LOCALITY PRESERVING CONSTRAINTS FOR SUPER-RESOLUTION WITH NEIGHBOR EMBEDDING

Bo Li[1]*, Hong Chang*[23]*, Shiguang Shan*[23]*, Xilin Chen*[23]

[1]School of Computer Science and Technology, Harbin Institute of Technology, Harbin, 150001, China
[2]Key Lab of Intelligent Information Processing, Chinese Academy of Sciences, Beijing, 100190, China
[3] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China
{bli, hchang, sgshan, xlchen}@jdl.ac.cn

ABSTRACT

In this paper, we revisit the manifold assumption which has been widely adopted in the learning-based image super-resolution. The assumption states that point-pairs from the high-resolution manifold share the local geometry with the corresponding low-resolution manifold. However, the assumption does not hold always, since the one-to-multiple mapping from LR to HR makes neighbor reconstruction ambiguous and results in blurring and artifacts. To minimize the ambiguous, we utilize Locality Preserving Constraints (LPC) to avoid confusions through emphasizing the consistency of localities on both manifolds explicitly. The LPC are combined with a MAP framework, and realized by building a set of cell-pairs on the coupled manifolds. Finally, we propose an energy minimization algorithm for the MAP with LPC which can reconstruct high quality images compared with previous methods. Experimental results show the effectiveness of our method.

Index Terms— Super-resolution, Neighbor embedding, Manifold assumption, Locality preserving constraints

1. INTRODUCTION

Super-resolution (SR) aims to estimate a high-resolution (HR) image from one or multiple given low-resolution (LR) counterparts. Among existing methods, interpolation-based techniques usually predict a target pixel using its spatial neighbors. However, this simple method results in blurring because the prior of the smoothness does not fit for the area of high-frequency. Thus, some edge preserving methods [1, 8] have been proposed to address this problem. Another class of methods is reconstruction-based approaches, which claims that the down-sampled HR estimation should be as close as possible to the input LR image. Back-projection (BP) algorithm [6] is proposed to minimize the reconstruction error in an iterative manner. Besides, some other methods [9, 11] regularized by priors (e.g., edges-direct) are proposed to avoid artifacts.

Following the early works [2, 4], learning-based SR methods have been extensively studied. Baker and Kanade [2] developed a method named "face hallucination" to infer HR face image from an LR input. Freeman et al. [4] use a Markov network to model the relationship between the LR images and the HR counterparts. In [7], authors introduce the primal sketch priors (e.g., edges) into the learning-based SR framework to further improve the performance.

Inspired by locally linear embedding (LLE) [10], a series of SR methods based on a manifold assumption (called LLE-like methods) have been presented. The assumption claims that point-pairs from the LR manifold (LRM) and the corresponding HR manifold (HRM) possess similar local geometry. Chang et al. [3] propose an algorithm based on NE which has fairly good performance. Researchers combine the benefits from the priors of image primitives and the LLE in [5]. Recently, authors of [12] propose the sparse representation into NE by L1-norm regularization. Compared with previous learning-based methods, the LLE-like synthesis manner greatly improves the SR performance.

Nevertheless, most of the previous work, does not discuss whether the coupled manifolds, LRM and HRM, meet the manifold assumption. Authors of [5] just validate that the assumption is more consistent with primitive patches than general ones. In this paper, we revisit the manifold assumption and find that the assumption does not hold well on the coupled manifolds. The intrinsic reason is that the one-to-multiple mapping from LR to HR makes the points selected by k nearest neighbors (k-NN) algorithm on LRM are non-local on HRM. To avoid this confusion caused by non-local reconstruction, we introduce constraints, called as Locality Preserving Constraints (LPC), to guarantee the locality and the consistency on the coupled manifolds. Furthermore, we propose a synthesizing algorithm with LPC to improve the SR results.

2. MANIFOLD ASSUMPTION

In this paper, we present SR problem in patch-wise as follows: to estimate a HR image given a LR image with the help of a training set, where LR and HR images are represented by two patch sets, $\{\mathbf{l}_x\}_{x=1}^{K}$ and $\{\mathbf{h}_x\}_{x=1}^{K}$, respectively. x indicates the patch site in the image and K is the number of patches. The raw training set consists of a collection of HR natural images and LR counterparts which

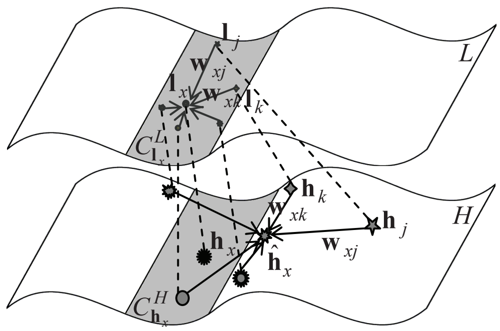

Fig. 1. Neighbor embedding based on manifold assumption.

are generated by degrading HR images using blurring and sub-sampling. Then, we cut the HR-LR image pairs into two patches set, $\mathbf{L}=\{\mathbf{l}_i\}_{i=1}^N$ and $\mathbf{H}=\{\mathbf{h}_i\}_{i=1}^N$, which are considered to be sampled from LRM and HRM. Here, *i* represents the patch index and *N* is the number of patches.

Generally, the original LLE-like algorithm, illustrated in Fig. 1, consists of following steps: (1) finding its *k* nearest neighbors in $\mathbf{L}$ for each query LR patch $\mathbf{l}_x$,; (2) calculating the weights $\mathbf{w}_x$ by reconstructing $\mathbf{l}_x$ using its LR neighbors; (3) transferring the weights to the corresponding HR patches on $\mathbf{H}$, and estimate the target HR patch $\mathbf{h}_x$ by the linear combination of its HR neighbors using the calculated weights. The last two steps can be formulated as the following mathematical equations:

$$\mathbf{w}_x = \arg\min_{\mathbf{w}_x} \| \mathbf{l}_x - \sum_{j\in N(x)} \mathbf{w}_{xj}\mathbf{l}_j \|^2, \qquad (1)$$

and

$$\hat{\mathbf{h}}_x = \sum_{j\in N(x)} \mathbf{w}_{xj}\mathbf{h}_j . \qquad (2)$$

Here, $N(x)$ consists of the neighbor indexes of patch $\mathbf{l}_x$.

Based on the manifold assumption, the reconstruction weights of a LR patch should be similar with the weights of reconstructing the HR counterpart with the corresponding HR neighbors. However, this is not always the case since an implicit precondition for the manifold assumption is ignored. In this paper, we refer this pre-condition as Locality Preserving Constraints (LPC), which states that neighbor embedding should be carried out on corresponding locally linear regions on the coupled manifolds. But, original NE algorithms do not consider the locality issues of two manifolds at the same time (especially the locality on HRM).

As shown in Fig. 1, the gray regions, $C^{\mathbf{L}}_{\mathbf{l}_x}$ and $C^{\mathbf{H}}_{\mathbf{h}_x}$, represent locally linear regions on manifolds, which are referred as cells. We can see that the LR neighbors of patch $\mathbf{l}_x$ lie in one cell while the corresponding HR patches scatter in different cells. This situation violates the manifold assumption. As a result, the estimate $\hat{\mathbf{h}}_x$ using neighbor embedding is apparently different from the ground-truth $\mathbf{h}_x$. Essentially, this confusion is mainly caused by the one-to-multiple mapping from LR to HR samples.

In SR, the one-to-multiple mapping means that one LR patch may correspond to several different HR patches. Fig. 2 illustrates the one-to-multiple mapping and the neighbor reconstruction results for 1-D signal. Markers and curves of different styles denote signals in low frequency (LF) and corresponding high frequency (HF) domains, respectively. More specifically, gray circles represent LF input signals. Fig. 2(a) shows the neighbor signals may be selected by original NE, which are close in LF, but very different in HF. The reason is that the locality on the coupled signal manifolds does not be considered simultaneously. In comparison, the neighbors selected with LPC guarantee consistency in both domains, as shown in Fig. 2(b). Fig. 2(c) shows the reconstruction results of NE (blue dotted curve) and LPC (red dashed curve). Obviously, the result of LPC is much more stable and accurate than that of the original NE.

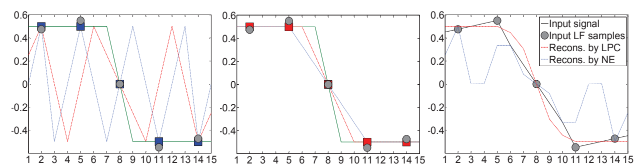

Fig. 2. Illustration of the one-to-multiple mapping and the neighbor reconstruction. (a) three signals selected by original NE; (b) three signals selected with LPC; (c) reconstructed signals in HF. In (c), Solid curve denotes the ground-truth; Blue dotted curve denotes result signal by original NE; Red dashed curve denotes result signal by LPC.

From the analysis above, LPC reduces the confusion caused by one-to-multiple mapping and constrains NE to be processed in cell-pairs. In next section, we will formulate the problem as a MAP with LPC and present a SR algorithm.

3. SUPER RESOLUTION BASED ON LPC

We propose a probabilistic model, where the SR problem is formulated in a MAP framework as follows:

$$\hat{I}_H = \arg\max p(I_H \mid I_L). \qquad (3)$$

When some priors or constraints such as manifold assumption are introduced, the model will transfer into

$$\hat{I}_H = \arg\max p(I_H \mid I_L, C). \qquad (4)$$

Here, C denotes the priors, and LPC is employed as the constraints in our scheme. Following Bayes rule, we transform (4) into

$$p(I_H \mid I_L, C) \propto p(C \mid I_H, I_L)\, p(I_L \mid I_H)\, p(I_H). \qquad (5)$$

We define the last two terms of (5) as

$$p(I_L \mid I_H) \propto \prod_x \exp(-E_r(\mathbf{l}_x, \mathbf{L})), \qquad (6)$$

and

$$p(I_H) \propto \prod_{x,y} \exp(-E_s(\mathbf{h}_x, \mathbf{h}_y)). \qquad (7)$$

In (6), $E_r(\mathbf{l}_x, L) = \| \mathbf{l}_x - \sum_{j\in N(x)} \mathbf{w}_{xj}\mathbf{l}_j \|^2$ computes the error of reconstructing patch $\mathbf{l}_x$ with its neighbors on the LRM $\mathbf{L}$. As for $p(I_H)$, we use the smoothness of the estimate HR image as the priors and define the smoothness function E_s using the Sum Squared Difference function of the overlapping regions between adjacent HR patches. As in (5), we introduce LPC term $p(C \mid I_H, I_L)$ into the MAP framework. Based on previous analysis, we propose LPC as

$$p(C \mid I_H, I_L) \propto \prod_x \exp(-E_l(\mathbf{l}_x; \mathbf{L})) \exp(-E_l(\mathbf{h}_x; \mathbf{H})). \quad (8)$$

In this formula, $E_l(\mathbf{l}_x; \mathbf{L})$ and $E_l(\mathbf{h}_x; \mathbf{H})$ are used to measure the locality on LRM and HRM respectively. Locality functions E_l are specialized as

$$E_l(C_{\mathbf{p}}^M, \mathbf{p}') = \exp\left(-\sum_{\mathbf{p}_i \in C_{\mathbf{p}}^M} (\mathbf{p}_i - \mathbf{p}')^T (\mathbf{p}_i - \mathbf{p}') / K_{\mathbf{p}}^M\right). \quad (9)$$

$C_{\mathbf{p}}^{\mathbf{M}}$ represents a cell that consists of the neighborhood set of point $\mathbf{p}$ on manifold $\mathbf{M}$. We define locality using a Gaussian function of point $\mathbf{p}'$ w.r.t. the points in cell $C_{\mathbf{p}}^{\mathbf{M}}$ on manifold $\mathbf{M}$. The locality function can measure to what degree $\mathbf{p}'$ belongs to cell $C_{\mathbf{p}}^{\mathbf{M}}$ on manifold $\mathbf{M}$. Here, $K_{\mathbf{p}}^{\mathbf{M}}$ denotes the number of points involved in $C_{\mathbf{p}}^{\mathbf{M}}$ and we set $K_{\mathbf{p}}^M = 64$ in our experiments.

From above, the main step of implementation of LPC in the proposed method is to build a set of cell-pairs on the coupled manifolds. Each cell-pair should satisfy the locality property on the coupled manifolds simultaneously. We establish the cell-pairs from training set by the following simple steps. Firstly, we define augmented vectors $\mathbf{P} = \{\mathbf{p}_i\}_{i=1}^N$ with $\mathbf{p}_i = [\mathbf{l}_i^T\ \mathbf{h}_i^T]^T$. Then, cells $C_{\mathbf{p}_i}^{\mathbf{P}}$ for point $\mathbf{p}_i$ are calculated by its k nearest neighbors in $\mathbf{P}$. Finally, $C_{\mathbf{p}_i}^{\mathbf{P}}$ is separated into LR part $C_{\mathbf{l}_i}^{\mathbf{L}}$ and HR part $C_{\mathbf{h}_i}^{\mathbf{H}}$. It is clear that each cell-pair contains two cells which achieve locality on the manifold $\mathbf{L}$ and the manifold $\mathbf{H}$. In the end, we learn a set of coupled cell-pairs $\mathbf{T} = \{C_{\mathbf{p}_i}^{\mathbf{P}}\}_{i=1}^N$, and $C_{\mathbf{p}_i}^{\mathbf{P}} = (C_{\mathbf{l}_i}^{\mathbf{L}}, C_{\mathbf{h}_i}^{\mathbf{H}})$.

The proposed SR algorithm based on LPC consists of two phases, training phase and reconstruction phase. At training phase, we combine the preprocessing in [4] and [7]. The original train set consists of LR images I_{L} and HR counterparts I_{H}. Firstly, I_{L} is magnified into I_H^l with the same size of I_{H} by interpolation. I_H^h is the differences between I_{H} and I_H^l, which denotes the high-frequency (HF) part of HR image. In our system, actually, I_H^l just saves the medium-frequency (MF) part of HR image, which is the convolution of I_H^l and a band-pass filter. As the work in [7], we focus our capability on the regions near the image contours which are close to the missing HF part. Finally, the MF and HF patches along the contours are collected to form the patch-wise training set for our system. Besides, the normalization method in [4] is adopted to avoid influence of the illumination. After the preprocessing steps, we carry out the foregoing scheme to learn LPC on training set.

In reconstruction phase, the MAP estimation in (4) is transformed into an energy minimization problem. For each patch, our algorithm is to minimize the function as

$$\begin{aligned} E(\mathbf{h}_x) = {} & E_l(\mathbf{l}_x, C_{\mathbf{l}_i}^{\mathbf{L}}) + \lambda_l E_l(\mathbf{h}_x, C_{\mathbf{h}_i}^{\mathbf{H}}) \\ & + \lambda_r E_r(\mathbf{l}_x, C_{\mathbf{l}_i}^{\mathbf{L}}) + \lambda_s \sum_y E_s(\mathbf{h}_x, \mathbf{h}_y) \end{aligned} \quad (10)$$

Here, parameters $\lambda_l = 1$, $\lambda_r = 1$ and $\lambda_s = 0.3$ serve to balance different terms in the energy function. If we set $\lambda_l = 0$, our method will degrade to the original NE method. Since the iterative optimization for each patch is time-consuming, we develop a non-iterative approximate algorithm described in Algorithm I.

Algorithm I. Neighbor Embedding Algorithm based on LPC

Input: LR patches $I_L = \{\mathbf{l}_x\}_{x=1}^K$ and cell-pairs $\mathbf{T} = \{C_{\mathbf{p}_i}^{\mathbf{P}}\}_{i=1}^N$.

Output: HR patches $I_H = \{\mathbf{h}_x\}_{x=1}^K$.

For each query primitive $\mathbf{l}_x$,

1. Sample 16 cell candidates $C_{\mathbf{p}_i}^{\mathbf{P}}$ (for speed-up) according to the first 16 minimums of
$$E_l(C_{\mathbf{l}_i}^{\mathbf{L}}, \mathbf{l}_x) = \sum_{\mathbf{l}_j \in C_{\mathbf{l}_i}^{\mathbf{L}}} (\mathbf{l}_j - \mathbf{l}_x)^T (\mathbf{l}_j - \mathbf{l}_x) / K_{\mathbf{l}_i};$$
2. Use each LR cell candidate $C_{\mathbf{l}_i}^{\mathbf{L}}$ and the corresponding HR cell $C_{\mathbf{h}_i}^{\mathbf{H}}$ to calculate reconstruction weights $\mathbf{w}_x$ by
$$\arg\min_{\mathbf{w}_x} \left(\| \mathbf{l}_x - \sum_{\mathbf{l}_j \in C_{\mathbf{l}_i}^{\mathbf{L}}} \mathbf{w}_{xj} \mathbf{l}_j \|^2 + \sum_y E_s(\mathbf{h}_y, \sum_{\mathbf{h}_j \in C_{\mathbf{h}_i}^{\mathbf{H}}} \mathbf{w}_{xj} \mathbf{h}_j) \right)$$
s.t. $\sum_j \mathbf{w}_{xj} = 1$. Where, y is x's adjacent sites;
3. Reconstruct each HR candidate $\hat{\mathbf{h}}_x = \sum_{\mathbf{h}_j \in C_{\mathbf{h}_i}^{\mathbf{H}}} \mathbf{w}_{xj} \mathbf{h}_j$;
4. Compute $\mathrm{E}(\hat{\mathbf{h}}_x)$ for each candidate $\hat{\mathbf{h}}_x$ by (10);
5. Output $\mathbf{h}_x = \arg\min_{\hat{\mathbf{h}}_x} \mathrm{E}(\hat{\mathbf{h}}_x)$;

Repeat Step 1 to 5 until all HR patches are estimated.

In the proposed algorithm, for efficiency, we firstly sample cell candidates by the locality function on LRM at Step 1 for efficiency. Then, at Step 2 and 3, we optimize reconstruction weights and predict HR patch on each candidate cell pair. Finally, we select the HR patch which minimizes the energy function (10) as final estimation. In our algorithm, we just use 5 nearest neighbor as [3, 5] in Step 2. The experimental results will indicate that our algorithm is fairly effective.

4. EXPERIMENTAL RESULTS

A patch-wise training set with size of 21,000 is extracted from 12 HR images and LR counterparts. In our experiments, the all patches with size of 7×7 pixels are expanded by rotations to different orientations (90°, 180°, and 270°). To achieve the translation invariance, we align the centers of patches on contours. In this evaluation, we only test our approach on gray images or luminance channel (interpolation is carried out on the color channels).

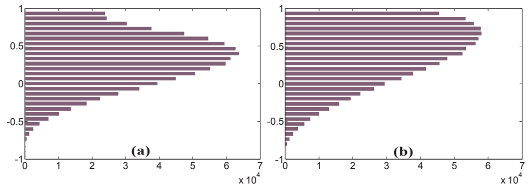

Fig. 3. Histograms of correlation coefficient between reconstruction weights of LR and HR patches. (a) original NE algorithm; (b) LPC-based algorithm.

Firstly, we verify the manifold assumption on the training set. Essentially, the assumption means that the reconstruction weights of a LR patch on LRM and that of its HR counterpart on HRM should be the same. Here, we use the standard linear correlation coefficient to measure this consistency. Fig. 3 shows two histograms of the correlation coefficients between the reconstruction weight vectors. Fig. 3(a) and 3(b) are calculated for NE [3] and our LPC-based method on training set, respectively. It is clear that the correlation coefficients in (b) are closer to 1 than that in (a), indicating that the local geometry is more consistent on the neighborhoods under LPC. In Fig. 4, we give an example to compare the original NE and the LPC-based method. It is easy to see that with LPC, the neighbors selected are more compact, and the estimate target patch is more accurate.

Fig. 4. An example of SR reconstruction. From top to bottom, (a) input LR patch, LR reconstructions by NE, and by LPC; (b) five LR neighbors selected by NE, and by LPC; (c) five HR neighbors selected by NE, and by LPC; (d) HR ground-truth, HR estimates by NE, and LPC;

In the second experiment, we compare our method with spline-interpolation and NE in Chang et al.'s work [9]. In Fig. 5, we can see that our HR results seem to be sharper and smoother without blurring and artifacts. These effects are more obvious at places with sharp edges: the petals and stamens of the flower and so on. We also calculate the RMS on a test set of 14 images which are different with training images. The average RMS errors of spline-interpolation, NE and our method are 5.90, 5.71 and 5.31, respectively. More results will be found in our website[1]. Our method improves the performance of LLE-like SR methods.

5. CONCLUSION

In this paper, we revisit the manifold assumption which is the basis of some learning-based SR algorithms. We analyze the limitations of these methods and propose a novel SR algorithm based on Locality Preserving Constraints. Our method calculates LPC from the training set simply and then processes neighbor embedding with the constraints on the coupled cell-pairs learned from the manifolds. Encouraging results verify the effectiveness of our method for single image SR problems. The proposed LPC can be viewed as general constraints for a class of methods with manifold assumption. Possibilities of applying LPC to other coupled manifolds learning problems will be pursued in future.

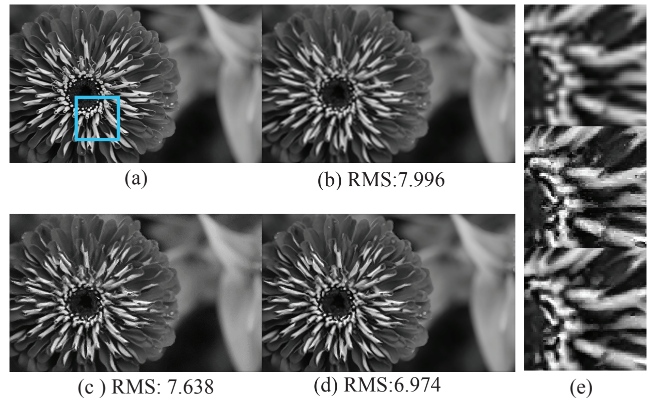

Fig. 5. 3X magnification of flower. (a) The original HR image. (b) The results by spline interpolation. (c) The result by Neighbor Embedding in Chang et al.'s work [3]. (d) The result by our method. (e) Patches from top to bottom correspond to the blue box regions in (b), (c) and (d) respectively.

6. ACKNOWLEDGEMENT

This paper is partially supported by Natural Science Foundation of China under contracts No.U0835005 and No.60803084; Grand Program of International S&T Cooperation of Zhejiang Province S&T Department under contract No. 2008C14063; Co-building Program of Beijing Municipal Education Commission; and ISVISION Technology Co. Ltd.

7. REFERENCES

[1] J. Allebach and P. W. Wong, "Edge-directed Interpolation," ICIP, 1996.
[2] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," PAMI, 24(9), pp. 1167-1183, 2002.
[3] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through Neighbor Embedding," CVPR, pp. 275-282, 2004.
[4] W. Freeman, E. Pasztor, and O. Carmichael, "Learning Low-level Vision," IJCV, vol. 40, no. 1, pp. 25–47, 2000.
[5] W. Fan, D. Yeung, "Image Hallucination Using Neighbor Embedding over Visual Primitive Manifolds," CVPR, 2007.
[6] M. Irani and S. Peleg, "Motion Analysis for Image Enhancement: Resolution, Occlusion and Transparency," JVCI, 1993.
[7] J. Sun, N. Zheng, H. Tao, and H. Y. Shum, "Image Hallucination with Primal Sketch Priors," CVPR, pp. 729-736, 2003.
[8] B. S. Morse and D. Schwartzwald, "Isophote-based Inter-polation," ICIP, pp. 227–231, 1998.
[9] B. S. Morse and D. Schwartzwald, "Image Magnification Using Level Set reconstruction," CVPR, pp. 333-340, 2001.
[10] S. T. Roweis, L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," Science, vol. 290, 2000.
[11] Y. W. Tai, W. S. Tong, and C. K. Tang, "Perceptually-inspired and Edge-directed Color Image Super-resolution," CVPR, pp. 1948-1955, 2006.
[12] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image Super-Resolution as Sparse Representation of Raw Image Patches," CVPR, pp. 1-8, 2008.

[1] http://www.jdl.ac.cn/user/bli/publication