VEHICLE LOGO SUPER-RESOLUTION BY CANONICAL CORRELATION ANALYSIS

Le An, Ninad Thakoor, Bir Bhanu

Center for Research in Intelligent Systems, University of California, Riverside
lan004@ucr.edu, ninadt@ucr.edu, bhanu@cris.ucr.edu

ABSTRACT

Recognition of a vehicle make is of interest in the fields of law enforcement and surveillance. In this paper, we develop a canonical correlation analysis (CCA) based method for vehicle logo super-resolution to facilitate the recognition of the vehicle make. From a limited number of high-resolution logos, we populate the training dataset for each make using gamma transformations. Given a vehicle logo from low-resolution source (*i.e.*, surveillance or traffic camera recordings), the learned models yield super-resolved results. By matching the low-resolution image and the generated high-resolution images, we select the final output that is closest to the low-resolution image in the histogram of oriented gradients (HOG) feature space. Experimental results show that our approach outperforms the *state-of-the-art* super-resolution methods in qualitative and quantitative measures. Furthermore, the super-resolved logos help to improve the accuracy in the subsequent recognition tasks significantly.

***Index Terms*—** Super-resolution, vehicle make recognition, subspace learning

1. INTRODUCTION

Recognition of the make and model of vehicles has generated interest in recent years [1, 2]. However, majority of the work has been focused on appearance based methods. Vehicle logos provide an alternate approach for the recognition [3]. However, the resolution of surveillance videos is insufficient for direct recognition of logos. This work proposes a super-resolution approach for vehicle logos to improve the recognition rate.

Super-resolution (SR) algorithms produce high-resolution (HR) image from low-resolution (LR) input. Accurate SR reconstruction is usually difficult and is an ill-posed image processing problem. The existing SR algorithms can be roughly categorized into two classes based on the types of input.

The SR methods in the first class take multiple images as input [4, 5]. Usually, registration is performed first to align the input images. Super-resolution or interpolation is carried out subsequently to fuse multiple aligned LR images to get a HR output. These methods are based on the assumption that the LR inputs can be generated by warping and downsampling the super-resolved image. However, when the magnification factor increases, this assumption becomes weaker [6].

The SR methods in the second class use single LR image as input to infer the HR output. With the general idea that the relationship between the HR images and the LR images can be learned from examples, many methods in this class require a training step [7, 8]. Glasner *et al.* [9] avoided using a set of training images by exploring the rich patterns in a single image. Besides, advanced interpolation algorithms without training have also been proposed which outperform the conventional interpolation techniques [10].

For highly structural images such as vehicle logos, it is natural to develop a learning based SR approach where the model is trained from a set of similar images. Inspired by the recent success in super-resolving face images (which are also highly structural) using manifold learning techniques [11], we also choose to work in the subspaces to cater to our specific application for vehicle logo super-resolution. The main assumption is that the HR and LR manifolds have similar structure which is locally linear and smooth [12]. Specifically, canonical correlation analysis (CCA) [13] is applied upon the PCA coefficients of HR and LR logo images to enhance the coherence of their neighborhood structure. To the authors' best knowledge, this is the first paper that addresses vehicle logo SR problem.

The rest of the paper is organized as follows. Technical details are provided in Section 2. Section 3 shows the experimental results. Finally, we draw conclusions in Section 4.

2. TECHNICAL APPROACH

The proposed vehicle logo super-resolution algorithm consists of two steps as illustrated in Figure 1. In the training step, for each make, we use a set of HR and LR logo image pairs to learn a model that maximize their correlation in the CCA subspace. To test a LR image, each model will produce an output and the final output we select is the one that is closest to the input LR image in the histogram of oriented gradients (HOG) feature space [14]. Before we delve into the details of the proposed method, we first briefly review CCA.

2.1. Canonical Correlation Analysis (CCA)

CCA finds basis vectors for two sets of random variables such that the correlation between the projections of these two sets of random variables is maximized [13]. Given two centered (zero mean) datasets, $X = \{x_i \in \mathbb{R}^m, i = 1, 2, ..., N\}$ and $Y = \{y_i \in \mathbb{R}^n, i = 1, 2, ..., N\}$, CCA aims at obtaining two basis vectors $W_X \in \mathbb{R}^m$ and $W_Y \in \mathbb{R}^n$ such that the correlation coefficient ρ of $W_X^T X$ and $W_Y^T Y$ is maximized.

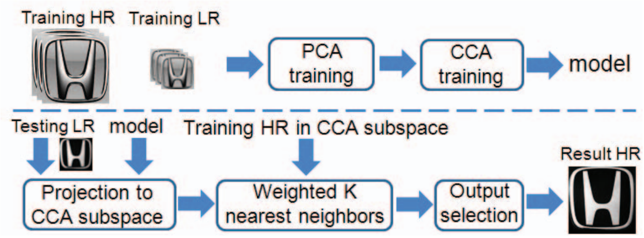

Fig. 1. Training (top) and testing (bottom) of the proposed approach. During training for each make, the correlation of HR and LR image pairs are maximized in the CCA subspace. Given a LR image, the outputs are reconstructed with K nearest-neighbors in different CCA subspaces. Final output is selected by analyzing its HOG features.

The objective function to be maximized is given by

$$\rho = \frac{W_X^T C_{XY} W_Y}{\sqrt{W_X^T C_{XX} W_X W_Y^T C_{YY} W_Y}} \tag{1}$$

where $C_{XX(YY)}$ is the autocovariance matrix of $X(Y)$. C_{XY} denotes the covariance matrix of X and Y.

Equivalently, the CCA can be formulated as a constrained optimization problem by

$$\underset{W_X, W_Y}{\operatorname{argmax}}\, W_X^T C_{XY} W_Y \tag{2}$$

subject to $W_X^T C_{XX} W_X = 1$ and $W_Y^T C_{YY} W_Y = 1$. The solution to this optimization problem can be obtained by solving a set of generalized eigenvalue problems [15].

2.2. Training for Super-Resolution

In the training step, given N HR images $\mathbf{I}^H = \{I_i^H\}_{i=1}^N$ and their corresponding LR images $\mathbf{I}^L = \{I_i^L\}_{i=1}^N$ for each vehicle make, we first apply PCA to obtain the projection coefficients of HR and LR images by

$$\begin{aligned} \mathbf{X} &= (P_H)^T(\mathbf{I}^H - u^H) \\ \mathbf{Y} &= (P_L)^T(\mathbf{I}^L - u^L) \end{aligned} \tag{3}$$

where P_H and P_L are the projection matrices, u^H and u^L are the mean HR and mean LR images. $\mathbf{X}$ and $\mathbf{Y}$ are further mean-centered to $\mathbf{X_0}$ and $\mathbf{Y_0}$ by extracting the mean values u^X and u^Y respectively. In the next step, CCA maximizes the correlation coefficient ρ in Eq. (1) using centered data $\mathbf{X_0}$ and $\mathbf{Y_0}$. The resulting two basis vectors $W_\mathbf{X}$ and $W_\mathbf{Y}$ transform the PCA coefficients into a coherent subspace where the correlation between the projected data is maximized.

2.3. Super-Resolution with Trained Models

Given a LR logo image i^L, for each model we trained for each make, the PCA coefficients are first computed by

$$y^L = (P_L)^T(i^L - u^L) \tag{4}$$

Then y^L is projected into the CCA subspace by

$$m^L = (W_Y)^T(y^L - u^Y) \tag{5}$$

In the CCA subspace, we find the weights $\{w_j\}_{j=1}^K$ for K nearest neighbors that minimize the reconstruction error

$$\underset{\{w_j\}_{j=1}^K}{\operatorname{argmin}} \left\| m^L - \sum_{j=1}^K w_j Y_j \right\| \tag{6}$$

subject to the constraint $\sum_{j=1}^K w_j = 1$. Y_j denotes the representation of a LR training image in the CCA subspace.

We then apply the same weighted neighborhood in the subspace for the HR training images. The reconstructed HR image in the CCA subspace is given by

$$m^H = \sum_{j=1}^K w_j X_j \tag{7}$$

where X_j is the HR version of Y_j in the CCA subspace.

We transform m^H to PCA subspace by

$$x = ((W_X)^T)^\dagger m^H + u^X \tag{8}$$

where $\dagger$ denotes pseudo-inverse. The final output of the reconstructed HR image is obtained by

$$i_q^H = ((P_H)^T)^\dagger m^H + u^H \tag{9}$$

here i_q^H is the super-resolved image using the q-th model. Suppose we have Q models (makes) in total, the HOG features $\{H_q\}_{q=1}^Q$ are extracted from all the possible candidate images $\{i_q^H\}_{q=1}^Q$. The final output i_q^H is selected by

$$\underset{i_q^H}{\operatorname{argmin}} \left\| H_{i^L} - H_{i_q^H} \right\| \tag{10}$$

where H_{i^L} is the HOG features from the LR image.

3. EXPERIMENTAL RESULTS

3.1. Vehicle Logo Data Collection

In the experiments, we select five of the most popular vehicle makes: Honda, Toyota, Nissan, Chevrolet, and Ford. For each make, we collected 20 HR images from Internet. All the logos are frontal view. To account for varying body colors and illuminations, gamma adjustment [16] is performed 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. In total, for each make 600 HR images are available. The HR images are normalized to size 120×120 and the LR images with size 30×30 are generated through downsampling. The magnification factor in our experiments is 4. In the PCA implementation, 95% of the variance is retained. Each logo image is divided into 9 blocks and in each block 15 histogram bins are used. Thus the length of the HOG feature vector is 135. In the neighborhood reconstruction, all of the HR and LR image pairs used for training are considered. Figure 2 shows some samples of the generated images.

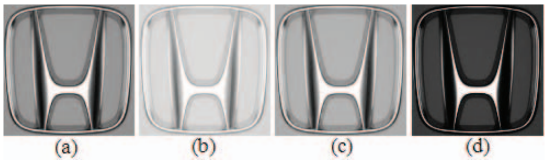

Fig. 2. (a) Original logo image. (b)-(d) Generated logo images by gamma adjustment.

To evaluate the proposed SR method, we collect 15 LR logos for each make from a surveillance camera recording from a highway. In the video, the vehicles are viewed from the rear. As soon as the license plate (LP) of a vehicle is detected, a region of interest (ROI) of size $4LP\ height \times LP\ width$ immediately above the LP is selected. The ROI is segmented to detect and crop the logos automatically. The LR logos are normalized to the same size as those in the training set.

3.2. Vehicle Logo Super-Resolution

We compare our results to bicubic interpolation and three *state-of-the-art* methods: kernel regression based SR (KR) [17], iterative curve based interpolation (ICBI) [10], and adaptive sparse domain selection based SR (ASDS) [18]. We use the default parameters in these methods. Note that in the proposed method it is not necessary to know the blurring kernel as contrary to [17] and [18]. Figure 3 shows some sample results using different super-resolution methods.

The input LR logo images suffer from severe noise and blurriness due to the low quality of the surveillance images. Also strong specular reflection is present due to highly reflective surface of the logos. In this scenario, the improvement from bicubic interpolation to ICBI is not explicit and all the image artifacts are retained. KR performs better in noise reduction. However the super-resolved results are over-smoothed and no high-frequency details are compensated to generate sharp outputs. Although ASDS recovers some details, the noise in the LR images is also exaggerated. The proposed method performs significantly better than the other methods and the details are faithfully reconstructed with the elimination of noise, specularities and blurriness. The results also suggest that to super-resolve highly structural images, manifold learning based method is superior compared to other kinds of approaches.

Since we do not have the original HR images for the detected logos, the quantitative measures used for comparison are non-reference based. We use three measures: The first is the distortion measure (DM) [19] that evaluates the image quality in the frequency domain. The second metric we use is based on the anisotropy (ANIS) [20] which well correlates to the classical reference based metrics such as PSNR. We also apply the recently proposed metric of Cumulative Probability of Blur Detection (CPBD) [21] that focuses on the image sharpness evaluation. Table 1 shows the average scores using different metrics.

Method	bicubic	KR	ICBI	ASDS	Proposed
DM	28.96	31.40	30.46	31.44	37.27
ANIS	0.47	0.57	0.52	0.56	7.886
CPBD	15.22	22.86	11.80	116.88	337.74

Table 1. Average metric scores by DM(in dB) [19], ANIS($\times 10^{-3}$) [20], and CPBD($\times 10^{-3}$) [21] for all the logo images. The higher is the better.

From Table 1 we see that the results by our method yield highest scores for all the metrics. Besides, all the LR images are super-resolved correctly using HOG as the output selection measure. Despite the poor LR image quality due to artifacts such as specularities, noise and blurriness, HOG successfully differentiates makes using gradient information which is not sensitive to the image degradation.

3.3. Effects on Vehicle Logo Recognition

One of the motivations for logo SR is to improve the performance of the subsequent recognition. We use the super-resolved images by different methods as inputs to different classifiers. Table 2 shows logo recognition performance. We compare the recognition performance of HOG and PCA compressed features combined with nearest neighbor and linear SVM classifier. While the recognition performance improves marginally between the other SR methods, the proposed method significantly improves the recognition accuracies irrespective of the feature or the classifier used.

Method	LR	bicubic	KR	ICBI	ASDS	Proposed
HOG+1-NN	53.33	72	77.33	76	64	100
PCA+1-NN	49.33	48	49.33	53.33	50.67	80
PCA+15-NN	50.67	52	54.67	57.33	56	90.67
PCA+SVM	30.67	30.67	34.67	33.33	36	100

Table 2. Average logo recognition accuracies in %.

The average size of the detected LR logo images used in the experiments was 26×46. Table 3 shows the effects of further downsampling the LR logos on recognition accuracy. As expected, the recognition performance deteriorates with reduced resolution. As the structural details of the logos differ and are affected by lower resolution differently, the recognition performance varies significantly among different makes.

Size to LR	Honda	Toyota	Nissan	Ford	Chevy	Overall%
0.75	15	13	13	5	15	81.33
0.5	15	11	13	0	13	69.33
0.25	13	1	7	0	14	46.67

Table 3. Number of logos recognized correctly (out of 15) with downsampled LR image using HOG+1-NN classifier.

4. CONCLUSIONS

In this paper, a manifold learning based super-resolution method for vehicle logos is developed. HR and LR logos are first projected into PCA subspace and canonical correlation analysis (CCA) is applied to create another subspace where the coherence between the projected PCA coefficients of the HR and LR image pairs is enhanced. For each vehicle make a specific model is learned. Given a low-resolution image, it is projected into the CCA subspace and its K nearest neighbors in the CCA subspace of the HR images are used to reconstruct the super-resolved image. We select the final result by finding

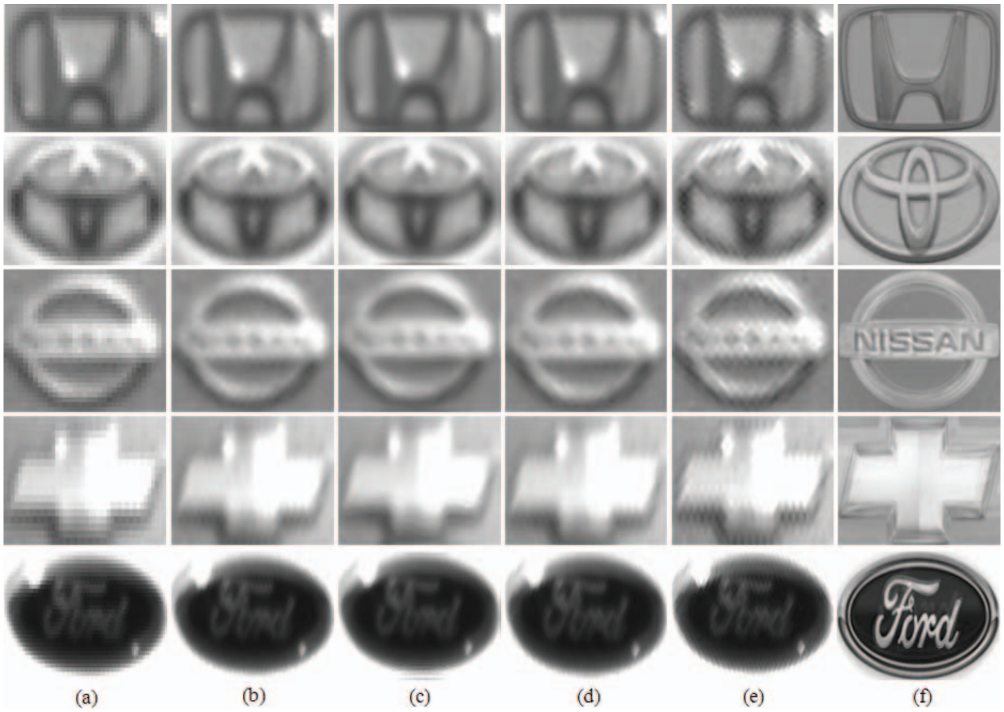

Fig. 3. Sample super-resolution results. (a) Low-resolution images (enlarged by pixel replication). (b) Bicubic interpolation. (c) KR [17]. (d) ICBI [10]. (e) ASDS [18]. (For display purpose all the images are normalized to the same size.) (f) Proposed.

the best match in the HOG feature space between the LR input and the generated HR images using different models. The logo detection and the subsequent super-resolution are automated to ensure practical applicability. Compared with the *state-of-the-art* methods, the proposed method performs the best in both qualitative and quantitative evaluations, and help to achieve the highest recognition accuracy.

Acknowledgment This work was supported in part by NSF grant 0905671 and a grant from Federal Signal Technologies.

5. REFERENCES

[1] V. S. Petrovic and T. F. Cootes, "Analysis of features for rigid structure vehicle type recognition," in *Proc. BMVC*, 2004.

[2] G. Pearce and N. Pears, "Automatic make and model recognition from frontal images of cars," in *Proc. AVSS*, 2011.

[3] A. P. Psyllos, C.-N. E. Anagnostopoulos, and E. Kayafas, "Vehicle logo recognition using a sift-based enhanced matching scheme," *IEEE ITS*, vol. 11, no. 2, pp. 322–328, June 2010.

[4] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: adaptive filtering approach," *IEEE TIP*, 1999.

[5] R.R. Schultz and R.L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE TIP*, 1996.

[6] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," in *Proc. CVPR*, 2000.

[7] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-based super-resolution," *IEEE CG&A*, 2002.

[8] K. I. Kim and Y.H. Kwon, "Example-based learning for single-image super-resolution," in *Proc. DAGM*, 2008.

[9] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. ICCV*, 2009.

[10] A. Giachetti and N. Asuni, "Real time artifact-free image upscaling," *IEEE TIP*, 2011.

[11] H. Huang, H. He, X. Fan, and J. Zhang, "Super-resolution of human face image using canonical correlation analysis," *Pattern Recognition*, 2010.

[12] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face recognition using Laplacianfaces," *IEEE TPAMI*, 2005.

[13] Harold Hotelling, "Relations between two sets of variates," *Biometrika*, 1936.

[14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, 2005.

[15] M. Borga, "Canonical correlation a tutorial," 2001.

[16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2001.

[17] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE TIP*, 2007.

[18] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE TIP*, 2011.

[19] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, and A.C. Bovik, "Image quality assessment based on a degradation model," *IEEE TIP*, 2000.

[20] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," *J. Opt. Soc. Am. A*, 2007.

[21] N.D. Narvekar and L.J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE TIP*, 2011.