White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/115963/

Version: Accepted Version

Proceedings Paper:
Zhang, Yiping, Zhang, Z., Hu, Guosheng et al. (2016) Face image super-resolution via weighted patches regression. In: Davis, L., Bimbo, A. Del and Lovell, B., (eds.) 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE Computer Society, Los Alamitos, CA, USA, pp. 3892-3897.

https://doi.org/10.1109/ICPR.2016.7900242

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Face Image Super-resolution Via Weighted Patches Regression

Yiping Zhang
School of Information Science and Engineering
Xiamen University
Xiamen, China

Zhihong Zhang*
Software School
Xiamen University
Xiamen, China

Guosheng Hu
LEAR team
Inria Grenoble Rhone-Alpes
France

Edwin R. Hancock
Department of Computer Science
University of York
York, UK

***Abstract*—Recently sparse representation has gained great success in face image super-resolution. The conventional sparsity-based methods enforce sparse coding on face image patches and the representation fidelity is measured by ℓ_2-norm. Such a sparse coding model regularizes all facial patches equally, which however ignores the natures of facial patches, where the facial patches in the different regions (patch positions) of human face may have distinct contributions to face image reconstruction. In this paper, we propose to weight facial patches based on their discriminative abilities in regression for robust face hallucination reconstruction. Specifically, we learn the weights for facial patches according to the information entropy in each face region, so as to highlight higher frequency details in face images and the facial discriminability can be well retrieved. Various experimental results on standard face databased show that our proposed method outperforms state-of-the-art methods in terms of both objective metrics and visual quality.**

I. INTRODUCTION

One of the most common challenges to practical face recognition system is that most face images captured in the wild are of low resolutions. The low-resolution (LR for short) face images not only bring down the human visual experience but also adversely affect the performance of the followed face recognition and analysis. To alleviate this problem, image super-resolution (SR) attempts to increases high-frequency components and removes the undesirable effects, e.g., the resolution degradation, blur and noise. For an observed LR image $\mathbf{y}$, the problem of image SR is generally modeled as $\mathbf{y} = \mathbf{SHx} + e$ with the goal of recovering an high-resolution (HR) image $\mathbf{x}$ from $\mathbf{y}$, where e is a small noise term, $\mathbf{H}$ is a blur filter, and $\mathbf{S}$ represents a down-sampling operator. The dimension of $\mathbf{y}$ is significantly smaller than that of $\mathbf{x}$; thus there are an infinite number of possible HR images $\mathbf{x}$ that can generate the same LR image $\mathbf{y}$. To obtain a unique and good HR image, additional information is imperative and of great important to eliminate the uncertainty of recovery.

The problem of face image super-resolution was first studied by Baker and Kanade [1] which developed a Bayesian approach to infer the missing high-frequency components of face images. Based on the parent-training images, it generates the high frequency details by learning the gradient prior from a parent image pyramid. Subsequently, Liu et al. [2] present a two-step approach to hallucinate faces, which integrates global structure reconstruction with local detail adjustment. Firstly, the method generates global face image keeping the main characteristics of the original high-resolution face. Secondly, it produces residual image containing the high-frequency image information to compensate the results of the first step. Both of the methods incorporated the degradation function into the formulation to solve the final hallucinated result. Wang and Tang [3] propose a face hallucination method by eigen transformation, which views hallucination as a transformation between different image styles. However, this can hardly maintain the global smoothness and visual rationality, especially at locations around the face contour and margin of the mouth. Inspired by the well-known locally linear embedding idea [4] in manifold learning, Chang et al. [5] developed a super-resolution method through Neighbor Embedding algorithm, which assuming that the training low- and high-resolution images from manifolds share similar local geometric structure. Motivated by Chang's work [5], a number of face hallucinations methods [6], [7], [8] were developed based on Neighbor Embedding or using neighbor patch. However, fixed number of neighbors for reconstruction may lead to blurred and unwanted edges dur to under- or over-fitting.

Due to the success of sparse representation used in incomplete signal recovering, a series of methods based on sparse representation are developed for face image super-resolution, which can fundamentally avoid the blurring effects of super-resolved faces. By forcing LR patch and the corresponding HR patch to have the same sparse coefficient, Yang et al. [9] are the first to introduce the idea of sparse representation to the face image SR. The method offline trains a HR and LR dictionary to sparsely decompose HR and LR image patches, respectively. Given a LR patch input, a sparse coefficient vector is computed using the LR dictionary by solving an ℓ_1-norm minimization problem. The desired HR patch is reconstructed by combining the HR dictionary. The similar intuitive is used in [10]. Chang et al. [10] used coupled over-complete dictionaries and sparse representation to synthesize face sketch which obtained better result. 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. The prior of face position can be incorporated into face super-

*Corresponding author: zhihong@xmu.edu.cn.

Preprint submitted to 23rd International Conference on Pattern Recognition.
Received April 22, 2016.

resolution. Ma et al. [12] take face position information as a feature and proposed a position-patch based face hallucination method. It estimate a HR image patch using the image patches at the same positions of all training face image. Specifically, the coding coefficients estimated via constrained least square (CLS) in each face region are used to generate the HR patch of the corresponding position. However, when the number of the training images is much larger than the dimension of the patch, the CLS problem is underdetermined and the solution is not unique. To address the biased solution problem caused by least square estimation, Jung et al. [13] provided a position-patch based face hallucination method using convex optimization, which obtained the optimal weights for face hallucination and achieved a better results than Ma's [12]. However, this sparse coding based method [13] fail to consider the manifold geometric structure of the face data that is important for image representation and analysis. Zhang et al. [14] presented a dual-dictionary learning method to recover more image details, in which both the main and the residual dictionaries are learned by sparse representation. These SR based methods gave impressive improvements for experimental noise free faces. However, due to the under sparse nature of noisy images, they usually perform unsatisfactorily in the presence of noise.

The above sparse representation based super-resolution approaches are based on the minimization of mean-squared-error (MSE) between the input LR image patches and the reconstructed SR image patches (i.e., $\|\mathbf{y}-\mathbf{D}\alpha\|_2^2$). In fact, the fidelity term has a high impact on the final coding results because it ensures that the given signal $\mathbf{y}$ can be faithfully represented by the dictionary $\mathbf{D}$. From the view of maximum likelihood estimation (MLE), defining the fidelity term with ℓ_2-norm actually assumes that the coding residual $e=\mathbf{y}-\mathbf{D}\alpha$ follows Gaussian distribution. Such coding fidelity treats all the face image patches equally, and it ignores the natures of facial patches, where the facial patches in the different regions (patch positions) of human face may have distinct contribution to face image reconstruction. Intuitively, face regions (such as mouth, eyes, nose) rich in texture contain more high-frequency detail; thus, facial patches in this regions are expected to be assigned with high weight values to ensure very small residuals. While some face regions with variational noise or corruptions, their corresponding image patches are given lower weight values to reduce their effects on the regression estimation so that sensitiveness to these regions can be greatly reduced.

To improve the robustness and effectiveness of face hallucination, we propose to incorporated the discriminative ability of pixel locations into the regression procedure. Such weight values are determined through the information entropy in each face region. In order to construct a robust weights to fully exploit structure information of each face region, we employed external data (not just limit to training data) to learn the weights. As the external data can cover all possible face image variants of different persons, so the robustness of obtained weights can be guarantee.

II. Related Work

We review some of the related previous works in this section, which will lay the foundation for the derivation of our approach later.

A. Super-resolution via Couple Dictionaries and Sparse Coding

Yang et al. [9] proposed an approach for super-resolution based on sparse representation. Given $\mathbf{D_l} \in \Re^{M\times K}$ be an over-complete LR dictionary of K prototype signal-atoms, $\mathbf{D_h} \in \Re^{N\times K}$ be the corresponding over-complete HR dictionary of K prototype signal-atoms, where N and M are the dimensions of a HR image patch and LR image patch, respectively. Yang et al. [9] start from a large collection of low resolution (LR) and high resolution (HR) training patch pairs and use a sparsity constraint to jointly train the LR and HR dictionaries by assuming that LR patches and their corresponding HR counterparts shares the same sparse coding vector. The optimal dictionary pair $\{\mathbf{D_h},\mathbf{D_l}\}$ is obtained by minimizing

$$\min_{\mathbf{D_h},\mathbf{D_l},Z} \frac{1}{N}\|X_h-\mathbf{D_h}Z\|_2^2+\frac{1}{M}\|Y_l-\mathbf{D_l}Z\|_2^2+\lambda(\frac{1}{N}+\frac{1}{M})\|Z\|_1 \ , \tag{1}$$

Once the dictionaries are trained, the input LR image is divided into overlapped patches, and each patch $\mathbf{y}$ can be sparsely encoded by a learned LR dictionary $\mathbf{D_l}$ using the following formulation:

$$\min_{\alpha} \|F\mathbf{D_l}\alpha - F\mathbf{y}\|_2^2 + \lambda\|\alpha\|_1 \ , \tag{2}$$

where F is a feature extraction operator, α is the sparse representation and λ is a weighting factor. The corresponding HR patch is reconstructed by $\mathbf{D_h}$ and α with $\mathbf{D_h}\alpha$. Finally, the HR image can be obtained by aggregating all the estimated HR patches into a whole image. One problem of Yang's work [9] is that the dictionary training process is time-consuming. Therefore, it will be much efficient if we can project the patch vectors into a lower subspace while preserving most of their average energy.

Zeyde et al. [15] improved the work of Yang et al. [9] with less computation time and better estimation result. They perform dimensionality reduction of LR image patches via Principal Component Analysis (PCA) to improve the execution speed. With the training patch pairs $\{\mathbf{P_h},\mathbf{P_l}\}$ prepared, they firstly learn the LR dictionary as:

$$\alpha = \min_{\alpha} \|\mathbf{P_l} - \mathbf{D_l}\alpha\|_2^2 + \lambda\|\alpha\|_1 \ , \tag{3}$$

The above optimization formula can be solved by K-SVD [16] and Orthogonal Matching Pursuit [16]. By the same assumption with Yang's work [9], the sparse code α trained from above can be utilized in constructing the HR dictionary $\mathbf{D_h}$. And the $\mathbf{D_h}$ training can be formulated as a least square regression problem:

$$\mathbf{D_h} = \min_{\alpha} \|\mathbf{P_h} - \mathbf{D_h}\alpha\|_2^2 \ , \tag{4}$$

Preprint submitted to 23rd International Conference on Pattern Recognition.
Received April 22, 2016.

Hence, a straightforward least-square solution of $\mathbf{D_h}$ can be obtained by:

$$\mathbf{D_h} = \mathbf{P_h}\alpha^{\mathbf{T}}(\alpha\alpha^{\mathbf{T}})^{-1} \ , \tag{5}$$

From above, it is clear that this method can only train for LR dictionary and its corresponding sparse code, leading to more time saving and less computation complexity. Despite the improvements, the use of OMP during sparse coding is clearly the bottleneck.

B. Anchored Neighborhood Regression

Starting from the same dictionaries training by K-SVD with OMP algorithms in [15], the Anchored Neighborhood Regression (ANR) approach [17] proposes to relax the sparsity constraint in Eq. 2 and reformulates the patch representation problem as a least squares (LS) ℓ_2-norm regression. The method uses the local neighborhoods of dictionary (i.e. $\mathbf{N_l}$ and $\mathbf{N_h}$) with a specific size instead of the entire dictionary used in [9]. Compared with solving ℓ_1-norm minimization which is computationally demanding, the ℓ_2-norm regression turns the problem into Ridge Regression [18] and a closed-form solution can be obtained.

$$\min_{\beta} \|\mathbf{y} - \mathbf{N_l}\beta\|_2^2 + \lambda\|\beta\|_2 \ , \tag{6}$$

where $\mathbf{N_l}$ is the LR neighborhood of input patch $\mathbf{y}$ chosen from $\mathbf{D_l}$. The algebraic solution of the coefficient vector β can be written as:

$$\beta = (\mathbf{N_l^T N_l} + \lambda\mathbf{I})^{-1}\mathbf{N_l^T y} \ , \tag{7}$$

the coefficients of β are then applied to the corresponding HR neighborhood $\mathbf{N_h}$ to reconstruct the HR patch $\mathbf{x}$,

$$\mathbf{x} = \mathbf{N_h}(\mathbf{N_l^T N_l} + \lambda\mathbf{I})^{-1}\mathbf{N_l^T y} = \mathbf{P_j y} \ , \tag{8}$$

where $\mathbf{P_j}$ is the projection matrix for dictionary atom $\mathbf{d_{lj}}$. Given the trained couple dictionaries, for each LR dictionary atom $\mathbf{d_l}$, we search for its K nearest neighborhoods of dictionary $\mathbf{N_l}$ by correlation between the whole dictionary atoms. Then, based on the neighborhoods of $\mathbf{d_l}$, a separate projection matrix $\mathbf{P_j}$ can be computed. Therefore, the projection matrix can be obtained offline and the procedure of SR for ANR at test time becomes mainly a nearest neighbor search followed by a matrix multiplication for each input patch.

Although the effectiveness of sparse representation has been proven, the spatial information is lost during the coding phase. We believe that the amount of information in different face regions is different and the spatial information should also be included in the face image reconstruction.

III. The Proposed Algorithm

Inspired from the recent development on regions division problem [19], [20], we learn the weight for image patches in each facial region according to the information entropy of the region. More specifically, we firstly partitioned each face image into overlapped patches according to the different regions (patch positions) of human face and then the importance of patches in each region is measured by information entropy. Figure.1 shows the details for learning the weight values in each face region using information entropy.

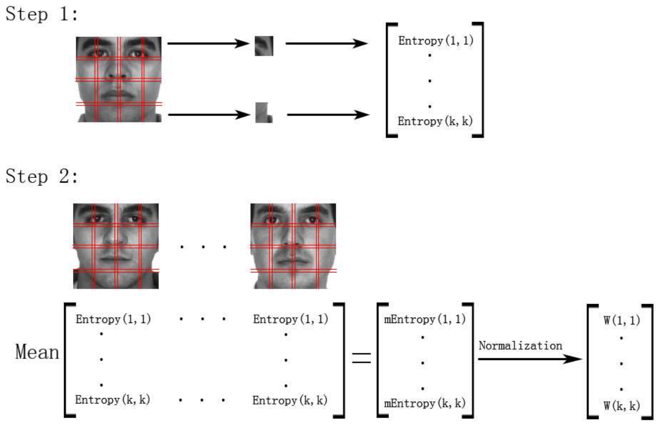

Fig. 1. The details for learning the weight values in each face region using information entropy.

The entropy is a term defined in information theory as a measurement of the uncertainty associated with a random variable [21]. It is relevant to the quantity and variability of the information. Here, we assume that the pixel intensity value is a random variable; thus, we can use the histogram of intensities in each face image patch to approximate the probability density function (PDF) for computing the information entropy. Applied to our case, the larger the entropy values is, the more information a face image patch should contain, and thus smaller residuals should be set for this patch.

The entropy value of the face patch i can be then defined as

$$H_i = \sum_{k=1}^{n} p(v_k)\log_2\left(\frac{1}{p(v_k)}\right) = -\sum_{k=1}^{n} p(v_k)\log_2 p(v_k) \tag{9}$$

where $p(v_k)$ is the probability of the pixel v with intensity value k in the histogram of the region.

In our method, the entropy following Eq. 9, is computed from the intensity histograms of the coarse divided regions for all face images in the training set. Then, the average entropy value of patches in a region for all images is used as the corresponding regional patch entropy. Although some images in the training set might be affected by noise, the average entropy values can still reflect the information quantity differences among different facial regions. Finally, we normalized these entropy values as the final weight values

$$W_i = \frac{H_i}{\sum_{k=1}^{n} H_k} \tag{10}$$

where W_i is the weight value of patch in face region i.

By imposing the weighted spatial information into the ANR [17] scheme, our method can be formulated as follows:

$$\min_{\alpha} ||W_i(y_i - N_l^i\alpha)||_2^2 + \lambda||\alpha||_2 \tag{11}$$

where y_i is the feature vectors of the i-th patch of input low-resolution face image, N_l^i is its corresponding neighborhood in LR space, W_i is its corresponding weight value and λ is a balance parameter. Eq. 11 is a convex formulation and its algebraic solution is give by

$$\alpha = ((W_i N_l^i)^{\mathrm{T}}(W_i N_l^i) + \lambda I)^{-1}(W_i N_l^i)(W_i y_i) \tag{12}$$

the HR patch x_i can then be computed using the same coefficient on the corresponding HR neighborhood N_h^i

$$x_i = N_h^i \alpha \tag{13}$$

Finally, we put all the HR patches together into a HR face image.

IV. Experiment

In this section, we conduct several experiments to evaluate the effectiveness of the proposed method, in terms of both objective metrics and visual quality. We compare the SR estimation results with several classical as well as state of the art SR methods including Bicubic interpolation, the sparse coding algorithms of Yang et al. [9] and Zeyde et al. [15], Anchored Neighborhood Regression (ANR) approach [17]. These methods were configured using the same patch size and overlap as indicated below and configured using the optimal parameters provided in their respective papers. The face super-resolution performance is quantified by the Peak Signal to Noise Ratio (PSNR) [22] between the ground truth face images and the super-resolved ones.

A. Experimental configurations

Database: The experiments conducted in this paper use different publicly available face datasets: i) FERET [23], ii) CAS-PEAL-R1 [24], iii) AR [25]. All these face images were aligned by an automatic alignment algorithm using the eye positions, and then cropped to the size of 64 $\times$ 64 pixels. The LR images are formed by blurred and down-sampling (by a factor of 4 resulting the size of LR face images to be 16 $\times$ 16 pixels) the corresponding HR images. The 450 face images from FERET were used as HR dictionary training images and LR dictionary training images. Here, we adopted the same way of dictionary learning procedure as Zeyde et al. [15], which combining GOMP and K-SVD to train the dictionary pair. Then, we randomly select 50 face images from AR (referred to as **SetA**) and 50 face images from CAS-PEAL (referred to as **SetB**) as the testing image sets.

Parameter Setting: Empirically, we set the size as 16 $\times$ 16 pixels for HR patch and the overlap between neighbor patches as 4 pixels. The corresponding LR patch size is set to 4 $\times$ 4 with overlap of one pixel. After learning the dictionaries for high-resolution and low-resolution image patches, we can group the dictionary atoms into neighborhoods. Specifically, for each atom in the dictionary we find its K nearest neighbors based on the correlation between the dictionary atoms, which will represent its neighborhood. In our experiments, we set K to 40 and parameter λ to 0.0001.

B. Comparison of Subjective and Objective Quality Results

In this section, we perform experiments on two representation benchmark databases (AR and CAS-PEAL) to demonstrate the effectiveness of the proposed method. Since the ground truth HR face images are available, we compare not only the visual quality but also the quantitative results of the reconstructed face images.

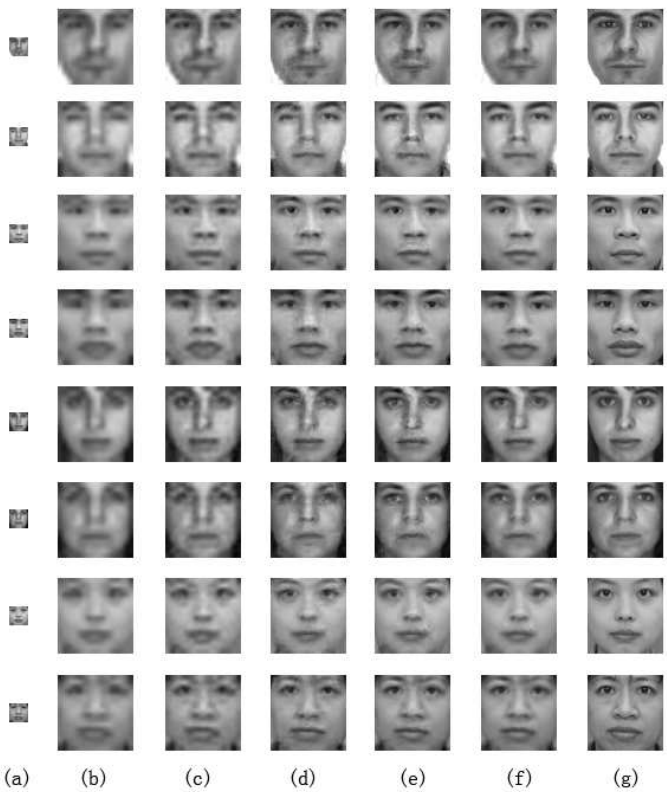

Fig. 2. Examples of our algorithm compared to other methods. From left to right columns: (a) low-resolution input; (b) Bicubic interpolation; (c) Yang's method; (d) Zeyde's method; (e) ANR (f) Our proposed method; (g) Original HR face images

In Fig. 2, some reconstructed face images by different methods are compared. It can be seen that the reconstructed face images by Bicubic interpolation is very blurry. Yang's results lost too many details and have many jaggy artifacts. The textures of ANR's results are heavily smoothed. Our method achieves better visual quality with more detail information and less artificial effect. More specifically, our proposed method can recover better high frequency components of facial features like nose and mouth (see 5th and 6th row in Fig. 2). Overall, the proposed method improved the visual effect significantly and the performance is better than others in subjective.

To further validate the effect of the proposed method, objective evaluation of PSNR are carried out too. The average values of PSNR from 100 testing images (**SetA** and **SetA**) are showed in Table 1. Seen from Table 1, we found that the proposed method obtained the highest PSNR values. It demonstrates that the reconstructed face images by our method are closest

Preprint submitted to 23rd International Conference on Pattern Recognition. Received April 22, 2016.

CONFIDENTIAL. Limited circulation. For review only.

Table 1
The average results of PSNR by different methods on the SetA and SetB.

TestSet	Bicubic	Yang	Zeyde	ANR	Proposed method
SetA	25.73	26.38	27.20	27.74	**28.09**
SetB	26.05	27.21	27.98	28.16	**28.57**

Table 2
Part of the experimental results on the SetA

SetA	Bicubic	Yang	Zeyde	ANR	Proposed method
	25.8	26.2	27.7	28.2	**28.6**
	25.3	25.7	26.8	27.6	**28.4**
	25.2	25.6	26.6	26.9	**27.2**
	27.1	27.1	27.2	28.7	**29.3**
	25.7	26.2	27.6	26.9	**27.9**
	28.6	29.0	30.1	30.5	**31.2**
	26.6	26.8	27.5	27.4	**28.1**
	25.3	25.5	26.2	27.0	**27.4**
	26.5	27.1	28.6	28.9	**29.2**
	26.9	27.4	28.3	29.0	**29.7**

to the ground truth HR images. The results between subjective quality and objective quality are consistent. It validates the effectiveness and advancement of the proposed method.

For clear comparison, we randomly selected 10 face images from **SetA** and **SetA** respectively for testing. The experimental results in Table 2 and Table 3 also clearly show that the proposed method outperforms each of the competing methods all face images studied. The results further verify the effective of proposed method.

V. Conclusion

In this paper, we have proposed a novel weighted patches regression method to face hallucination. This idea is motivated by the observation that face regions (such as mouth, eyes, nose) rich in texture contain more high-frequency detail; thus, facial patches in this regions are expected to be assigned with high weight values to ensure very small residuals. Experimental results on three benchmark datasets demonstrate its superiority over state-of-the-art methods in terms of PSNR values and visual quality.

Acknowledgment

This work is supported by National Natural Science Foundation of China (Grant No.61402389).

References

[1] S. Baker and T. Kanade, "Hallucinating faces," in *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition.* IEEE, 2000, pp. 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 *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*, vol. 1. IEEE, 2001, pp. I–192.

[3] X. Wang and X. Tang, "Hallucinating face by eigentransformation," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 35, no. 3, pp. 425–434, 2005.

[4] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

Preprint submitted to 23rd International Conference on Pattern Recognition.
Received April 22, 2016.

Table 3
Part of the experimental results on the SetB

SetB	Bicubic	Yang	Zeyde	ANR	Proposed method
	25.5	27.0	29.1	29.2	**29.7**
	27.0	28.1	29.1	28.7	**29.3**
	26.6	27.3	28.8	28.7	**29.2**
	25.4	26.9	28.3	29.0	**29.4**
	25.1	26.3	27.2	28.0	**28.2**
	26.3	26.5	27.0	26.9	**27.6**
	26.7	27.1	28.6	28.8	**29.4**
	27.4	27.5	27.8	28.0	**28.4**
	27.2	27.8	28.4	29.0	**29.6**
	26.9	27.4	28.9	29.2	**29.5**

[5] H. Chang, D.-Y. Yeung, and Y. Xiong, “Super-resolution through neighbor embedding,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2004, pp. 1275–1282.

[6] Y. Zhuang, J. Zhang, and F. Wu, “Hallucinating faces: Lph super-resolution and neighbor reconstruction for residue compensation,” *Pattern Recognition*, vol. 40, no. 11, pp. 3178–3194, 2007.

[7] W. Liu, D. Lin, and X. Tang, “Neighbor combination and transformation for hallucinating faces,” in *IEEE International Conference on Multimedia and Expo*. IEEE, 2005, pp. 478–484.

[8] ——, “Hallucinating faces: Tensorpatch super-resolution and coupled residue compensation,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2005, pp. 478–484.

[9] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[10] L. Chang, M. Zhou, Y. Han, and X. Deng, “Face sketch synthesis via sparse representation,” in *20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 2146–2149.

[11] M. Aharon, M. Elad, and A. Bruckstein, “The k-svd: An algorithm for designing of overcomplete dictionaries for sparse representation,” *IEEE transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[12] X. Ma, H. Q. Luong, W. Philips, H. Song, and H. Cui, “Sparse representation and position prior based face hallucination upon classified over-complete dictionaries,” *Signal processing*, vol. 92, no. 9, pp. 2066–2074, 2012.

[13] C. Jung, L. Jiao, B. Liu, and M. Gong, “Position-patch based face hallucination using convex optimization,” *IEEE Signal Processing Letters*, vol. 18, no. 6, pp. 367–370, 2011.

[14] J. Zhang, C. Zhao, R. Xiong, S. Ma, and D. Zhao, “Image super-resolution via dual-dictionary learning and sparse representation,” in *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2012, pp. 1688–1691.

[15] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in *Curves and Surfaces*. Springer, 2010, pp. 711–730.

[16] M. Aharon, M. Elad, and A. Bruckstein, “K-svd: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[17] R. Timofte, V. Smet, and L. Gool, “Anchored neighborhood regression for fast example-based super-resolution,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1920–1927.

[18] R. Timofte and L. Van Gool, “Adaptive and weighted collaborative representations for image classification,” *Pattern Recognition Letters*, vol. 43, pp. 127–135, 2014.

[19] Z. Chai, H. Mendez-Vazquez, R. He, Z. Sun, and T. Tan, “Semantic pixel sets based local binary patterns for face recognition,” in *ACCV*. Springer, 2012, pp. 639–651.

[20] Z. Zhang, Y. Liang, L. Bai, and E. R. Hancock, “Discriminative sparse representation for face recognition,” *Multimedia Tools and Applications*, vol. 75, no. 7, pp. 3973–3992, 2016.

[21] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[23] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The feret evaluation methodology for face-recognition algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[24] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, “The cas-peal large-scale chinese face database and baseline evaluations,” *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 38, no. 1, pp. 149–161, 2008.

[25] A. M. Martinez, “The ar face database,” *CVC Technical Report*, 1998.