Face Image Superresolution via Locality Preserving Projection and Sparse Coding

Kebin Huang[1,2], Ruimin Hu[1], Zhen Han[1], Tao Lu[1], Junjun Jiang[1], Feng Wang[2]
1National Engineering Research Center on Multimedia Software, Wuhan University, Wuhan 430072, China
2Department of Digital Media Technology, Huanggang Normal University, Huangzhou, 438000, China
E-mail:{ kebin_huang, lut}@whu.edu.cn, hurm1964@gmail.com,
{hanzhen_1980, junjun0595}@163.com, mwfwf@hgnu.edu.cn

***Abstract*—It is important to enhance the resolution of face images from video surveillance for recognization and other post processing. In this paper, a novel sparse representation based face image superresolution (SR) method is proposed to reconstruct a high resolution (HR) face image from a LR observation. First, it gets a HR-LR dictionary pair for certain input LR patch via position patch clustering and locality preserving projection (LPP). LPP is used to further refine the training samples by taking into account the geometry structure between training patches. It can make the dictionary more expressive. Second, it sparse codes the input patch over LR dictionary by solving L1-regularized least squares optimization. Feature-sign algorithm is used to get the stable solution. Third, it maps the sparse coding coefficitons obtain from the LR dictionary to the HR dictionary, and reconstruct the HR face image patch with the coefficitions and HR dictionary. By integrating the hallucinated HR patches together with overlapped in adjacent, we can obtain the final HR face image. Experiments conducted on CAS-PEAL-R1 database validate the proposed method both in subjective and objective quality.**

***Index Terms*—Superresolution, Sparse representation, Locality preserving projection.**

I. Introduction

Human faces in surveillance video images usually have low resolution (LR) and poor quality due to equipment, environment, shooting and other factors. They cannot meet the requirements for recognition and post processing. Face image super-resolution (SR) technology is also called face hallucination[1], which reconstruct a high resolution (HR) face image using one or more LR inputs from a scene. It can improve the clarity of face images without changing hardware conditions and has important application in security monitoring, computer vision and etc.

Baker and Kanade [1] was the first to propose the term of face hallucination. They proposed to infer the high frequency details with probabilistic models from training samples. Followed by Baker's work, Liu et al. [2] proposed a two-step face hallucination approach by integrating a global parametric model and a local nonparametric model, which aims to constrain the reconstructed image, not only has common characteristics of a human face, such as eyes, mouth and nose, symmetry, etc, but also has specific characteristics of a certain face image. Based on Baker and Liu's [1-2] pioneering work on face hallucination, many face image SR algorithms are proposed. Such as subspace based methods[3-4], neighbor embedding based methods[5-6,21] and sparse representation based methods[7-9,22,24].

Due to the success of sparse representation using in incomplete signal recovering, sparse representation based SR method draw enormous attention in the SR research community. Based on the assumption that LR and HR patch pair might share the same sparse distribution, Yang et al. [7] introduced the idea of sparse representation to the face image SR to further enhance the detailed facial information by training coupled dictionaries. The ideal of coupled overcomplete dictionaries and sparse representation is also used to synthesize face sketch by Chang et al. [8]. Both obtained better results. In order to fully use the structure information of facial images, Elad et al. [11] used sparse representation for photo-ID image compressing by adapting to the image content. Ma et al. [10] take face position information as a feature and proposed a position-patch based face hallucination method, which reported better result too. Jung et al. [9] provided a position-patch based face hallucination method using convex optimization instead of using least square estimation in Ma's method, which obtained the optimal weights for face hallucination and achieved a better results than Ma's.

However, most of the existing approaches to sparse coding fail to consider the geometrical structure of the data space. Cai et al. [12,13,23] proposed a graph regularized sparse coding method for image classification and clustering. The method explicitly takes into account the local manifold structure of the data, and achieved better discrimination ability. Inspired by Cai's work, we proposed a novel sparse representation based face image SR method using locality preserving projection(LPP) and sparse coding.

The rest of the paper is organized as follows: in section 2, we provide a brief review of locality preserving projection and sparse coding. Our face image SR method is introduced in section 3 in detail. The experimental results are presented in section 4. Finally, we conclude the work in section 5.

doi:10.4304/jsw.8.8.2039-2046

II. Related Work

In this section, we will briefly review the locality preserving projection algorithm [14], sparse coding algorithm [15] and sparse coding based SR framework, which are important to our work.

A . Locality Preserving Projection

LPP is one of the manifold learning algorithms. It aims to find a linear projection for dimensionality reduction with local structure of the data space preserved. Compared to other manifold learning algorithms, LPP methods can generate mapping functions for new test images which are absent from a training set[14]. The LPP algorithm contains the following steps:

Step1: Construct the adjacency graph. Let $X = [x_1, x_2, \cdots, x_N]$ be a set of face images. Let G denote a graph with N nodes where ith node corresponds to the image x_i .We put an edge between nodes i and j if $\|x_i - x_j\|_2 < \varepsilon$, where $\varepsilon \in R$.

Step2: Compute the weight matrix. Let W be the weight matrix of G. If x_i is among the k-nearest neighbors of x_j or x_j is among the k-nearest neighbors of x_i , $W_{ij} = 1$; otherwise, $W_{ij} = 0$.

Step3: Compute the eigenvectors and eigenvalues of matrix W as follows:

$$XLX^T a = \lambda XDX^T a \tag{1}$$

where D is a diagonal N × N matrix whose entries are column sums of W , $D_{ii} = \sum_j W_{ij}, L = D - W$ is the Laplacian matrix. The projection matrix A can be founded as $A = [a_1, a_2, \cdots, a_l], l \ll N$. After project all the face image samples on the projection matrix A , we can get the reduced feature vector y_i of each face image sample.

$$y_i = A^T x_i \tag{2}$$

where y_i is l dimensions vector.

B. Sparse Coding

The objective of signal sparse representation is to represent a signal as a sparse linear combination with respect to an overcomplete dictionary. Suppose that the matrix $X \in R^{d \times N}$ is an overcomplete dictionary, in which each column vector is a d -dimensional atom. The number of columns in the matrix X is far greater than the number of rows, i.e., $N \gg d$, which ensures that the dictionary is overcomplete. Given a signal $x \in R^d$, its sparse representation can be seen as finding a sparse vector $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_N]^T \in R^N$ by solving the optimization problem

$$\min \|\alpha\|_0 \quad s.t. \quad \|x - X\alpha\|_2^2 \le \varepsilon \tag{3}$$

where ε is a small positive constant number; $\|\cdot\|_0$ denotes the l_0-norm, which counts the number of non-zero elements in a vector. Although the optimization problem (3) is NP-hard, they can be efficiently recovered by instead minimizing the l_1-norm [7,8]. Eq. (3) can be viewed as approximate error constrained sparse representation model. Lagrange multipliers offer an equivalent formulation, as

$$\min \|x - X\alpha\|_2^2 + \lambda \|\alpha\|_1 \tag{4}$$

In Eq.(4), the first item is approximation error and the second item is sparsity constrain item, λ is the regularization parameter, used to balance the weight between the approximation error and the sparseness. Eq.(4) is known as LASSO or basis pursuit[15]. There are many existing methods can be used to solve the problem, such as LASSO l_1 -norm regularization [16], feature sign search algorithm [15], BP algorithm or FOCUSS [17].

C. Superresolution via Couple Dictionaries and Sparse Coding

Here gives a brief introduction to sparse representation based SR method.

Suppose $\{D_H, D_L\}$ are HR and LR dictionary pair, I_L^p denotes a LR patch from LR image I_L , the sparse coding coefficient of I_L^p under LR dictionary D_L is

$$\alpha_L^p = \arg\min \|I_L^p - D_L\alpha_L^p\|_2 + \lambda\|\alpha_L^p\|_1 \tag{5}$$

In Eq. (5), the l_1-norm is used to replace l_0-norm for efficient convex optimization. λ is the regularization parameter, used to balance the approximation error and the sparsity. Then LR patch I_L^p can be represented sparsely as

$$I_L^p = D_L \alpha_L^p \tag{6}$$

In the scenario of image SR, it assumes that the LR and HR image patches share the same underlying sparse representations. Therefore, the sparse representation of a LR image patch can be applied with the HR image patch dictionary to generate a HR image patch. That is

$$I_H^p = D_H \alpha_H^p \approx D_H \alpha_L^p . \tag{7}$$

In order to make α_H^p keep in consistent with α_L^p , it trains two dictionaries jointly for the LR and HR image patches.

$$\{D_H, D_L, \alpha\} = \arg \min_{D_H, D_L, \alpha} \|I - D\alpha\|_2^2 + \lambda\|\alpha\|_1 \tag{8}$$

where $I = [I_H / \sqrt{W1}, I_L / \sqrt{W2}]^T$ is the concatenated HR-LR patches. $D = [D_H / \sqrt{W1}, D_L / \sqrt{W2}]^T$ is the coupled HR-LR dictionary trained from α . $W1$ and $W2$ are the dimensions of the HR and LR patches.

$\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_i]$ is the matrix collecting sparse representation vectors as columns.

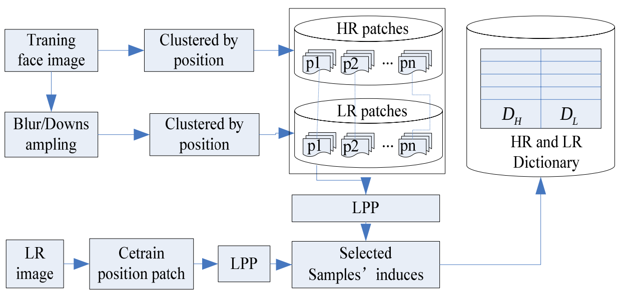
III. The Proposed Algorithm

A. Motivation and Basic Ideal

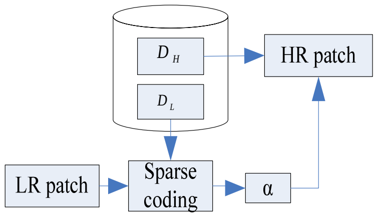
Image sparse representation is a novel image representation model and widely used for image SR [7-9]. Compared to other existing SR methods, such as neighbor embedding method [5] and soft edge prior method [7], the sparse representation based method could achieve superior results. However, they can not capture the intrinsic geometric structure in the training image data. Recently researches have shown that the geometrical structure of the data can improve the performance of sparse coding [12, 13]. Also the existing sparse representation based methods are facing to learn a pair of overcomplete dictionaries, which usually need to take a huge amount of calculation. In order to fully excavate the potential of the training data set and promote the efficiency of the dictionary training, LPP method is used at the dictionary training stage. As LPP is an efficient manifold learning method to reveal the non-linear structure hidden in the high-dimensional image space. It is used to analyze the local intrinsic features on the manifold of local facial areas. By searching out patches online in the LPP sub-space, it makes the resultant training set more expressive to a given testing patch.

Inspired by the recent developments on manifold learning that, LPP methods can generate mapping functions for new test images which are absent from a training set, this paper proposes a novel sparse representation based face image SR method which can capture the intrinsic geometric structure of the face image space and adapted to human visual cortex.

The algorithm mainly contains two stages: training and reconstruction. In the training stage, two dictionaries D_H^p and D_L^p are gained according to the position p, just like the test image. D_H^p is high resolution dictionary, D_L^p is LR dictionary. In the reconstruction stage, D_L^p is used to sparse coding the input patch first, then mapping the coefficients to D_H^p and reconstruct the HR patch of the input LR face image. At last, by merging all HR patches together with overlapped in adjacent patches, we can obtain the final HR face image. The framework of the algorithm is shown in Figure.1.



(a)The training stage



(b) The reconstruction stage

Fig. 1. The framework of the proposed face image SR algorithm

B. The Training Stage

In training stage, it aims to get a HR-LR dictionary pair for certain input LR patch. The choice of the dictionary is crucial to the success of the sparse representation application. Usually, there are two kind of dictionaries: the analytical dictionary and the learning dictionary. As the analytically designed dictionaries are universal dictionaries, they may lack sufficient flexibility to sparsely represent a given local patch. The learning dictionary is more flexible to sparsely represent different image patches. Here we take the learning dictionary.

In order to keep the global similarity information of face image and local geometric structure information between patches, two measures are adopted. Position feature of face images is used to obtain the global similarity information of face image. Manifold method of LPP is used to keep local geometric structure information between patches.

Let $I_H = \{I_H^q\}_{q=1}^Q = [I_H^1,\ I_H^2, \cdots, I_H^Q] \in R^{M\times Q}$ and $I_L = \{I_L^q\}_{q=1}^Q = [I_L^1,\ I_L^2, \cdots, I_L^Q] \in R^{N\times Q}$ represent the HR and LR training face images, respectively, Q is the number of training images. M and N are the dimension of each HR and LR face image by reshaping pixels to a column, $M = s^2 N$, s is down sampling factor. Each image in the training sets I_H and I_L is divided into a set of P small overlapping patch sets. According to the position of face, all training patches are classified into P groups as $\{I_{H,q}^1\}_{q=1}^Q, \{I_{H,q}^2\}_{q=1}^Q, \cdots, \{I_{H,q}^P\}_{q=1}^Q$ and $\{I_{L,q}^1\}_{q=1}^Q, \{I_{L,q}^2\}_{q=1}^Q, \cdots, \{I_{L,q}^P\}_{q=1}^Q$ respectively. Since all faces are aligned, the contents of the patches group in the same position are very similar. When we reconstruct the high resolution image patch via the patches in all training images that have the same position with certain patch in the input face image, there is more prior information, which may lead to a better result. Therefore, instead of using a universal dictionary for sparsely representing all patches in a face image, we use different dictionary for certain patch. It can keep the global similarity information of face image, such as eyes, mouth and nose, symmetry, etc.

Because faces are varied and the alignment algorithm can not align all face images so accurate, the position patch set for a certain input LR face patch needed to be further refined by K nearest neighbors' selection. As suggested by Cai et al. [12,13], takes into account the local manifold structure of the training data, it can learn a more distinguishable sample set. We adopt LPP method to extract the local manifold structure feature for performing an accurate patch pursuit. After position patch clustering and K nearest neighbor selection, we can obtain two dictionary sets online according to the patch number of a face image divided. That is

$$D_H = \{D_H^p \left| I_{H,q}^p, 1 \le p \le P, 1 \le q \le Q \right.\}$$

$$D_L = \{D_L^p \left| I_{L,q}^p, 1 \le p \le P, 1 \le q \le Q \right.\} \qquad (5)$$

By compressed sensing theory, we can regard the refined training patch set as over-complete dictionary like the choice in [9,10].This representation is naturally sparse if the size of the K nearest neighboring position-patches is reasonably large. In the reconstruction stage, we will use the dictionary of the region p to generate the high-resolution version of the region p adaptively.

C. The Reconstruction Stage

In reconstruction stage, we divide the image into small patches and apply SR algorithm on each patch and combine the patches to form the HR image. L1-regularized least squares optimization is used for sparse representation the input LR image patch over the LR dictionary and then HR face image patch is reconstructed by the sparse coding and HR dictionary.

Suppose $\{D_H^p, D_L^p\}$ are HR and LR dictionary pair in position p, $p = \{1, 2, \cdots, P\}$, I_L^p denotes a LR patch from LR image I_L in position p, the sparse coding coefficient of I_L^p under LR dictionary D_L^p is

$$\alpha_L^p = \arg\min \left\| I_L^p - D_L^p \alpha_L^p \right\|_2 + \lambda \left\| \alpha_L^p \right\|_1 \qquad (6)$$

In Eq. (6), the first item is approximation error and the second item is sparsity constrain item, the l_1-norm is used to replace l_0-norm for efficient convex optimization. λ is the regularization parameter, used to balance the approximation error and the sparsity. When D_L^p is fixed, Eq.(6) is essentially a L1-regularized least squares problem on the coefficients, well known as Lasso in the statistical literature. There are many algorithms are proposed to solve the problem above[11,15-17]. When tested for the benchmark data, the feature-sign search algorithm outperforms many other existing algorithms [15]. Here we take feature-sign search algorithm to solve the optimization. The main step of feature-sign algorithm is:

Step 1: Initialize the active set of potentially nonzero coefficients and its corresponding signs.

Step 2: Activate $\alpha_{L,q}^p$ (add q to the active set) from zero coefficient of α_L^p, obtaining the signs set and active set.

Step 3: Proceed in feature-sign steps. (a) compute the analytical solution with a given current guess for the active set and the signs. (b) update the solution, the active set and the signs using an efficient discrete line search between the current solution and the new solution.

Step 4: Check the optimality conditions for nonzero and zero coefficients and decided whether it would go to Step 3 (without any new activation) or go to Step 2, otherwise return the final solution.

Then LR patch I_L^p can be represented sparsely as $I_L^p = D_L^p \alpha_L^p$. In the scenario of image SR, it assumes that the LR and HR image patches share the same underlying sparse representations. It is validated in literatures [7][8][9]. Therefore, the sparse representation of a LR image patch can be applied with the HR image patch dictionary to generate a HR image patch. That is

$$I_H^p = D_H^p \alpha_H^p \approx D_H^p \alpha_L^p \qquad (7)$$

After getting all the input LR face image patches reconstructed, we integrate the hallucinated HR patches together with an arithmetic averaging operation in adjacent patches. This operation plays a similar role as the compatibility matrix of Markov network in [2]. The merging of the overlapped patches is shown in Fig.2. It improved the effect of reconstructed HR face image obviously.

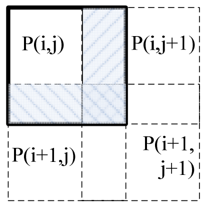


Fig. 2. The merging of the overlapped patches

IV. EXPERIMENTAL RESULTS

A. Database

In the experiments, we use the frontal and pre-aligned images of CAS-PEAL-R1 database[18]. CAS-PEAL-R1 is a subset of the entire CAS-PEAL face database. It contains 30 863 images of 1040 subjects. These images belong to two main subsets: the frontal and nonfrontal subsets. Our experiments are conducted on 1040 images from the CAS-PEAL-R1 database with neutral expression and frontal pose. Some of the images in CAS-PEAL-R1 are shown in Fig.3.

Fig.3. Some training faces in CAS-PEAL-R1 database

We randomly select 1000 images for training and leave the other 40 for testing. The images are cropped to 112×100 pixels and aligned by five manually selected feature points. All images are degraded by blurred(with an averaging filter of size 8×8), down-sampled(by a factor of 4 times), and added Gaussian noise with standard deviation 12. The degraded procedure is formulated as follows

$$I_L = DBI_H + n \tag{9}$$

Where I_H and I_L are high resolution face image and low resolution face image separately. B is the fuzzy matrix, D is the down sampling operation, n is the additive noise. The degraded LR face images are used for forming HR-LR training face image patch pairs and used for testing LR face image. Testing image is like in Fig.5.(a). While using testing image as input, we firstly amplify it to 112×100.

B. Parameter Settings

In order to validate the performance of the proposed algorithm, we compared our method with bicubic interpolation method, Ma et al. [10]'s method, and Jung et al. [9]'s methods based on the same training set. The size of reconstruction patch is 8×8 pixels with 32 pixels overlap in adjacent patches. The error tolerance in [9] was set to 1. In the proposed method, the selection of parameter λ has great affection to the reconstructed results. λ is used to balance the weight between the approximation error and the sparseness. It can be affected by the level of noise. Literatures [7,20]have improved that λ is in direct proportion to the level of noise. J.C. Yang, etc. take one tenth of the noisy deviation as the value of λ. In this paper we take $\lambda = 12$ empirically. So $K = 300$.

C. Results

Due to space limitation, we don't provide all the results here. Some of the experimental results are shown in Fig.4. The result of bicubic interpolation method(Fig.4.(b)) is very smooth, but it is not clear and hard to be recognized. There are obvious ghost effects in the results of Ma's method(Fig.4.(c)) and Jung's method(Fig.4.(d)). Also, artificial effects in the results of Ma's method are obvious. Compared with the reference methods, the results of the proposed method remained more detail information and less ghost and artificial effect. The proposed method improved the visual effect significantly and the performance is better than others in subjective.

(a) (b) (c) (d) (e) (f)

Fig. 4. Subjective results (a) are the LR input face images. (b)-(e) are the results reconstructed by bicubic interpolation method, Ma method[10], Jung method[9] and the proposed method. (f) are the HR face images.

To further validate the effect of the proposed method, objective evaluation of RMSE (root mean square error), PSNR (peak signal to noise ratio) and SSIM (structural similarity) are carried out too. These values are calculated between the reconstructed images and the original HR images. PSNR and RMSE are defined as follows:

$$PSNR = 10\log_{10}(\frac{255^2}{MSE})$$

$$MSE = \sum_{n=1}^{N}\frac{(I_n - P_n)^2}{N}$$

$$RMSE = \sqrt{MSE}$$

Where I_n is the n-th pixels of the original HR image. P_n is the n-th pixels of the reconstructed image. N is the total number of the compared image. SSIM[19] is defined as follows:

$$SSIM(I_n, P_n) = l(I_n, P_n)\cdot c(I_n, P_n)\cdot s(I_n, P_n)$$

Where, I_n and P_n are the images to be compared, $l(I_n, P_n)$ is luminance comparison function, $c(I_n, P_n)$ is contrast comparison function, $s(I_n, P_n)$ is structure comparison function. RMSE, PSNR and SSIM are the common criterions to evaluate a picture's quality.

The average values of RMSE, PSNR and SSIM from 40 testing images are showed in Table 1. Seen from Table 1, we found that the proposed method obtained the lowest RSME value(the lower the better) and the highest PSNR and SSIM values(both are the higher the better). It demonstrates that the proposed method can reconstruct an image closest to the original HR image. The results between subjective quality and objective quality are consistent. It validates the effectiveness and advancement of the proposed method.

TABLE I.
RSME, PSNR (DB) AND SSIM RESULTS

Method	Bicubic	Ma method[10]	Jung method[9]	Proposed method
RMSE	22.5374	11.3645	10.5769	**10.3413**
PSNR	21.1096	27.1390	27.7531	**27.9718**
SSIM	0.5941	0.8605	0.8724	**0.8812**

D. The Affection of LPP

In order to check the ability of LPP's picking out the consistent sample sets for the HR and LR testing patches, we conduct the samples selecting experiments on HR and LR testing patches in LPP subspace. The matching ratio is adopted to measure the consistency, just like in literature [22]. It is defined as follows

$$MatchRation = \frac{MaNum}{SeNum}$$

Where $MaNum$ is the number of the selected samples' indices both in HR and LR ones, $SeNum$ is the number of the selected samples. In these experiments, 100 samples were selected for each testing patch extracted from 40 images, and additional 1000 face images were used to construct the training samples. The average matching ratio is 73% by using LPP, however, the average matching ratio without using LPP is 66%. Seen from the matching ratio, we can found that the LPP algorithm can preserve the structure feature of the patch well. We presented such an example in Fig. 5.

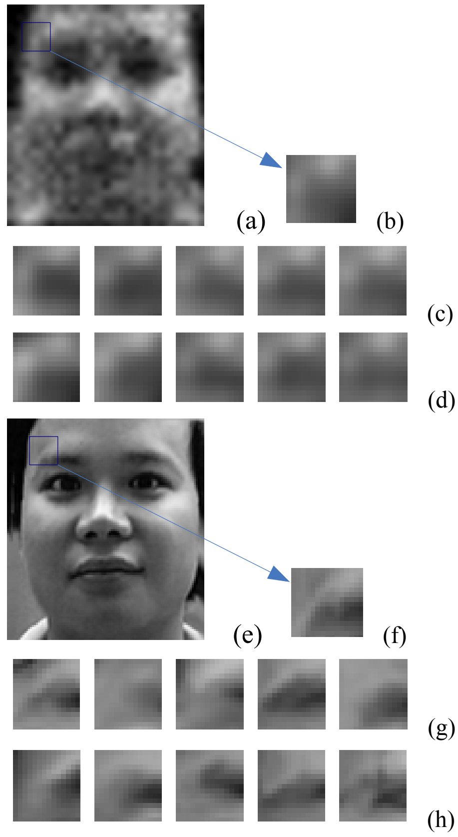


Fig.5: An example of samples selection by using LPP dimensions reduction. (a) and (e) are HR and LR test image pair. (b) and (f) are HR and LR patch pair extracted from the HR and LR test image separately. (c) (d) and (g) (h) are the first ten selected HR and LR patches according to (b) and (f) after LPP dimension reduction.

Also, we validate the affection of LPP dimensions via experiments. We calculated the average PSNR of ten testing face images with different LPP dimensions(the dimensions are 3,5,7,10,20,30, 50,64). Fig.6. shows the affection of LPP dimensions.

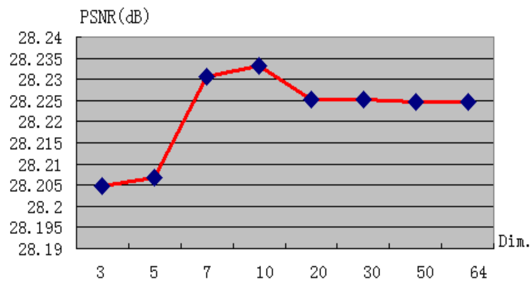


Fig.6. the affection of LPP dimensions

It can be seen that the LPP dimensions affect the reconstructed image quality greatly. When the dimension is lower than 10, the average PSNR of the reconstructed face image is bigger as the dimension increasing. However, when the dimensions is higher than 10, the average PSNR of the reconstructed face image decreasing as the dimension increasing, until the dimension 20. When the dimension is 10, it achieved the highest PSNR. So, in our algorithm, the dimension of LPP is set to 10. It achieved better result with LPP dimension reduction.

V. Conclusion

In this paper, it proposed a face image super-resolution method using locality preserving projection and sparse coding. First, it obtains a couple of HR and LR over-complete dictionary using patches in the same position as input LR face image from all training samples. Then, it refined the dictionaries with LPP based K nearest neighboring position-patch selection. The patches in LPP transformed space have better distinguish ability, which lead to more expressive dictionaries. Feature-sign search algorithm is used for solving the L1-regularized least squares problem. The HR face image patch is reconstructed by the sparse coding and HR dictionary. By integrating the hallucinated HR patches together with overlapped in adjacent patches, we can obtain the final HR face image. Experiments conducted on CAS-PEAL-R1 database validate the proposed method both in subjective and objective quality.

Acknowledgements

The research was supported by the major national science and technology special projects (2010ZX03004-003-03), the National Natural Science Foundation of China (61172173, 61070080, 60970160, 61003184), the Natural Science Foundation of Hubei Province of China (2009CDB404, 2009CDA134, 2010CDB05103), the Fundamental Research Funds for the Central Universities (3101011).

References

[1] S. Baker and T. Kanade. “Limits on super-resolution and how to break them”. IEEE Transactions onPattern Analysis and Machine Intelligence, 24(9), pp.1167-1183, 2002.

[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 Proc. 2001 Computer Vision and Pattern Recognition, pp. 192–198.

[3] X.Wang and X. Tang. “Hallucinating face by eigentransform”. IEEE IEEE Transactions on Systems, Man and Cybernetics part C, 34(3), pp.425-434, 2005.

[4] Jeong-Seon Park and Seong-Whan Lee. “An example-based face hallucination method for single-frame, low-resolution facial images” . IEEE Transactions on Image Processing, 17(10), pp.1806-1816, 2008.

[5] H. Chang, D. Yeung, and Y. Xiong. “Super-resolution through neighbor embedding”. in Proc. 2004 Computer Vision and Pattern Recognition, pp. 275-282.

[6] W. Fan and D. Y. Yeung, “Image hallucination using neighbor embedding over visual primitive manifolds”, in Proc. 2007 Computer Vision and Pattern Recognition, pp. 1-7

[7] J.C. Yang, H. Tang, Y. Ma and T. Huang. “Face hallucination via sparse coding”. in Proc. IEEE Conf. Image. Processing2008, pp. 1264-1267.

[8] L. Chang, M.Zhou, “Face Sketch Synthesis via Sparse Representation”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,2010, pp.2146-2149

[9] C. Jung, L. Jiao, B. Liu and M. Gong. “Position-Patch Based Face Hallucination Using Convex Optimization”. IEEE Signal Processing Letters, 18(6), pp. 367-370, 2011

[10] X. Ma, J. Zhang, and C. Qi, “Hallucinating face by position-patch, Pattern Recognition”, 43(1), pp. 2224-2236, 2010.

[11] M. Aharon, M. Elad, AM. Bruckstein, “The K-SVD: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representation”. IEEE transactions on signal processing, 54(11), pp.4311-4322, 2006

[12] D. Cai, H. Bao, X. He, “Sparse Concept Coding for Visual Analysis”, In Proc. 2011 Computer Vision and Pattern Recognition, pp.2905-2910

[13] S. Gao, I. Tsang, L. Chia, and P. Zhao, “Local features are not lonely-Laplacian sparse coding for image classification”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2010, pp. 3555–3561.

[14] X.F He, Y.P. NI, “Locality Preserving Projections”. Neural Information Processing Systerm , pp.290-307, 2003.

[15] R. Tibshirani, “Regression shrinkage and selection via the Lasso”, Journal of the Royal Statistical Socity, Series B,(Methodological), 58(1), pp. 267-288, 1996.

[16] H. Lee, A. Battle, R. Raina, and A. Y. Ng, “Efficient sparse coding algorithms”, Advances in Neural Information Processing Systems, vol.19, pp. 801-808, 2007.

[17] I.F. Gorodnitsky, and B.D. Rao “Sparse Signal Reconstruction from Limited Data Using FOCUSS: A Re-weighted Minimum Norm Algorithm”, IEEE transactions on signal processing, VOL. 45, NO. 3, MARCH 1997

[18] 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 Transactions on Systems, Man and Cybernetics Part A: Systems and Humans, 38 (1):149-161, 2008.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. “Image quality assessment: From error measurement to structural similarity”. IEEE Transactions on Image Processing, 13(4), pp. 600-612, 2004.

[20] W. Dong, L. Zhang, G. Shi. “Centralized Sparse Representation for Image Restoration”, In Proc. 13th International Conference on Computer Version (ICCV 2011), pp. 1259-1266 ,.

[21] T. Lu, R. Hu, Z. Han, J. Jiang, and J. Chang, “Face hallucination based on sample selection bias correction," International Journal of Advancements in Computing Technology, vol. 4, pp. 91-98, 2012

[22] Xuesong Zhang, Silong Peng, Jing Jiang “An Adaptive Learning Method for Face Hallucination Using Locality Preserving Projections”, 8th IEEE International Conference on Automatic Face & Gesture Recognition, Sept. 2008, pp.1-8

[23] G. Lin, Y. Bie, G. Wang, M. Lei. A Novel Clustering Algorithm Based on Graph Matching. Journal of Software, Vol 8, No 4 (2013), pp.1035-1041

[24] X. Jin, J. Bao, J. Du. Image Enhancement Based on Selective - Retinex Fusion Algorithm. Journal of Software, Vol 7, No 6 (2012), pp.1187-1194

Kebin Huang, received the bachelor's degree in educational technology from Huanggang Normal University, Huanggang, China, 2003, the Master's degree in educational technology from Central China Normal University, Wuhan, China, 2006. Now he is a PhD candidate in communication and information system in Wuhan University. His primary research is on computer vision, superresolution, face hallucination, video and image process in sparse representation.

Ruimin Hu (M'09–SM'10) received the B.S. and M.S. degrees in communication and electronic system from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 1984 and 1990, respectively, and the Ph.D. degree in communication and electronic systems from the Huazhong University of Science and Technology, Wuhan, China, in 2000. He is currently a Professor with Wuhan University, Wuhan. He also serves as the Dean of the School of Computer Science and the Director of the National Engineering Research Center on Multimedia Software. His current research interests include audio and video signal processing, multimedia network communication, security and surveillance techonology, digital multimedia content management, and protection.