Single Image Super Resolution via Manifold Approximation

Chinh T. Dang, Student Member, IEEE, and Hayder Radha, Fellow, IEEE

Abstract—Image super-resolution remains an important research topic to overcome the limitations of physical acquisition systems, and to support the development of high resolution displays. Previous example-based super-resolution approaches mainly focus on analyzing the co-occurrence properties of low-resolution and high-resolution patches. Recently, we proposed a novel single image super-resolution approach based on linear manifold approximation of the high-resolution image-patch space [1]. The image super-resolution problem is then formulated as an optimization problem of searching for the best matched high resolution patch in the manifold for a given low-resolution patch. We developed a novel technique based on the \( \ell_1 \) norm sparse graph to learn a set of low dimensional affine spaces or tangent subspaces of the high-resolution patch manifold. The optimization problem is then solved based on the learned set of tangent subspaces. In this paper, we build on our recent work as follows. First, we consider and analyze each tangent subspace as one point in a Grassmann manifold, which helps to compute geodesic pairwise distances among these tangent subspaces. Second, we develop a min-max algorithm to select an optimal subset of tangent subspaces. This optimal subset reduces the computational cost while still preserving the quality of the reconstructed high-resolution image. Third, and to further achieve lower computational complexity, we perform hierarchical clustering on the optimal subset based on Grassmann manifold distances. Finally, we analytically prove the validity of the proposed Grassmann-distance based clustering. A comparison of the obtained results with other state-of-the-art methods clearly indicates the viability of the new proposed framework.

Index Terms—Image super-resolution, tangent space estimation, \( \ell_1 \) norm sparse graph, Grassmann manifold.

I. INTRODUCTION

IMAGE super-resolution (SR) reconstruction is a process of generating a higher spatial resolution image, along with increasing the high frequency components, from one or several low resolution (LR) images of the same scene. The problem remains an important research topic to overcome the limitations of physical acquisition systems, such as consumer videos captured by LR camcorders (e.g. cellphone or surveillance cameras). Generally speaking, SR techniques can be broadly grouped into three categories:

- Interpolation-based methods [21], [22]: (e.g. bilinear or bicubic convolution interpolation) have advantages in simplicity and fast implementation. However, they suffer from inherent defects, including over smooth, block effects, blurred details and ringing artifacts around edges.
- Reconstruction-based methods [23], [24]: require multiple LR images of the same scene, in which each LR image imposes a set of linear constraints on the unknown high resolution (HR) intensity values [25]. These methods have limitations of small magnification factors (smaller than 2) [26], and the quality of reconstruction is up to the number of input LR images. In addition, multiple LR images are not always available in general.
- Example based SR or image hallucination [3], [5], [27], [42] methods: recover missing HR details from a training dataset by learning the co-occurrence or non-linear mapping functions between LR and HR patch spaces. These methods achieve a high quality reconstructed image with large magnification factors (up to 4 and higher).

Our approach belongs to the example based category. However, our proposed framework represents a significant departure from prior example-based approaches due to the following: Prior example-based SR methods focus on learning the mapping functions between HR and LR patch spaces [3]-[5], [19], [27]-[28], [30], [32]. Typically, these methods demand a huge set of training example patch pairs. In addition, learning joint dictionaries of two spaces endures a fundamental problem of no guarantee of low coherence in two dictionaries simultaneously [31], which is a condition for a viable sparse representation. Instead of learning the mapping functions, our proposed approach focuses on exploring the real manifold structure of HR patches, without relying on creating a set of LR images or patches.

Moreover, this paper differs from our recent works [1][41] in many ways. To elaborate further about these differences, we first highlight the salient aspects of our recent work. In [1], we introduced an off-line manifold learning approach based on sparse subspace clustering. We employed a linear tangent-space approximation for the HR manifold. In [41], we proposed a self-learning based framework that requires only the input image and its different down-sampled scales, without relying on training dataset. In particular, our initial method (the SLAM algorithm) [1] includes two main phases:

1) **Training phase**: the step is performed off-line for one time. We assume that the HR patch space has a low dimensional manifold structure. Given HR training images, a set of patches are extracted from the training data. These patches are now considered as samples taken from the underlying low dimensional manifold. We exploit the self-expressiveness property [38] to create a sparse graph, and iteratively perform...
unsupervised clustering for a set of tangent spaces from the graph.

ii) HR image reconstruction (testing phase): For a given LR patch, the HR image reconstruction is formulated as an optimization problem that looks for the closest HR patch in the manifold. The problem is solved by looking for the nearest tangent space among those learned tangent spaces mentioned above, and then projecting the LR patch onto that nearest space.

The reconstruction part involves searching for the exact nearest tangent space, which is very time consuming. In particular, the number of learned tangent spaces directly impacts the SLAM algorithm [1]. For example, increasing the number of approximated tangent spaces generally leads to a higher quality image reconstruction due to a better manifold approximation, however at the cost of higher amount of computation. Our previous work [1] did not take into account the number of training patches, the number of learned tangent spaces that well approximate the HR patch manifold, as well as analyze the structure of the learned tangent spaces for a better nearest tangent space search. Therefore, it decreases the overall algorithm performance in terms of HR image quality and the time computation.

**Contribution of the paper:** In this paper, we build on our prior work [1], by addressing the following questions:

*Question 1*: what is the sufficient number of tangent spaces that need to approximate the HR patch manifold and how to find the subset of tangent spaces? Naturally, one should target a smaller number of tangent spaces while still preserving the HR image quality.

We address this question by considering each tangent space as one point in a Grassmann manifold that allows measuring geodesic distances among every pair of the tangent spaces. Based on the measured distances, we employed the min-max algorithm [16][36] to select an optimal set of tangent spaces that still preserve the approximated manifold structure.

*Question 2*: how to analyze the structure of the selected tangent spaces such that for a given LR patch from the testing phase, the algorithm could quickly detect the approximated nearest tangent spaces?

We address this question based on a technique from the problem area of nearest-neighbor search. In particular, we construct the hierarchical k-means tree, a typical method in solving approximated nearest neighbor problem. The set of tangent spaces is recursively divided into clusters based on geodesic Grassmann manifold distances. We note that, in the traditional nearest neighbor problem, a set of points and a query point belong to the same (distance) space. In our problem, the set of tangent spaces and a given LR patch are not in the same (distance) space. One is the Euclidean distance from the LR patch to a tangent space; the other is the geodesic distance in the Grassmann manifold among the query set of tangent spaces. In this paper, we prove that for a set of clusters with high inter-cluster geodesic dissimilarities and low within-cluster geodesic dissimilarities, the proposed algorithm could detect the “nearest” tangent space to a given LR patch, where “nearest” here is in terms of Euclidean distances. Fig.1 shows the general proposed SR framework using manifold approximation.

The remainder of the paper is organized as follows. Section II describes our proposed image SR framework based on manifold approximation including three main sub-sections. First, we briefly review the SR problem from HR patch manifold approximation perspective [1]. Second, we introduce these two mentioned questions and solutions. Finally, we present the HR image reconstruction (testing phase) based on the set of estimated tangent subspaces. Section III shows some experimental results and visual comparisons with related methods, and Section IV outlines some concluding remarks.

II. IMAGE SUPER-RESOLUTION VIA HIGH-RESOLUTION PATCH MANIFOLD APPROXIMATION

In this section, we first briefly review the SR problem using a novel manifold perspective that we introduced in [1].

A. Problem formulation

The problem of SR can be cast as an ill-posed inverse problem. Given a LR patch $l_p$, an automated algorithm needs to recover its corresponding HR patch $h_p$ that satisfies the underdetermined linear equation:

$$l_p = DBh_p + \varepsilon \quad (1)$$

Here, $D$ and $B$ are the decimation and blur (low pass filter) operators, respectively. $\varepsilon$ denotes an additive noise, normally i.i.d. white Gaussian noise. To eliminate the complexities of dealing with different spatial resolution (number of pixels) between $l_p$ and $h_p$, image/patch $l_p$ is scaled-up to the target HR by a simple interpolation operator, e.g. bicubic interpolation [31].

$$\bar{l}_p = Ql_p = Q(DBh_p + \varepsilon) = (QDB)h_p + \bar{\varepsilon} \quad (2)$$

Many example-based approaches focus on learning the co-occurrence model, specially coupled dictionaries for both LR and HR features (e.g. $A_l$ and $A_h$ respectively) [1][3][37]. However, as pointed out in [31], even if the dictionary $A_h$ of the HR space has low coherence, the corresponding LR dictionary $A_l = (QDB)A_h$ could not guarantee of having low coherence due to the undefined and unknown multiplication $(QDB)$.

Our approach considers the space of HR patches as a low dimensional manifold in a high dimensional ambient space. Given a LR patch $l_p$, its HR patch $h_p$ will be determined as the closest HR patch in the HR space. Denote the space of HR patches for a given patch size as $M \subset \mathbb{R}^D$, then, the process of finding $h_p$ can be defined as follows:

$$h_p = \arg\min_{v \in M} d(l_p, v) \quad (3)$$

Here, $d(l_p, v)$ denotes a distance measure. In general, $M$ is infinite; hence, we need a practical approach for searching that space for the best result using moderate computation. We note that in our experiment, $\bar{l}_p$ will be used instead of $l_p$. In general, $M$ is infinite; hence, we need a practical approach for searching that space for the best result using moderate computation. Our proposed solution to (3) includes two main steps. The underlying is to create a set of tangent spaces that
are approximation of $M$ from which we can infer the optimum solution for (3).

**B. Training phase**

1) The tangent space estimation: In our problem, we want to estimate the set of tangent spaces for $M$ via the set of training points $H = \{h_j \in \mathbb{R}^D\}_{j=1}^N \subseteq M$. The key question that impacts the tangent space estimation is how to partition $H$ into unknown $k$ subsets, denoted by $\{H_j\}_{j=1}^k$, such that each subset contains sample points from the same neighborhood, and then estimate the tangent space from each subset. Due to the number of samples in a neighborhood varies according to the local characteristics of the manifold. For example, a circle in $\mathbb{R}^2$ need two neighbor points to approximate a tangent space, while a sphere in $\mathbb{R}^3$ need three neighbor points for a tangent space. Hence, we define the upper bound constrain on the number of samples per neighborhood. The tangent space estimation problem can be summarized as follow:

**Problem 1:** How to partition the set of training data points $H \in M$ into unknown $k$ subsets of samples, $\{H_j\}_{j=1}^k$, and the corresponding $k$ neighborhood regions approximated by tangent spaces of $M$, denoted by $\{T(H_j)\}_{j=1}^k$, such that $\dim(H_j) \leq r_0$ for $1 \leq j \leq k$.

Here, $r_0$ is the upper bound for the dimension of these approximated tangent spaces. The underlying factor that impacts the partition result, and then tangent space estimation, is how to define neighbors for each datum. The intuitive reasonable approach is using pairwise Euclidean distance, and one point will be connected with other points via nearest neighbor method or $\varepsilon$-ball based method. The former connects one point with exactly a predetermined nearest points, while in the latter, the samples within its surrounding $\varepsilon$-ball are connected. This approach endures a problem of using a predetermined number of neighbors or fixed radius for each ball. It does not fit well with our problem, since the number of necessary samples in each neighborhood varies according to the local characteristics of the manifold, along with the set of training sample points are subject to change depending on the input LR image. In addition, as pointed out by [3], a sparse representation over a learned dictionary leads to a better prediction of a HR patch, while in most cases; elements in the sparse representation do not belong to its Euclidean neighborhood set. This implies the method of defining a neighborhood using Euclidean distance does not perform as well as sparse representation based neighborhood.

Here, we develop a novel algorithm based on creating $\ell_1$-graph for the set of HR patch samples, $H = \{h_j \in \mathbb{R}^D\}_{j=1}^N \subseteq M$. In the set $H$ (normally $N \gg D$), each patch sample is shifted to zero mean, and then normalized. The underlying idea of the $\ell_1$ norm sparse graph is the self-expressiveness property [38] that considers data matrix $H$ itself as a dictionary for sparse representation. The set of data points can be written in a matrix form $H = [h_1, h_2, \ldots, h_N] \in \mathbb{R}^{D \times N}$. Let $H_i \in \mathbb{R}^{D \times (N-1)} = H/\{h_i\}$ be the matrix obtained from $H$ by removing its $i^{th}$ column. In our case, the self-expressive property implies looking for the sparsest representation of $h_i$ from its corresponding dictionary $H_i$:

$$\min \|c_i\|_0 \text{ subject to } h_i = H_i c_i$$

Here, $\|\cdot\|_0$ is the $\ell_0$ norm of a vector that counts the number of nonzero elements. Although the problem is NP hard, recent results [9] has concluded that if the solution is sparse enough, then it could be found via solving the $\ell_1$ norm minimization.

$$\min \|c_i\|_1 \text{ subject to } h_i = H_i c_i$$

The problem can be solved in a polynomial time via linear programming method. A sparse graph is generated from these coefficients $c_i$ and an algorithm from spectral graph theory, e.g. normalized cut algorithm [10], can be exploited for data segmentation. Details of implementation are shown in Algorithm 1 (step 1). The normalized cut algorithm iteratively segments input dataset into two subsets of data points and check for the upper bound rank condition. If the dimension of
a subset does not satisfy the upper bound rank condition, it is iteratively bi-partitioned into smaller subsets. For a given set of sufficient HR patch samples \( H \), an affine subspace created by a group of neighbored data points approximates the non-linear space \( M \) around these points. Therefore, the union of these subspaces created by all these sufficient samples \( H \) approximates \( M \). The quality of approximation is controlled by the upper bound rank \( r_0 \). In an extreme case, for example \( r_0 = \text{rank}(H) \), and hence, \( M \approx \text{Aff}(H) \) represents a coarse approximation.

2) Select an optimal subset of tangent subspaces: The iterative process solving Problem 1 leads to a set of \( k \) tangent spaces. Since the HR image reconstruction step (the testing phase, which we will discuss in detail latter) involves searching for the nearest tangent space from the \( k \) tangent-subspaces set, the value of \( k \) becomes an important parameter in term of balancing image reconstruction quality and the computational time. In particular, a larger value of \( k \) implies a finer approximation of \( M \), which normally leads to higher quality of image reconstruction with the tradeoff of higher computational complexity. In this part, we investigate the impact of the number of tangent subspaces on the quality of HR image reconstruction. More importantly, we propose a method to select an optimal subset of tangent spaces from the input original set of \( k \) tangent subspaces, which is developed based on geodesic Grassmann manifold distance. The problem of selecting an optimal subset of tangent subspaces can be stated as follows:

**Problem 2**: Given a set of \( k \) tangent spaces and an integer \( k_0 \leq k \), how to select a subset of \( k_0 \) key elements from the original set such that the overall quality of the HR image reconstruction using only these \( k_0 \) tangent spaces is as good as using the original set of \( k \) tangent spaces?

We note that our final objective for creating a set of tangent spaces is to approximate the HR patch manifold \( M \), and then to reconstruct the HR image. Therefore, for a predetermined number \( k_0 \), we prefer a subset of tangent spaces with the following properties:

1) The subset of selected tangent spaces needs to approximate the whole manifold \( M \), not only focus on a particular region of \( M \) (locally) [41].
2) Overall redundancies (dissimilarity) between any pair of tangent spaces should be minimized.

Choosing two close tangent spaces does not significantly improve the quality of HR image reconstruction when compared to choosing one of these two tangent spaces (since the projected points from a given LR patch onto them does not differ by much); meanwhile, choosing two tangent spaces in comparison to selecting one of them naturally increases the computational cost. Consequently, we need to measure the distance between two tangent spaces. Below, we present the concept of geodesic Grassmann manifold distance.

**Grassmann manifold**: given \( n, p \ (p \leq n) \) are positive integers, denote \( \text{Grass}(p, n) \) and \( \mathbb{R}_+^{n \times p} \) are the set of all \( p \)-dimensional subspaces of \( \mathbb{R}^n \), and the set of all \( n \times p \) matrices whose columns are linear independent, respectively. \( \mathbb{R}_+^{n \times p} \) is an open subset of \( \mathbb{R}^{n \times p} \). The subset admits a structure of an open sub-manifold of \( \mathbb{R}^{n \times p} \) where its differential structure is created using the chart \( \Phi : \mathbb{R}_+^{n \times p} \to \mathbb{R}^{np} : X \to \text{vec}(X) \). Therefore, this manifold is referred to as non-compact Stiefel manifold of full rank \( n \times p \) matrices. The manifold \( \mathbb{R}_+^{n \times p} \) is equipped with an equivalence relation \( \sim \) that is defined as follows:

\[
X \sim Y \text{ if and only if } \text{span}(X) = \text{span}(Y)
\] (6)

Here, \( X, Y \in \mathbb{R}_+^{n \times p} \) and \( \text{span}(X) \) denotes the subspaces spanned by the columns of matrix \( X \). The quotient manifold defined on the non-compact Stiefel manifold \( \mathbb{R}_+^{n \times p} \) with the above equivalence relation \( [X] := \{ Y \in \mathbb{R}_+^{n \times p} : Y \sim X \} \) is the equivalence class that contains element \( X \), and the set \( \mathbb{R}_+^{n \times p} / \sim := \{ [X] : X \in \mathbb{R}_+^{n \times p} \} \) is a quotient space that has one-to-one correspondence to \( \text{Grass}(p, n) \), where each point in \( \text{Grass}(p, n) \) is one \( p \)-dimensional subspace. The distance between two subspaces is now mapped to geodesic distance between two points in the manifold, which is mainly computed using the concept of principal angles.

Denote \( H_1 \) and \( H_2 \) be two subspaces (assuming that \( \text{dim}(H_1) = d_1 \geq \text{dim}(H_2) = d_2 \)), the principal angles between two subspaces, \( 0 \leq \theta_1 \leq \ldots \leq \theta_t \leq \pi/2 \), are defined recursively for \( t = 1, \ldots, d_2 \) as follows [12-13]:

\[
\cos \theta_k = \max_{\text{subject to } u_j^T v_j = 0} \left( \frac{d_1 \theta_k}{\sqrt{d_1^2 - \theta_k^2}} \right) = d_1 \theta_k^2
\] (7)

These vectors \( (u_1, \ldots, u_{d_1}) \) and \( (v_1, \ldots, v_{d_2}) \) are called principal vectors of these two subspaces \( H_1 \) and \( H_2 \). The principal angle \( \theta_k \) is the angle between two principal vectors \( u_k \) and \( v_k \). There are several methods of computing the principal angles and principal vectors; one efficient stable method has been developed using singular value decomposition on the product of two basis matrices \( H_1^T H_2 \) (the subspace \( H_1 \) and its basis matrix are used interchangeably in this context). In particular,

\[
H_1^T H_2 = U S V^T
\] (8)

where \( U = [u_1, \ldots, u_{d_2}] \), \( V = [v_1, \ldots, v_{d_2}] \) are matrices of these principal vectors and \( S = \text{diag}(\cos \theta_1, \ldots, \cos \theta_{d_2}) \).

There are several methods of computing Grassmann manifold distance based on these obtained principal angles, i.e. projection distance, Binet-Cauchy distance, etc. Some additional properties and applications of these distances could be found at [13]. In this paper, we exploit the geodesic Grassmann manifold distance (arc length) in the form:

\[
G(H_1, H_2) = \sqrt{\sum_{j=1}^{d_2} \theta_j^2}
\] (9)

This distance has been exploited successfully in previous work on image search problem to manipulate leaf nodes in the data partition tree [14]. It has some desired properties of a metric, such as symmetric, triangular properties. In addition, the geodesic distance is derived from the intrinsic geometry of Grassmann manifold [15], which is the length of the geodesic curve connecting two subspaces (two points) on the Grassmann manifold. Next, we discuss the technique exploiting the geodesic Grassmann manifold distance and the
min-max algorithm to select a subset of tangent spaces. 

The min-max algorithm: in the training step, a set of abundant HR patches has been exploited to create a set of tangent spaces. The obtained set is not optimal in the following senses. There are tangent spaces which are very close with each other, or too many tangent spaces distributed around a local region of the HR patch manifold $M$. Here, we make use of the geodesic Grassmann manifold distance to select an optimal subset of tangent spaces. The affinity matrix that contains pairwise distance of every pair of tangent spaces is computed using the geodesic distance. The min-max algorithm, which has been used successfully in the problem of key frame extraction from consumer videos [16], [36], is exploited in our work. We note that the min-max algorithm has been used here for a set of tangent spaces, not for a set of frames in a video. Details and an example of the min-max algorithm could be found in [5]. Algorithm 2 outlines the min-max algorithm. In this case, the benefit of the min-max algorithm is twofold. First, it is optimum in each step. The algorithm would select the next tangent space as the best space given the previous selected tangent spaces. Second, it is consistent and synergetic with our problem in the following sense. It preserves previously selected tangent spaces when the number of desired important tangent spaces increases. For example, initially we might prefer to select four important tangent spaces from the input; then, we might wish to increase the number of selected tangent spaces to five (or any number larger than the original four). In this case, the four previously selected tangent spaces will be a subset of the set of five tangent subspaces selected afterward.

3) The approximated nearest tangent space detection: Example-based SR methods achieve state-of-the-art results when compared to other approaches. However, these methods endure high computational cost due to processing hundreds of thousands of patches, while each one involves an $\ell_1$-norm minimization problem [3],[30]. Under the proposed framework, the demanding computational part is to look for the closest tangent space for a given input LR patch (we would discuss in more details about this step in the testing phase). The problem can be formally stated as follows:

**Problem 3:** For a given set of tangent spaces, \(\{T(H_j)\}_{j=1}^k \subset \mathbb{R}^D\), and an input vector of LR patch \(l_p \in \mathbb{R}^D\), how to find the tangent space \(T(H_{j_0}) \in \{T(H_j)\}_{j=1}^k\) such that the distance \(d_E(l_p, T(H_{j_0}))\) is minimum. Here, \(d_E(l_p, T(H_{j_0}))\) is the $\ell_2$ norm orthogonal projection of vector \(l_p\) onto the subspace \(T(H_j)\), which has the closed form:

\[
d_E(l_p, T(H_{j_0})) = \left\|l_p - T(H_j)\left(T(H_j)^T T(H_j)\right)^{-1} T(H_j)^T l_p\right\|_2
d\]

The nearest tangent space problem is related to the nearest neighbor search that can be stated as follows: given a set of points in a space $S$, and a query point $q \in S$, find the point in $S$ that is closest to $q$. There are no known algorithms that are faster than linear search for solving the nearest-neighbor problem with exact solution. On the other hand, approximated algorithms provide a significant speedup with a small loss of accuracy. In this paper, we adopt the hierarchical $k$-means tree method in solving the nearest tangent space problem. In particular, our proposed algorithm iteratively clusters the set of tangent spaces based on the geodesic Grassmann manifold distance, and defines the centers of the obtained clusters at each level. The point with minimum sum of distances to all other points in the same cluster will be assigned as the center of that cluster. Detail steps are shown in Algorithm 3.
In the nearest neighbor search problem, a query point \( q \) belongs to the same neighborhood as the given set of points. The dissimilarity measure between \( q \) to the set of points is also identical to the dissimilarity measure among the given set of points. In our case, there are two different dissimilarity measures: one is the Euclidean distance from the given input vector (black point in Fig.3) of LR patch to a tangent space (black arrows); and the other is the geodesic distance being used for dissimilarity measure among set of tangent spaces (red line in Fig.2). For an obtained hierarchical set of clusters, and a given LR patch \( l_p \), the algorithm measures the Euclidean distance from \( l_p \) to every cluster center (at the first level), and then select the closest group. An iterative process is performed for the next level in that selected group until reaching the final level. Since our framework exploits two different dissimilarity measures, the following result guarantees that for a set of clusters with high inter-cluster geodesic dissimilarities and low within-cluster geodesic dissimilarities, the proposed algorithm does detect the nearest tangent space to a given LR patch in terms of a Euclidean distance.

| \( \ell _2 \) norm orthogonal projection distance | Arc Grassmann manifold distance |
|-------------------------------------------------|-------------------------------|

Fig. 2: Illustration of searching for the nearest tangent spaces. For a given vector input \( l_p \), the distances between \( l_p \) and two cluster centers \((h_1 \text{ and } h_2)\), are measured, and the group with a smaller distance will be further considered. In the figure, since \( d(l_p, h_1) > d(l_p, h_5) \), the algorithm measures distance between \( l_p \) and every other elements \((h_6, h_7, h_8, h_9)\) and select the smallest one.

**Theorem 1.** Let \( U_1 \) and \( U_2 \) be two clusters of tangent spaces with centers \( q_1 \) and \( q_2 \), respectively, and let \( v \in \mathbb{R}^D (||v||_2 = 1) \). If \( d_E(v, q_2) - d_E(v, q_1) \geq \sqrt{2} (\sin (r_1) + \sin (r_2)) \), then we have:

\[
\min_{u \in U_2} d_E(v, u_2) \geq \max_{u_1 \in U_1} d_E(v, u_1) \tag{11}
\]

Here, \( r_1, r_2 \) is a radius of a clusters of tangent spaces \( U_1, U_2 \), which is defined as the maximum geodesic distance from the center of the cluster to all other tangent spaces. In particular, \( r_1 = \max_{u_1 \in U_1} d_G(u_1, q_1) \) and \( r_2 = \max_{u_2 \in U_2} d_G(u_2, q_2) \). The proof of Theorem 1 is based on the following results that show the relation between geodesic Grassmann manifold distance and Euclidean distance:

**Lemma 1.** Given two matrices or subspaces \( A \in \mathbb{R}^{D \times n}, B \in \mathbb{R}^{D \times m} \), and a vector \( v \in \mathbb{R}^D (||v||_2 = 1) \). Prove that:

\[
\sin^{-1} \left[ \frac{d_G(v, A)}{\sqrt{2}} \right] + G(A, B) \geq \sin^{-1} \left[ \frac{d_G(v, B)}{\sqrt{2}} \right] \tag{12}
\]

The detail proofs are shown in APPENDIX. The condition \( d_E(v, q_2) - d_E(v, q_1) \geq \sqrt{2} (\sin (r_1) + \sin (r_2)) \) implies that the set of clusters has low within-cluster geodesic dissimilarities, \( \sin (r_1) + \sin (r_2) \) is small, while the Euclidean distances from input vector to two cluster centers are sufficient different. The result indicates that if the condition is satisfied, then the distance from every other tangent space in the non-selected cluster \((U_2)\) to the input vector is greater than the distance from every other tangent space in the selected cluster \((U_1)\) to the input vector. Hence, the algorithm could skip all elements from \( U_2 \) in looking for the nearest tangent space.

**Algorithm 3.** Hierarchical clustering structure for set of tangent spaces

**Inputs:** Set of tangent spaces \( \{T(H_j)\}_{j=1}^{k} \), number of regions each level \( n_0 \), desired number of hierarchical level.

**Outputs:** Hierarchical data structure for the set of tangent spaces, set of center clusters at each level.

1. **Create** a dissimilarity graph from set of tangent spaces
   - Graph construction \( G \) in which each tangent space corresponds to one vertex.
   - Measure dissimilarity between every pair of tangent spaces using geodesic Grassmann manifold distance. The distance becomes the weight between two vertices.
2. **Perform** normalized cut for \( G \) to partition the set of tangent spaces into \( n_0 \) non-overlapped regions.
3. **Detect** the center of each region as the one with minimum sum of distances to all other points in the same region.
4. **Repeat** (until reach desired hierarchical level)
   - If (number of elements in a subgroup is smaller than \( n_0 \)), cluster the subgroup into next hierarchical level

**C. Training phase**

For a given set of tangent spaces \( \{T(H_j)\}_{j=1}^{k} \), the original problem (4) for HR patch reconstruction is now approximated as:

\[
h_p \equiv \arg \min_{v \in \{T(H_j)\}_{j=1}^{k}} d(l_p, v) \tag{13}
\]

The reconstruction step considers \( \ell_2 \) distance from LR patch to the set of tangent spaces, and selects the smallest one.

\[
j_0 = \arg \min_{1 \leq j \leq k} d(l_p, T(H_j)) \tag{14}
\]

The co-occurrence HR patch is approximated by a projected vector into that closest tangent space:

\[
h_p \equiv \Pi_{T(H_{j_0})}(l_p) \tag{15}
\]

Here, \( \Pi_{T(H_{j_0})}(l_p) \) is the vector projection of \( l_p \) onto space \( T(H_{j_0}) \).

Some points about the computation of the orthogonal projection of one vector in a high-dimension space into a lower-dimension vector subspace need to be considered. Let \( A \in \mathbb{R}^{D \times n} \) be a full rank matrix with fewer number of columns than rows \((n < D)\), and let \( b \in \mathbb{R}^D \) be a vector that we want to find its orthogonal projection onto the subspace spanned by columns of \( A \). The solution is given by:

\[
\tilde{b} = A(A^T A)^{-1} A^T b \tag{16}
\]
The obtained solution minimizes the distance between input vector \( b \) and every other vectors belonging to the space spanned by columns of \( A \). Using (16) directly could lead to numerical instability, as it is analogous to using normal equations for computing the solution to a linear least squares regression [35]. The main reason arises in the inversion of \( A^T A \), since the direct use of normal equations squares the condition number \( \kappa(A) \) due to the production \( A^T A \). For numerical stability, then the condition number must satisfy: \( \kappa(A) = \sqrt{\kappa(A^T A)} < \sqrt{1/\varepsilon} \approx 1E8 \). Hence, in our work, we discard every subspace with a large condition number value. In addition, in the testing phase, each time the algorithm projects one LR patch \( f_p \) onto the space \( T(H_{j_o}) \), it computes the term \( T(H_{j_o}) (T(H_{j_o})^T T(H_{j_o}))^{-1} T(H_{j_o})^T \). Therefore, this term will be computed during the training phase, and save the value for the testing phase to avoid the repeated computation for each input LR patch.

In the final step, overlapping pixels from corresponding set of HR patches will be averaged and the last condition (3) will be treated using back projection as a simple denoising method [3][4].

III. EXPERIMENTAL RESULTS

In this section, we validate the effectiveness of our proposed algorithm for single image SR using a set of standard images. The gallery of tested images including seven images is shown on Fig. 4. We compared quantitatively our method with Bicubic (as a reference method), Feedback Control based Super Resolution (FCSR) [20], Neighbor Embedding based Super Resolution (NESR) [19], and ScSR [3], using both peak-signal-to-noise ratio (PSNR), and the structure similarity index (SSIM) [18].

A. Experimental settings

Training dataset: In this work, we conduct our experiments on a set of 24 HR training images. The set of images could be found at the author webpage, which is a subset of training dataset from prior approaches [1][30][37]. The proposed approach does not require creating a set of corresponding LR images, only focusing on HR patch space. Since the human visual system is more sensitive to luminance channel than other chrominance channels, all input images are converted into YCbCr color space, and the proposed super resolution works with luminance channel only. The two other chrominance channels are simply interpolated using bicubic interpolation. All computations of PSNR, SSIM values in our paper are for the luminance channel. An input image is interpolated into image of the same size as desired. We refer to the interpolated image as LR image, and patch extracted from that image is LR patch.

In the experiment, we work with \( p = 8 \), then extracted patches belong to the Euclidean space \( \mathbb{R}^{64} \). The dimension of the ambient Euclidean space is smaller than dimension of feature vectors being used in prior approaches [37]. The upper bound rank has been chosen to be \( r_0 = 3 \), and the underlying dimension \( d \) is then select to be equal to the upper bound rank. Hence, the tangent space simply becomes the affine space spanned by samples in a neighborhood. It is reasonable since all input training HR patches are assumed to be noiseless, so they should belong to the estimated tangent space. To create a set of patches for training, the algorithm scans all patches from the training dataset with overlap two pixels. Each patch is then shifted into zero mean, and normalized. For each image, only 100 patches having maximum variances are extracted for further processing. Since these extracted patches are redundant within and across different images, we discard identical elements due to they could lead to the potential instability of sparse decomposition in creating the \( \ell_1 \) norm sparse graph, which could finally leads to reconstruction artifacts [33]. Algorithm 1 is performed to create a set of tangent spaces. The total number of obtained tangent spaces (after step 4 in Algorithm 1) is 908.

To select an optimal number of tangent spaces, we experimented with different numbers of tangent spaces. In particular, we run our simulation using the following set of numbers for the tangent spaces: \{100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 908\}. Next, we evaluate the obtained results based on PSNR and SSIM values and then plot the averaged values as a function of the number of tangent spaces as shown in Fig. 4. Based on these observations, we concluded that the averaged PSNR increases sharply from 100 to 300. However, the PSNR does not improve much when increasing the number of tangent spaces beyond 300. Therefore, we select 300 for the number of tangent subspaces for the rest of our experiment.

B. Experimental Results

1) Using exact nearest tangent space search: In this experiment, we exploit the full search of the nearest tangent spaces, which leads to a better quality of HR image reconstruction when compared to the approximated nearest tangent space search that we would show its result later. Naturally, the full search comes with a higher computational cost. To evaluate objectively the quality of the proposed super-resolution using manifold approximation, Table I shows the quantitative comparison (PSNR and SSIM values) with some other methods including bicubic, NESR [19], FCSR [20], and ScSR [3]. The
Table I. Comparisons (PSNR and SSIM) among current methods.

| Image (x3) | Bicubic | NESR [19] | FCSR [20] | ScSR [3] | Proposed Method |
|------------|---------|-----------|-----------|---------|-----------------|
| Lena       | 30.099  | 29.856    | 30.885    | 31.052  | 31.300          |
|            | 0.843   | 0.822     | 0.851     | 0.866   | 0.871           |
| Flower     | 37.198  | 36.08     | 37.473    | 37.953  | 38.309          |
|            | 0.918   | 0.899     | 0.913     | 0.924   | 0.931           |
| Mountain   | 27.052  | 26.342    | 27.4490   | 27.751  | 27.901          |
|            | 0.7080  | 0.6524    | 0.7206    | 0.756   | 0.762           |
| House      | 24.417  | 24.093    | 24.777    | 24.852  | 24.910          |
|            | 0.6944  | 0.663     | 0.7103    | 0.728   | 0.728           |
| BabyFace   | 32.709  | 31.87     | 32.8415   | 33.14   | 33.473          |
|            | 0.799   | 0.768     | 0.7935    | 0.8164  | 0.826           |
| Lion       | 28.392  | 27.568    | 28.478    | 28.860  | 29.056          |
|            | 0.725   | 0.671     | 0.7167    | 0.765   | 0.776           |
| Car        | 27.423  | 26.181    | 28.432    | 28.183  | 28.503          |
|            | 0.841   | 0.8307    | 0.8684    | 0.8655  | 0.8703          |
| Avg.       | 29.613  | 28.856    | 30.0479   | 30.270  | **30.490**      |
|            | 0.7898  | 0.7542    | 0.7962    | 0.8173  | 0.8219          |

image. Although ScSR recovers details information, it contains more ringing and jaggy artifacts than our method as indicated in Table I.

2) (Optional) approximated nearest tangent space detection: We apply the approximated nearest tangent space detection technique for speeding up the proposed algorithm. Under the exact nearest tangent space detection, the only way for a given LR patch is to compute every distance to a single tangent space, which is time consuming. Here, the set of tangent spaces are partitioned into a set of several groups based on the pairwise Grassmann manifold distance among them. The distance from one point to every other point in the same group will be computed, and the one with the smallest distance will be selected as the center of that group. Since the set of tangent spaces contains only 300 elements, our algorithm partitions it table shows that the obtained results are superior to others in term of PSNR values. We also report the SSIM values, which are based on human visual system, to fully evaluate the effectiveness of the proposed method.

Fig. 5 shows the visual comparison of these methods for several images: BabyFace, Flower, and Lena. The bicubic interpolation method is shown as a reference method in which the results are normally blurred lacking high frequency details. On the other hand, NESR [19] adds some high frequency details along with artifacts. We keep the same training data as appeared in [19] (using 2 images for training purpose), so the obtained results in some cases have lower PSNR than the bicubic interpolation. However, NESR produces a sharper image than the bicubic interpolation method. The results by FCSR [20] are better than those of NESR and bicubic interpolation, however not as good as ScSR [3] in terms of quantitative comparison. In some cases (e.g. BabyFace image), result from FCSR lacks more details compared with results from NESR. Our proposed method and sparse representation based method (ScSR) [3] keep the naturalness of the output
TABLE II. AVERAGED PSNR, SSIM, AND TIME COMPUTATION

| Image  | ScSR [3] | SLAM | Exact Nearest Tangent Space | Approximated Nearest Tangent Space |
|--------|----------|------|----------------------------|-----------------------------------|
| PSNR   | 30.270   | 30.4537 | 30.490                     | 30.4180                           |
| SSIM   | 0.8173   | 0.8211  | 0.8219                     | 0.8197                            |
| Time (s)| 219.8121 | 153.3825 | 107.4557                  | 20.3804                           |

IV. CONCLUSION

In this paper, we proposed a novel single image super resolution approach based on manifold approximation of HR patch space. The method does not require learning the relationship between LR and HR spaces using a huge amount of training data, but focus on learning properties of HR patch manifold only via a small number of training HR patch samples. The corresponding HR patch is then located via an orthogonal projection of a given LR patch into its nearest tangent space. In addition, we consider a novel technique to select a very small number of tangent spaces using the Grassmann manifold distance, and approximate the nearest tangent space detection to reduce computational cost.

APPENDIX

Proof of Lemma 1:

Denote \( v_A = \Pi_A(v) \) be the orthogonal projection of vector \( v \) onto \( \text{span}(A) \). We consider an vector \( u_m \in \text{span}(A) \) such that \( u_m = \arg \max_{u_2=1} v^T u \). Denote \( v_p = \Pi_{u_m}(v) \), we have:

\[
v^T u_m = \langle v, u_m \rangle = \langle v_p + (v - v_p), u_m \rangle \\
= v^T u_m = \|v_p\|^2
\]

Since \( v^T u_m \) is maximum, \( \|v_p\|^2 \) is maximum. On the other hand, \( \|v\|^2 = \|v_p\|^2 + \|v - v_p\|^2 = 1 \), so \( \|v - v_p\|^2 \) is minimum. Based on the definition of \( v_A \), we know that \( \|v - v_p\|^2 \geq \|v - v_A\|^2 \Rightarrow v_p = v_A \) is the orthogonal projection of vector \( v \) and the subspace \( \text{span}(A), \text{span}(B) \). We have: \( d_E(v, A) = \sin^2 \theta_A \) and \( d_E(v, B) = \sin^2 \theta_B \). Using the definition of principal angle between vector \( v \) and \( \text{span}(A) \), we got:

\[
\cos \theta_A := v^T u_m = \|v_A\|^2 = \cos^2 \theta_A \\
= d_E(v, A)^2 = 1 - \cos^2 \theta_A = 2\sin^2 \left( \frac{\theta_A}{2} \right) \\
= d_E(v, A) = \sqrt{2} \sin \left( \frac{\theta_A}{2} \right) \\
= G(v, A) = 2\sin^{-1} \left( \frac{d_E(v, A)}{\sqrt{2}} \right)
\]

In a similar produce, we got \( G(v, B) = 2\sin^{-1} \left( \frac{d_E(v, B)}{\sqrt{2}} \right) \). Using the triangular inequality property of geodesic Grassmann manifold distance [13], we got the final result.

Proof of Theorem 1:

Based on the proof of Lemma 1, we got the maximum value of \( \sin^{-1} \left( \frac{d_E(v, q_2)}{\sqrt{2}} \right) \) is \( \frac{\pi}{4} \) since \( d_E(v, q_2) \leq 1 \). If \( G(q_2, u_2) \geq \frac{\pi}{4} \) then we have:

\[
\frac{d_E(v, q_2)}{\sqrt{2}} \leq \frac{d_E(v, q_2)}{\sqrt{2}} + \sin G(q_2, u_2) \leq \frac{d_E(v, q_2)}{\sqrt{2}} + \sin(\pi/2)
\]

Here, we note that \( G(q_2, u_2) \leq \frac{\pi}{7} \) by the definition. On the other hand, if \( G(q_2, u_2) \leq \frac{\pi}{4} \) and we have (from Lemma 1):

\[
0 \leq \sin^{-1} \left( \frac{d_E(v, q_2)}{\sqrt{2}} \right) \leq \sin^{-1} \left( \frac{d_E(v, q_2)}{\sqrt{2}} \right) + G(q_2, u_2) \leq \frac{\pi}{4}
\]
\[
\sin\left(\sin^{-1}\left(\frac{d_E(v, u_2)}{\sqrt{2}}\right)\right) = \frac{d_E(v, u_2)}{\sqrt{2}} \\
\leq \sin\left(\sin^{-1}\left(\frac{d_E(v, u_2)}{\sqrt{2}}\right) + G(q_2, u_2)\right) \\
\leq \sin\left(\sin^{-1}\left(\frac{d_E(v, u_2)}{\sqrt{2}}\right) + \sin(G(q_2, u_2))\right) \\
\leq \frac{d_E(v, u_2)}{\sqrt{2}} + \sin(r_2) (\ast)
\]

On a similar approach, we obtain:
\[
\frac{d_E(v, u_1)}{\sqrt{2}} \leq \frac{d_E(v, u_1)}{\sqrt{2}} + \sin(r_1) (\ast\ast)
\]

From (\ast), (\ast\ast) and the given condition we got the proof of the theorem.

REFERENCES

[1] C. T. Dang, M. Aghagolzadeh, A. A. Moghadam, and H. Radha. “Single image super-resolution via manifold linear approximation using sparse subspace clustering,” In IEEE Global Conf. Signal Information Process., pp. 949-952, 2013.

[2] He, Li, Hairong Qi, and Russell Zaretski. “Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 345-352, 2013.

[3] Yang, Jianchao, John Wright, Thomas S. Huang, and Yi Ma. “Image super-resolution via sparse representation,” IEEE Trans. Image Process., Vol.19, no. 11 (2010): 2861-2873.

[4] Chen, Xiaoxuan, and Chun Qi. “Low-Rank Neighbor Embedding for Single Image Super-Resolution,” Signal Process. Letters, IEEE Vol.21, no.1, Jan. 2014.

[5] Yang, Jianchao, Zhe Lin, and Scott Cohen. “Fast image super-resolution based on in-place example regression,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1059-1066, 2013.

[6] Guillenmin, Victor, and Alan Pollack. “Differential topology,” American Mathematical Soc., Vol. 370, 2010.

[7] Tyagi, Hemant, Elf Vural, and Pascal Frossard. “Tangent space estimation for smooth embeddings of Riemannian manifolds,” Information and Inference, Vol.2, no.1, pp.69-114, 2013.

[8] Deshpande, Amit, Madhur Tulsiani, and Nisheeth K. Vishnoi. “Algorithms and hardness for subspace approximation,” In Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 482-496. SIAM, 2011.

[9] D. Donoho. “For most large underdetermined systems of linear equations the minimal $\ell_1$-norm solution is also the sparsest solution,” In Communications on pure and applied mathematics, 2006.

[10] Shi, Jianbo, and Jitendra Malik. “Normalized cuts and image segmentation,” In IEEE Trans. Pattern Anal. Mach. Intell., no. 8 (2000): 888-905.

[11] Absil, P. A., Mahony, R., & Sepulchre, R. (2009). “Optimization algorithms on matrix manifolds,” Princeton University Press.

[12] Björck, A., & Golub, G. H. (1973). “Numerical methods for computing angles between linear subspaces,” Mathematics of Computation, 27(123), 579-594.

[13] J. Hamm, “Subspace-based learning with Grassmann kernels,” Ph.D. dissertation, 2008.

[14] Wang, X., Li, Z., & Tao, D. (2011). “Subspaces indexing model on Grassmann manifold for image search,” IEEE Trans. Image Process., 20(9), 2627-2635.

[15] Wang, T., & Shi P. (2009). “Kernel Grassmannian distances and discriminant analysis for face recognition from image sets,” Pattern Recognition Letters, 30(13), 1161-1165.

[16] Dang, C. T., M. Kumar, and H. Radha. “Key frame extraction from consumer videos using epitome,” In Proc. IEEE Int. Conf. Image Processing, pp. 93-96, 2012.

[17] Jafarpour, Sina, Volkan Cevher, and Robert E. Schapire. “A game theoretic approach to expandable-based compressive sensing,” In IEEE International Symposium on Information Theory Proceedings, pp. 464-468, 2011.

[18] Wang Zhou, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. Image Process., no. 4 (2004): 600-612.

[19] Chang, Hong, Dit-Yao Yeung, and Yimim Xiong. “Super-resolution through neighbor embedding,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 1, 2004.

[20] Shan, Q., Zhaorong Li, Jiaya Jia, and Chi-Keung Tang. “Fast image/video upsampling,” ACM Trans. Graphics, vol. 27, no. 5, p. 153. ACM, 2008.

[21] Carew, W. Knox, Daniel B. Chung, and Sheila S. Hemami. “Regularity-preserving image interpolation,” IEEE Trans. Image Process., no. 9 pp.1293-1297, 1999.

[22] Li, Xin, and Michael T. Orchard. “New edge-directed interpolation,” IEEE Trans. Image Process., no. 10, pp.1521-1527, 2001.

[23] Farsiu, Sina, M. Dirk Robinson, Michael Elad, and Peyman Milanfar. “Fast and robust multiframe super-resolution,” IEEE Trans. Image process., no. 10, pp.1327-1344, 2004.

[24] Su, Heng, Liang Tang, Ying Wu, Daniel Tretter, and Jie Zhou. “Spatially adaptive block-based super-resolution,” IEEE Trans. Image Process., no. 3, pp.1031-1045, 2012.

[25] Glesner, Daniel, Shai Bagon, and Michal Irani. “Super-resolution from a single image,” In Proc. IEEE Int. Conf. Comput. Vis., pp. 349-356, 2009.

[26] Lin, Zhouchen, and Heung-Young Shum. “Fundamental limits of reconstruction-based superresolution algorithms under local translation,” IEEE Trans. Pattern Anal. Mach. Intell., no. 1, pp.83-97, 2004.

[27] Freeman, William T., Thouis R. Jones, and Egon C. Pasztor. “Example-based super-resolution,” IEEE Comput. Graphics and Applications, no. 2, pp.56-65, 2002.

[28] Li, Bo, Hong Chang, Shiguang Shan, and Xilin Chen. “Aligning coupled manifolds for face hallucination,” IEEE Signal Process. Letters, no. 11 (2009): 957-960.

[29] Roweis, Sam T., and Lawrence K. Saul. “Nonlinear dimensionality reduction by locally linear embedding,” Science 290.5500 (2000): 2323-2326.

[30] Yang, Jianchao, Zhaowen Wang, Zhe Lin, Scott Cohen, and Thomas Huang. “Coupled dictionary training for image super-resolution,” IEEE Trans. Image Process., no. 8 (2012): 3467-3478.

[31] Zeyde, Roman, Michael Elad, and Matan Proctor. “On single image scale-up using sparse-representations,” Proc. 7th Int. Conf. Curves Surf., pp. 711-730, 2010.

[32] Yu, Guoshen, Guillermo Sapiro, and Stéphane Mallat. “Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity,” IEEE Trans. Image Process., no. 5 (2012): 2481-2499.

[33] Mairal, Julien, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. “Non-local sparse models for image restoration,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 2272-2279, 2009.

[34] Zontak, Maria, and Michal Irani. “Internal statistics of a single natural image,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 977-984, 2011.

[35] Coakley, E. S., Vladimir Rokhlin, and Mark Tygert. “A fast randomized algorithm for orthogonal projection,” SIAM Journal on Scientific Computing, 33, no. 2 (2011): 849-868.

[36] Dang, C.; Radha, H.; “Heterogeneity image patch index and its application to consumer video summarization,” IEEE Trans. Image Process., vol.23, no.6, pp.2704-2718, Jun. 2014.

[37] Gao, Xinbo, Kaibing Zhang, Dacheng Tao, and Xuelong Li. “Joint learning for single-image super-resolution via a coupled constraint,” IEEE Trans. Image Process., vol.21, no. 2 (2012): 469-480.

[38] Elhamifar, Ehsan, and Ren Vidal. “Sparse subspace clustering,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 2790-2797, 2009.

[39] Yang, Jianchao, Zhaowen Wang, Zhe Lin, Xianbiao Shu, and Thomas Huang. “Non-local sparse models for image restoration,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2012.

[40] Wang, Shenlong, D. Zhang, Yan Liang, and Quan Pan. “Semi-coupled dictionary learning with applications to image super-resolution and photosketch synthesis,” In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 2216-2223, 2012.

[41] Chinh Dang, M. Aghagolzadeh, and Hayden Radha. “Image super-resolution via local self-learning manifold approximation,” IEEE Signal Process. Letters, Vol. 21, No. 10, October 2014.

[42] Peleg, T., and M. Elad. “A statistical prediction model based on sparse representations for single image super-resolution,” IEEE Trans. Image Process., vol.23, No.6 (2014): 2569-2582.