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/145309/

Version: Accepted Version

Article:
Zhang, Zhihong, Xu, Chen, Zhang, Zhonghao et al. (2019) Single Image Super Resolution via Neighbor Reconstruction. Pattern Recognition Letters. ISSN: 0167-8655

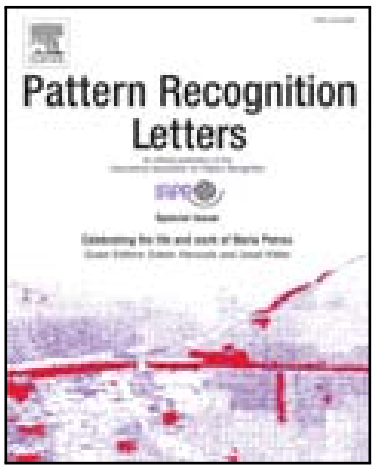
https://doi.org/10.1016/j.patrec.2019.04.021

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

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

Accepted Manuscript



Single Image Super Resolution via Neighbor Reconstruction

Zhihong Zhang, Chen Xu, Zhonghao Zhang, Guo Chen, Yide Cai, Zeli Wang, Heng Li, Edwin R. Hancock

PII: S0167-8655(19)30136-9
DOI: https://doi.org/10.1016/j.patrec.2019.04.021
Reference: PATREC 7503

To appear in: *Pattern Recognition Letters*

Received date: 17 January 2019
Revised date: 3 April 2019
Accepted date: 22 April 2019

Please cite this article as: Zhihong Zhang, Chen Xu, Zhonghao Zhang, Guo Chen, Yide Cai, Zeli Wang, Heng Li, Edwin R. Hancock, Single Image Super Resolution via Neighbor Reconstruction, *Pattern Recognition Letters* (2019), doi: https://doi.org/10.1016/j.patrec.2019.04.021

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights

- We present a novel regression-based SR method that is built on neighbor reconstruction.
- We designed a new projector which has better numerical stability to adapt to our new problem.
- When the harvested samples are sparse on the manifold, our method can still construct much closer points.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ELSEVIER

Pattern Recognition Letters
journal homepage: www.elsevier.com

Single Image Super Resolution via Neighbor Reconstruction

Zhihong Zhang[a], Chen Xu[b], Zhonghao Zhang[c], Guo Chen[c], Yide Cai[d], Zeli Wang[e,**], Heng Li[e], Edwin R. Hancock[f]

[a]*Xiamen University, Xiamen, China*
[b]*China Energy Engineering Group Shaanxi Electric Power Design Institute Co., LTD, Xian, China*
[c]*State Grid Shaanxi Information and Telecommunication Company, LTD, Xian, China*
[d]*Beijing Normal University, Beijing, China*
[e]*The Hong Kong Polytechnic University, Hongkong, China*
[f]*University of York, York, UK*

ABSTRACT

Super Resolution (SR) is a complex, ill-posed problem where the aim is to construct the mapping between the low and high resolution manifolds of image patches. Anchored neighborhood regression for SR (namely A+ (Timofte et al., 2014)) has shown promising results. In this paper we present a new regression-based SR algorithm that overcomes the limitations of A+ and benefits from an innovative and simple Neighbor Reconstruction Method (NRM). This is achieved by vector operations on an anchored point and its corresponding neighborhood. NRM reconstructs new patches which are closer to the anchor point in the manifold space. Our method is robust to NRM sparsely-sampled points: increasing PSNR by 0.5 dB compared to the next best method. We comprehensively validate our technique on standardised datasets and compare favourably with the state-of-the-art methods: we obtain PSNR improvement of up to 0.21 dB compared to previously-reported work.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The purpose of single image super-resolution (SR) is to estimate a high resolution (HR) image from a single low resolution (LR) image. It provides a way to enhance the existing images which were generated by delayed imaging equipment or limited imaging conditions, and have been widely studied in recent years. Acquiring a HR estimation from an LR observation is an ill-posed problem and so priors of high quality images are normally relied on in the estimation process. Based on the different priors, existing single image SR methods can be broadly classified into three categories: interpolation-based methods (Irani and Peleg, 1991; Duchon, 1979; Li and Orchard, 2001; Fattal, 2007; Freeman et al., 2002), reconstruction-based methods (Chang et al., 2004; Glasner et al., 2009; Protter et al., 2009) and example learning-based methods (Dai et al., 2015; Dong et al., 2011; Cui et al., 2014; Kim and Kwon, 2010; Zhang et al., 2015; Timofte et al., 2013; Dong et al., 2016; Timofte et al., 2014).

Interpolation-based methods use priors based on rigid models of the imaging process. The unknown pixel values are estimated by interpolation (i.e. bilinear, bicubic and cubic spline interpolation). Representative methods include Iterative Back Projection (IBP) (Irani and Peleg, 1991), Lanczos upsampling (Duchon, 1979) and New Edge Directed Interpolation (NEDI) (Li and Orchard, 2001). Although such generative methods are able to capture some of the characteristics of high quality images, they cannot recover the high-frequency information in texture regions and also produce many ringing and jaggy artifacts along edges since no new information is added in the procedure.

Reconstruction-based methods view the SR problem as an inverse problem and impose reconstruction constraints on the HR image estimation. Such constraints aim to find a downsampled and blurred HR image which is well approximated by the LR input image. However, artifacts like jaggies and ringing may be introduced in SR results because of the ill-conditioned nature of deconvolution of the blur operation. To stabilise the estimated image and suppress artifacts, a prior knowledge is combined with the reconstruction constraint to regularize the reconstruction results. Representative priors, such as the soft

**The first two authors contribute equally to this work.
**Corresponding author
e-mail: jerry.wang@connect.polyu.hk. (Zeli Wang)

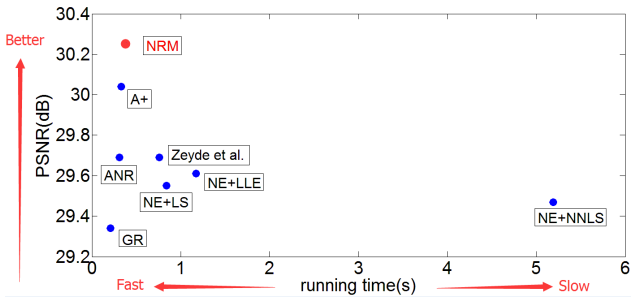
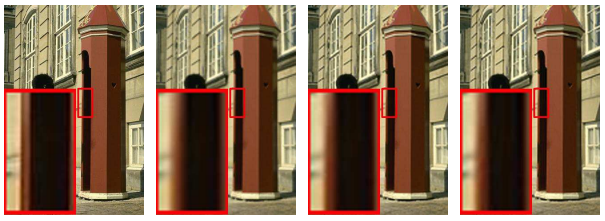
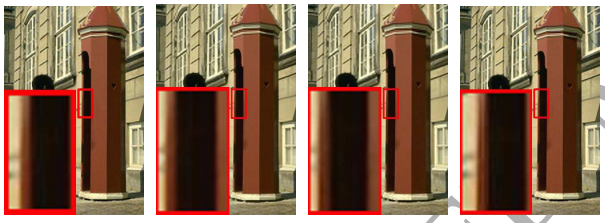


Fig. 1: Average PSNR (dB) vs time (s) of our algorithm (NRM) compared to other SR methods. We largely improve (red) over the original example based single image super-resolution methods (blue), i.e. our NRM method is 0.21dB better than A+(Timofte et al., 2014) and 0.91dB better than the Global Regression (GR)(Timofte et al., 2013). Results reported on Set5 with magnification 4. Details in Section 4.



(a) Original (b) Bicubic, 23.2(dB) (c) Zeyde et al, 24.1(dB) (d) NE+LLE, 24.0(dB)



(e) A+, 24.4(dB) (f) SRCNN, 24.5(dB) (g) RFL, 24.4(dB) (h) Proposed, 24.5(dB)

Fig. 2: Visual qualitative assessment for soldier image with magnification 3.

edge smoothness prior proposed in (Dai et al., 2007), similarity redundancy priors (Zhang et al., 2012) and total variation regularization (Marquina and Osher, 2008) are widely used in reconstruction-based methods. Although the prior knowledge can produce sharp edges and suppress aliasing effects, they do not add new high-frequency details that are lost in degradation, especially at high magnification (e.g. grater than $\times$ 2).

Example learning-based SR methods are superior to reconstruction-based methods since they are able to produce novel details that cannot be found in the LR input. These approaches exploit the information from a training dataset composed of millions of co-occurring LR and HR image patch pairs or a learned LR-HR overcomplete dictionary pair to estimate the relationship between the LR and HR image patches for SR reconstruction. One of the most successful learning approaches is the sparse representation-based approach. For example, Yang *et al* proposed to use a pair of LR and HR dictionaries to model the relationship between LR and HR patches in (Yang et al., 2010), this leads to a family of sparse representation methods, including the efficient K-SVD/OMP method of Zeyde (Zeyde et al., 2010). Apart from building relationships in the sparse coefficient domain, different regression methods have been utilized to model the relationship between LR and HR images; These include local regression (e.g. the Neighbor Embedding with Locally Linear Embedding (NE+LLE) (Chang et al., 2004), the Simple Functions (SF) method of Yang and Yang (Yang and Yang, 2013), the Anchored Neighborhood Regression (ANR) method introduced by Timofte *et al* (Timofte et al., 2013) and the Adjusted Anchored Neighborhood Regression (A+) by the same authors (Timofte et al., 2014)) and the convolutional neural network method (SR-CNN) of Dong *et al* (Dong et al., 2016).

Among the above mapping-based methods, neighbor embedding approaches have achieved great research interests. In (Timofte et al., 2013), Timofte *et al* proposed a highly efficient and effective SR algorithm called ANR, which maps the LR patches onto the HR domain using the projections learned form neighborhoods. Specifically, it relaxes the ℓ_1-norm regularization commonly used in most of the neighbor embedding and sparse coding approaches (Zeyde et al., 2010; Yang et al., 2010) to a ℓ_2-norm regularized regression which can be solved offline and stored for each dictionary atom/anchor. This results in large speed benefits. Subsequently, those authors proposed an improved variant of the ANR method called A+ (Timofte et al., 2014) that learns the regressors from the locally nearest training LR and HR patches instead of the small dictionary. It thus better utilizes the prior data to achieve improved performance. Under the framework of A+, many notable methods such as the Half Hypersphere Confinement Regression (HHCR) (Salvador et al., 2016), the Patch Symmetry Collapse (PSyCo) (Prez-Pellitero et al., 2016) and RFL (Schulter et al., 2015) were proposed.

Although the A+ method (Timofte et al., 2014) has achieved great success in delivering high quality HR estimation, it has two serious limitations: First, to obtain dense sample patches, A+ needs to harvest data images with different scales repeatedly, resulting in a large amount of computation and storage; Second, even if A+ does a so-called densely harvesting, we find that these patches are still too sparse for the high dimension space.

In this paper, we propose a novel and simple neighbor reconstruction method and extend the concept of A+ resulting in a significant improvement.

1. Compared with A+, our method utilizes fewer features to construct a closer neighbor and that results in a more accurate reconstruction coefficient vector $\mathbf{x}$. Specifically, we present a new neighbor reconstruction method which adds an anchor point and its corresponding neighbor features together and divides the result by a scalar to generate a much closer neighbor. Compared with the A+ method, our method requires fewer features to generate a closer neighbor set.
2. Meanwhile, we have also designed a new projector which has much better numerical stability to adapt to our new

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

problem. As in A+, to obtain the low resolution reconstruction coefficient vector $\mathbf{x}$, we solve a regularized and over-completed least-squares problem detailed in Eq.(4). We present a numerically stable projector Eq. (6) to supplement our method.

3. In this case, by benefiting from closer neighbor we obtain a more accurate reconstruction coefficient vector $\mathbf{x}$ leading to an improvement circa $0.1 \sim 0.21$ dB over A+. Moreover, with fixed memory, more anchor points can be trained leading to much better generalization. Fig.1 shows improved quantitative performance, and Fig.2 gives an illustrative qualitative output.

The remainder of this paper is organized as follows. In Section 2, we present related work and discuss the relationship between our proposed model NRM and alternative methods. In Section 3, we present a formal definition of the model, including descriptions of neighbor reconstruction method and its optimization procedures. In Section 4, we investigate why NRM is useful to generate better neighbor set. This is followed by an experimental evaluation in Section 5 which explores the performance of NRM at single image super resolution task. Finally, conclusions are presented in Section 6.

2. Related work

Neighbor Embedding (NE) approaches assume that features which are drawn from small low- and high-resolution patches lie on two local geometrically similar manifolds (Wang et al., 2019; Bai et al., 2014, 2018). Based on this assumption NE approaches reconstruct high-resolution features with local geometric structure recording coefficients which are shared in low-resolution space (Liu and Bai, 2012; Cui et al., 2017, 2019). A representative NE approach is A+ method proposed by Timofte *et al* (Timofte et al., 2014). A+ has succeeded in reducing the time complexity and has achieved improved performance. However in training phase it can not handle large databases and its anchor points do not generalize well. Under the framework of A+, many notable methods like HHCR (Salvador et al., 2016), PSyCo (Prez-Pellitero et al., 2016) and RFL (Schulter et al., 2015) were proposed. HHCR and PSyCo utilizes symmetric prior over the manifolds to collapse the redundant variability of the neighbor of anchor points. When employed with random forests RFL directly maps from low to high-resolution patches to avoid tedious parameter tweaking. Although all of these methods give further improvements, they all suffer from limitations caused by framework A+.

Deep Learning has been applied to SR with remarkable success. A representative deep learning based SR method is SRCNN (Dong et al., 2016) which consists of three layers: a) patch representation. b) non-linear mapping. c) reconstruction with filters of spatial sizes $9\times9, 1\times1, 5\times5$ respectively. However, to achieve a result which surpasses A+'s, SRCNN needs to be fed with a large database, like ILSVRC2013 ImageNet which contains 395,909 images. Following SRCNN, more CNN methods were proposed: like CSCN (Wang et al., 2015), VDSR (Kim et al., 2016). They utilize more effective priors, such as the sparse prior and the deep learning structure prior, to surpass SRCNN.

3. Analysis of manifold-based single image SR

We analyse in more detail the A+ technique and explain the limitations of their method. All of our analysis is based on a basic property of the manifold: if an assigned neighbor is close enough then the local manifold subspace can be well described by the observed coordinates of the neighbor. Namely, if the neighbor of aimed anchor point is close enough, we can use our coordinated points to describe the inherent property of the manifold. The well-known Local Linear Embedding (LLE) (Roweis and Saul, 2000) was proposed based on this property and A+ method was, in turn, motivated by LLE. There are two major deficiencies of A+ method.

1. To harvest dense sample patches, the A+ method samples patches at different scales. If we generate dense patches with the A+ method on a large database, it is massively expensive in both computation and memory. For example, for a 91-image dataset, to obtain dense pathes around the anchored point, A+ method attempts to harvest 12 times at different scales resulting in about 5 millions patches.
2. A simple estimation shows that the patches harvested with the A+ method are not close enough. In practice the dimension of features drawn from the low dimensional patches is around 30. We aim to find a neighbor which lies within an anchor point centred hypersphere whose radius is 0.1. Without loss of generality, supposing that features are normalized and uniformly distributed, at least 10^{30} features are needed to reconstruct that required neighbor while only 5 million features are used in A+.

3.1. A manifold-based model

We analyse the generalisation capacity of manifold-based single image SR. Firstly, some notation is introduced. Suppose $\mathbf{p}_h$ are small sampled patches which are directly cropped from raw training images. $\mathbf{p}_l$ is downsampled patches from $\mathbf{p}_h$. And that $\mathbf{f}_l$ and $\mathbf{f}_h$ are normalized features extracted from $\mathbf{p}_l$ and $\mathbf{p}_h$ respectively by feature extractors,

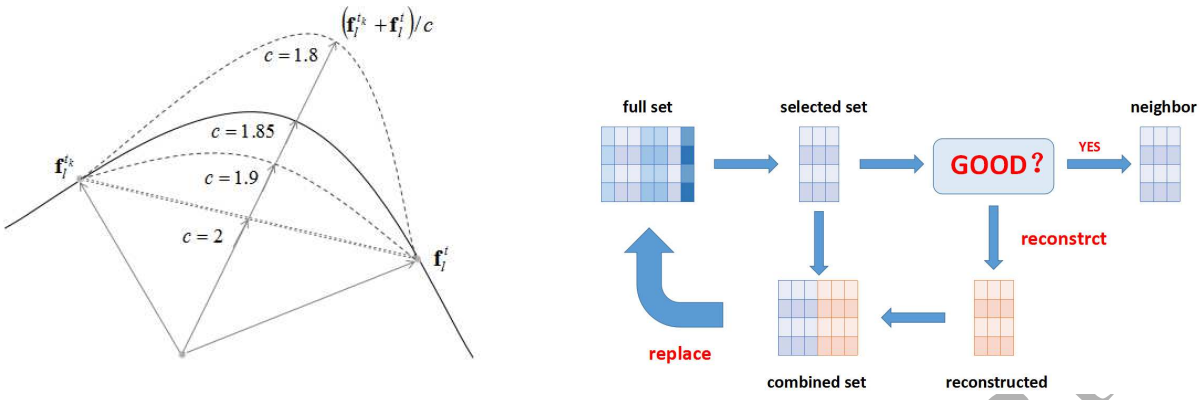
$$\mathbf{f_l} = K_l(\mathbf{p_l}),$$
$$\mathbf{f_h} = K_h(\mathbf{p_h}).$$

where K_l and K_h is linear feature extractors.

Further suppose that $\widehat{M}_l$ and $\widehat{M}_h$ are sampled manifolds corresponding to low-dimensional and high-dimensional feature spaces, namely,

$$\widehat{M}_l = \{\mathbf{f}_l^{(i)}\}_{i=1}^n,$$
$$\widehat{M}_h = \{\mathbf{f}_h^{(i)}\}_{i=1}^n,$$

where n is the number of extracted features in the low-dimensional or high-dimensional feature space. Suppose M_l and M_h are continuous ground truth manifolds corresponding to the LR and HR feature spaces. These two manifolds are



(a) (b)

Fig. 3: Illustration of sample reconstruction. (*a*) geometric interpretation of neighborhood reconstruction. The figure shows how to create a cosine similarity closer point $(\mathbf{f}_l^{t_k}+\mathbf{f}_l^t)/c$ by using $\mathbf{f}_l^t$ and its neighbor $\mathbf{f}_l^{t_k}$. c is an adjustable parameter to make $(\mathbf{f}_l^{t_k}+\mathbf{f}_l^t)/c$ be close to the intrinsic manifold, namely the solid line. In this figure, when $c=1.85$, $(\mathbf{f}_l^{t_k}+\mathbf{f}_l^t)/c$ can fall on the intrinsic manifold. (*b*) shows how to do neighbor reconstruction process iteratively.

structurally similar at local subspace. The relationship between the sampled manifolds and ground truth manifolds is:

$$M_l = \lim_{n\to\infty} \widehat{M}_l,$$
$$M_h = \lim_{n\to\infty} \widehat{M}_h,$$

There is an important one-to-one mapping, $H(\mathbf{p}_h)=\mathbf{f}_l(\in \widehat{M}_l)$, which is a naturally formed result when we are preparing the low and high patches. In practice we firstly train an LR dictionary $\mathbf{D}_l$,

$$\mathbf{D}_l, \alpha^i = arg\min_{\mathbf{D}_l,\alpha^i} \Sigma_i \|\mathbf{f}_l^{(i)} - \mathbf{D}_l\alpha^i\|_2^2 + \lambda^2\|\alpha^i\|_2^2. \quad (1)$$

Each column of D_l is called as an atom, $\mathbf{d}_l$. In A+ researchers use atoms as anchor points in $\widehat{M}_l$ to anchor offline projectors. Given a target low dimensional feature $\mathbf{f}_l^t$ researchers use a neighbor set of its nearest atom to reconstruct $\mathbf{f}_l^t$. This reconstruction leads to a reconstruction parameter $\mathbf{x}$. The reconstruction process can be formulated as,

$$\mathbf{x} = arg\min_{\mathbf{x}} \|\mathbf{f}_l^t - \mathbf{N}_l(\mathbf{d}_l)\mathbf{x}\|_2^2 + \lambda^2\|\mathbf{x}\|_2^2 \quad (2)$$

where $\mathbf{N}_l(\mathbf{d}_l)$ is a neighbor set of $\mathbf{d}_l$. The Eq.(2) can be solved with a closed-form,

$$\mathbf{x} = \mathbf{P}\mathbf{f}_l^t,$$

where $\mathbf{P} = (\mathbf{N}_l^T\mathbf{N}_l + \lambda^2\mathbf{I})^{-1}\mathbf{N}_l^T$. Obviously for each atom its corresponding $\mathbf{P}$ can be prepared offline. With parameter $\mathbf{x}$ and the one-to-one mapping $H(\mathbf{p}_h) = \mathbf{f}_l(\in \widehat{M}_l)$ high-dimensional patch $\mathbf{p}_l$ can be reconstructed in the way used in LLE (Roweis and Saul, 2000).

The SR problem in the NE framework is to construct a generalized function $G(\mathbf{f}_l) \approx \mathbf{p}_h : M_l \to P_h$ where P_h is continuous high-dimensional image patches manifold space. referring to the former one-to-one mapping H. During testing, a given evaluation criterion is used, such as PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index) and IFC (Information Fidelity Criterion), to estimate the performance of G. The estimator is,

$$C(I(G(\mathbf{f}_l^{(i)})) - I(\mathbf{p}_h^{(i)})),$$

where C is a chosen image evaluation criterion, I is a patch combining function which generates final patch-combining images. And $\mathbf{f}_l^{(i)} \in \widehat{M}_l, \mathbf{p}_h^{(i)} \in \widehat{P}_h$, $\widehat{P}_h$ are HR patch sets harvested from the training database.

The object fun of SR is,

$$\max_G \sum_i C(I(G(\mathbf{f}_l^{(i)})) - I(\mathbf{p}_h^{(i)})),$$

3.2. The neighbor reconstruction method

As in A+ when we are training the function G, given a target feature $\mathbf{f}_l^t$, we want to obtain a reconstruction coefficient vector $\mathbf{x}$. Then we directly transfer the coefficient vector into HR patch space, and construct the interest $\mathbf{p}_h^t$ with one-to-one mapping H. In the HR patch space we use the coefficient vector $\mathbf{x}$ and the corresponding neighbor to reconstruct target $\mathbf{p}_h^t$. So it is crucial to choose a good neighbor. Inspired by a Euclidean theorem in plane space, namely the parallelogram axiom of vectors, we have designed a neighbor reconstruction method denoted NRM, more detailed in Fig.3(a). Based on the cosine similarity metric we construct a closer, or more highly correlative, neighbor set for $\mathbf{f}_l^t$ which will be beneficial in generating a more accurate reconstruction coefficient $\mathbf{x}$.

Denote the neighbors $\mathbf{N}_l(\mathbf{d}_l)$ of $\mathbf{f}_l^t$ as the set of vectors $[\mathbf{f}_l^{t_1}, \mathbf{f}_l^{t_2}, ..., \mathbf{f}_l^{t_k}]$. We concatenate the central point and its corresponding neighbors together as column in the matrix $\bar{\mathbf{F}} = [\mathbf{f}_l^{t_1}, \mathbf{f}_l^{t_2}, ..., \mathbf{f}_l^{t_k}, \mathbf{f}_l^t]$. We induce a reconstruction operator,

$$\mathbf{R} = \begin{bmatrix} \frac{1}{c} & 0 & \dots & 0 & 0 \\ 0 & \frac{1}{c} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{c} & 0 \\ \frac{1}{c} & \frac{1}{c} & \frac{1}{c} & \frac{1}{c} & 1 \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)} \quad (3)$$

ACCEPTED MANUSCRIPT

where $c(>1)$ is an adjustable parameter. For the jth$(1 \le j < k+1)$ column R_j, it can generate the jth reconstructed neighbor $\frac{1}{c}\mathbf{f}_l^t + \frac{1}{c}\mathbf{f}_l^{t_j}$ by the right multiplication $\bar{\mathbf{F}}R_j$. For the $(k+1)$th column, it is used to preserve central point $\mathbf{f}_l^t$ for the next iteration. In NRM, reconstruction manipulation is achieved in parallel by right multiplying $\mathbf{R}$ by $\bar{\mathbf{F}}$. This manipulation can be done achieved iteratively. $\bar{\mathbf{F}}^{(r)} = \bar{\mathbf{F}}\mathbf{R}^r (r \in \{0, 1, 2, 3, ..., s\})$ where s is a truncation number. After operating on $\bar{\mathbf{F}}$ for s times, NRM collects $\pm\bar{\mathbf{F}}^{(r)}$ as a large set $\Im = \{\pm\bar{\mathbf{F}}^{(r)}\}_{r=0}^{s}$. The final step in NRM is to select k the nearest points for $\mathbf{f}_l^t$ from $\Im$ to replace the original neighbor set. Further details of the iterative approach are shown in Fig. 3(b). And the complete NRM algorithm is summarized in Alg. 1

Algorithm 1 NRM

Require:
Target low-dimensional feature $\mathbf{f}_l^t$;
Low-dimensional dictionary $\mathbf{D}_l$;
Truncation number s;
Adjustable parameter c.

Ensure:
Reconstructed neighbor set $\mathbf{N}_l^{new}$;

1: Find nearest atom $\mathbf{d}_l^t$ from $\mathbf{D}_l$ for $\mathbf{f}_l^t$;
2: Find k neighbors $[\mathbf{f}_l^{t_1}, \mathbf{f}_l^{t_2}, ..., \mathbf{f}_l^{t_k}]$ from $\widehat{M}_l$ for $\mathbf{d}_l^t$;
3: **for** r=1,2, ..., s **do**
4: Put low-dimensional feature and neighbors together $\bar{\mathbf{F}} = [\mathbf{f}_l^{t_1}, \mathbf{f}_l^{t_2}, ..., \mathbf{f}_l^{t_k}, \mathbf{f}_l^t]$;
5: Do the manipulation $\bar{\mathbf{F}}^{(r)} = \bar{\mathbf{F}}\mathbf{R}^r$;
6: Collect $\bar{\mathbf{F}}^{(r)}$ into $\Im = [\pm\bar{\mathbf{F}}^{(0)}, \pm\bar{\mathbf{F}}^{(1)}, ..., \pm\bar{\mathbf{F}}^{(r-1)}]$;
7: **end for**
8: In $\Im$ select another k nearest neighbors except $\pm\mathbf{f}_l^t$ to be $\mathbf{N}_l^{new}$;
9: **return** $\mathbf{N}_l^{new}$;

-1 before $\bar{\mathbf{F}}^{(r)}$ reverse the sign, if we want to employ the parallelogram axiom of vectors to efficiently generate a closer neighbor feature, we must ensure $\mathbf{f}_l^t$ and $\mathbf{f}_l^{t_j}$ lie on the same side of the anchor. Considering the existence of antipodal points we reverse the neighbor set by multiplying a negative one (-1) on its features, and utilize these reversed antipodal points to generate reconstructed points.

3.3. Solving the model

First, given a target feature $\mathbf{f}_l^t$, we employ NRM to generate a corresponding neighbor set $\mathbf{N}_l$. To obtain reconstruction coefficients $\mathbf{x}$ in a low resolution space, we need to solve the optimization problem,

$$\min_{\mathbf{x}} \|\mathbf{f}_l^t - \mathbf{N}_l\mathbf{x}\|_2^2 + \lambda^2\|\mathbf{x}\|_2^2. \tag{4}$$

For the problem, in A+, the solution is,

$$\mathbf{x} = \mathbf{P}\mathbf{f}_l^t,$$

where the projector $\mathbf{P} = (\mathbf{N}_l^T\mathbf{N}_l + \lambda^2\mathbf{I})^{-1}\mathbf{N}_l^T$.

In our method, we reconstruct a closer neighbor leading to a greater condition number of $\mathbf{N}_l$. If we still apply the projector

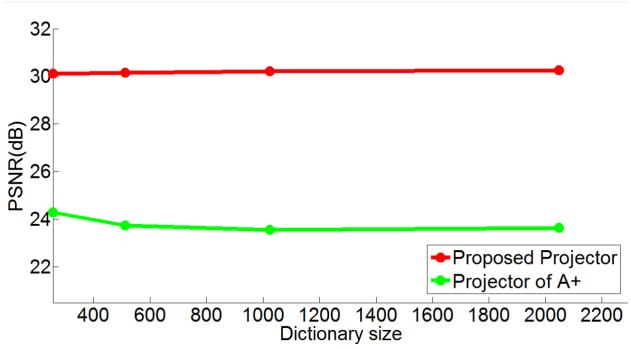


Fig. 4: PSNR results of proposed projector and original projector in A+. The red line shows PSNR performance of our method employing with proposed projector. The green one shows the performance of our method employing with original projector.

$\mathbf{P}$ which is deduced with normal equation method to obtain $\mathbf{x}$ in Eq.(4), this will lead to poor results. Because in normal equation method an inverse of matrix is needed to be computed, a large condition number will lead to a big numerical error which can be a deviation from our best results about 6dB as shown in Fig. 4.

To regular this great condition number problem we design a new projector based on matrix QR decomposition in which we do not have to compute a inverse of matrix. Rewriting Eq.(4) in the least-squares form:

$$min\| \begin{bmatrix} \lambda\mathbf{I} \\ \mathbf{N}_l \end{bmatrix}_{(m+n,n)} \mathbf{x} - \begin{bmatrix} \mathbf{O} \\ \mathbf{f}_l^t \end{bmatrix}_{(m+n,1)} \|_2^2, \tag{5}$$

where m is the dimension of the features in $\mathbf{N}_l$, n is the number of neighbor features, $(m \ll n)$. And $\mathbf{N}_l \in \mathbb{R}^{m\times n}$, $\lambda\mathbf{I} \in \mathbb{R}^{n\times n}$, $\mathbf{O} \in \mathbb{R}^{n\times 1}$, $\mathbf{f}_l^t \in \mathbb{R}^{m\times 1}$.

Applying the QR decomposition method to Eq.(5) gives:

$$\begin{bmatrix} \lambda\mathbf{I} \\ \mathbf{N}_l \end{bmatrix}_{(m+n,n)} = \mathbf{Q}\mathbf{R},$$

where $\mathbf{Q}$ is unitary, $\mathbf{R}$ is upper-triangular, $\mathbf{Q} \in \mathbb{R}^{(m+n)\times(n+m)}$, $\mathbf{R} \in \mathbb{R}^{(m+n)\times(n)}$.

Our problem now becomes:

$$(\mathbf{Q}\mathbf{R})\mathbf{x} = \begin{bmatrix} \mathbf{O} \\ \mathbf{f}_l^t \end{bmatrix}_{(m+n,1)},$$

$$\begin{aligned} \mathbf{R}\mathbf{x} &= \hat{\mathbf{Q}} \begin{bmatrix} \mathbf{O} \\ \mathbf{f}_l^t \end{bmatrix}_{(m+n,1)} \\ &= \begin{bmatrix} \hat{\mathbf{Q}}_n & \hat{\mathbf{Q}}_m \end{bmatrix} \begin{bmatrix} \mathbf{O} \\ \mathbf{f}_l^t \end{bmatrix}_{(m+n,1)} \\ &\Rightarrow \mathbf{R}\mathbf{x} = \hat{\mathbf{Q}}_m\mathbf{f}_l^t, \end{aligned}$$

$$\begin{aligned} \mathbf{y} &= \hat{\mathbf{Q}}_m\mathbf{f}_l^t, \\ \mathbf{R}\mathbf{x} &= \mathbf{y}, \end{aligned} \tag{6}$$

where, $\hat{\mathbf{Q}} = \mathbf{Q}^*$, and $\hat{\mathbf{Q}}_m$ is the last mth columns of $\hat{\mathbf{Q}}$, $\mathbf{Q}^*$ is conjugate transpose of $\mathbf{Q}$, and $\mathbf{R}\mathbf{x} = \mathbf{y}$ can be solved by substitution

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

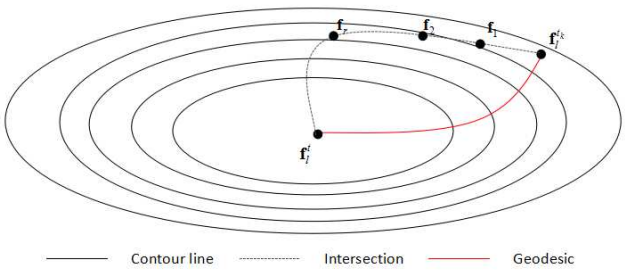


Fig. 5: The trace of reconstructed points. Reconstructed neighbor point $\mathbf{f}_i$ walk from $\mathbf{f}_l^{t_k}$ to $\mathbf{f}_l^t$ alone the dashed intersection line. With enough steps, $\mathbf{f}_i$ always fall on a contour line on which points have shorter geodesic to $\mathbf{f}_l^t$.

method. The performance comparison between normal equation method based and our method based projector is shown in Fig. 4.

4. Advantages

In this section we investigate why NRM is useful to generate better neighbor set. To quantify the meaning of better, we define a metric. Specially we define a distance function on intrinsic manifold M which we suppose it is a Riemann manifold:

$$L : \mathbb{R}^{d\times(k+1)} \rightarrow \mathbb{R}$$
$$L(\bar{\mathbf{T}}) = \sum_{i=1}^{k} \int_{t_j}^{t_{j_i}} |l_i'(t)|dt \qquad (7)$$

where t is a d dimensional intrinsic parameter point in M, t_j and t_{j_i} are the central point and the neighbor points respectively. Like $\bar{\mathbf{F}}, \bar{\mathbf{T}} = [\mathbf{t}_{j_1}, \mathbf{t}_{j_2}, ..., \mathbf{t}_{j_k}, \mathbf{t}_j]$. And $l_i(t_j, t_{j_i})$ is the geodesic between t_j and t_{j_i}.

Given a local coordinate system, Eq.(7) can be rewritten in a coordinate form:

$$L(\bar{\mathbf{T}}) = \sum_{i=1}^{k} \int_{t_j}^{t_{j_i}} \sum_{a,b=1}^{D} \sqrt{g_{a,b}(l_i(t))\frac{dx^a}{dt}\frac{dx^b}{dt}}dt \qquad (8)$$

where $g_{a,b} = g(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b})$, and g is a Riemann metric on M. According to the properties of Riemann metric which is an 2 order covariant tensor, Eq.(8) is invariant to the choice of coordinates. So the defined metric can describe the intrinsic relationships on M with different coordinates. We use $L(\bar{\mathbf{T}})$ to measure the distance between the central point and its corresponding neighbors. To its neighbors when the value of L is small, the central point is closer, namely the neighbors are closer.

Based on the defined metric, for comparison, we can firstly focus on one neighbor point. In Figure 5, a target low-dimensional feature $\mathbf{f}_l^t$ and a neighbor point $\mathbf{f}_l^{t_k}$ are given. The black curves are contours on which the points have the same geodesic length to $\mathbf{f}_l^t$. $H =< \mathbf{f}_l^t, \mathbf{f}_l^{t_k} >$ is a hyperplane spanned by $\mathbf{f}_l^t$ and $\mathbf{f}_l^{t_k}$. The dashed curve is a subset of the intersection between H and M. All of the reconstructed feature points $\mathbf{f}_i (i \in (1, 2, ..., r))$ are lie in the intersection subspace. The solid red curve is the geodesic between $\mathbf{f}_l^t$ and $\mathbf{f}_l^{t_k}$. The NRM reconstructed neighbor points $\mathbf{f}_i \in H$ will walk along the dashed curve from $\mathbf{f}_l^{t_k}$ to $\mathbf{f}_l^t$ until $\mathbf{f}_i$ fall on a contour on which points have smaller geodesic distance than that of $\mathbf{f}_l^{t_k}$. For two or more neighbors, on the meaning of former proposed metric we can always find closer neighbors by carefully fine-tuning parameters.

5. Experiments

We now comprehensively analyze the performance of our proposed NRM in relation to its design parameters and benchmark it in quantitative and qualitative comparison with A+ and other state-of-the-art methods.

We use the training set of images as proposed by Yang *et al* (Yang et al., 2010), Timofte *et al* (Timofte et al., 2014) and by Zeyde *et al* (Zeyde et al., 2010). However we use a different way to harvest patches from these images. Timofte *et al* (Timofte et al., 2014) repeatedly harvested dense patches by means of image pyramid. Because NRM can group a set of dense patches by reconstruction, which is shown in Fig. 6(c), we employ the Augmented Data set proposed by Timofte *et al* in (Timofte et al., 2016), which is a more general sparse data set, and harvest it once. To compare with A+ as fairly as possible, we also trained A+ on the Augmented Data set with the same harvest configuration. However, this configuration degraded A+s quality results. So in the following we use the original configurations of A+.

Note that Set5 and Set14 contain respectively 5 and 14 commonly used images for super-resolution evaluation. B100 aka Berkeley Segmentation Dataset is the B100 data set proposed by Timofte *et al* in (Timofte et al., 2014). We use the same LR path features as Zeyde *et al* (Zeyde et al., 2010) and Timofte *et al* (Timofte et al., 2014).

We compare with the following six methods which share the same training data set: standard bicubic upsampling method, the efficient sparse coding method of Zeyde *et al* (Zeyde et al., 2010), neighbor Embedding with Locally Linear Embedding (referred to as NE+LLE) (Chang et al., 2004), Adjusted Anchored Neighborhood Regression (referred to as A+) of Timofte *et al* (Timofte et al., 2014), Convolutional Neural Network Method (referred to as SRCNN) of Dong *et al* (Dong et al., 2016) and Fast and Accurate Image Upscaling with Super-Resolution Forest (referred to as RFL) of Schulter *et al* (Schulter et al., 2015).

5.1. Parameters

We analyze the main parameters of our proposed method, and at the same time compare with A+ which is the most related method. The standard settings we use are upscaling factor x4, 5000000 training samples of LR and HR patches which were sampled from the Augmented Data set, a dictionary size of 1024, a neighborhood size of 2048 and for NRM iteration time is 2. A+ is set up with the same parameters as reported in its original work. To verify our method is benefited by sample reconstruction rather than the choice of cosine metric like HHCR

ACCEPTED MANUSCRIPT

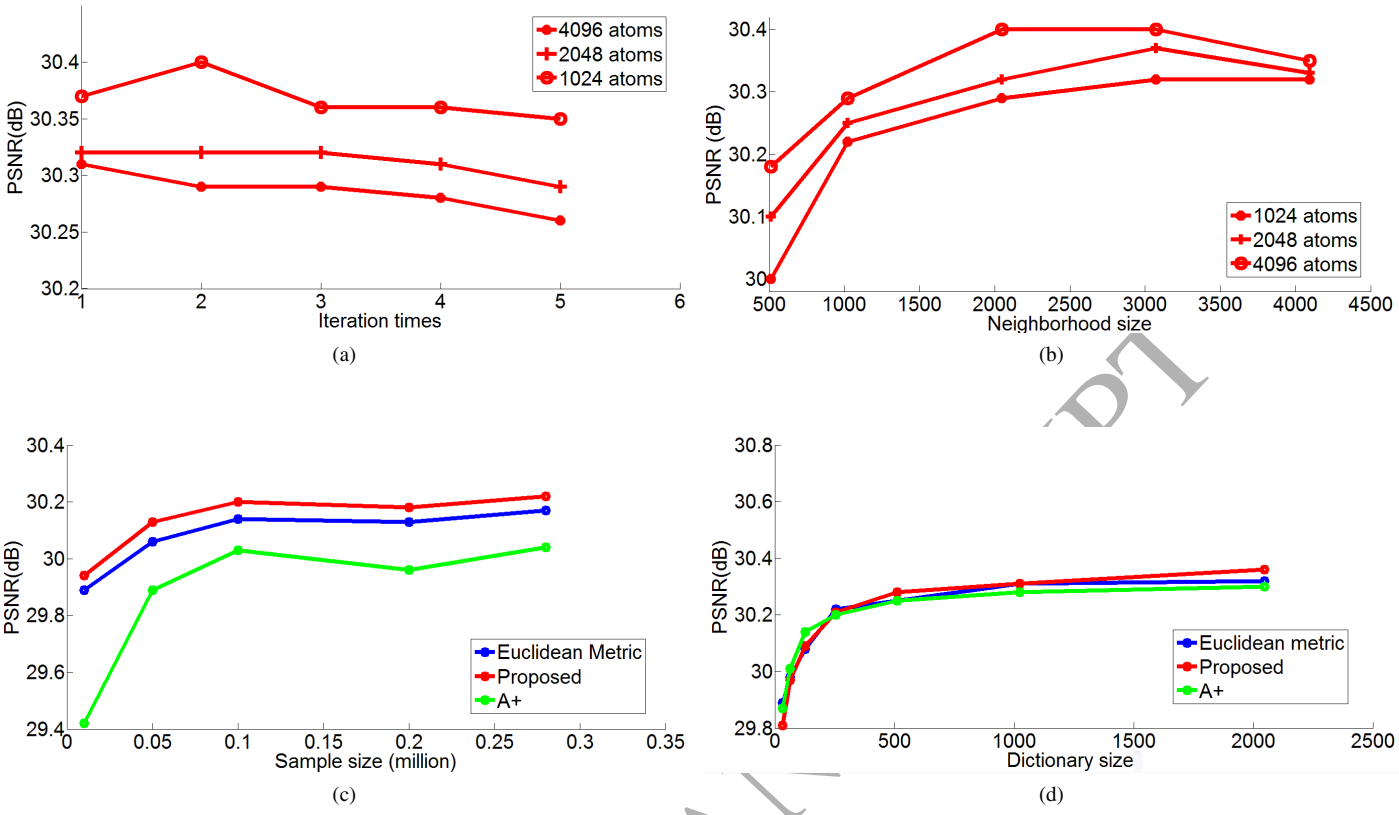
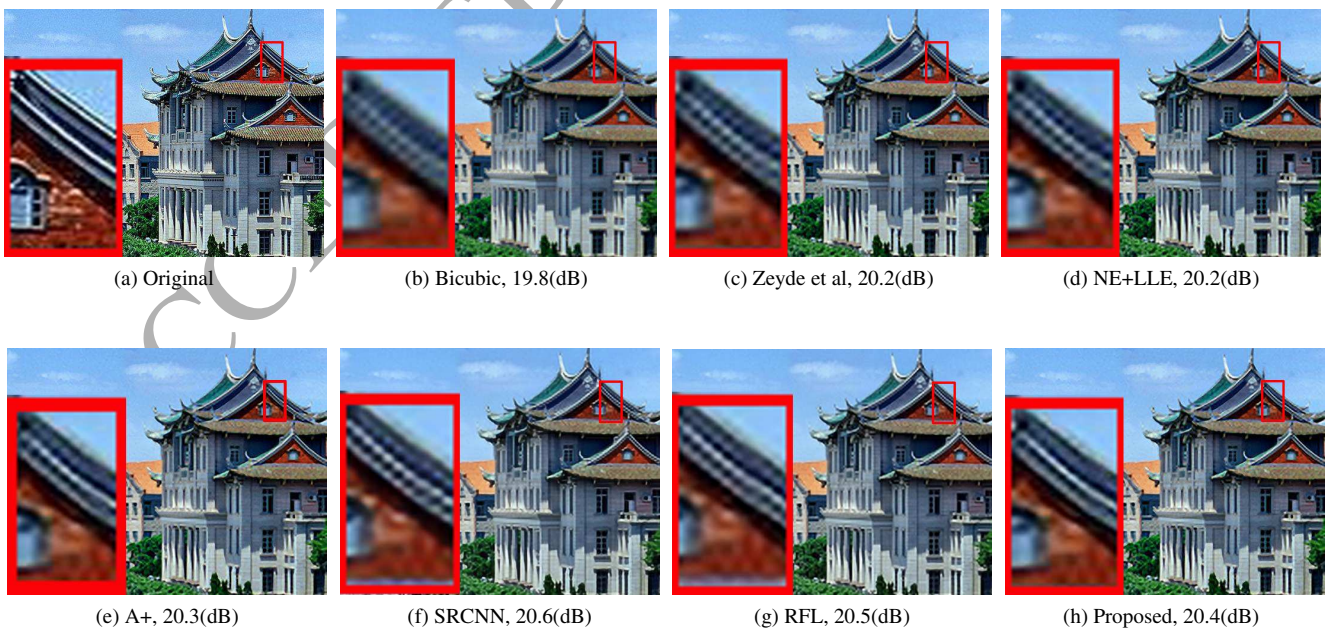


Fig. 6: Parameter influence and method comparison on average for Set5. (*a*) Iteration times for proposed method versus PSNR, with three kind of dictionary sizes; (*b*) Neighborhood size for proposed method versus for PSNR, with different dictionary sizes; (*c*) Number of training samples versus PSNR for A+, proposed and compared version of proposed method; (*d*) Dictionary size for A+, proposed method and its compared version.



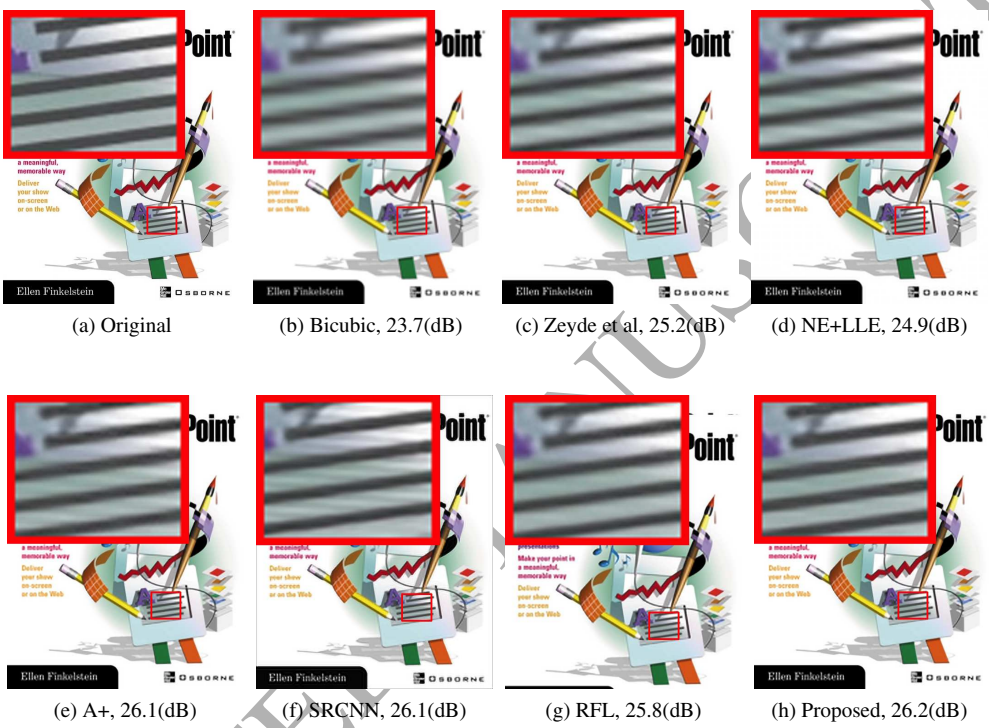
(a) Original (b) Bicubic, 19.8(dB) (c) Zeyde et al, 20.2(dB) (d) NE+LLE, 20.2(dB)

(e) A+, 20.3(dB) (f) SRCNN, 20.6(dB) (g) RFL, 20.5(dB) (h) Proposed, 20.4(dB)

Fig. 7: Illustrative output qualitative assessment for building image with magnification 3.

Table 1: Performance of x2, x3,and x4 magnification in terms of averaged PSNR (dB),SSIM and execution time (s) on data set Set5, Set14 and BSD100. Best results in red and runner-up in blue.

data set	s	Bicubic PSNR SSIM Time	Zeyde PSNR SSIM Time	NE+LLE PSNR SSIM Time	A+ PSNR SSIM Time	SRCNN PSNR SSIM Time	RFL PSNR SSIM Time	Proposed PSNR SSIM Time
Set5	2	33.66|0.9382|0.0	35.78|0.9563|2.5	35.77|0.9560|4.1	36.55|0.9611|0.8	36.34|0.9590|3.0	36.55|0.9585|1.1	36.65|0.9614|1.6
	3	30.39|0.8802|0.0	31.90|0.9075|1.1	31.84|0.9064|1.9	32.59|0.9139|0.5	32.39|0.9141|3.0	32.45|0.9162|1.0	32.67|0.9202|0.8
	4	28.42|0.8246|0.0	29.69|0.8565|0.7	29.61|0.8540|1.1	30.28|0.8737|0.3	30.09|0.8669|3.2	30.13|0.8680|0.8	30.40|0.8760|0.5
Set14	2	30.23|0.9415|0.0	31.81|0.9611|5.0	31.76|0.9620|8.5	32.28|0.9649|1.6	32.18|0.9637|4.9	32.27|0.9442|2.3	32.39|0.9649|3.6
	3	27.54|0.8587|0.0	28.67|0.8859|2.4	28.60|0.8868|3.9	29.13|0.8940|0.9	29.00|0.8910|5.0	29.03|0.8923|1.8	29.20|0.8946|1.7
	4	26.00|0.7838|0.0	26.88|0.8159|1.5	26.81|0.7322|2.4	27.32|0.8281|0.6	27.20|0.8210|5.2	27.21|0.8247|1.3	27.42|0.8300|1.1
B100	2	29.32|0.8338|0.0	30.40|0.8682|3.6	30.41|0.8708|6.1	30.77|0.8773|1.1	31.14|0.8847|3.4	31.13|0.8838|2.5	30.83|0.8772|2.3
	3	27.15|0.7364|0.0	27.87|0.7695|1.8	27.85|0.7713|2.9	28.18|0.7791|0.6	28.21|0.7800|3.4	28.21|0.7805|2.3	28.23|0.7820|1.1
	4	25.92|0.6673|0.0	26.51|0.6968|1.0	26.47|0.6974|1.5	26.77|0.7085|0.4	26.71|0.7022|3.5	26.74|0.7054|2.1	26.83|0.7105|0.7



(a) Original (b) Bicubic, 23.7(dB) (c) Zeyde et al, 25.2(dB) (d) NE+LLE, 24.9(dB)

(e) A+, 26.1(dB) (f) SRCNN, 26.1(dB) (g) RFL, 25.8(dB) (h) Proposed, 26.2(dB)

Fig. 8: Visual qualitative assessment for ppt3 image with magnification 3.

(Salvador et al., 2016), we also evaluate a compared NRM version based on Euclidean metric. Fig. 6 depicts the most relevant results of the parameter settings and comparisons.

In Fig. 6 (a) we compare the behavior of three kinds of dictionary size setup NRM, while varying the iteration time. Note that NRM's peak only at the first or second iteration and decrease with more iterations. It means more iterations lead to an overfitting on truth ground neighborhood. Fig. 6(b) shows the influence of neighborhood on NRM. The same as A+ (Timofte et al., 2014), NRM faces a plateau at 2048. As indicated in A+ this plateau limitation is caused by our training pool. In Fig. 6(c) we present the comparison between A+ and our NRM and NRMs comparison version while varying sample sizes from 0.01 million to 0.28 million. The quality is around 0.5 dB higher than A+ for a small sample size (e.g. 0.01 millions) and around 0.2 dB for 0.28 millions (which is the size set in A+). This is a notable fact that NRM can reconstruct useful neighborhoods and perform well even if the ground truth manifold is sampled sparsely. At the same time compared with the original NRM, the Euclidean version has slightly decreased (i.e 0.05 dB). This fact makes a difference between the influence of sample reconstruction and HHCRs. Fig. 6(d) shows the influence of dictionary size on the algorithms. Limited by the training pool, the algorithms still face a plateau that our NRM and its Euclidean version do not suffer from.

5.2. *Results*

In order to assess the quality of our proposed method, we tested on 3 datasets (Set5, Set14, B100) used by Timofte *et al* (Timofte et al., 2014) for 3 upscaling factors (x2, x3, x4) in the same CPU (Intel Core i7 4750HQ 2GHz) and memory (8Gb). Considering quality and time cost, we use dictionary with 4096 atoms and a neighborhood size of 2048. The method of Zeyde *et al*, NE+LLE, the similarity to Chang *et al* (Chang et al., 2004), and A+ is set up with its common parameters. SRCNN

and RFL are training on the same training data set proposed by Timofte *et al* leading to a decrease compared to their best performance reported in articles. We report quantitative PSNR and (structural similarity) SSIM results, as well as running times for our bank of methods. In Table 1 we summarize the quantitative results, while in Figs. 7, 2, 8 we provide visual assessments.

In Table 1 we show the averaged PSNR, SSIM and execution times of the benchmark. NRM almost obtains the best PSNR values, around 0.12dB higher across all scale and data set when compare to the most related algorithm A+. We also outperform some very recent methods (SRCNN and RFL) which are less competitive when trained on the same 91 images training data set. In the terms of computation time, our algorithm is very slightly slower than A+ but still faster than all other methods. In Figs. 7, 2, 8 as can be seen, the proposed NRM produces a visual quality comparable or superior to the other compared methods.

6. Conclusion

In this paper we present a new method for regression-based SR that is built on a novel neighbor reconstruction method (NRM). Via manipulations on anchored points and corresponding neighborhoods, NRM can reconstruct new points which are more closer to anchor point on the assumed manifold. Our contributions are: (1) a new sample reconstruction method with application to regression-based SR; (2) Supported by matrix QR decomposition, we design a more condition-number-stable regressor to compute effective result under closer neighborhood situation. Our results confirm the effectiveness of this approach using various accepted benchmarks, where we clearly outperform the current state-of-the-art. Finally, when the harvested samples are sparse on the manifold, NRM can still construct much closer points and perform well.

Acknowledgment

This work is supported by the Research Funds of State Grid Shaanxi Electric Power Company and State Grid Shaanxi Information and Telecommunication Company.

Appendix A. The Proof Under High Dimensional Condition

Theorem 1. *Given vectors* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$*, which satisfy* $||\mathbf{a}||_2 = ||\mathbf{b}||_2 = 1$ *and* $\langle \mathbf{a}, \mathbf{b} \rangle > 0$*, then we have* $\langle \mathbf{a}, \mathbf{b} \rangle \geq \langle \mathbf{a}, \mathbf{a} + \mathbf{b} \rangle$*, where* $\langle \cdot, \cdot \rangle$ *represents the angle between a pair of vectors.*

Proof. Employed with cosine similarity,

$$f(\mathbf{a}, \mathbf{b}) = |\frac{(\mathbf{a}, \mathbf{b})}{||\mathbf{a}||_2 ||\mathbf{b}||_2}| = |(\mathbf{a}, \mathbf{b})| \tag{A.1}$$

$$f(\mathbf{a}, \mathbf{a} + \mathbf{b}) = |\frac{(\mathbf{a}, \mathbf{a} + \mathbf{b})}{||\mathbf{a}||_2 ||\mathbf{a} + \mathbf{b}||_2}| = |\frac{(\mathbf{a}, \mathbf{a} + \mathbf{b})}{||\mathbf{a} + \mathbf{b}||_2}| \tag{A.2}$$

$$\geq |(\mathbf{a}, \mathbf{a} + \mathbf{b})|/2 \tag{A.3}$$

with $||\mathbf{a} + \mathbf{b}||_2 \leq ||\mathbf{a}||_2 + ||\mathbf{b}||_2 = 2$.

From Eq.(1)-(3), we have,

$$f(\mathbf{a}, \mathbf{a} + \mathbf{b}) - f(\mathbf{a}, \mathbf{b}) = |\frac{(\mathbf{a}, \mathbf{a} + \mathbf{b})}{||\mathbf{a} + \mathbf{b}||_2}| - |(\mathbf{a}, \mathbf{b})| \tag{A.4}$$

$$= (|(\mathbf{a}, \mathbf{a}) + (\mathbf{a}, \mathbf{b})| - 2|(\mathbf{a}, \mathbf{b})|)/2 \tag{A.5}$$

$$= (|1 + (\mathbf{a}, \mathbf{b})| - 2|(\mathbf{a}, \mathbf{b})|)/2 \tag{A.6}$$

because $\langle a, b \rangle > 0$, then

$$f(\mathbf{a}, \mathbf{a} + \mathbf{b}) - f(\mathbf{a}, \mathbf{b}) \tag{A.7}$$

$$= (1 - (\mathbf{a}, \mathbf{b}))/2 \geq 0 \tag{A.8}$$

$$\tag{A.9}$$

with $(\mathbf{a}, \mathbf{b}) = \cos\langle \mathbf{a}, \mathbf{b} \rangle \in (0, 1)$.
Finally we get,

$$f(\mathbf{a}, \mathbf{a} + \mathbf{b}) - f(\mathbf{a}, \mathbf{b}) \tag{A.10}$$

$$\Rightarrow |\cos\langle \mathbf{a}, \mathbf{a} + \mathbf{b} \rangle| \geq |\cos\langle \mathbf{a}, \mathbf{a} \rangle| \tag{A.11}$$

$$\Rightarrow \langle \mathbf{a}, \mathbf{a} + \mathbf{b} \rangle \leq \langle \mathbf{a}, \mathbf{b} \rangle \tag{A.12}$$

□

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

Bai, X., Yan, C., Yang, H., Bai, L., Zhou, J., Hancock, E.R., 2018. Adaptive hash retrieval with kernel based similarity. Pattern Recognition 75, 136–148.

Bai, X., Yang, H., Zhou, J., Ren, P., Cheng, J., 2014. Data-dependent hashing based on p-stable distribution. IEEE Transactions on Image Processing 23, 5033–5046.

Chang, H., Yeung, D.Y., Xiong, Y., 2004. Super-resolution through neighbor embedding, in: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition., IEEE. pp. I–I.

Cui, L., Bai, L., Zhang, Z., Wang, Y., Hancock, E.R., 2019. Identifying the most informative features using a structurally interacting elastic net. Neurocomputing 336, 13–26.

Cui, L., Jiao, Y., Bai, L., Rossi, L., Hancock, E.R., 2017. Adaptive feature selection based on the most informative graph-based features, in: International Workshop on Graph-Based Representations in Pattern Recognition, Springer. pp. 276–287.

Cui, Z., Chang, H., Shan, S., Zhong, B., Chen, X., 2014. Deep network cascade for image super-resolution. European Conference on Computer Vision , 49–64.

Dai, D., Timofte, R., Van Gool, L., 2015. Jointly optimized regressors for image super-resolution, Wiley Online Library. pp. 95–104.

Dai, S., Han, M., Xu, W., Wu, Y., Gong, Y., 2007. Soft edge smoothness prior for alpha channel super resolution, in: IEEE Conference on Computer Vision and Pattern Recognition, IEEE. pp. 1–8.

Dong, C., Loy, C.C., He, K., Tang, X., 2016. Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence 38, 295–307.

Dong, W., Zhang, L., Shi, G., Wu, X., 2011. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. IEEE Transactions on Image Processing 20, 1838–1857.

Duchon, C.E., 1979. Lanczos filtering in one and two dimensions. Journal of Applied Meteorology 18, 1016–1022.

Fattal, R., 2007. Image upsampling via imposed edge statistics. ACM Transactions on Graphics (TOG) 26.

Freeman, W., Jones, T., Pasztor., E., 2002. Example-based super resolution. IEEE Trans. Computer Graphics and Applications 22(2), 56–65.

Glasner, D., Bagon, S., Irani, M., 2009. Super-resolution from a single image, IEEE. pp. 349–356.

Irani, M., Peleg, S., 1991. Improving resolution by image registration. Graphical Models and Image Processing 53(3), 231–239.

Kim, J., Lee, J., Lee, K., 2016. Accurate image super-resolution using very deep convolutional networks. CVPR .

Kim, K.I., Kwon, Y., 2010. Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 1127–1133.

Li, X., Orchard, M.T., 2001. New edge-directed interpolation. IEEE transactions on image processing 10, 1521–1527.

Liu, S., Bai, X., 2012. Discriminative features for image classification and retrieval. Pattern Recognition Letters 33, 744–751.

Marquina, A., Osher, S.J., 2008. Image super-resolution by tv-regularization and bregman iteration. Journal of Scientific Computing 37, 367–382.

Protter, M., Elad, M., Takeda, H., Milanfar, P., 2009. Generalizing the nonlocal-means to super-resolution reconstruction. IEEE Transactions on image processing 18, 36–51.

Prez-Pellitero, E., Salvador, J., Torres, I., 2016. Psyco: Manifold span reduction for super resolution. CVPR .

Roweis, S., Saul, L., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500), 2323–2326.

Salvador, J., Ruiz-Hidalgo, J., Rosenhahn, B., et al., 2016. Half hypersphere confinement for piecewise linear regression. IEEE Winter Conference on Applications of Computer Vision (WACV) , 1–9.

Schulter, S., Leistner, C., Bischof, H., 2015. Fast and accurate image upscaling with super-resolution forests. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition , 3791–3799.

Timofte, R., De Smet, V., Van Gool, L., 2013. Anchored neighborhood regression for fast example-based super-resolution. Proceedings of the IEEE International Conference on Computer Vision , 1920–1927.

Timofte, R., De Smet, V., Van Gool, L., 2014. A+: Adjusted anchored neighborhood regression for fast super-resolution. Asian Conference on Computer Vision , 111–126.

Timofte, R., Rothe, R., Gool, L.V., 2016. Seven ways to improve example-based single image super resolution. CVPR .

Wang, C., Bai, X., Wang, S., Zhou, J., Ren, P., 2019. Multiscale visual attention networks for object detection in vhr remote sensing images. IEEE Geoscience and Remote Sensing Letters 16, 310–314.

Wang, Z., Liu, D., Yang, J., Han, W., Huang, T., 2015. Deep networks for image super-resolution with sparse prior. ICCV .

Yang, C.Y., Yang, M.H., 2013. Fast direct super-resolution by simple functions. ICCV , 561–568.

Yang, J., Wright, J., Huang, T.S., Ma, Y., 2010. Image super-resolution via sparse representation. IEEE transactions on image processing 19, 2861–2873.

Zeyde, R., Elad, M., Protter, M., 2010. On single image scale-up using sparse-representations. International conference on curves and surfaces , 711–730.

Zhang, K., Gao, X., Tao, D., Li, X., 2012. Single image super-resolution with non-local means and steering kernel regression. IEEE Transactions on Image Processing 21, 4544–4556.

Zhang, K., Tao, D., Gao, X., Li, X., Xiong, Z., 2015. Learning multiple linear mappings for efficient single image super-resolution. IEEE Transactions on Image Processing 24, 846–861.

ACCEPTED MANUSCRIPT