear view, parking
Peng et al.(2015)	VLR, SRSD, multi-scale	Front-view, 2126 images, 56 class

Chapter 3

Technical Approach

The proposed system contains two parts: logo detection and recognition. The overall architecture of the proposed system is depicted in Fig. 3.1.

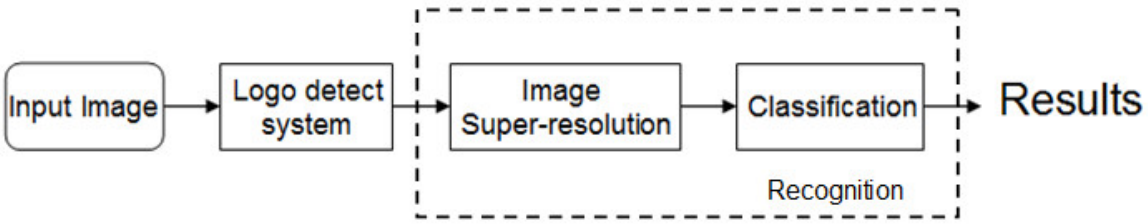


Figure 3.1: Flowchart of Logo Detection System.

The proposed vehicle logo detection system consists of two parts, which are mathematical morphology analysis and multi-scale sliding window technique. In the mathematical morphology analysis part, there are mainly four processes: image enhancement including filtering, morphological transforms, morphological operations or post processing, resulting in the vehicle logo candidate area. Whether the following sliding window module will switch

on will be judged by the decision value obtained from an SVM classifier. If the decision value is lower than a pre-set threshold, the sliding window module will be switched on. If not, the decision value from candidate area combined with decision value obtained from areas above and below the candidate area to form a new descriptor. Finally, the new descriptor is fed into a pre-trained binary SVM classifier to see if the system has detected the logo or not. The overall logo detection process can be illustrated in Fig. 3.2.

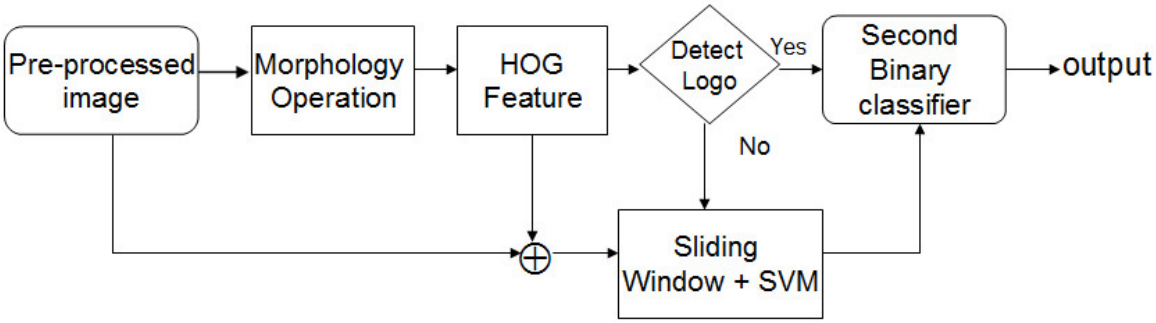


Figure 3.2: Flowchart of Logo Detection System.

3.1 Image Enhancement

In this step, we use methods that include adjusting the intensity of the image and enhancing the contrast in the image. The technique used for intensity adjustment is known as histogram equalization. The contrast can be enhanced by several methods including the top/bottom hat transformations.

(a) Oringinal (b) Top Hat

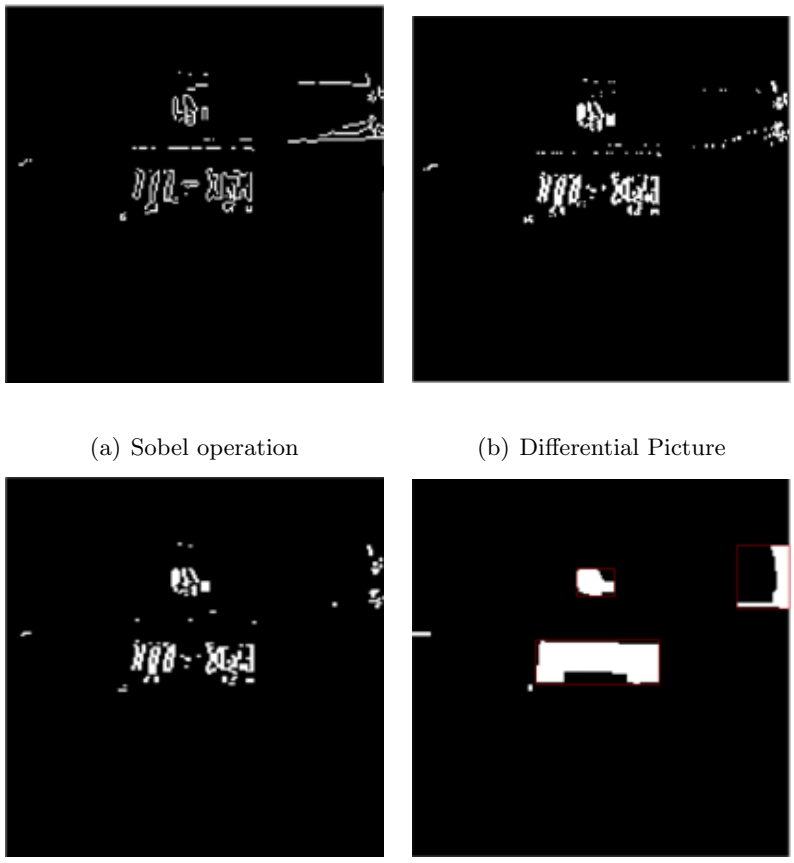
(c) Bottom Hat (d) After enhancement

Figure 3.3: Image Enhancement

3.1.1 Hat Transformation

Hat transforms are used for contrast enhancement. There are two operations and are known as top hat and bottom hat transformations [6]. Top hat operation is the result of subtraction of an opened image from the original one whereas in the case of bottom hat the closed image minus the original image. The top hat operation suppresses the dark background and highlights the foreground objects. So this operation can highlight the characters and suppresses the irrelevant background. If we covert the resulting image in

a binary image and remove all the small connected areas, there are only a few foreground areas left and most of the irrelevant objects have been moved.



(a) Sobel operation

(b) Differential Picture

(c) Closed image

(d) Opened image

Figure 3.4: Morphological operations

3.2 Morphological Operations

Mathematical morphological operations refers to a broad set of image processing operations that process images based on shapes. In this project, we just use only opening operation and closing operation for the purpose of logo extraction. Erosion shrinks the

objects by eroding the boundaries. Closing is the dilation that allows objects to expand, followed by erosion and vice versa is the opening operation. These operations can be modified by proper choice of the structuring element which determines how many objects will be dilated or eroded. In Fig. 5.9(a), we use rectangle shaped structuring element for Sobel operator. The we do the column by column subtraction over the Sobel image, to remove those horizontal lines which may disturb the detection of vehicle logo area. In Fig. 5.2(c) and Fig. 3.4(d), we use the same rectangle structuring element for closing operation and opening operation. These four steps produce an image with candidate areas for the vehicle license plate and vehicle logo. Next, we apply some checks and conditions, which are based on the properties of the vehicle license plate, for all the remaining objects in the image, such as the area and aspect ratio to locate the vehicle license plate.

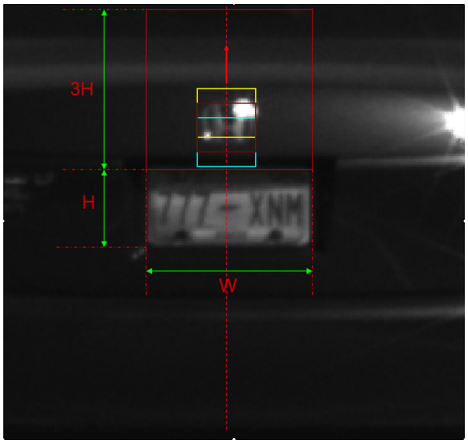


Figure 3.5: Flowchart of Logo Detection System.

3.3 Multi-scale Sliding Window

The vehicle logo is located just above the vehicle license plate in most cases. Thus, once we detected the vehicle license plate in the images, we can further crop the region just above license plate which most likely contains vehicle logo. Suppose the detected license plate has a width W and height H, then a rough location of logo region, namely new ROI, can be determined with a width $3W$ and height $3H$. The new ROI segmentation is illustrated in Fig. 3.5. We then extract HOG features from all the candidate areas in new ROI and feed them into SVM classifier. We set a threshold value to decide whether the sliding window module will switched on or not. If SVM decision value is greater than 0.95, we consider that area to be a vehicle logo and the candidate area will directly go into the final binary SVM classifier. If not, sliding window will be applied. Once the sliding window module switch on, the sliding window will slide along the vertical axis that separates the license plate in two equal sized regions. Due to the variable size of different vehicle logos, the window will be scaled by a factor s which has a range from 0.8 to 1.2 with a step of 0.2. At each scale step, we find the maximum SVM decision value when we extract HOG features from all the windows and feed them into the SVM classifier. Comparing the maximum value obtained at each scale, finally, we choose the window with the utmost value as the final output of detection.

3.4 Binary SVM Classifier

In order to determine whether the detected area contains the relevant information to carry out vehicle logo classification, we create a new descriptor which is fed into a binary

SVM classifier. The new feature is shown in Fig. 3.6. The Black area is the maximum decision value which is obtained by the previous SVM classifier through sliding window, and then we collect the decision values of adjacent area, above and below by 4 pixels obtained by the previous SVM classifier. Finally, we combined them together to form this new descriptor shown in Fig. 3.6.

Figure 3.6: The new descriptor

We collect 100 positive samples with logo images and 200 images sampled from arbitrary positions in images which do not contain a logo as negative samples. With the radial basis function(RBF) kernel, we train the SVM classifier to get model. Given a detected area from previous stage, after several procedures to get the new feature, we can use the trained model to obtain the final result.

Chapter 4

2D CCA for Super-resolution

4.1 1D CCA

we now briefly introduce the main concepts of CCA which has been discussed first in [7]. CCA is a way to find the maximum value of linear correlation between two set of random variables. Suppose we are given two sets of m vectors P and Q with zero means, $P = \{P_j\}_{j=1}^{m} = [p_i \in \mathbb{R}^m, P_1, ..., P_m]$, $Q = \{Q_j\}_{j=1}^{n} = [q_i \in \mathbb{R}^n, Q_1, ..., Q_n]$. 1D CCA is applied to get two basis vectors W_P and W_Q both of which have dimension are m such that the correlation coefficient ρ of $W_P^T P$ and $W_Q^T Q$ is maximized. So the objective function is given by

$$\rho = \frac{W_P^T C_{PQ} W_Q}{\sqrt{W_P^T C_{PP} W_P W_Q^T C_{QQ} W_Q}} \tag{4.1}$$

where C_{PP} and C_{QQ} represents the auto-covariance matrices of P and Q respectively. Finally, we can formulate the 1D CCA problem as a constrained optimization

$$\underset{W_P, W_Q}{\operatorname{argmax}} \; W_P^T C_{PQ} W_Q \tag{4.2}$$

$$subject\ to\ W_P^T C_{PP} W_P = 1\ and\ W_Q^T C_{QQ} W_Q = 1$$

.

4.2 2D CCA

Different from traditional CCA method, 2D CCA directly extracts the features from two dimensional matrix. For some particular data type, such as image, the data are represented as two dimensional matrix. So it is valuable if we use the data in its original space without reshaping the data into 1D vectors. 2D CCA was developed by [9], which was motivated by 2D Principal Component Analysis(2D PCA) [23]. Suppose we are given two centered datasets, $P = \{p_i \in \mathbb{R}^{m_p \times n_p},\ i = 1, 2, ..., N\}$ and $Q = \{q_i \in \mathbb{R}^{m_q \times n_q},\ i = 1, 2, ..., N\}$, different from the CCA technique, 2D CCA trys to find two left projection matrices and two right projection matrices, which are $L_P, L_Q, R_P\ and\ R_Q$. So we can find the correlation coefficient ρ between the two projected datasets $L_P^T P R_P$ and $L_Q^T Q R_Q$ is maximized, we can finally obtain the ρ by

$$\rho = \frac{Cov(L_P^T P R_P, L_Q^T Q R_Q)}{\sqrt{Var(L_P^T P R_P)}\sqrt{Var(L_Q^T Q R_Q)}} \tag{4.3}$$

ρ can be divided in two parts

$$\rho_L = \frac{L_P^T C_{PQ}^R L_Q}{\sqrt{L_P^T C_{PP}^R L_P L_Q^T C_{QQ}^R L_Q}} \tag{4.4}$$

$$\rho_R = \frac{R_P^T C_{PQ}^L R_Q}{\sqrt{R_P^T C_{PP}^L R_P R_Q^T C_{QQ}^L R_Q}} \tag{4.5}$$

where C_{PP}^R and C_{QQ}^R are the auto-covariance matrix of PR_P and QR_Q respectively. C_{PQ}^R is the cross-covariance matrix of PR_P and QR_Q. Similarly, C_{PP}^L and C_{QQ}^L are the auto-covariance matrix of $L_P^T P$ and $L_Q^T Q$ respectively. C_{PQ}^L is the cross-covariance matrix of

$L_P^T P$ and $L_Q^T Q$ respectively. Thus, in a similar way, we can transform the problem into a constrained problem

$$\underset{L_X, L_Y, R_X, R_Y}{\operatorname{argmax}} \quad Cov(L_P^T P R_P, L_Q^T Q R_Q) \tag{4.6}$$

$$subject\ to\ Var(L_P^T P R_P) = 1\ and\ Var(L_Q^T Q R_Q) = 1$$

.

4.3 Training for Super-Resolution

In our super-resolution context, the variables are vectors each of which represents a single Low Resolution (LR) image or High Resolution (HR) image, such as projection coefficients of HR and LR images. In the training step, 2D CCA is applied to find the left and right projection matrices that project HR and LR images into a subspace in which the correlation between the projections is maximized.Given N HR images $I^H = \{I_i^H \in \mathbb{R}^{m_p \times n_p}\}_{i=1}^N$ and corresponding LR images $I^L = \{I_i^L \in \mathbb{R}^{m_q \times n_q}\}_{i=1}^N$ for each vehicle make, the mean images μ_X and μ_Y are subtracted to obtain the centered datasets $\widehat{p}$ and $\widehat{Q}$, respectively. We mark I^H as X, i^L as Y for the efficiency.

The left transforms $L_{\widehat{X}}$ and $L_{\widehat{Y}}$, right transforms $R_{\widehat{X}}$ and $R_{\widehat{Y}}$ are obtained by maximizing in Eq.(4.6). Then we can get $P_X = L_{\widehat{X}}^T \widehat{X} R_{\widehat{X}}$ and $P_Y = L_{\widehat{Y}}^T \widehat{Y} R_{\widehat{Y}}$ from the image datasets $\widehat{X}$ and $\widehat{Y}$, respectively.

The 2D CCA based super-resolution (SR) approach consists of two main steps: training and reconstruction. Through manifold learning, the SR will further refine vehicle logo extracted in the first step which often does not contain sufficient details.

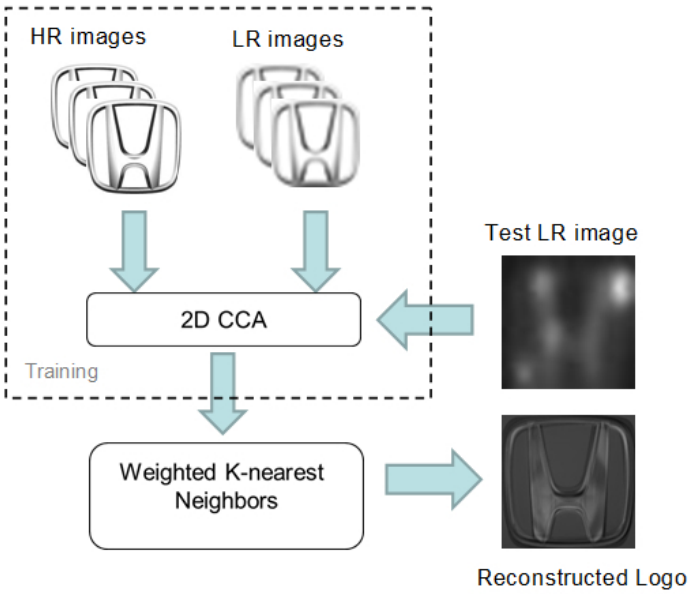


Figure 4.1: System diagram for 2D CCA super-resolution.

4.4 Super-Resolution with Trained Models

In order to get the super-resolved image of the input LR image, the LR image is projected to the subspace by

$$P_i^{LR} = L_{\widehat{Y}}^T (i_{LR} - \mu_Y) R_{\widehat{Y}}. \tag{4.7}$$

Suppose P_i^{LR} can be reconstructed by a linear combination of its K nearest neighbours in P_Y. In 2D CCA space, we can find weights $\{\omega_j\}_{j=1}^K$ for K nearest neighbours that minimize the reconstruction error

$$\begin{aligned} &\underset{\{\omega_j\}_{j=1}^K}{\text{argmin}} \ \|P_i^{LR} - \sum_{j=1}^K \omega_j P_{Y_j}\|_F \\ &\textit{subject to} \ \sum_{j=1}^K \omega_j = 1. \end{aligned} \tag{4.8}$$

where P_{Y_j} denotes representation of a sample from LR dataset in the 2D CCA space, and $\|.\|_F$ is the Frobenius norm. The methods of solving the above constrained least square

problems can be found in [19].

We can then apply the same weighted neighbourhood in the subspace for the HR training images. The reconstructed HR image i_{HR} in the 2D CCA space is given by

$$P_i^{HR} = \sum_{j=1}^{K} \omega_j P_{X_j} \tag{4.9}$$

where P_{X_j} is the reconstructed HR image from P_{Y_j} in 2D CCA space. Then we can get the i_{HR} through Eq.(4.6), so the i_{HR} is derived as

$$i_q^{HR} = L_{\widehat{X}}^{T+} P_i^{HR} R_{\widehat{X}}^{+} + \mu_X \tag{4.10}$$

where + denotes the Moore-Penrose pseudoinverse operation. i_q^{HR} is the super-resolved image using q-th model. Suppose we have Q models in total, the HOG features $\{H_q\}_{q=1}^{Q}$ are extracted from all the possible candidate images $\{i_q^{HR}\}_{q=1}^{Q}$. we can select the final output by using the nearest neighbour

$$\underset{i_q^{HR}}{\operatorname{argmin}} \ \|H_{i^{LR}} - H_{i_q^{HR}}\| \tag{4.11}$$

where $H_{i^{LR}}$ is the HOG features from LR images.

Chapter 5

Experiments and Results

5.1 Vehicle Logo Datasets Collection

In our experiment, the rear-view images are obtained by the previous work by Thakoor *et al* [21]. In their work, the vehicle images are extracted as the vehicle enters the view of the camera and is followed by symmetric detection. We normalize all the detected vehicles to 300 $\times$ 400. Vehicle image samples are shown in Fig. 5.1. In most of the samples, the logo resolutions are approximately 25 $\times$ 25, but the dimension varies from different classes due to their shapes. Take Audi as an instance, the dimension of it are usually 30 $\times$ 60. The whole database contains 1070 images which covers over 10 logo categories, as shown in Fig. 5.1. In the database, the images captured under varying lighting conditions. We categorize the whole database into three datasets under different conditions using two measurement: one is the distortion measure (DM) [4] that evaluates the image quality in frequency domain. The other one is cumulative probability of blur detection (CPBD) [12]

Figure 5.1: Examples of captured picture from surveillance camera.

which focuses on the image sharpness evaluation. The criterion of categorization is shown in 5.1. The two thresholds in the table are the average scores computed over whole dataset. The samples of three different datasets are shown in Fig. 5.2.

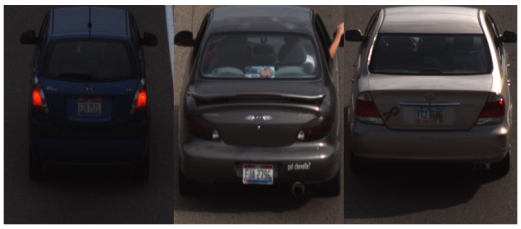
Table 5.1: Image Database

Dataset	DM	CPBD	Image Amount
Good	>15	>6	711
Fair	>15	<6	278
	<15	>6	
Bad	<15	<6	82

5.2 Vehicle Logo Detection

In logo detection stage, sufficient representative training data is prepared to build a reliable and accurate logo detection system, which is composed of two types of images, negative ones and positive ones. Negative samples correspond to non-logo images which were randomly cropped from vehicle images excluding logo as shown in Fig. 5.9(b). The positive samples contains two parts : one is 700 HR logo images obtained from internet, another part is 2/3 of the logo samples from each dataset which have been manually labeled.

(a) Good samples



(b) Fair samples

(c) Bad samples

Figure 5.2: Data set samples

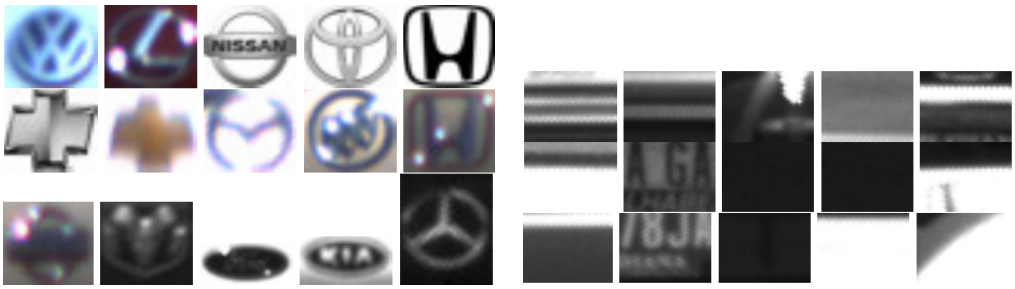
The rest 1/3 of the samples are used to do the test.

In order to have a stable result, a three-fold cross validation experiment were conducted to get classification accuracy. The results are shown in Table 5.2.

Table 5.2: Three-fold Cross Validation

Dataset	subset 1	Accuracy	subset 2	Accuracy	subset 3	Accuracy	Average
Good	237/224	94.5	237/216	91.1	237/223	94.1	92.7
Fair	93/81	87.1	93/72	77.4	92/77	83.7	82.7
Bad	27/16	59.3	28/15	53.5	27/17	59.3	58.5
Overall	357/321	89.7	358/303	84.6	356/317	89.0	87.8

According to the table, we can observe that this detection method yields a good performance on good and fair dataset. However, the method performs badly on the last dataset. The illumination of the image and some strong reflection of the vehicle logo may



(a) Positive logo samples (b) Negative logo samples

Figure 5.3: Training Logo Samples

(a) Too Bright (b) Too Dark

Figure 5.4: Detection failure examples

be the main reasons. Some failure examples are sown in Fig. 5.4

The Table 5.3 shows the times that sliding window module is triggered for each subset, as we can see, the effectiveness of this combination can be demonstrated. When it is not triggered, the detection rate of just using morphological operation is also very good. This combination can combine the strength of these two methods.

We also use the receiver operating characteristic (ROC) curve to evaluate the final

Table 5.3: Sliding window module trigger times

Dataset	subset 1	subset 2	subset 3
hit times/others	88/91	133/134	100/102
Times	266	224	254
Total	744/1071	Percentage	69.5

stage of logo detection. As we can see in Fig. 5.5, the area under curve (AUC) is 0.86 which can demonstrate the good performance of the final binary classifier.

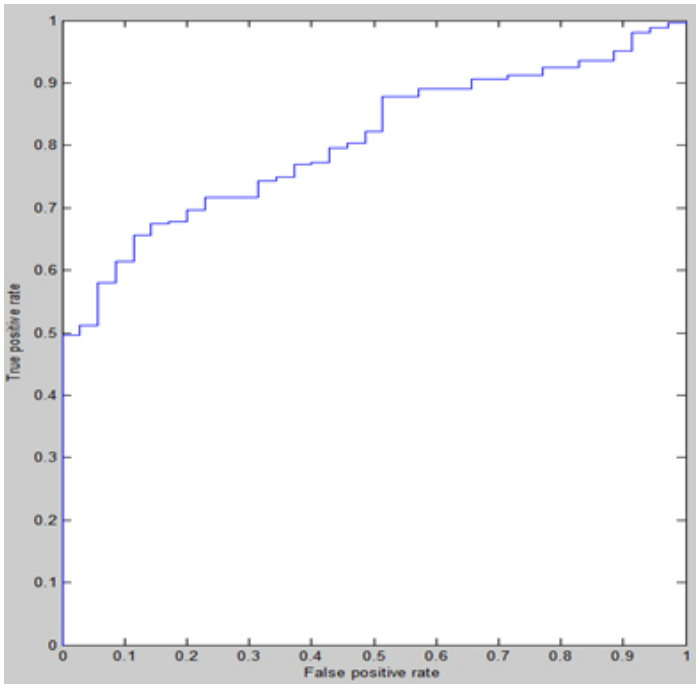
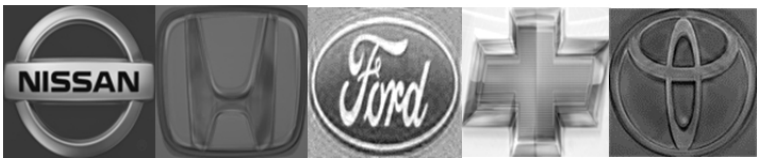


Figure 5.5: The ROC curve

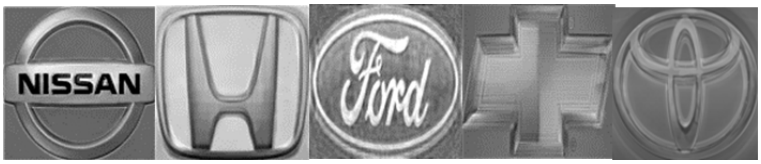
5.3 Vehicle Logo Recognition

To build a reliable training database for image super-resolution, we collect 20 HR images from internet. To account for varying body colors and illuminations, we use gamma adjustment on each HR image to generate 30 images of different contrast by varying gamma value from 0.1 to 3 with a step of 0.1. we normalized all the HR images to 120 $\times$ 120. And then we just simply down-sampled HR images to 30 $\times$ 30 to obtain our LR images for training. The magnification factor in our experiment is 4. We divide each logo image into

9 blocks and each block has 15 histogram bins, so the length of HOG feature vector is 135.



(a) 2D CCA



(b) 1D CCA

Figure 5.6: Compare with 1D CCA super-resolution

In Fig. 5.6, we can see the comparison of super-resolved image between by using 1D CCA and 2D CCA. Visually, we can say the the edges in 2D CCA super-resolved images are clearer and the noise is much less compared to 1D CCA super-resolved images.

Predicted Class ⟶	Honda	Toyota	Ford	Chevrolet	Nissan	Lexus	Mazda	Hyundai	Buick	Volkswagen	Accuracy (%)
Honda	119	14	0	10	2	1	0	0	0	0	81.5
Toyota	3	54	2	0	1	1	1	0	0	1	87.09
Ford	10	7	136	13	7	11	7	8	0	0	73.11
Chevrolet	5	0	0	99	3	0	0	0	0	0	92.52
Nissan	0	7	0	0	45	5	2	3	0	1	72.58
Lexus	1	1	0	0	1	12	1	0	0	0	75
Mazda	0	1	0	0	0	1	12	0	0	0	85.7
Hyundai	4	2	0	0	1	1	0	30	0	0	81.1
Buick	0	0	0	0	0	0	0	0	18	0	100
Volkswagen	0	0	0	0	0	0	0	0	0	25	100
Overall	672/552										82.14

Figure 5.7: Confusion matrix for recognition.

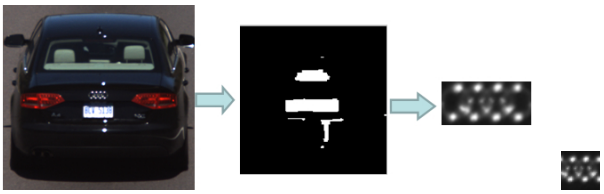
From the confusion matrix shown in Fig 5.7, we can observe that some specific logos perform bad, for example, Ford, Kia and Chevrolet. They have lowest classification accuracy as their is similarity in their shapes and the noise makes some character features contained in the logo more obscure, leading to incorrect classification. Other than the illumination, another reason for the bad performance is incorrect alignment. The accurate alignment is needed for the high-quality super-resolved image. However, sliding window technique cannot guarantee proper alignment, some examples are shown in Fig. 5.8.

Figure 5.8: Some examples without proper alignment.

The proposed detection method can alleviate this situation to some extent. For example, Audi logo has a rectangular shape of 25×45, whereas the average dimension of most logos is 26×26. Just using traditional sliding window which is designed for most of the logos will cause the Audi logo to be partially detected. Even tough we can adjust the window size to fit shape the Audi logo, a large irrelevant area will be account in the logo area too, which will cause the poor alignment. It can illustrated in Fig. 5.9.

5.3.1 Comparison with Other Methods

We compare the results with three other methods, in which SIFT and HOG are well known features in traditional recognition tasks and both of them yield very good



(a) Accurate alignment (b)

Figure 5.9: Advantages using combined detect system

performance on high-resolution image.

Table 5.4: Comparison with other other methods

Different Method	VLR Accuracy Rate			
	2D CCA	1D CCA	SIFT/SVM	HOG/SVM
Accuracy	82.14	67.16	42.74	54.79

Table. 5.4 shows the classification accuracy by testing the whole dataset in our experiment. From the table, we can clearly see 2D CCA method performs much better than other methods in the table. The reason that SIFT and HOG performance is has to do with the resolution of image due to which enough features can not be extract effectively to ensure the discriminative power of the classifier. For 1D CCA, this method was not designed specifically for the image data. In order to fit the image data into 1D CCA formulation, the image has to be first converted into a 1D vector. However, an image is inherently represented in a 2D matrix. For the highly structured images like vehicle logo and human face, the appearance becomes obscured when reshaped into a vector and it affects the classification performance to a certain degree.

Chapter 6

Conclusions

In this thesis, a novel vehicle logo detection and recognition system is proposed. First, a morphological operations combined with sliding window technique to detect the vehicle logo is adopted. Then we use the 2D CCA to reconstruct the low-resolution image captured by traffic surveillance camera. Finally, HOG features from super-resolved image are extracted and fed in to an SVM classifier with a radial basis function kernel to get the final classification result. Compared with traditional recognition methods, the proposed method yields a much better performance in dealing with low-resolution image and thus is suitable for vehicle logo recognition in real world. Future work will involve aligning the logo in detected region more accurately. Since high-quality super-resolved image needs an accurate alignment and training images and testing images should be aligned in the same manner. Without proper alignment, the quality of super-resolved image will degrade and it will influence the classification performance.

Bibliography

[1] Le An, Ninad Thakoor, and Bir Bhanu. Vehicle logo super-resolution by canonical correlation analysis. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2229–2232. IEEE, 2012.

[2] Travis Burkhard, AJ Minich, and Christopher Li. Vehicle logo recognition and classification: Feature descriptors vs. shape descriptors. 2011.

[3] Bogusław Cyganek and Michał Woźniak. Vehicle logo recognition with an ensemble of classifiers. In *Intelligent Information and Database Systems*, pages 117–126. Springer, 2014.

[4] Niranjan Damera-Venkata, Thomas D Kite, Wilson S Geisler, Brian L Evans, and Alan C Bovik. Image quality assessment based on a degradation model. *Image Processing, IEEE Transactions on*, 9(4):636–650, 2000.

[5] Louka Dlagnekov. *Video-based car surveillance: License plate, make, and model recognition*. PhD thesis, Citeseer, 2005.

[6] Rafael C Gonzalez and Richard E Woods. Digital image processing, 2002.

[7] Harold Hotelling. Relations between two sets of variates. *Biometrika*, pages 321–377, 1936.

[8] Andrew HS Lai, George SK Fung, and Nelson HC Yung. Vehicle type classification from visual-based dimension estimation. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pages 201–206. IEEE, 2001.

[9] Sun Ho Lee and Seungjin Choi. Two-dimensional canonical correlation analysis. *Signal Processing Letters, IEEE*, 14(10):735–738, 2007.

[10] DF Llorca, R Arroyo, and MA Sotelo. Vehicle logo recognition in traffic images using hog features and svm. In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, pages 2229–2234. IEEE, 2013.

[11] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.

[12] Niranjan D Narvekar and Lina J Karam. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *Image Processing, IEEE Transactions on*, 20(9):2678–2683, 2011.

[13] Greg Pearce and Nick Pears. Automatic make and model recognition from frontal images of cars. In *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*, pages 373–378. IEEE, 2011.

[14] Haoyu Peng and Xun Wang. Recognition of low-resolution logos in vehicle images based on statistical randon sparse distribution. In *IEEE transactions on intelligent transportation systems*, pages 681–691. IEEE, 2015.

[15] Vladimir S Petrovic and Timothy F Cootes. Analysis of features for rigid structure vehicle type recognition. pages 1–10, 2004.

[16] A Psyllos, Christos-Nikolaos Anagnostopoulos, and Eleftherios Kayafas. Vehicle model recognition from frontal view image measurements. *Computer Standards & Interfaces*, 33(2):142–151, 2011.

[17] Apostolos P Psyllos, Christos-Nikolaos E Anagnostopoulos, and Eleftherios Kayafas. Vehicle logo recognition using a sift-based enhanced matching scheme. *Intelligent Transportation Systems, IEEE Transactions on*, 11(2):322–328, 2010.

[18] Kam-Tong Sam and Xiao-Lin Tian. Vehicle logo recognition using modest adaboost and radial tchebichef moments. In *International Conference on Machine Learning and Computing (ICMLC 2012)*, 2012.

[19] Lawrence K Saul and Sam T Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *The Journal of Machine Learning Research*, 4:119–155, 2003.

[20] Sayanan Sivaraman and Mohan Manubhai Trivedi. A general active-learning framework for on-road vehicle recognition and tracking. *Intelligent Transportation Systems, IEEE Transactions on*, 11(2):267–276, 2010.

[21] Ninad S Thakoor and Bir Bhanu. Structural signatures for passenger vehicle classification in video. *Intelligent Transportation Systems, IEEE Transactions on*, 14(4):1796–1805, 2013.

[22] Wei Wu, Zhang QiSen, and Wang Mingjun. A method of vehicle classification using models and neural networks. In *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, volume 4, pages 3022–3026. IEEE, 2001.

[23] Jian Yang, David Zhang, Alejandro F Frangi, and Jing-yu Yang. Two-dimensional pca: a new approach to appearance-based face representation and recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(1):131–137, 2004.

[24] Wang Yunqiong, Liu Zhifang, and Xiao Fei. A fast coarse-to-fine vehicle logo detection and recognition method. In *Robotics and Biomimetics, 2007. ROBIO 2007. IEEE International Conference on*, pages 691–696. IEEE, 2007.

[25] Bailing Zhang. Reliable classification of vehicle types based on cascade classifier ensembles. *Intelligent Transportation Systems, IEEE Transactions on*, 14(1):322–332, 2013.

[26] Bailing Zhang, Yifan Zhou, and Hao Pan. Vehicle classification with confidence by classified vector quantization. *Intelligent Transportation Systems Magazine, IEEE*, 5(3):8–20, 2013.