

# A SUPER-RESOLUTION METHOD FOR LOW-QUALITY FACE IMAGE THROUGH RBF-PLS REGRESSION AND NEIGHBOR EMBEDDING

Junjun Jiang, Ruimin Hu, Zhen Han, Tao Lu, Kebin Huang

National Engineering Research Center for Multimedia Software  
School of Computer, Wuhan University, Wuhan, 430072, China

## ABSTRACT

In this paper, a new two-step method is proposed to infer a high-quality and high-resolution (HR) face image from a low-quality and low-resolution (LR) observation based on training samples in the database. First, a global face image is reconstructed based on the non-linear relationship between LR and HR face images, which is established according to radial basis function and partial least squares (RBF-PLS) regression. Based on the reconstructed global face patches manifold (formed by the image patches at the same position of all global face images), whose *local geometry* is more *consistent* with that of original HR face patches manifold than noisy LR one is, the Neighbor Embedding is applied to induce the target HR face image by preserving the similar local geometry between global face patches manifold and the original HR face patches manifold. A comparison of some state-of-the-art methods shows the superiority of our method, and experiments also demonstrate the effectiveness both under simulation and real conditions.

**Index Terms**—Manifold learning, Super-resolution, Face hallucination, RBF-PLS, Neighbor Embedding

## 1. INTRODUCTION

Nowadays, surveillance cameras have been widely used in security and protection systems. They can provide very important clues about objects for solving a case, such as criminals. However, the object is so far away from the camera that the resolution of the interested face in the picture is too low to provide helpful information. Additionally, in many real surveillance scenarios, the qualities of the surveillance images are very poor because of the influences of many factors, such as underexposure, optical blurring, defocusing and so on. Due to the low resolution and noise disturbance, the face images of interest lose too many detailed facial features to be identified by human. Therefore, in order to obtain enough facial feature details for recognition, it is necessary to infer a high-quality and HR face image from a low-quality and LR one and this technique is called face super-resolution (SR) or face hallucination [1-11].

---

The research in this paper use the CAS-PEAL-R1 face database collected under the sponsor of the Chinese National Hi-tech Program and ISVISION Tech. Co. Ltd. The research was supported by the major national science and technology special projects (2010ZX03004-003-03), the National Basic Research Program of China (973 Program) (2009CB320906), the National Natural Science Foundation of China (61172173, 60832002, 60970160, 61070080, 61003184), the National Science Foundation of Hubei Province of China (2009CDA134, 2010CDB05103), the Fundamental Research Funds for the Central Universities (201121102020009, 3101011), the Hubei Province Key Laboratory of Intelligent Robot (HBRI 200902).

Up to now, many face SR methods have been proposed. And they can be divided into two categories: reconstruction-based methods and learning-based methods. Between them, learning-based methods have received more attention because they can achieve higher magnification factor and produce better super-resolved results. In this paper, we only focus on the learning-based method whose input is a single frontal face.

The common idea of learning-based methods is to infer HR face by training the relationship between LR and HR image pairs. For example, Baker and Kanade [1] propose a learning-based SR method named “*face hallucination*”, and it is the first SR method targeted at face images specially. Liu et al. [2] propose to integrate a global parametric model and a local nonparametric model for face super-resolution. Following the work of [1-2], learning-based methods draw enormous attention in the SR research community.

In general, face images are a class of highly structured objects and have similar appearances. Machine learning theory suggests that face images reside on a non-linear low-dimensional manifold and span a small subset in the high-dimensional image space [13]. Inspired by the manifold learning results, a series of SR algorithms with a *manifold assumption* have been proposed. The manifold assumption states that the LR and HR image manifolds share similar local geometry. Chang et al. [6] first use Neighbor Embedding algorithm for SR of general images. Then, such manifold techniques have been used for global face reconstruction [4-5], local detail enhancement [7] or both [8]. However, the projection of LR to HR image is “one-to-multiple” mapping, so the manifold assumption will not hold well. Li et al. [9] design two projection matrixes to project original coupled manifolds (LR and HR image manifold) to a common manifold for face super-resolution. Huang et al. [10] propose a manifold learning based two-step method. They apply canonical correlation analysis (CCA) [17] to maximize the correlation between the local neighbor relationships of LR and HR images both in global face reconstruction and residual face compensation. However these methods do not consider other degrading cause, such as blurring and noise, they simply use down-sampled images as input. From the core idea of manifold learning based approaches, we learn that they tend to explore the global features from local geometry structure of the training data. Therefore, when the training sets are contaminated by noise, even small noise, the consistency of local geometric structure of LR and HR image manifolds will be affected badly [13], which will result in the invalidation of manifold learning based face super-resolution methods.

In order to reveal the underlying relationship between noisy LR images and HR images, we have analyzed noisy LR images to find out that there are three kinds of information in the noisy LR

images: useful information to expression of HR images, redundant information and noise. What we want is to extract the useful information for modeling the relationship and eliminate the redundant information as well as noise.

Combining advantages of linear nature in parameters of RBF and the ability of PLS to maximize the squared covariance between training sets, we recover the non-linear relationship between LR and HR training pairs using RBF-PLS regression [14] in this paper. Motivated by the two-step framework, we firstly introduce RBF-PLS regression to learn the relationship between LR and HR training pairs, and the inference (global face) can be done by substituting a LR observation into the established model. In the second step, we learn the similar local geometry between the original HR face manifold and the obtained global face manifold, whose local geometry is more consistent with that of original HR face manifold than noisy LR one is, thereby leading to better results. We conduct experiments on the CAS-PEAL-R1 database [15] to verify the effectiveness of our method under simulation and real conditions.

## 2. RADIAL BASIS FUNCTION-PARTIAL LEAST SQUARES (RBF-PLS) REGRESSION

RBF-PLS is a regression technique, and it can be used to establish the non-linear relationship between two training sets [14]. Compared with PCA method used by Liu [2] and Wang [4], RBF-PLS regression not only effectively generalizes the information of independent variables system (LR faces), but also correctly explains the dependent variables system (HR faces) while excludes the impact of noise in the training sets by maximizing the covariance between elements in independent variables system and dependent variables system. In this section, we briefly introduce the RBF-PLS technique.

Define  $x_1, \dots, x_m$  as a set of  $n$ -dimension independent variables, then the data matrix is  $X = [x_1, \dots, x_m]$ ,  $X \in \mathfrak{R}^{m \times n}$ ; and  $y_1, \dots, y_m$  as a set of  $l$ -dimension dependent variable, and the data matrix is  $Y = [y_1, \dots, y_m]$ ,  $Y \in \mathfrak{R}^{m \times l}$ . RBF-PLS aims at establishing the non-linear relationship between the observation  $X$  and the target  $Y$ . Specially, it carry out the non-linear transformation of the observation  $X$ , forming an activation matrix  $A$ , and then the PLS is applied to the activation matrix  $A$  and the target  $Y$ .

### 2.1 RBF Network

In RBF network, Gaussian function is the most commonly used radial basis function to carry out the non-linear transformation of  $X$  and form the activation matrix  $A$ :

$$A = \begin{Bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{Bmatrix} \quad (1)$$

The elements of matrix  $A$  are defined as follows:

$$a_{ij} = \exp(-\|x_i - c_j\|^2 / \sigma_j^2), \quad i, j = 1, \dots, m, \quad (2)$$

where  $x_i$  is the  $i$ -th column of the observation  $X$ ,  $\|\bullet\|$  is a distance measurement (e.g., the Euclidean Distance),  $c_j$  denotes the center of the  $j$ -th radial basis function and  $\sigma_j$  is the corresponding width. We limit our considerations to the width  $\sigma_j$  with a

constant width  $\sigma = 32$ , and the center of the  $j$  th RBF equals to the  $j$  th observation of  $X$ ,

$$c_j = x_j, \quad j = 1, \dots, m \quad (3)$$

### 2.2 PLS Method for the RBF Network

After the activation matrix  $A$  is calculated, the PLS method can be applied to model the relationship between activation matrix  $A$  and the target  $Y$ . For the centered  $A$  and  $Y$ , we can set up the following linear PLS model:

$$Y = TC + F = AWC + F, \quad (4)$$

where  $T$  represents the low-dimensional score matrix of  $A$  with the dimension of  $m \times m_\tau$  ( $m_\tau < m$ ), and the scores of  $A$  are the linear combinations of the RBF maximizing the covariance between  $A$  and  $Y$ .  $C$  is the regression coefficient matrix,  $W$  is the transformation matrix between  $A$  and  $T$ ,  $F$  is the error matrix. Here, the value of  $m_\tau$  is a key factor for making a balance between reducing the affection of noise and improving the preciseness of the model. For more details of the PLS modeling, we refer the reader to [18].

The non-linear relationship between  $X$  and  $Y$  is turned into the linear algebra problem after applying PLS regression to construct the RBF network, and the resulting RBF-PLS regression can be used to calculate the generalization. Given the practical observation  $X_i$ , the RBF-PLS regression can be used to predict the response  $Y_i$ . The centered  $X_i$  is used to calculate the activation matrix  $A_i$ . Then, the response  $Y_i$  is

$$Y_i = A_i W C \quad (5)$$

## 3. A TWO-STEP FACE SUPER-RESOLUTION METHOD

### 3.1 Global Face Reconstruction

In this subsection, we apply the RBF-PLS to model the non-linear relationship between the LR and HR images. Given a LR input face, we use the model to predict the global face. Let  $I^L = \{I_i^L\}_{i=1}^m = [I_1^L, \dots, I_m^L] \in \mathfrak{R}^{l \times m}$  and  $I^H = \{I_i^H\}_{i=1}^m = [I_1^H, \dots, I_m^H] \in \mathfrak{R}^{n \times m}$  represent the LR and HR training face images, respectively, where  $m$  is the number of training images. Every column of  $I^L$  and  $I^H$  is the observation after reshaping all the pixels of LR and HR face image.  $l$  and  $n$  are the dimension of LR and HR face image, where  $l = s^2 n$ , and  $s$  is the down-sampling factor. Based on RBF-PLS technique presented in Section 2, we establish the non-linear regression relationship between LR and HR images by replacing  $X$  and  $Y$  with  $I^L$  and  $I^H$ , respectively. When inputting a LR observation  $I_i^L$ , the global face  $I_i^G$  can be reconstructed by the established regression model. The main steps of the reconstruction of global face are as follows:

- Construct the activation matrix  $A$ , whose elements are defined as follows:

$$a_{ij} = \exp(-\|I_i^L - I_j^L\|^2 / \sigma^2), \quad i, j = 1, \dots, m \quad (6)$$

- Mode matrix  $A$  and matrix  $I^H$  using PLS,  $A$  and  $I^H$  project to the low-dimension matrix  $T$  whose dimension is  $m \times m_\tau$  ( $m_\tau < m$ ), respectively:

$$I^H = TC + F = AWC + F \quad (7)$$

c) Predict the HR global image from the LR observation  $I_i^L$ : After constructing the activation matrix  $A_i$ , where  $A_i = \exp(-\|I_i^L - I_j^L\|^2 / \sigma^2)$   $i, j = 1, \dots, m$ , we can predict the global face:

$$I_i^G = A_i W C \quad (8)$$

If we select a proper value of  $m_\tau$ , the reconstructed global face will maintain primitive facial features and be immune to noise.

### 3.2 Position Patch Based Face Super-resolution via Neighbor Embedding

In the training phase, we gain the global faces  $I^G = \{I_i^G\}_{i=1}^m$  for LR training set based on the method presented above. Each image in the training sets,  $I^G$  and  $I^H$ , is divided into a set of  $p$  small overlapping patch sets  $\{I_{i,j}^G | I_{i,j}^G \in I_i^G, 1 \leq i \leq m, 1 \leq j \leq p\}$  and  $\{I_{i,j}^H | I_{i,j}^H \in I_i^H, 1 \leq i \leq m, 1 \leq j \leq p\}$  using the same division configuration.

The global face image patches manifold, which is formed by the image patches of all global faces at the same position, is more consistent with that of original HR patches manifold than noisy LR one is. We introduce the *neighborhood preservation rate*<sup>1</sup> [12] to measure the consistency between coupled manifolds. In Fig. 1, we can see that neighborhood preservation rate is increased after global face reconstruction.

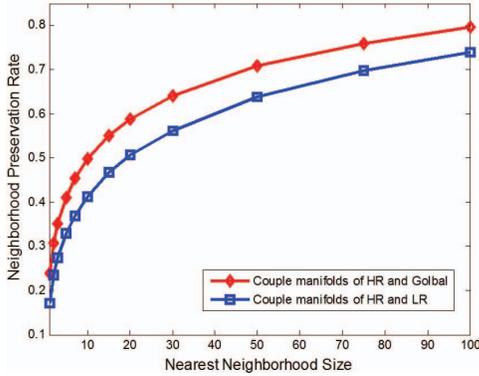


Fig.1 neighborhood preservation rates between couple manifolds of LR patches and HR patches, and global patches and HR patches.

In the testing phase, for a novel LR face input  $I_i^L$ , we first reconstruct the global face  $I_i^G$  using the algorithm proposed in Section 3.1. And then, divide it into overlapping patches,  $\{I_{i,j}^G | I_{i,j}^G \in I_i^G, 1 \leq j \leq p\}$ . For each patch  $I_{i,j}^G$ , we apply Neighbor Embedding to determine the HR patch  $I_{i,j}^H$  using the LR and HR training pairs of position patches [11] as  $I_{i,j}^G$ . For more detail of this Neighbor Embedding method, please refer to [6].

## 4. EXPERIMENTS AND RESULTS

### 4.1 Database

Our experiments are conducted on 640 images with neutral expression and frontal pose of the CAS-PEAL-R1 database [15]. We randomly select 600 images for training and leave the other 40 for

<sup>1</sup> The higher neighborhood preservation rate value is, the more consistent between coupled manifolds are.

testing. The images are aligned by five manually selected feature points and cropped to  $112 \times 96$  pixels. All images are smoothed (an averaging filter  $H$  of size  $4 \times 4$ ) and down-sampled by a factor of 4 to LR  $28 \times 24$  pixels images with the Poisson noise added (which well models real noise).



Fig.2. Simulation results of different methods. (a) Input LR faces. (b) Wang's Results. (c) Huang's global faces. (d) Huang's results after residual compensation. (e) Global faces based on RBF-PLS. (f) Our final results. (g) Original HR faces. (Note that the effect is more pronounced if you resize the figure yourself on the electronic version, and so does Fig.3.)

### 4.2 Simulation Results

In this experiment, we first construct the global images using the RBF-PLS method. The parameters of the regression model are set as:  $\sigma = 32$ ,  $m_\tau = 50$ . Next, we perform Neighbor Embedding algorithm to inform the target HR images. The number of the nearest neighbors  $K$  is set to 50. The patch size is  $8 \times 8$  with an overlap of 6 pixels between adjacent patches. Note that all parameters are set empirically. Due to space limitations, we only exhibit the results of eight test images in Fig.2. Our global faces remove most of the noise of the input faces, while maintaining the primitive facial feature information as illustrated in Fig.2 (e). The final results (Fig. 3(f)) have much clearer detail features and fewer artifacts at edges, and they are good approximations to the original HR images. We also compare our method with Wang's eigentransformation method [4] and Huang's manifold learning based two-step method [10]. In order to obtain the best performance, we adjust parameters to the optimal for each comparative method. Fig.2 (b) is the results of Wang's method. Fig.2 (c) and Fig.2 (d) are Huang's global and

residual faces, respectively. It is observed that ghost effect exists in the hallucinated faces reconstructed by Wang’s method. Meanwhile, they are similar to the mean face for the PCA bases are holistic. Compared with Huang’s method, our method can generate better performance both in global faces and final results, especially in face contour and eyes. Simulation experiments demonstrate that our approach is able to generate HR face images with visually satisfactory global face appearance and local detailed features. Comparative experiments show the superiority of our method over Wang and Huang’s approaches.

Table 1. RMSE, PSNR (dB) and SSIM comparisons

	Bicubic	Wang’s	Huang’s method		Our method	
		method	Step1	Step2	Step1	Step2
RMSE	7.248	7.255	6.817	7.011	6.613	<b>6.356</b>
PSNR	21.84	23.79	24.73	24.35	25.14	<b>25.68</b>
SSIM	0.6456	0.7504	0.8099	0.7800	0.8149	<b>0.8268</b>

We calculate the root mean square error (RMSE), peak signal-to-noise ratio (PSNR) and SSIM index [16] for the 40 test images for each method, and give the results in Table. 1. Again, we can see that our method consistently outperforms all the other methods. Note that because Huang’s method fails to explore the manifold structure when affected by noise, and in some cases performs even worse than its first global face reconstruction results.

### 4.3 Results of Real Condition

In order to further testify the effectiveness of our method, we perform experiments according to real surveillance camera condition. Fig.3 (a) is a picture with a CIF-size ( $352 \times 288$ ) taken by a surveillance camera in the condition of underexposure. We extract the interested face and crop it to  $56 \times 48$  pixels, as can be seen in Fig.3 (b). Fig.3 (c) is the input LR face by converting Fig.3 (b) to grayscale and adjusting its levels. We set the values of all the parameters equal to those mentioned in Section 4.1 except for down-sampling factor, which is set to 2. Fig.3 (d)-(i) are the subjective results comparison of different methods. Actual results show that our method is effective.

## 5. CONCLUSION

In real video surveillance scenarios, the quality of the surveillance images is affected by many factors, and the interested objects (such as faces) are low-quality and LR. In this paper, we propose a new two-step method to hallucinate a high-quality and HR face from a single low-quality and LR frontal face image. First, a non-linear regression method based on RBF-PLS technique is proposed to render a global HR face, which is robust to noise. Then we construct the target HR face through position patch based Neighbor Embedding. Experimental results show that the proposed method is effective not only in simulation condition but also in the real surveillance cameras conditions.

## 6. REFERENCES

[1] S. Baker and T. Kanade. Hallucinating faces. In *FG*, Grenoble, France, Mar. 2000, 83-88.  
[2] C. Liu, H.Y. Shum, and C.S. Zhang. A two-step approach to hallucinating faces: global parametric model and local nonparametric model. In *CVPR*, pp. 192–198, 2001.

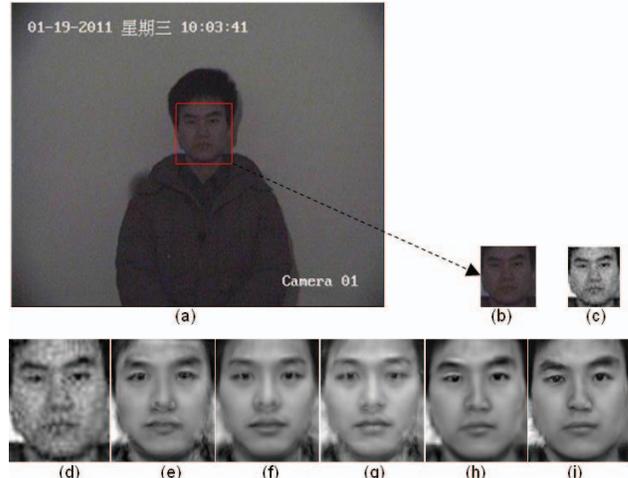


Fig.3. Results of real condition. (a) A picture from surveillance camera. (b) The interested face. (c) Input LR image. (d) Bicubic interpolation result of (c). (e) Wang’s Result. (f) Huang’s global face. (g) Huang’s result after residual compensation. (h) Global face based on RBF-PLS. (i) Our final result.

[3] W. Freeman, E. Pasztor, and O. Carmichael. Learning low-level vision. In *IJCV*, 40(1):25–47, 2000.  
[4] X.G. Wang and X.O. Tang. Hallucinating face by eigen-transformation. *IEEE TSMC-C*, 35 (3):425–434, 2005.  
[5] B. Kumar and R. Aravind. Face hallucination using OLPP and Kernel ridge regression. In *ICIP*, pp. 353–356, 2008.  
[6] H. Chang, D.Y. Yeung, and Y.M. Xiong. Super-resolution through neighbor embedding. In *CVPR*, pp. 275–282, 2004.  
[7] S.W. Park and M. Savvides. Breaking the Limitation of Manifold Analysis for Super-Resolution of Facial Images. In *ICASSP*, pp.573–576, 2007.  
[8] Y.T. Zhuang, J. Zhang, and F. Wu. Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation. *Pattern Recognition*, 40 (11):3178–3194, 2007.  
[9] B. Li, H. Chang, S. G. Shan and X. L. Chen. Aligning Coupled Manifolds for Face Hallucination. *IEEE signal process. Lett.*, 16(11):957-960, 2009.  
[10] H. Huang, H. He, X. Fan, and J. Zhang. Super-resolution of human face image using canonical correlation analysis. *Pattern Recognition*, 43(7):2532–2543, 2010.  
[11] X. Ma, J.P Zhang, and C. Qi. Hallucinating face by position-patch. *Pattern Recognition*, 43 (6):3178–3194, 2010.  
[12] K. Su, Q. Tian, N. Sebe and J. Ma. Neighborhood issue in single-frame image super-resolution. In *ICME*, 2005.  
[13] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290 (5500):2323–2326, 2000.  
[14] B. Walczak and D. L. Massart. The Radial Basis Functions–Partial Least Squares Approach as a Hexible Non-Linear Regression Technique. *Anal. Chim. Acta*, 331(177):177-185, 1996.  
[15] W. Gao, B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H. Zhang, and D.B. Zhao. The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. *IEEE TSMC-A*, 38 (1):149-161, 2008.  
[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error measurement to structural similarity. *IEEE TIP*, 13(4):600-612, 2004.  
[17] H. Hotelling. Relations between two sets of variates, *Biometrika* 28, 312–377, 1936.  
[18] H. Wold. Partial least squares. In S. Kotz and N. Johnson, ed-itors, *Encyclopedia of Statistical Sciences*, volume 6, pp.581–591, 1985.