Very Low Resolution Face Recognition Problem

Wilman, W.W. Zou and Pong C. Yuen

Abstract—This paper addresses the very low resolution (VLR) problem in face recognition in which the resolution of face image to be recognized is lower than 16x16. The VLR problem happens in many surveillance camera-based applications and existing face recognition algorithms are not able to give satisfactory performance on VLR face image. While face super-resolution (SR) methods can be employed to enhance the resolution of the images, the existing learning-based face SR methods do not perform well on such a very low resolution face image. To overcome this problem, this paper models the SR problem under VLR case as a regression problem with two constraints. First, a new data constraint is design to perform the error measurement on high resolution image space which provides more detailed and discriminative information. Second, discriminative constraint is proposed and incorporated in the training stage so that the reconstructed HR image has higher discriminability. CMU-PIE, FRGC and surveillant camera face (SCface) databases are selected for experiments. Experimental results show that the proposed method outperforms the existing methods, in terms of image quality and recognition accuracy.

I. INTRODUCTION

A. Background and Motivation

While face recognition research has been studying for more than three decades and many promising practical face recognition systems have been developed, it is assumed the face region is large enough and contains sufficient information for recognition [3]. Empirical studies [4] showed that minimum face image resolution between 32x32 and 64x64 is required for existing algorithms. However, a wide range of applications cannot provide enough resolution of face for recognition, such as small-scale stand-alone camera applications in banks and supermarkets, large-scale multiple networked close-circuit television (CCTV) in law enforcement applications in public streets , etc. As shown in Figure 1, when the person is not close to the camera, the face region will be smaller than 16x16 pixels. Recognition of such a very low-resolution face image is called very low resolution (VLR) face recognition problem.

Super-resolution (SR) algorithms have been proposed [5], [6], [7] to increase the resolution of the face images so that the reconstructed higher resolution image is used for recognition. Existing learning-based face SR algorithms can be roughly categorized into two approaches, namely example-based approach [9] [11] [10] [12] [18] and maximum a posterior (MAP) based approach[1] [2] [8]. These two approaches reconstruct the high resolution face images by find the most similar face image from the face image subspace learned by the training face images. They verify the reconstruction

Wilman W.W. Zou and Pong C. Yuen are with Department of Computer Science, Hong Kong Baptist University, Hong Kong {wwzou,pcyuen}@comp.hkbu.edu.hk
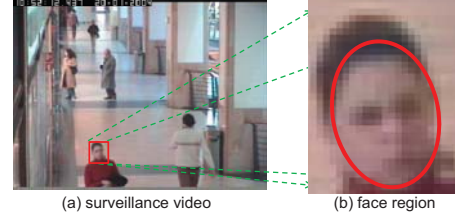
Fig. 1. A typical frame from a surveillance video (CAVIAR database)

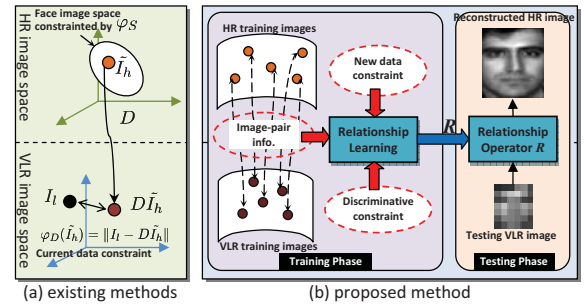(a) existing methods (b) proposed method

Fig. 2. The overview of existing methods vs. the proposed method. The circles represent images.

HR image by comparing the reconstruction HR images to the input low resolution (LR) image in (LR) image space as shown in Figure 2(a). This is always done by minimizing the data constraint [8]. However, this data constraint may not work well under the VLR problem. (Please refer to Section II for more details.)

From recognition perspective, a few methods were developed. Gunturk *et al.* [13] applied MAP-based SR method to reconstruct the eigenface coefficients for face recognition. Hennings-Yeomans *et al.* [14] proposed to perform SR and feature extraction from LR image simultaneously. Wang *et al.* [15] employed example-based approach to reconstructed HR images for face recognition, while Li *et al.* [16] reconstructed the image features, instead of HR images, for face recognition. However, under the VLR problem, these algorithm may not work well because the input LR image does not carry much useful information.

B. Novelty and Contributions of This Paper

To overcome the VLR problem, this paper model the super-resoluion (SR) as a regression problem with two new constraints on the given training data, as shown in Figure 2(b). Unlike the existing methods that recover the images directly from the LR image(s) and the training images, our

method first determines the face image mapping pattern (relationship R) between VLR and HR image spaces which can make use of the useful information from the training data, and then recovers the HR images by applying the relationship operator R on LR images. The proposed new framework two additional advantages:

- With the use of VLR-HR image pair information in the training phase, a new data constraint which performs the error measurement in HR image space, is designed. This helps the proposed SR better handling the VLR problem.
- With the use of class label information, discriminative constraint can be easily integrated when determining the relationship operator R, which is generic for VLR face images in testing phase. To the best of our knowledge, this is the first learning-based SR algorithm which is able to make use of the label information to enhance the discriminability of the face image.

II. LIMITATIONS ON EXISTING LEARNING-BASED FACE SUPER-RESOLUTION ALGORITHMS

Most, if not all, of the existing face SR algorithms are learning-based and are formulated as a two-constraint optimization problem as follows,

$$\tilde{I}_h = \arg\min_{I_h}(\varphi_D(I_h) + \varphi_S(I_h)) \tag{1}$$

where $\varphi_S(.)$ is called algorithm-specific constraint, while $\varphi_D(.)$ is the data constraint[8].

To model the algorithm-specific constraint, Baker and Kanade [1] made use of recognition prior while Liu *et al.* [8] employed Markov random field to do that. Some other researches made use of subspace method to model the algorithm-specific constraint, such as kernel prior [2], Eigentransformation [18]. To model the data constraint, both MAP-based and example-based approaches shares the same data constraint as,

$$\varphi_D(I_h) = \|DI_h - I_l\|^2 \tag{2}$$

MAP-based approach employs the data constraint to model the condition probability $P(I_l|\tilde{I}_h)$, while example-based methods implicitly use the data constraint to determine the weights for the reconstructed HR image.

In VLR problem, such a data constraint may not work well because of the limited information carried by the input VLR image due to low-dimension of the VLR image space. Let the solution space (set) of equation e be $\mathbf{U}(e)$. It can be shown that the solution of Eq.(1), $\tilde{I}_h$, locates in the intersection of two solution spaces of the constraints, as follows:

$$\{\tilde{I}_h\} = \mathbf{U}(\varphi_D \le c_1) \cap \mathbf{U}(\varphi_S \le c_2) \tag{3}$$

where c_1 and c_2 are two positive error terms to control the dimension of the solution space.

Under the VLR face recognition problem, even though c_1 is set to 0, the solution space $\mathbf{U}(\varphi_D = 0)$ is still very large. Therefore, in determining the high resolution image, the algorithm specific constraint $\mathbf{U}(\varphi_s \le c_2)$ will
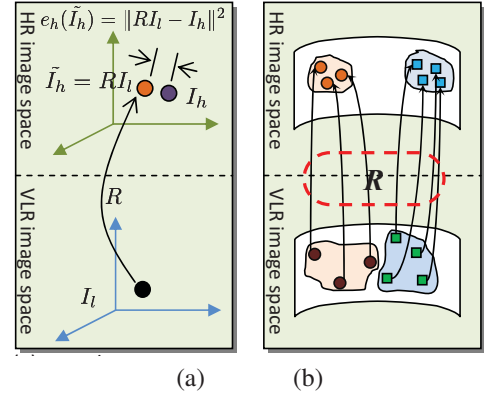

Fig. 3. The illustration of two new constraints. (a) New data constraint. (b) Discriminative constraint.

dominate during the optimization process. In turn, there is a possibility that reconstructed HR image may have serious artifacts and/or not look like the original person.

Moreover, optimization-based algorithms do not fully use all the information of the training data, such as label information which is very important from recognition perspective. This may restrict the performance of the reconstructed HR images in recognition system.

III. PROPOSED RELATIONSHIP LEARNING BASED SUPER-RESOLUTION

As illustrated in Figure 2(a), in this paper, super-resolution problem is modeled by learning the the relationship R from the given training data, and the reconstructed HR image is recovered by applying R on the testing image. The relationship R, namely mapping pattern, is learnt by a regression model which minimizing the proposed new data constraint and discriminative constraint. Unlike existing data constraint verifying the reconstructed HR image in LR image space (Figure 2(a)), the new data constraint perform this in HR image space which help the proposed algorithm better make use of the information from HR training images. Discriminative constraint is designed for recognition purpose, so that the proposed SR algorithm can use the class label information to boost the performance.

A. Super-resolution with New Data Constraint

In the VLR problem, the input VLR image space contains very little useful information, so existing data constraint may not be able to estimate the reconstruction error well. A better method is to estimate such error in HR image space. But the challenge is that given a VLR query image, the corresponding HR image is not available. Instead of recovering the HR image directly, we propose a new SR algorithm to learn the relationship R (in form of matrix) between HR image space and VLR image space. After determining the relationship R, HR image can be reconstructed by R.

Given a set of training HR and VLR image pairs ($\{I_h^i, I_l^i\}_{i=1}^N$), let R be the relationship between the HR

image space and the VLR image space, so that the HR image can be reconstructed from its VLR image and R as follows,

$$\tilde{I}_h = RI_l + n \qquad (4)$$

where n is the noise. So the reconstruction error in HR image space, $e_h(\tilde{I}_h)$, is estimated by

$$e_h(\tilde{I}_h) = \|\tilde{I}_h - RI_l\|^2 \qquad (5)$$

To determined the best R, the reconstruction error should be minimized. So we have,

$$R = \arg\min_{R'} \sum_{i=1}^{N} \|I_h^i - R'I_l^i\|^2 \qquad (6)$$

It can be shown that R is unique if the number of training image pairs N is larger than the data dimension of VLR image space d_L. $N > d_L$ is true in many VLR face recognition applications. Eq.(6) can be minimized by using gradient-descent manner effectively.

In query stage, given a testing VLR image I_l, the corresponding HR image is recovered as follows,

$$\tilde{I}_h = RI_l \qquad (7)$$

It can be seen that the reconstruction of HR images is equivalent to applying R on the input space. We call this method as relationship learning (RL) based super-resolution.

Error Analysis

Considering a HR image I_h consists of two components namely, the low frequent image component l and high frequent image details h; and they satisfy $Dh = 0$ and $Dl = I_l$. The HR image can then be represented as

$$I_h = l + h \qquad (8)$$

The reconstruction error can also be separated into two parts namely, low frequent image component error Δl and high frequent details error Δh. So the reconstructed HR $\tilde{I}_h$ is given by,

$$\tilde{I}_h = I_h + \Delta l + \Delta h \qquad (9)$$

For the current data constraint used in existing methods, the error is formulated as

$$\begin{aligned}\varphi_d(\tilde{I}_h) = e_l(\tilde{I}_h) &= \|D\tilde{I}_h - I_l\|^2 \\ &= \|D(I_h + \Delta l + \Delta h) - I_l\|^2 = \|\Delta I_l\|^2\end{aligned} \qquad (10)$$

so $e_l(\tilde{I}_h)$ reflects the error introduced by low frequent image component only. For our proposed method, we have

$$\begin{aligned}e_h(\tilde{I}_h) &= \|\tilde{I}_h - I_h\|^2 \\ &= \|I_h + \Delta l + \Delta h - I_h\|^2 = \|\Delta l + \Delta h\|^2\end{aligned} \qquad (11)$$

That means our proposed method can estimate the reconstruction error from both low frequent image component l and the high frequent details h.

In turn, both Δh and Δl will be minimized when $e_h(\tilde{I}_h)$ is minimized, while minimizing $e_l(\tilde{I}_h)$ only leads Δl to be minimized.

From the discussion in Section II, in VLR face recognition problem, the reconstruction HR image $\tilde{I}_h$ easily satisfies

$$\|\Delta l\|^2 = 0 \qquad (12)$$

due to the large solution space of $\mathbf{U}(\|D\tilde{I}_h - I_l\|^2 = 0)$. This implies that the major reconstruction error is caused by $\|\Delta h\|^2$. To get better HR image quality, error caused by both Δl and Δh should be minimized. Our proposed data constraint can properly estimate the reconstructed error.

B. Discriminative Super-Resolution

It can be seen that R restricts the reconstructed HR images locating in a subspace which minimizes the reconstruction error $e_h(\tilde{I}_h)$. This inspires us to find a better subspace induced by R with other additional constraint(s), so that the reconstructed HR images have more discriminative features. In order to further boost the discriminability of the reconstructed HR image, discriminative constraint is designed to the relationship learning based SR algorithm to determine the "optimal" R. Discriminative super-resolution (DSR) algorithm is proposed.

A natural step is to make use of the class information of the training data. From recognition perspective, we expect the reconstructed HR images should be clustered with the images from the same class, and far away from the images from other classes, as shown in Figure 3. Therefore, inspired by maximum margin criterion (MMC) [17], we design a discriminative constraint as follows:

$$\begin{aligned}d(R) = \ &\text{mean}(\{\|I_h^i - RI_l^j\|^2 | \Omega(I_h^i) = \Omega(I_l^j)\}) \\ &- \text{mean}(\{\|I_h^i - RI_l^j\|^2 | \Omega(I_h^i) \neq \Omega(I_l^j)\})\end{aligned} \qquad (13)$$

where $\Omega(u)$ is the class label of u. Integrating Eq.(13) with Eq.(6), the new discriminative super resolution formula can be written as :

$$\dot{R} = \arg\min_{R'} \frac{1}{N} \sum_{i=1}^{N} \|I_h^i - R'I_l^i\|^2 + \gamma d(R') \qquad (14)$$

where γ is a constant to balance the above two terms. In our experiments, γ is set to 1. The HR image can be reconstructed after $\dot{R}$ is determined.

The subspace induced by $\dot{R}$ is a subspace which is optimized for recognition with respect to MMC. That means the HR images reconstructed by $\dot{R}$ locates in a subspace where they can be better linear separable. Therefore, the HR image reconstructed by $\dot{R}$ will contain more discriminability and be better for recognition purpose.

IV. EXPERIMENTS AND ANALYSIS

Two experimental results are reported in this section. In the first experiment, we would like to evaluate the effectiveness of the proposed new data constraint as described in Section 3.1 (RL). The quality of reconstructed HR images was evaluated by both objective measurement in terms of mean squared error (MSE), information entropy, as well as subjective human visual quality. The second experiment

is to evaluate the performance of the reconstructed HR image for face recognition in terms of recognition accuracy. Comparison with existing face SR methods is also reported.

Three databases, two public available face databases (CMU PIE and FRGC 2.0) and one surveillant camera face database, SCface [19] are used for experiments. For CMU-PIE database, a subset of 68 classes with 21 different illuminations is considered. For FRGC database, the subset of 311 classes with 10 images (with illumination, expression and mild pose variations) per class is used. For SCface database, 10 images (5 LR images and 5 HR images) per class are used in our experiments, totally 130 classes. The HR images and LR images are extracted from the surveillant cameras. The face images of SCface have very challenging illumination and pose variations, which cause the very low benchmark performance in which the recognition rate is ranging from 0.7% to 7.7% as reported in [19].

In experiments on CMU-PIE and FRGC, the VLR image(7x6) is generated by downsampling the HR image (56x48). To determine the relationship R and train the recognition engine, 13 and 8, HR and VLR image pairs from each person are randomly selected from CMU PIE and FRGC databases for training, respectively, while the rest are used for testing. In experiments on SCface database, we manually extract the face images and normalize to the resolution of 64x56 (HR images) and 16 x 14 (LR images). HR images are used as training data and gallery data, while the LR images are used as testing data.

Three existing face SR methods, namely Hallucination Face (HF) method [1], Eigentransfromation based Face SR (EF) method [18] and KPCA-based Face SR (KF) method [2] are selected for comparison.

A. Experiment 1: Effectiveness of the New Data Constraint

To estimate the effectiveness of proposed data constraint, we apply the proposed RL method on CMU-PIE and FRGC images. Because we do not have original HR and LR image pairs information in SCface database, we do not carry out experiments on SCface database in this Section.

Figure 4 shows some of the reconstructed images using proposed RL method and existing methods. Figures 4(a) and (g) show the input 7x6 query image and original 56x48 HR image. Figures 4 (b) - (e) display the results using bicubic interpolation method, HF [1] method, EF [18] method and KF [2] method. It can be seen that both bicubic interpolation (BC) method and KF method give a relatively blur image and high frequency details cannot be recovered. Both HF method and EF method could recover some high frequent details. However, HF method generates some artifacts which degrade the human visual quality. The visual quality of the reconstructed images from EF method are good. However, when comparing with the original HR image, these HR image does not look like the original HR image. Figure 4(f) shows the results using our proposed method. It can be seen that the proposed method gives a good visual quality image which also look like the original one.

Database	BC	HF [1]	EF [18]	KF [2]	Proposed RL
CMU PIE	424.4	475.6	291.9	1143.1	179.9
FRGC 2.0	1259.4	1838.8	1510.1	1707.6	870.5

TABLE I

THE MSE OF DIFFERENT SR METHODS

Fig. 4. SR results: (a)input VLR images (7 x 6), (b) SR results by Bicubic interpolation,(c) by Hallucination Face method (HF)[1], (d) Eigentransfromation based Face SR method (EF) [18], (e) KPCA-based Face SR method (KF) [2], (f) Our proposed method (RL), (g) original HR images. The resolution of reconstructed HR images is 56 x 48.

Moreover, two objective measurements, namely mean squared error (MSE) and Shannon information entropy are calculated and reported in Table I and Table II respectively. The error and entropy reported in the table are the average of all testing images. As shown in Table I, the proposed method give the smallest MSE. That means the proposed new data constraint can properly restrict the error. As shown in Table II, the proposed method recover the HR images with higher information entropy. This implies that the our proposed new data constraint helps the SR algorithm recover images with more information while the MSE is small.

B. Experiment 2: Image Discriminability

In this experiment, we would like to evaluate the performance of the proposed discriminative super resolution (DSR) algorithm in terms of the recognition result.

First, we conduct the experiments on CMU-PIE and FRGC databases. Recognition experiments are performed on (i) input VLR query images, (ii) original HR images, reconstructed HR images from (iii) HF [1], (iv) EF [18], (v) KF [2] and (vi) the proposed DSR method. Three popular face recognition algorithms, namely eigenface (linear method), KPCA (kernel method) and SVM (discriminative method), are used for experiments.

Second, we conduct the experiments on a surveillant video

Database	Original HR	BC	HF [1]	EF [18]	KF [2]	Proposed RL
CMU PIE	7.10	6.75	7.04	7.03	6.38	7.06
FRGC 2.0	7.92	7.57	7.82	7.75	7.40	7.83

TABLE II

THE IMAGE ENTROPY OF DIFFERENT SR METHODS (AVERAGE OF ALL TESTING IMAGES)
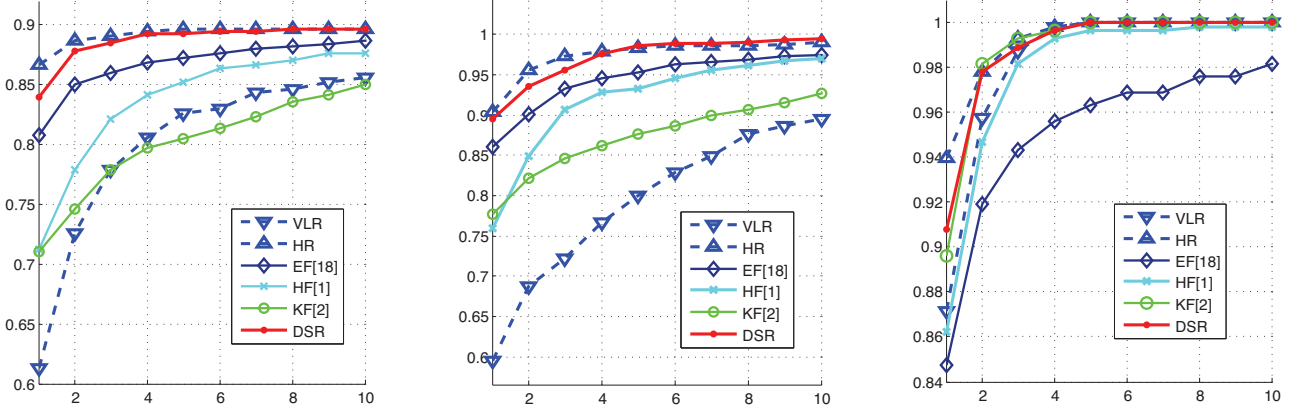

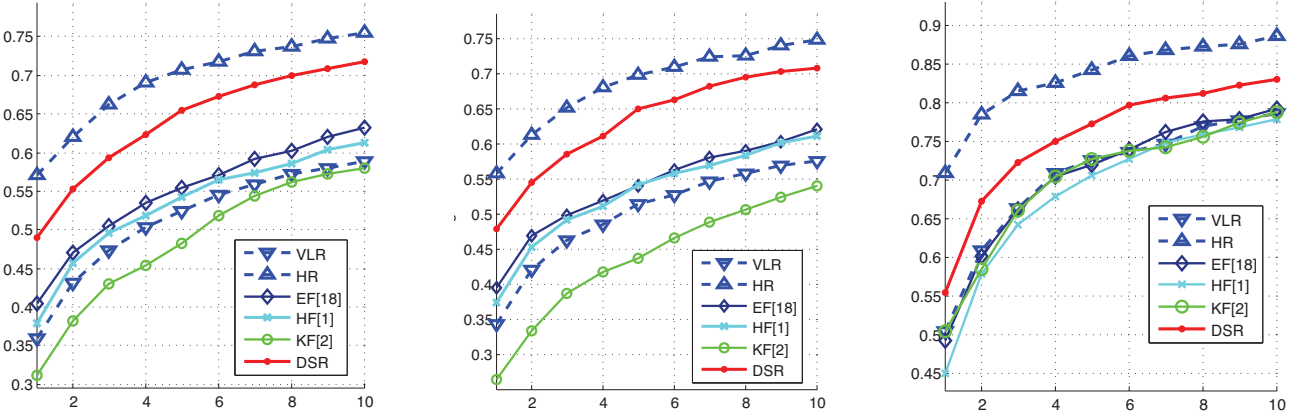Fig. 5.   Results on CMU PIE: Eigenface (left); Kernel PCA (middle); SVM (right)

Fig. 6.   Results on FRGC: Eigenface (left); Kernel PCA (middle); SVM (right)
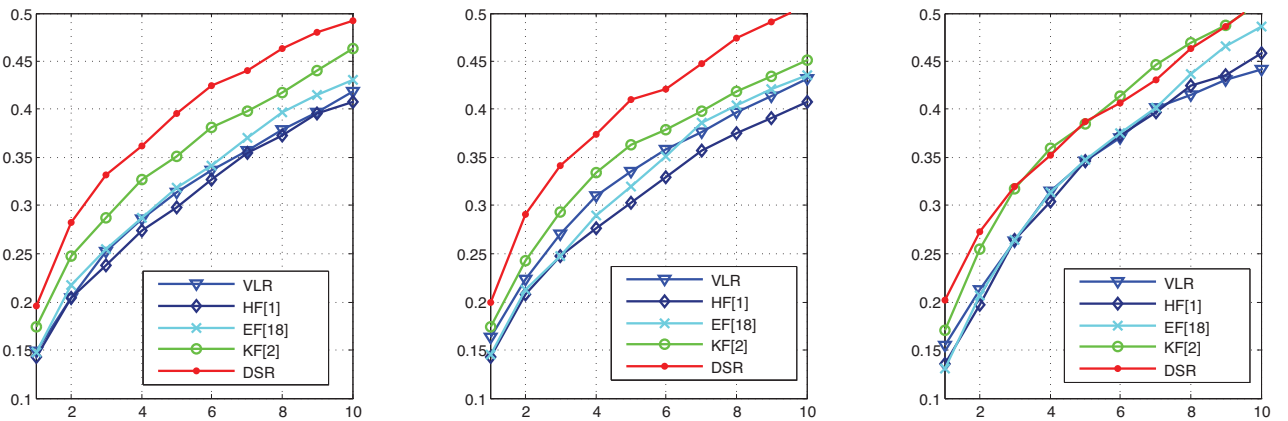

Fig. 7.   Results on SCface: Eigenface (left); Kernel PCA (middle); SVM (right)

Database	face recognition algorithm	VLR image	original HR image	HF [1]	EF [18]	KF [2]	Proposed DSR
CMU PIE	Eigenface	61.3	86.7	80.8	71.3	71.7	**83.9**
	Kernel PCA	59.5	90.4	86.0	75.9	77.8	**89.5**
	SVM	87.1	93.9	84.7	86.2	89.6	**90.8**
FRGC 2.0	Eigenface	36.0	57.1	40.5	37.9	31.1	**49.0**
	Kernel PCA	34.4	55.8	39.6	37.5	26.4	**47.9**
	SVM	50.4	70.9	49.2	45.0	50.4	**55.5**
SCface	Eigenface	14.9	/	14.3	14.7	17.3	**19.6**
	Kernel PCA	16.3	/	14.3	14.5	17.3	**20.0**
	SVM	15.5	/	13.5	13.1	17.0	**20.2**

TABLE III

RANK 1 RECOGNITION RATE (%).

SCface database. Comparison results are also reported. The SCface database is very challenging, and the benchmark results show that the recognition rate on low resolution images is only 0.7% to 3.1% [19], when one single HR image is used as gallery data per person. The results are plotted in Figures 5 for CMU PIE database, Figures 6 for FRGC v2.0 database and Figures 7 for SCface database, in terms of the CMC curves. The rank 1 recognition results are recorded in Table 2.

The experimental results show that:

- There is a significant drop of recognition accuracy (as high as 30%) for VLR image, comparing with the original HR image, for all recognition engines on both CMU-PIE and FRGC databases.
- The proposed method outperforms existing SR methods. It implies that the reconstructed HR image using the proposed method has high discriminability for recognition purpose.
- The recognition accuracy of the proposed method on real surveillance video also have considerable improvement, even though such database is very challenging.

V. CONCLUSION

The very low resolution face recognition problem is defined and discussed in this paper. To solve the problem, a new learning-based face super-resolution framework, a new data constraint and a learning-based discriminative super-resolution algorithm are developed and reported. Comprehensive experiments on three publicly available face databases are performed. Experimental results show that the proposed algorithm outperforms existing algorithms at VLR, in terms of both image quality and recognition accuracy.

Acknowledgements

REFERENCES

[1] Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 1167–1183

[2] Chakrabarti, A., Rajagopalan, A.N., Chellappa, R.: Super-resolution of face images using kernel pca-based prior. IEEE Transactions on Multimedia **9** (2007) 888–892

[3] Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Surveys (CSUR) **35** (2003) 399–458

[4] Lui, Y.M., Bolme, D., Draper, B.A., Beveridge, J.R., Givens, G., Phillips, P.J.: A meta-analysis of face recognition covariates. In: Proceedings of International Conference on Biometrics: Theory, Applications and Systems(BTAS). (2009)

[5] Cristobal, G., Gil, E., Sroubek, F., Flusser, J., Miravet, C., Rodriguez, F.B.: Superresolution imaging: a survey of current techniques. Advanced Signal Processing Algorithms, Architectures, and Implementations XVIII **7074** (2008) 70740C–70740C18

[6] van Ouwerkerk, J.: Image super-resolution survey. Image and Vision Computing **24** (2006) 1039 – 1052

[7] Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine **20** (2003) 21–36

[8] Liu, C., Shum, H.Y., Freeman, W.T.: Face hallucination: Theory and practice. International Journal of Computer Vision **75** (2007) 115–134

[9] Zhang, W., Cham, W.K.: Learning-based face hallucination in dct domain. In: Proceedings of IEEE International Conference on CVPR. (2008) 1–8

[10] Liu, W., Lin, D., Tang, X.: Hallucinating faces: Tensorpatch super-resolution and coupled residue compensation. In: Proceedings of IEEE International Conference on CVPR. Volume 2. (2005) 478 – 484

[11] Jia, K., Gong, S.: Generalized face super-resolution. IEEE Transactions on Image Processing **17** (2008) 873–886

[12] Park, J.S., Lee, S.W.: An example-based face hallucination method for single-frame, low-resolution facial images. IEEE Transactions on Image Processing **17** (2008) 1806–1816

[13] Gunturk, B., Batur, A., Altunbasak, Y., Hayes, M., Mersereau, R.: Eigenface-domain super-resolution for face recognition. IEEE Transactions on Image Processing **12** (2003) 597–606

[14] Hennings-Yeomans, P.H., Baker, S., Kumar, B.: Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In: Proceedings of IEEE International Conference on CVPR. (2008) 1–8

[15] Wang, X., Tang, X.: Face hallucination and recognition. LNCS: Advances in Neural Networks (2003) 486–494

[16] Li, B., Chang, H., Shan, S., Chen, X., Gao, W.: Hallucinating facial images and features. In: Proceedings of IEEE International Conference on Pattern Recognition. (2008) 1–4

[17] Li, H., Jiang, T., Zhang, K.: Efficient and robust feature extraction by maximum margin criterion. IEEE Transactions on Neural Networks **17** (2006) 157–165

[18] Wang, X., Tang, X.: Hallucinating face by eigentransformation. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews **35** (2005) 425–434

[19] Grgic, M., Delac, K., Grgic, S.: SCface - surveillance cameras face database. Multimedia Tools and Applications Journal DOI:10.1007/s11042-009-0417-2 (2009)