# BREAKING THE LIMITATION OF MANIFOLD ANALYSIS FOR SUPER-RESOLUTION OF FACIAL IMAGES

Sung Won Park and Marios Savvides
*Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213*
*sungwonp@andrew.cmu.edu, Marios.Savvides@ri.cmu.edu*

## ABSTRACT

A novel method for robust super-resolution of face images is proposed in this paper. Face super-resolution is a particular interest in video surveillance where face images have typically very low-resolution quality and there is a need to apply face enhancement or super-resolution algorithms. In this paper, we apply a manifold learning method which has hardly been used for super-resolution. A manifold is a natural generalization of a Euclidean space to a locally Euclidean space. Manifold learning algorithms are more powerful than other pattern recognition methods which analyze a Euclidean space because they can reveal the underlying nonlinear distribution of the face space; however, there are some practical problems which prevent these algorithms from being applied to super-resolution. Almost all of the manifold learning methods cannot generate mapping functions for new test images which are absent from a training set. Another factor is that super-resolution seeks to recover a high-dimensional image from a lower-dimensional one while manifold learning methods perform the exact opposite as they are applied to dimensionality reduction. In this paper, we break the limitation of applying manifold learning methods for face super-resolution by proposing a novel method using Locality Preserving Projections (LPP).

***Index Terms—*** super-resolution, face image analysis, locality preserving projections, manifold learning methods

## 1. INTRODUCTION

The face super-resolution task is to recover a high-resolution face image from a given low-resolution face image (e.g. captured from surveillance footage) by modeling the face image space. In video surveillance, it is often the case that the resolution of a captured facial image is not sufficient for face recognition even by a human being, so we need to recover higher-resolution images by super-resolution techniques. Baker et al. [1][2] developed a face hallucination method using a Bayesian formulation. This approach infers the high frequency components from a parent structure with the assistance of training samples. Rather than using the whole or parts of a face, the super-resolution is established based on training images (pixel by pixel) using Gaussian, Laplacian and feature pyramids.

A different model proposed by Liu et al. [3] describes a two-step approach integrating a global parametric model and a local nonparametric model. In the first step, the relationship between the high-resolution face images and their smoothed, down-sampled lower resolution ones is learned. In the second step, the residual between an original high-resolution image and a reconstructed one is re-compensated. Also, Wang et al. [4] developed an efficient face hallucination algorithm using an eigentransformation algorithm.

However, all these methods have not utilized the neighborhood relationship in the distribution of face images. Facial images change appearance due to multiple factors such as pose variations, lighting acquisition conditions and facial expressions. Previous work cited in literature has not paid attention to this distribution and has dealt with all the diverse images equally. Inspired by realizing the limitation of previous work, we propose a novel manifold learning approach using Locality Preserving Projections (LPP) for producing better super-resolution images.

In the previous work of face image analysis using manifold learning methods, it has been shown that face images lie on a manifold [5][6][7][8]. Also, it has been demonstrated that the variation of a certain facial factor such as pose or expression makes a sub-manifold in the manifold structure [6][9]. Thus, it is expected that manifold learning methods can improve the tasks demanding face image analysis, such as face recognition, super-resolution, or face synthesis. Based on this idea, Chang et al. developed [10] the Neighbor Embedding algorithm for super-resolution of general images. They assume that the local distribution structure in sample space is preserved in the down-sampling process, and apply one of the manifold learning methods, Locally Linear Embedding (LLE) [5].

However, two problems mainly prevent manifold analysis from applying to face super-resolution. First of all, most of the manifold learning methods such as LLE do not clearly define mapping functions for new test images which are absent from a training set. So, if we still want to use those methods for super-resolution, we define the way to generate mapping functions for unseen test images [10]. Moreover, super-resolution is to recover a high-dimensional image from a low-dimensional one whereas manifold learning methods are more suited for dimensionality reduction. In this paper, we propose a new method for face super-resolution, breaking the current limitation which hinders using manifold learning methods for super-resolution. We apply LPP [6] in this paper since it has an advantage over LLE and other manifold learning methods; LPP has a well-defined mapping for new test points. Additionally, we show that we can employ a MAP estimator to infer the LPP coefficients of a high-resolution image from a low-resolution image. Finally, in this paper, we break the two main obstacles to applying manifold analysis to face super-resolution.

In section 2, we introduce the concept of manifolds and LPP briefly, and in section 3, we explain the proposed super-resolution method for face images. In section 4, we demonstrate that the proposed method yields better results than other baseline algorithms for super-resolution. Section 5 summarizes the conclusions and discussions of the proposed method.

## 2. MODELING THE MANIFOLD STRUCTURE IN FACE IMAGE SPACE

PCA and LDA effectively see only the Euclidean structure; they fail to discover the underlying structure when the data lie on a nonlinear manifold. So, we need to detect and analyze the manifold structure underlying in the distribution of image samples. The analysis of manifolds reveals the characters of the distribution and can be applied to dimensionality reduction. Thus, to discover the nonlinear structure of manifolds, manifold learning techniques have been proposed.

### 2.1. Manifold Learning Methods Using Neighbor Embedding

In many real-world classification problems, the local manifold structure is more important than the global Euclidean structure. Thus, manifold learning techniques often use adjacency to preserve the local manifold structure. By manifold learning techniques, neighboring points should still be close after mapping, and the points far from each other should still be far from each other in the new mapping.

LPP is to find a linear projective mapping for dimensionality reduction. Compared to LPP, other manifold learning techniques such as Isomap [7], LLE [5], or Laplacian Eigenmap [8] define the mapping only on the training data. They successfully show the training data are distributed along manifolds, but it is unclear how to evaluate the maps for new test samples. On the other hand, by LPP, we attain the well-defined transformation matrix which is applicable to new test images absent from the training set.

**Algorithm 1. The Algorithm of LPP**

1. Constructing the adjacency graph: Let $G$ denote a graph with $m$ nodes. One node (or a training image) has K nearest neighbors in the meaning of Euclidean distance, and the neighbors are connected by edges.
2. Choosing the weights between neighbors: the weight between any two neighbors can be calculated by Gaussian kernel of the Euclidean distance. In this paper, binary kernel is used in this paper; $w_{ij}$ is set up as 1 if the two images $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ are connected by an edge, and otherwise $w_{ij}$ is set up as 0.
3. Eigenmaps: Compute the eigenvectors and eigenvalues for the generalized eigenvector problem:
$$\boldsymbol{XLX^T a} = \lambda \boldsymbol{XDX^T a} \quad (1)$$
where $\boldsymbol{D}$ is a diagonal matrix whose diagonal entries are $\boldsymbol{D_{ii}}=\sum_j \boldsymbol{W_{ij}}$, and $\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{W}$. Now, the projective matrix $\boldsymbol{A}$ has the eigenvectors $\boldsymbol{a_i}$ as column vectors.

### 2.2. Locality Preserving Projections

LPP is designed for optimally preserving the neighborhood structure of the data set while Principal Component Analysis (PCA) utilizes only a global basis. LPP is a novel method for dimensionality reduction by using both the local structure and the global basis of the data set. LPP aims to find a linear projection for dimensionality reduction such that the local structure of the data space is preserved. LPP utilizes the weight, which represents how close any two data points are in the data space. Using a set of these weights, we can extract a set of eigenvectors which represent both the global basis and the neighbor embedding in the data set.

When the high-dimensional data lies on a low-dimensional manifold embedded in data space, the locality preserving projections are obtained by finding the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the manifold. The algorithm of LPP is shown in Algorithm 1 [6].

## 3. FACE SUPER-RESOLUTION USING LPP

In this paper, we introduce super-resolution of facial images using a novel manifold method, Locality Preserving Projections (LPP). The goal of manifold methods is modeling the distribution of a set of data using both local and global structures. LPP is one of the novel manifold methods. Manifold methods have been used for image representation and recognition rather than super-resolution. Among the manifold methods, only LLE has been used for super-resolution problems. In this paper, we show LPP can be applied to super-resolution. LPP is better than other manifold methods, such as LLE, which define the mapping only on the training data.

### 3.1. Patch-based Modeling

Saul and Roweis [11] shows that local areas such as the mouths in face images also can be analyzed by manifold learning methods. Based on the idea, in this paper, we divide one image into multiple patches and perform super-resolution for each patch. Patch-based approaches for super-resolution using small image patches are less dependent on person-identity than global super-resolution approaches using a whole image.

In this paper, we select 24×24 pixel patches which overlap both horizontally and vertically with each other by two pixels, and remaining pixels compose one small patch. The size of patches and overlaps were empirically concluded optimal according to experimental results. To integrate all the patches, all the pixel values at the same position are added, and divided by the number of overlaps to normalize intensity. The reason why patches are overlapped is to remove strong difference of intensity at the boundary of two patches. If the overlapping area shrinks, effects of undesirable local distortions spread to neighboring patches.

### 3.2. Estimating the LPP Coefficients of High-Resolution

For every patch, its own LPP model is constructed to compute the mapping between a high-resolution patch $\boldsymbol{X_H}$ and a down-sampled patch $\boldsymbol{X_L}$. Given a high-resolution patch, the corresponding low-resolution patch is computed by down-sampling:

$$\boldsymbol{X}_L = \boldsymbol{B}\boldsymbol{X}_H \qquad (2)$$

where $\boldsymbol{B}$ is the transformation matrix for the mapping from high-resolution to low-resolution.

LPP aims to find a low dimensional embedding from a high dimensional patch, so it is proper to be used for dimensionality reduction. LPP has been applied to dimension reduction by projecting a high dimensional vector onto a low dimensional subspace. On the contrary, in this paper, we show LPP can be also applied to super-resolution problem; which is to map a low-resolution patch onto a high- resolution patch subspace.

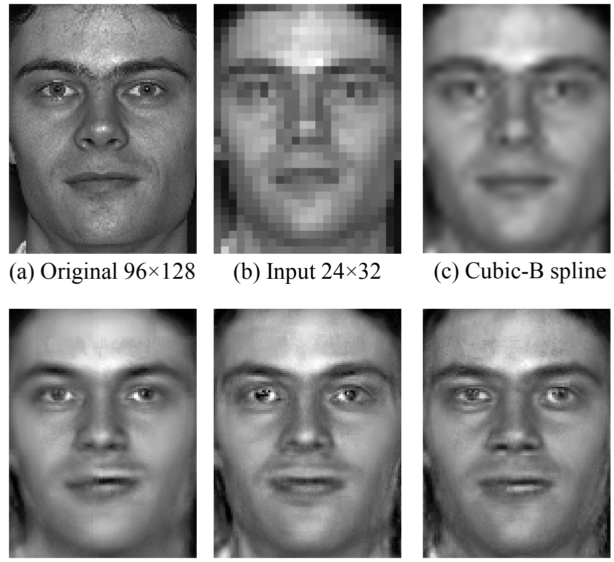
However, in the case of super-resolution, we must attain the coefficients in the high-resolution space from a given low-resolution patch. To do this estimation, various probability approaches such as Markov network or Belief Propagation have been employed to infer the coefficients of a high-resolution patch from a low-resolution patch [3][12][13].

We employ a Maximum *a posterior* (MAP) estimator to find the LPP coefficients of a high-resolution patch. A MAP method has been used to estimate the PCA coefficients of a high-resolution patch from a low-resolution one [3][13].

Given the patches taken from training images, $\{\boldsymbol{X}_H^{(i)}\}_{i=1}^{k}$, the LPP coefficients $\boldsymbol{Y}_H$ are calculated by

$$\boldsymbol{Y}_H = \boldsymbol{A}^T\boldsymbol{X}_H,\ \boldsymbol{X}_H = \boldsymbol{A}\boldsymbol{Y}_H \qquad (3)$$

where $\boldsymbol{A}$ is the projective matrix of LPP in eq.(3).



(a) Original 96×128 (b) Input 24×32 (c) Cubic-B spline

(d) Baker et al. (e) Chang et al. (f) Our method

Figure 1. Comparison between our method and others

Maximizing $p(\boldsymbol{X}_L \mid \boldsymbol{X}_H)p(\boldsymbol{X}_H)$ in eq. (3) is equivalent to maximizing $p(\boldsymbol{X}_L \mid \boldsymbol{Y}_H)p(\boldsymbol{Y}_H)$. The prior $p(\boldsymbol{Y}_H)$ is modeled by Gaussian distribution function:

$$p(\boldsymbol{Y}_H) = \frac{1}{\boldsymbol{Z}} exp(-\boldsymbol{Y}_H{}^T \Lambda^{-1} \boldsymbol{Y}_H), \qquad (4)$$

where $\Lambda = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_N^2)$ and $\boldsymbol{Z}$ is a normalization constant. The likelihood (4) is denoted by

$$p(\boldsymbol{X}_L \mid \boldsymbol{Y}_H) = \frac{1}{\boldsymbol{Z}} exp(-\|\boldsymbol{B}\boldsymbol{A}\boldsymbol{Y}_H - \boldsymbol{X}_L\|^2 / \lambda) \qquad (5)$$

To minimize $p(\boldsymbol{X}_L \mid \boldsymbol{Y}_H)p(\boldsymbol{Y}_H)$, the optimal $\boldsymbol{Y_H}$ is selected such that it satisfies the following objective:

$$\boldsymbol{Y}_H^* = arg\min_{\boldsymbol{Y}_H} \lambda \boldsymbol{Y}_H{}^T \Lambda^{-1} \boldsymbol{Y}_H + \|\boldsymbol{B}\boldsymbol{A}\boldsymbol{Y}_H - \boldsymbol{X}_L\|^2 \qquad (6)$$

Finally, the optimal solution is given by

$$\boldsymbol{Y}_H^* = (\boldsymbol{A}^T\boldsymbol{B}^T\boldsymbol{B}\boldsymbol{A} + \lambda\Lambda^{-1})^{-1}\boldsymbol{A}^T\boldsymbol{B}^T\boldsymbol{X}_L \qquad (7)$$

where $\lambda$ is decided empirically. If $\lambda$ is too small, $\boldsymbol{X}$ cannot be obtained because $\boldsymbol{A}^T\boldsymbol{B}^T\boldsymbol{B}\boldsymbol{A}$ is close to singular.

## 4. EXPERIMENTAL RESULTS

For experiments, a subset of the color FERET database [14] was used. We selected the images with neutral expression and frontal pose, and used 1500 images for training and 500 images for testing. Before experiments, the face images were aligned with given eye coordinates, cropped to 96×128 pixel images, and normalized by intensity. The high-resolution images were down-sampled to a low-resolution 24×36 pixel images. We choose $\lambda$=1000, and $K$=100.

The proposed method is compared to three baseline methods in Figure 1: cubic B-spline, face hallucination proposed by Baker et al. [1], and super resolution for

(a) 16×16 (b) 24×24 (c) 96×128

Figure 2. The effect of the size of a patch when *K* is 100.

(a) K=10 (b) K=100 (c) K=1000

Figure 3. The effect of the number of nearest neighborhoods when the size of a patch is 24×24.

general images proposed by Chang et al. [10]. Cubic B-spline is one of baseline methods of super-resolution, and the other two methods are state of the art. For implementing the algorithm in [10], LLE code obtained from [16] is used. Figure 1.(e) shows the proposed method recovers high-frequency and details more than the other methods.

Figure 2 shows that the size of a patch is important for getting reliable results. When a patch is too small, it loses the geometrical information of a human face, so the super-resolution reconstruction image becomes blurry as we can see in Figure 2.(a). On the other hand, as a patch becomes larger, it needs much more training images to extract reliable generalized basis. In particular, when a new test image is totally different from the training image, (for example, a new person absent from the training set is given in the test image) a large patch cannot generalize to reconstruct it. Figure 2.(c) shows the result when one image is used as one patch; it is significantly noisy. Thus, it is necessary to find the optimal patch size empirically.

The number of nearest neighbors for each patch also has significant impact on the super-resolution results. Figure 3 shows the manifold structure cannot be analyzed as *K* becomes too small or large. When *K* is too large, we cannot analyze neighborhood embedding while too small *K* makes it impossible to analyze the global structure in data space.

## 5. CONCLUSION

In this paper we proposed a novel approach for robust super-resolution by employing Locality Preserving Projections (LPP), one of the newest manifold learning algorithms. LPP is typically employed for dimensionality reduction applications that are mainly used to represent an image using a local structure as well as the global basis of an image set. However, we show in this paper that LPP can be successfully applied to super-resolution problem and provide superior reconstruction results. Manifold learning techniques are powerful tools that can analyze the adjacency relationship by preserving the local manifold structure. In particular, LPP calculates the basis of the training image by giving bigger weights to neighboring images such as those captured from similar lighting conditions. According to our experiments, we demonstrate that the proposed method produces reliable results and work well with untrained face data. The proposed method produces more robust visual reconstruction results with more facial details than other baseline methods. Future work will include examining the face recognition performance.

## REFERENCES


[1] S. Baker and T. Kanade, “Hallucinating Faces,” in Proc. of Inter. Conf. on Automatic Face and Gesture Recognition, pp.83-88, 2000.

[2] S. Baker and T. Kanade, “Limits on Super-Resolution and How to Break the,” IEEE Trans. on PAMI, Vol. 24, No. 9, pp.1167-1183, 2002.

[3] C. Liu, H. Shum, and C. Zhang, “A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Patter Recognition, Vol. 1, pp. 192-198, 2001.

[4] X. Wang and X. Tang, “Hallucinating face by eigentransformation,” IEEE Trans. on Systems, Man, and Cybernetics, Part-C, Special issue on Biometrics Systems, 2005.

[5] S. Roweis and L. K. Saul, “Nonlinear Dimensionality Reduction by Locally Linear Embedding,” Science, vol. 290(5500):2323-2326, 2000.

[6] X. He and O. Niyogi, “Locality preserving projections”, Advances in Neural Information Processing Systems, 16, 2003.

[7] J. B. Tenenbaum, V. Silva, and J.C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” Science, vol 290(12);2319-2323, 2003.

[8] M. Belkin and P. Niyogi, “Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering,” Advances in Neural Information Precessing Systems 14, pp. 585-591, MIT Press, Cambridge, MA, USA, 2002.

[9] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, “Face Recognition Using Laplacianfaces,” IEEE trans. on Pattern Analysis and Machine Intelligence, Vol. 27, No. 3, March 2005.

[10] H. Chang, D.Y. Yeung, and Y.Xiong, “Super-resolution through Neighbor Embedding,” in Proc. of CVPR, Vol.1, pp.275-282, 2004.

[11] L. K. Saul and S. T. Roweis, “Think Globally, Fit Locally: Unsupervised Learning of Low Dimensional Manifolds,” Journal of Machine Learning Research, Vol. 4, pp. 119-155, 2003.

[12] D. Capel and A. Zisserman, “Super-resolution from multiple views using learnt image models,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Patter Recognition, Vol. 2, pp. 627-634, 2001.

[13] W. T. Freeman, E. C. Pasztor, “Learning low-level vision,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Patter Recognition, Vol. 2, pp. 1182-1189, 1999.

[14] http://www.nist.gov/humanid/colorferet

[15] P. Philips, H. Moon, P. Pauss, and S.Rivzvi. “The FERET Evaluation Methodology for Face-Recognition Algorithms,” in Proc. of CVPR, pp. 137 - 143, 1997.

[16] http://www.cs.toronto.edu/~roweis/lle/code.html