Image Interpolation via Graph-Based Bayesian Label Propagation

Xianming Liu, *Member, IEEE*, Debin Zhao, *Member, IEEE*, Jiantao Zhou, *Member, IEEE*, Wen Gao, *Fellow, IEEE*, and Huifang Sun, *Fellow, IEEE*

***Abstract*— In this paper, we propose a novel image interpolation algorithm via graph-based Bayesian label propagation. The basic idea is to first create a graph with known and unknown pixels as vertices and with edge weights encoding the similarity between vertices, then the problem of interpolation converts to how to effectively propagate the label information from known points to unknown ones. This process can be posed as a Bayesian inference, in which we try to combine the principles of local adaptation and global consistency to obtain accurate and robust estimation. Specially, our algorithm first constructs a set of local interpolation models, which predict the intensity labels of all image samples, and a loss term will be minimized to keep the predicted labels of the available low-resolution (LR) samples sufficiently close to the original ones. Then, all of the losses evaluated in local neighborhoods are accumulated together to measure the global consistency on all samples. Moreover, a graph-Laplacian-based manifold regularization term is incorporated to penalize the global smoothness of intensity labels, such smoothing can alleviate the insufficient training of the local models and make them more robust. Finally, we construct a unified objective function to combine together the global loss of the locally linear regression, square error of prediction bias on the available LR samples, and the manifold regularization term. It can be solved with a closed-form solution as a convex optimization problem. Experimental results demonstrate that the proposed method achieves competitive performance with the state-of-the-art image interpolation algorithms.**

***Index Terms*— Image interpolation, graph, label propagation, local adaptation, global consistency, regression.**

I. Introduction

IMAGE interpolation, which is the art of rescaling a low-resolution (LR) image to a high-resolution (HR) version, has become a very active area of research in image processing. The interest in image interpolation is born not only in the great practical importance of enhancing resolution of images, such as in the fields of digital photography, computer vision, computer graphics, medical imaging and consumer electronics, but also the important theoretical value of using image interpolation to understand the validity of different image models in inverse problems. In the last several years, there have been a great deal of works on image interpolation. In general, image interpolation techniques can be categorized into three families: upscaling-based methods [1]–[13], reconstruction-based methods [14], [15], and learning-based methods [16]–[18].

Considering the underlying image model during interpolation, most of image interpolation algorithms could be categorized as globally nonadaptive and locally adaptive ones. A globally nonadaptive algorithm trains the interpolation model using the whole image sample set, while a locally adaptive algorithm trains the model by using only useful local information. The representative globally nonadaptive methods are those based on classical data-invariant linear filters, such as bilinear, bicubic [2], and cubic spline algorithms [3]. These methods have a relatively low computation complexity but suffer from the inability to adapt to varying pixel structures, which results in blurred edges and annoying artifacts.

The locally adaptive algorithms usually demonstrate better empirical results than globally nonadaptive ones since it is much easier to seek some functions that are capable of producing good predictions on some specified regions of images. In the literature, many locally adaptive learning methods have been proposed with great success. Li and Orchard [4] propose to estimate local covariance coefficients from a LR image, and then project the estimated covariance to the HR image to adapt the interpolation. The improved new edge-directed interpolation (INEDI) method [5] modifies NEDI by varying the size of the training window according to the edge size and achieves better performance. In [6], a method named ICBI is proposed to use local second order information to adapt the interpolation and an iterative refinement is further exploited to remove artifacts while preserving image features and texture. In [7], Zhang and Wu propose the named SAI algorithm, which learns and adapts varying scene structures using a locally linear regression model, and interpolates the missing pixels in a group by a soft-decision manner. Liu *et al.* in [8] propose an effective image interpolation algorithm based on regularized local linear regression (RLLR), in which the ordinary least squares error norm is replaced with the moving least squares error norm leading to a robust estimator of local image

Manuscript received November 22, 2012; revised October 11, 2013; accepted November 3, 2013. Date of publication December 11, 2013; date of current version January 27, 2014. This work was supported in part by the National Science Foundation of China under Grants 61300110 and 61272386, in part by the Macau Science and Technology Development Fund under Grant FDCT/009/2013/A1, and in part by the Research Committee at University of Macau under grant SRG023-FST13-ZJT and MRG021/ZJT/2013/FST

X. Liu and D. Zhao are with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: xmliu.hit@gmail.com; dbzhao@jdl.ac.cn).

J. Zhou is with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Taipa, Macau (e-mail: jtzhou@umac.mo).

W. Gao is with the National Engineering Laboratory for Video Technology, and Key Laboratory of Machine Perception, School of Electrical Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: wgao@pku.edu.cn).

H. Sun is with the Mitsubishi Electric Research Laboratories, Cambridge, MA 02139 USA (e-mail: hsun@merl.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2294543

structure. Similar to RLLR, Hung and Siu in [9] extend the SAI algorithm in a weighted manner to get more robust results. Although the local model based methods achieve wonderful performance, the description ability of this family of models is limited. It is difficult to guarantee that the interpolated image is best under the global view. Another problem for local learning is that in some cases the number of samples is not enough to sufficiently train the interpolation model, which will result in over-fitting the data. Therefore, how to improve the performances of the local learning algorithms is still a challenging problem.

After reviewing the global and local methods, a natural question is if combining the robustness of global models with the flexibility of local models can lead to a more effective solution. From a statistical perspective, image interpolation is an ill-posed problem. The key to this task is the prior assumption about the properties that the intensity labels should have over the image sample set. One common assumption for natural images is *intensity consistency* [4], [19], which means: 1) nearby points are likely to have the same or similar intensity labels; and 2) points on the same structure (manifold) are likely to have the same or similar intensity labels. Note that the first assumption means natural images are locally smooth, which defines the local consistency; while the second one means natural images possess the non-local self-similarity property, which defines the global consistency. Accordingly, it is a reasonable idea to consider both local and global information contained in images during model learning.

Alternatively, from a machine learning perspective, the available LR image pixels can be regarded as labeled samples while the missing HR pixels as unlabeled ones. In the field of machine learning, the success of semi-supervised learning [20]–[25] is plausibly due to effective utilization of the large amounts of unlabeled data to extract information that is useful for generalization. Therefore, it is reasonable to leverage both labeled and unlabeled data to achieve better predictions. This is especially useful for interpolation on some bridge regions (*i.e.*, regions connecting two different objects), where the number of labeled points is usually not enough to train a robust predictor. We model the whole image sample set as a undirected graph with the vertex set corresponding to the labeled and unlabeled samples, and the edge set representing the relationships between vertices. In this way, we can model the geometric relationships between all data points through the graph. Moreover, the reference [20] states that predicting the labels of the unlabeled data on a graph is equivalent to propagating the labels of the labeled data along the edges to the unlabeled ones. Therefore, the task of image interpolation converts to the problem of graph-based label propagation, that is, how to effectively propagate the label information from the known LR samples through the graph to the missing HR ones to get accurate and robust estimation.

According to the intuition stated above, in this paper, we propose a novel image *I*nterpolation algorithm via *G*raph-based *B*ayesian *L*abel *P*ropagation (IGBLP). We formulate the process of label propagation as a Bayesian inference. The aim is to obtain a labeling of the vertices that is both smooth over the graph and compatible with the labeled data. Specially, the proposed method predicts the intensity label of a data point according to its local neighborhood in a linear way, and then uses a global optimization to ensure robust predictions. In each neighborhood, an optimal model is estimated via regularized locally linear regression. With this model, the intensity labels of all samples in the neighborhood can be predicted. A loss term will be minimized to keep the predicted labels of available LR samples sufficiently close to the original ones. Then, all of the losses evaluated in local neighborhoods are accumulated together to measure the global consistency on the label and unlabeled data. Moreover, a graph-Laplacian based manifold regularization term is incorporated to penalize the global smoothness of intensity labels, such smoothing can alleviate the insufficient training of the local models and make them more robust. Finally, we propose a unified loss function to combine together the global loss of the locally linear regression, square error of intensity labels of the available LR samples and the manifold regularization term, which could be solved with a closed-form solution as a convex optimization problem. In this way, a graph-based label propagation algorithm with local and global consistency is developed.

The rest of this paper is organized as follows. Section II introduces the general framework of Bayesian label propagation for interpolation. Section III presents how transductive regression can be performed with local and global consistency. Section IV details the convex optimization solution of the proposed algorithm. Experimental results are presented in Section V. Section VI concludes the paper.

II. Graph-Based Bayesian Label Propagation for Interpolation

The problem of image interpolation could be defined as follows: given an image sample set including n pixels $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l, \mathbf{x}_{l+1}, \ldots, \mathbf{x}_n\} \in \Re^2$, of which $\mathbf{X}_L = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l\}$ are the available LR samples with intensity values labeled by $\mathbf{y}_L = \{y_i\}_{i=1}^{l}$; and $\mathbf{X}_U = \{\mathbf{x}_{l+1}, \ldots, \mathbf{x}_n\}$ are the missing HR samples, their intensity values are unlabeled. Given the dataset $D = [\mathbf{X} = \mathbf{X}_L \cup \mathbf{X}_U, \mathbf{y}_L]$, the task of image interpolation converts to infer the posteriori probability $p(\mathbf{y}_U|D)$, where $\mathbf{y}_U = \{y_i\}_{i=l+1}^{n}$ is the labels of the remaining $n - l$ missing samples in $\mathbf{X}_U$. By the fact that unlabeled samples are given beforehand and no other test samples will ever be considered, this is a *transductive regression* problem.

We model the whole image sample set as a undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the vertex set $\mathcal{V} = \mathbf{X}$ corresponding to the labeled and unlabeled pixels, and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ representing the relationships between vertices. Each edge is assigned a weight $\mathbf{W}_{ij}$ which reflects the affinity of $\mathbf{x}_i$ and $\mathbf{x}_j$. In this way, we can model the geometric relationships of all samples in the form of a graph. Through the defined graph, predicting the labels of the unlabeled data is equivalent to propagating the labels of the labeled data along the edges to the unlabeled ones.

To leverage both labeled and unlabeled samples, we assume that the hard intensity label y_i depends upon the hidden soft

label f_i for all i, then the posterior can be written as:

$$p(\mathbf{y}_U|D) = \int_{\mathbf{f}} p(\mathbf{y}_U|\mathbf{f})\, p(\mathbf{f}|D) d\mathbf{f}, \tag{1}$$

where $\mathbf{f} = \{f_1, f_2, \ldots, f_n\}$ are the soft labels, which are defined as outputs of the prediction functions, *i.e.*, the estimates of true labels of all samples in the graph. In this way, all samples in the graph including labeled and unlabeled ones are participated into the inference process. We first approximate the posterior $p(\mathbf{f}|D)$ and then use the result of Eq. (1) as the estimated intensity labels of missing HR samples. Using the Bayes rule we can write $p(\mathbf{f}|D)$ as:

$$p(\mathbf{f}|D) = p(\mathbf{f}|\mathbf{X}, \mathbf{y}_L) \propto p(\mathbf{f}|\mathbf{X})\, p(\mathbf{y}_L|\mathbf{f}). \tag{2}$$

The term $p(\mathbf{f}|\mathbf{X})$ is the probability of the soft labels given the sample set. It can be regarded as the prior. One commonly used prior assumption for natural images is *insistency consistency* [4], [19], which states: 1) nearby points are more likely to have the same or similar intensity labels; and 2) points on the same manifold are more likely to have the same or similar intensity labels. This prior term means that it should give higher probability to the labeling that respects the local smoothness and global consistency of the data graph. Equivalently, we can interpret this probability in an exponential form [25]:

$$p(\mathbf{f}|\mathbf{X}) \propto \exp\left(-\frac{1}{2} F(\mathbf{f}, \mathbf{X})\right), \tag{3}$$

where $F(\cdot, \cdot)$ is a general form of the regularization function.

The second term $p(\mathbf{y}_L|\mathbf{f})$ is the likelihood that incorporates the information provided by the LR intensity labels. Assuming conditional independence of the observed labels given the hidden soft labels, the likelihood can be written as:

$$p(\mathbf{y}_L|\mathbf{f}) = \prod_{i=1}^{l} p(y_i|f_i). \tag{4}$$

The likelihood models the probabilistic relation between the observed label y_i and the hidden label f_i, which can also be interpreted in an exponential form:

$$p(y_i|f_i) \propto \exp\left(-\frac{1}{2} \mathcal{L}(y_i, f_i)\right), \tag{5}$$

where $\mathcal{L}(\cdot, \cdot)$ is a general form of the loss function.

With the above definition, we can easily derive the following result:

$$p(\mathbf{f}|D) \propto \exp\left\{-\left(\sum_{i=1}^{l} \mathcal{L}(y_i, f_i) + \lambda F(\mathbf{f}, \mathbf{X})\right)\right\}. \tag{6}$$

III. Transductive Regression With Local and Global Consistency

As stated in the above section, maximum a posteriori probability (MAP) estimate is equivalent to minimizing an augmented optimization objective function. As a consequence, the problem of posterior inference converts to how to appropriately define and combine the loss and regularization terms, which should thoroughly respect the statistical property of the target image. In the following, we will show how to define the loss and regularization terms from the principle of local adaptation and global consistency.

A. The Principle of Global Consistency

Keep in mind that the desired soft labels should be both smooth over the graph and compatible with the labeled data. First, let us consider the definition of loss function (*i.e.*, the likelihood) from the principle of global consistency. Given labeled samples $\{\mathbf{X}_L, \mathbf{y}_L\} = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_l, y_l)\}$, we try to select a good interpolation model f by minimizing the following *global loss function*:

$$\mathcal{J}_g = \sum_{i=1}^{l} \mathcal{L}(y_i, f(\mathbf{x}_i, \mathbf{w})) + \lambda ||f||^2_{\mathcal{F}}, \tag{7}$$

where $\mathcal{L}(\cdot, \cdot)$ measures the bias of estimated soft labels from the original ones, $\mathbf{w}$ is the parameter vector of the interpolation model, $||f||_{\mathcal{F}}$ is the induced norm of f in the functional space $\mathcal{F}$ (*e.g.*, $\mathcal{F}$ can be a reproducing kernel Hilbert space (RKHS) induced by some kernel k), which reflects the complexity of the predictor f. Clearly, Eq. (7) is in a supervised and global manner, which only utilizes labeled samples to train one interpolation function for the whole image samples.

Then, let us consider the definition of regularization term (*i.e.*, the prior term). From a geometric perspective, there is a probability distribution p on $\mathbf{X} \times \Re$ to generate image samples. The available LR samples are $(\mathbf{x}, y)$ pairs generated according to $p(\mathbf{x}, y)$, the rest missing HR samples are simply drawn according to the marginal distribution $p(\mathbf{x})$ of p. One might hope that knowledge of the marginal $p(\mathbf{x})$ can be exploited for better model learning. In the field of machine learning, Belkin *et al.* [24] propose a geometric framework for learning from labeled and unlabeled examples. They state that there is a specific relationship between the marginal distribution $p(\mathbf{x})$ and the conditional distribution $p(y|\mathbf{x})$. It assumes that if two points $\mathbf{x}_1$ and $\mathbf{x}_2$ are close in the intrinsic geometry of $p(\mathbf{x})$, the conditional distribution $p(y|\mathbf{x}_1)$ and $p(y|\mathbf{x}_2)$ should be similar. In another word, $p(y|\mathbf{x})$ should vary smoothly along the geodesics in the intrinsic geometry of $p(\mathbf{x})$. According to the geometric intuition stated above, we extend the framework of model learning from supervised to semi-supervised by incorporating additional information about the geometric structure of the marginal $p(\mathbf{x})$, which seeks a global optimal interpolation function f by minimizing the following objective function:

$$\mathcal{R}_g = \sum_{i=1}^{l} \mathcal{L}(y_i, f(\mathbf{x}_i, \mathbf{w})) + \gamma_A ||f||^2_{\mathcal{F}} + \gamma_I ||f||^2_{\mathcal{I}}, \tag{8}$$

where the additional penalty term $||f||^2_{\mathcal{I}}$ reflects the intrinsic geometric information of the marginal distribution $p(\mathbf{x})$. In most cases, the marginal distribution $p(\mathbf{x})$ is unknown. Therefore, we have to attempt to get empirical estimates of $p(\mathbf{x})$ and $||f||^2_{\mathcal{I}}$. We assume the support of $p(\mathbf{x})$ is a compact submanifold $\mathcal{M} \in \Re^n$, then naturally $||f||^2_{\mathcal{I}}$ can be approximated by

$$\|f\|^2_{\mathcal{I}} = \int_{x \in \mathcal{M}} \|\nabla_{\mathcal{M}} f\|^2 \, dp(\mathbf{x}), \tag{9}$$

where $\nabla_{\mathcal{M}} f$ denotes the gradient of f with respect to $\mathcal{M}$. Intuitively, $||f||_{\mathcal{I}}^2$ measures how f varies on $\mathcal{M}$, *i.e.*, the smoothness of f.

In the case of digital image, we can derive the discretized form for gradient as follows:

$$\nabla_{\mathcal{M}} f = \frac{f(\mathbf{x}_i) - f(\mathbf{x}_j)}{d(\mathbf{x}_i, \mathbf{x}_j)}, \tag{10}$$

where $d(\mathbf{x}_i, \mathbf{x}_j)$ represents the distance between $\mathbf{x}_i$ and $\mathbf{x}_j$. Then we can rewrite $||f||_{\mathcal{I}}^2$ as:

$$||f||_{\mathcal{I}}^2 = \sum_{i,j=1}^{n} (f(\mathbf{x}_i) - f(\mathbf{x}_j))^2 \mathbf{W}_{ij} = \mathbf{f}^{\mathrm{T}} \mathbf{L} \mathbf{f}, \tag{11}$$

where $\mathbf{W}_{ij}$ is in inverse proportion to $d^2(\mathbf{x}_i, \mathbf{x}_j)$. $\mathbf{W}_{ij}$ is the edge weight in the data adjacency graph which reflects the affinity between $\mathbf{x}_i$ and $\mathbf{x}_j$. In graph construction, edge weights plays a crucial role. In this paper, we combine the edge-preserving property of bilateral filter [27] and the robust property of non-local-means weight [19] to design the edge weights, which are defined as follows:

$$\mathbf{W}_{ij} = \frac{1}{C} \exp\left\{-\frac{||\mathbf{x}_j - \mathbf{x}_i||^2}{\varepsilon^2}\right\} \times \exp\left\{-\frac{G \cdot ||SW(\mathbf{x}_j) - SW(\mathbf{x}_i)||^2}{\varepsilon^2}\right\}, \quad \varepsilon > 0, \tag{12}$$

where the first exponential term considers the geometrical nearby, and the second one considers the structural similarity. The structural similarity of $\mathbf{x}_i$ and $\mathbf{x}_j$ is computed by comparing the similarity of windows $SW(\mathbf{x}_i)$ and $SW(\mathbf{x}_j)$, which are the local patches centered on $\mathbf{x}_i$ and $\mathbf{x}_j$. G is a Gaussian kernel used to further take into account the distance between the central pixel and other pixels in the local patch. We define $\mathbf{L} = \mathbf{D} - \mathbf{W} \in \Re^{n \times n}$ as the graph Laplacian where $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}(i,i) = \sum_j \mathbf{W}_{ij}$. Through the graph Laplacian, the label information is propagated from labeled samples to unlabeled ones.

B. The Principle of Local Adaptation

The objective function defined above provides us an excellent framework to learn from both labeled and unlabeled samples. However, the structural loss is defined in a global way, *i.e.*, for the whole image, we only need to pursue one interpolation model f. Actually, as pointed out by [26], it is usually not easy to find a unique function which holds good predictability in the entire data space. But it is much easier to seek some models that are capable of producing good predictions on some specified regions of the input space. Accordingly, we resort to the local learning strategy to improve the accuracy of predictions. More specially, for each data point $\mathbf{x}_i \in \mathbf{X}$, we consider the linear affine transformation model $f(\cdot; \mathbf{w}_i, b_i)$ defined as follows:

$$f(\mathbf{x}_i; \mathbf{w}_i, b_i) = \mathbf{w}_i^T \Phi(\mathbf{x}_i) + b_i, \tag{13}$$

where $\mathbf{w}_i$ and b_i are the *weight vector* and *bias* of the linear estimator, $\Phi(\mathbf{x}_i) \in \Re^{d \times 1}$ is the intensity label vector of 8-connected neighboring samples of $\mathbf{x}_i$ [7]; $f(\mathbf{x}_i)$ is the estimated intensity label of $\mathbf{x}_i$, which is obtained by a weighted aggregation of the pixels in its local neighborhood.

To compute the model parameters, we split the whole input space $\mathcal{N}$ into c overlapped local neighborhoods and formulate model learning as a set of c optimization problems. It is usually more effective to minimize the following local cost function for each neighborhood $\mathcal{N}_i (1 \le i \le c)$:

$$\mathcal{J}_l^i = \sum_{j=1}^{l} \mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon) \mathcal{L}(y_j, f(\mathbf{x}_j; \mathbf{w}_i, b_i)) + \gamma_A ||f_{\mathbf{w}_i, b_i}||_{\mathcal{F}}^2, \tag{14}$$

where $\mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon)$ is the similarity kernel centered at $\mathbf{c}_i$ with width ε, which is defined in the same form as the edge weight $\mathbf{W}_{ij}$:

$$\mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon) = \frac{1}{C_i} \exp\left\{-\frac{||\mathbf{x}_j - \mathbf{c}_i||^2}{\varepsilon^2}\right\} \times \exp\left\{-\frac{G \cdot ||SW(\mathbf{x}_j) - SW(\mathbf{c}_i)||^2}{\varepsilon^2}\right\}, \quad \varepsilon > 0. \tag{15}$$

As shown in the above equation, $\mathcal{K}$ includes two parts, which play dual role in model learning: the first exponential function is to choose samples located in the local neighborhood $\mathcal{N}_i$ centered on $\mathbf{c}_i$; the second exponential function is actually the form of non-local-means weight [19], which is to choose samples with similar structure and therefore can reduce the influence of outlier in regression.

In this way, we define a function $f_{\mathbf{w}_i, b_i}$ with parameter $(\mathbf{w}_i, b_i)$ for each local region $\mathcal{N}_i$, that is, we define c local interpolation functions $\{f_{\mathbf{w}_i, b_i}\}_{i=1}^c$. Intuitively, similar local neighborhoods should share similar model parameters. Therefore, it is natural to add together the losses estimated on all of the c neighborhoods, and the *total local structural loss* is defined as

$$\mathcal{J}_l = \sum_{i=1}^{c} \mathcal{J}_l^i = \sum_{i=1}^{c} \sum_{j=1}^{l} \mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon) \mathcal{L}(y_j, f(\mathbf{x}_j; \mathbf{w}_i, b_i)) + \gamma_A ||f_{\mathbf{w}_i, b_i}||_{\mathcal{F}}^2. \tag{16}$$

By minimizing the above objective function, a collaborative model learning mechanism is achieved.

Now let us return to the semi-supervised learning scenario, which aims to learn from both labeled and unlabeled data samples. As shown in Eq. (1), we introduce a set of soft labels $\{f_1, f_2, \ldots, f_n\}$ into the loss function such that f_i directly determines the final estimated label of $\mathbf{x}_i$. Then we can redefine the total local loss as

$$\mathcal{J}_l = \sum_{i=1}^{c} \sum_{j=1}^{n} \mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon) \mathcal{L}(f_j, f(\mathbf{x}_j; \mathbf{w}_i, b_i)) + \gamma_A ||f_{\mathbf{w}_i, b_i}||_{\mathcal{F}}^2. \tag{17}$$

[3pt] In this way, interpolation models are trained locally using all samples in the neighborhood. Note that by minimizing $\mathcal{J}_l$ we can obtain the optimal $\{f_i\}_{i=1}^n$ and $\{\mathbf{w}_i, b_i\}_{i=1}^c$. The advantage of this approach is that while it may be nontrivial

to produce an image with the desired property everywhere, it is often easier to obtain the property locally.

C. Local and Global Consistency

Recalling the graph-Laplacian regularization framework introduced in the global principle, we also expect f_i to have some geometrical properties. More concretely, we hope $\{f_i\}_{i=1}^n$ to be sufficiently smooth with respect to the intrinsic data graph. Therefore, we finally combine the local and global principle and construct a unified objective function which uses both labeled and unlabeled data and achieves local and global consistency:

$$\mathcal{R} = \sum_{i=1}^{l} \mathcal{L}(y_i, f(\mathbf{x}_i, \mathbf{w})) + \lambda \left(\sum_{i=1}^{c} \sum_{j=1}^{n} \mathcal{K}(\mathbf{x}_j - \mathbf{c}_i, \varepsilon) \mathcal{L}(f_j, f(\mathbf{x}_j; \mathbf{w}_i, b_i)) + \gamma_A ||f_{\mathbf{w}_i, b_i}||_{\mathcal{F}}^2 \right) + \gamma_I \mathbf{f}^T \mathbf{L} \mathbf{f}, \tag{18}$$

where the first term is called the *prediction loss*, which measures the inconsistency between the predicted and initial labels on known LR samples $\mathbf{X}_L$. The second term is called the *total local structural loss*, which is defined to reflect the idea that we split the input space into c local neighborhoods and perform model learning collaboratively. This term punishes the predictability and complexity of the local prediction functions, which is therefore called the local regularization. The third term is called *manifold regularization*, which creates the dependency between different local models, and therefore, can tie the individual local model learning together. It penalizes the smoothness of the intensity labels over the entire data graph, thus is referred to as the global regularization.

IV. Convex Optimization

In the previous section, we introduce the global and local principal for image interpolation, and construct a unified framework to perform graph-based label propagation with local and global consistency. To derive a practical image interpolation algorithm, we should derive an efficient solution for the objective function defined in Eq. (18). In the following, let us take this issue into account.

A. The Derivation of Analytical Solution

Now return to the total local loss defined in Eq. (17). With the loss function $\mathcal{L}(\cdot, \cdot)$ defined as square loss like that in the least square regression, the total local loss can be further formulated as

$$\mathcal{J}_l = \sum_{i=1}^{c} \sum_{\mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)} \theta(\mathbf{x}_i, \mathbf{x}_j) \left(\mathbf{w}_i^T \Phi(\mathbf{x}_j) + b_i - f_j \right)^2 + \gamma_A ||f_{\mathbf{w}_i, b_i}||^2, \tag{19}$$

where

$$\theta(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{ -\frac{G \cdot ||SW(\mathbf{x}_j) - SW(\mathbf{x}_i)||^2}{\varepsilon^2} \right\} \tag{20}$$

is the non-local-means part of $\mathcal{K}$, which plays the same role as moving weights in moving least squares [28]; and $\mathcal{N}(\mathbf{x}_i)$ represents the local neighborhood centered at $\mathbf{x}_i$. Similarly, the local structural loss in each neighborhood define in Eq. (14) can be rewritten as

$$\mathcal{J}_l^i = \sum_{\mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)} \theta(\mathbf{x}_i, \mathbf{x}_j) \left(\mathbf{w}_i^T \Phi(\mathbf{x}_j) + b_i - f_j \right)^2 + \gamma_A ||f_{\mathbf{w}_i, b_i}||^2. \tag{21}$$

Let

$$\mathbf{G}_i = \begin{bmatrix} \Phi_i^T & \mathbf{1} \\ \sqrt{\gamma_A} \cdot \mathbf{I}_d & \mathbf{0} \end{bmatrix}, \Phi_i = [\Phi(\mathbf{x}_{i_1}), \Phi(\mathbf{x}_{i_2}), \ldots, \Phi(\mathbf{x}_{i_{n_i}})], \quad \widehat{\mathbf{f}}_i = [f_{i_1}, f_{i_2}, \ldots, f_{i_{n_i}}, \mathbf{0}^T]^T, \tag{22}$$

where $\mathbf{x}_{i_j}$ is the j-th neighbor of $\mathbf{x}_i$, n_i is the cardinality of $\mathcal{N}(\mathbf{x}_i)$, $\mathbf{1} = [1, 1, \ldots, 1]^T \in \Re^{n_i \times 1}$ and $\mathbf{0}$ is a $d \times 1$ zero vector, Eq. (21) can be formulated in the matrix form as

$$\mathcal{J}_l^i = \left\{ \mathbf{G}_i \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \widehat{\mathbf{f}}_i \right\}^T \cdot \mathbf{V} \cdot \left\{ \mathbf{G}_i \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \widehat{\mathbf{f}}_i \right\}, \tag{23}$$

where $\mathbf{V} = diag\left(\theta(\mathbf{x}_i, \mathbf{x}_{i_1}), \ldots, \theta(\mathbf{x}_i, \mathbf{x}_{i_{n_i}}), 1, \ldots, 1\right) \in \Re^{(n_i+d)\times(n_i+d)}$. $\mathbf{V}$ is a diagonal matrix, so the above equation can be further formulated as

$$\mathcal{J}_l^i = \left\{ \mathbf{V}^{\frac{1}{2}} \mathbf{G}_i \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \mathbf{V}^{\frac{1}{2}} \widehat{\mathbf{f}}_i \right\}^T \left\{ \mathbf{V}^{\frac{1}{2}} \mathbf{G}_i \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \mathbf{V}^{\frac{1}{2}} \widehat{\mathbf{f}}_i \right\}. \tag{24}$$

Let $\widetilde{\mathbf{G}_i} = \mathbf{V}^{\frac{1}{2}} \mathbf{G}_i$ and $\widetilde{\mathbf{f}_i} = \mathbf{V}^{\frac{1}{2}} \widehat{\mathbf{f}}_i$, the above equation becomes

$$\mathcal{J}_l^i = \left\{ \widetilde{\mathbf{G}_i} \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \widetilde{\mathbf{f}_i} \right\}^T \left\{ \widetilde{\mathbf{G}_i} \begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix} - \widetilde{\mathbf{f}_i} \right\}. \tag{25}$$

To derive the optimal transformation parameters $[\mathbf{w}_i \ b_i]^T$, we take the derivative of the loss function $\mathcal{J}_l^i$ with respect to $[\mathbf{w}_i \ b_i]^T$ and set the derivative to 0, then the optimal solution can be represented by

$$\begin{bmatrix} \mathbf{w}_i \\ b_i \end{bmatrix}^* = (\widetilde{\mathbf{G}_i}^T \widetilde{\mathbf{G}_i})^{-1} \widetilde{\mathbf{G}_i}^T \widetilde{\mathbf{f}_i}. \tag{26}$$

With this solution, the total structural loss defined in Eq. (19) becomes

$$\mathcal{J}_l = \sum_i \mathcal{J}_l^i = \sum_i \widetilde{\mathbf{f}_i}^T \widehat{\mathbf{G}_i}^T \widehat{\mathbf{G}_i} \widetilde{\mathbf{f}_i}, \tag{27}$$

where $\widehat{\mathbf{G}_i} = \mathbf{I} - \widetilde{\mathbf{G}_i} (\widetilde{\mathbf{G}_i}^T \widetilde{\mathbf{G}_i})^{-1} \widetilde{\mathbf{G}_i}^T$. It can be easily demonstrated that $\widehat{\mathbf{G}_i}$ is a orthogonal projection matrix. With the property of orthogonal projection matrix, $\mathcal{J}_l$ can be rewritten as:

$$\mathcal{J}_l = \sum_i \mathcal{J}_l^i = \sum_i \widetilde{\mathbf{f}_i}^T \widehat{\mathbf{G}_i} \widetilde{\mathbf{f}_i}. \tag{28}$$

We split the matrix $\widehat{\mathbf{G}_i}$ into four blocks after the n_i-th row and column:

$$\widehat{\mathbf{G}_i} = \begin{bmatrix} \mathbf{A}_i & \mathbf{B}_i \\ \mathbf{C}_i & \mathbf{D}_i \end{bmatrix}, \tag{29}$$

where $\mathbf{A}_i \in \Re^{n_i \times n_i}$. Let $\mathbf{f}_i = [f_{i_1}, f_{i_2}, \ldots, f_{i_{n_i}}]^T$ and $\mathbf{K} = diag(\theta(\mathbf{x}_i, \mathbf{x}_{i_1}), \theta(\mathbf{x}_i, \mathbf{x}_{i_2}), \ldots, \theta(\mathbf{x}_i, \mathbf{x}_{i_{n_i}})) \in \Re^{n_i \times n_i}$, then

$$\widetilde{\mathbf{f}_i} = \mathbf{V}^{\frac{1}{2}}\widehat{\mathbf{f}_i} = \begin{bmatrix} \mathbf{K}^{\frac{1}{2}} & \\ & \mathbf{I}_d \end{bmatrix} \begin{bmatrix} \mathbf{f}_i \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\frac{1}{2}}\mathbf{f}_i \\ \mathbf{0} \end{bmatrix}. \tag{30}$$

According to Eq. (29) and Eq. (30), we can derive

$$\begin{aligned} \widetilde{\mathbf{f}_i}^T \widehat{\mathbf{G}_i} \widetilde{\mathbf{f}_i} &= [\mathbf{f}_i^T \mathbf{K}^{\frac{1}{2}} \quad \mathbf{0}^T] \begin{bmatrix} \mathbf{A}_i & \mathbf{B}_i \\ \mathbf{C}_i & \mathbf{D}_i \end{bmatrix} \begin{bmatrix} \mathbf{K}^{\frac{1}{2}}\mathbf{f}_i \\ \mathbf{0} \end{bmatrix} \\ &= \mathbf{f}_i^T \mathbf{K}^{\frac{1}{2}} \mathbf{A}_i \mathbf{K}^{\frac{1}{2}} \mathbf{f}_i = \mathbf{f}_i^T \widehat{\mathbf{A}_i} \mathbf{f}_i. \end{aligned} \tag{31}$$

For the derivation of $\widehat{\mathbf{A}_i}$, please refer to Appendix.

Define the vector $\widehat{\mathbf{f}} = [\mathbf{f}_1^T, \mathbf{f}_2^T, \ldots, \mathbf{f}_c^T]$ as the concatenated label vector in all of c local neighborhoods, and define the selection matrix $\mathbf{S}$ as a 0-1 matrix with $\mathbf{S}_{ij} = 1$ if $\mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)$, so $\widehat{\mathbf{f}} = \mathbf{S}\mathbf{f}$. We further define the block-diagonal matrix

$$\mathbf{G} = \begin{pmatrix} \widehat{\mathbf{A}_1} & & 0 \\ & \ddots & \\ 0 & & \widehat{\mathbf{A}_c} \end{pmatrix}, \tag{32}$$

then Eq. (28) can be rewritten as

$$\mathcal{J}_l = \mathbf{f}^T \mathbf{S}^T \mathbf{G} \mathbf{S} \mathbf{f}. \tag{33}$$

Let

$$\mathbf{M} = \mathbf{S}^T \mathbf{G} \mathbf{S}, \tag{34}$$

so finally Eq. (19) can be rewritten as

$$\mathcal{J}_l = \mathbf{f}^T \mathbf{M} \mathbf{f}. \tag{35}$$

Finally, the objective function formulated in Eq. (18) can be rewritten as the following matrix form:

$$\mathcal{R} = (\mathbf{f} - \mathbf{y})^T \mathbf{J} (\mathbf{f} - \mathbf{y}) + \lambda \mathbf{f}^T \mathbf{M} \mathbf{f} + \gamma_I \mathbf{f}^T \mathbf{L} \mathbf{f}, \tag{36}$$

where $\mathbf{y} = [\mathbf{y}_L, \mathbf{0}]$ and $\mathbf{y}_L$ records the intensity labels of the labeled image samples, $\mathbf{J} \in \Re^{n \times n}$ is a diagonal matrix whose diagonal elements are one for labeled samples and zero for unlabeled data.

Theorem 1. Minimization of the objective function in Eq. (36) is a convex optimization problem.

Proof: First, we compute the first-order derivative of $\mathcal{R}$ with respect to $\mathbf{f}$:

$$\frac{\partial R}{\partial \mathbf{f}} = 2\mathbf{J}(\mathbf{f} - \mathbf{y}) + \lambda(\mathbf{M} + \mathbf{M}^T)\mathbf{f} + 2\gamma_I \mathbf{L}\mathbf{f} = 0, \tag{37}$$

Then, we compute the second order derivative of $\mathcal{R}$ with respect to $\mathbf{f}$:

$$\frac{\partial R}{\partial \mathbf{f}^2} = 2\mathbf{J} + \lambda(\mathbf{M} + \mathbf{M}^T) + 2\gamma_I \mathbf{L} \succ 0 \tag{38}$$

Due to the second-order derivative of $\mathcal{R}$ with respect to $\mathbf{f}$ is a positive definite matrix, the objective function in Eq. (36) is proved to be a convex optimization problem [32]. ∎

Setting the first-order derivative to zero:

$$\frac{\partial R}{\partial \mathbf{f}} = 2\mathbf{J}(\mathbf{f} - \mathbf{y}) + \lambda(\mathbf{M} + \mathbf{M}^T)\mathbf{f} + 2\gamma_I \mathbf{L}\mathbf{f} = 0, \tag{39}$$

the optimal $\mathbf{f}$ can be finally represented as:

$$\mathbf{f} = 2\left(2\mathbf{J} + \lambda(\mathbf{M} + \mathbf{M}^T) + 2\gamma_I \mathbf{L}\right)^{-1} \mathbf{J}\mathbf{y}. \tag{40}$$

Since $\mathbf{f}$ in Eq. (40) is linear with respect to all known values, the final solution is actually in an analytical form, and can be computed efficiently without any iterative process. After deriving $\mathbf{f}$, we choose the values corresponding to the unlabeled samples in $\mathbf{f}$ as the estimated intensity values of the missing HR samples, and keep the original values of LR samples unchanged.

B. Complexity Analysis and Discussion

The result in Eq. (40) gives a closed-form solution based on the inversion of a matrix in $\Re^{n \times n}$. Let $T(n)$ be the complexity of computing the inverse of a matrix in $\Re^{n \times n}$, and $T(n) = \mathrm{O}(n^3)$ using standard method or $T(n) = \mathrm{O}(n^{2.376})$ with the method of Coppersmith and Winograd, where n is the number of labeled and unlabeled samples in the input space. To achieve good tradeoff between effectiveness and efficiency, we use the same strategy as non-local-means [19], *i.e.*, not use the whole image but a large neighborhood as the input space $\mathcal{N}$.

To gain some insight into our method, we investigate the proposed method from a co-regularized perspective. The last two terms in Eq. (36) can both be reviewed as regularization terms contained with different regularization matrices, one is derived from the view of local adaptation and the other is from global consistency. Benefiting from the co-regularization terms, our method achieves good tradeoff between adaptation and robustness. The interpolation models induced from the local view enjoy the flexibility to adapt to the local statistics of images. However, local adaptation alone is not enough. It may run into the risk of over-fitting the data. To alleviate this issue, the global regularization term is introduced to enforce that a good labeling should not change too much between similarity points (similarity is measured by the edge weights in the graph). Consequently, the global term ensures the interpolation functions vary smoothly with respect to the intrinsic structure collectively revealed by both measured and unmeasured pixels. In a nut shell, different types of regularization matrices can better reveal complementary information and thus could provide a more accurate and robust predictor. For clarity, the algorithm flow is shown in Algorithm 1.

An intuitive understanding about why our method works well is shown as follows. In NEDI and SAI, only one local neighborhood is used as the training window to learn model parameters. As illustrated in Fig. 1, the red window represents such a single local neighborhood. However, it usually runs into the risk of over-fitting the data, which reduces the accuracy of model and further degrades the quality of reconstructed HR images. To alleviate this issue, we split the whole input space into a set of overlapped local neighborhoods to carry out models learning collaboratively. As illustrated in Fig. 1, the blue window denotes the input space, where multiple local neighborhoods are utilized for model learning including not only the red window but also many black ones. More specially, similar local neighborhoods are enforced to share similar model parameters. Such a collaborative learning mechanism is achieved by the global regularization term, in which the edge weight (defined in Eq. (12)) is computed to reflect the affinity of any vertices even that they are far away from

Algorithm 1 Algorithm of Image Interpolation via Graph-Based Bayesian Label Propagation

Input:

$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_l, \mathbf{x}_{l+1}, \cdots, \mathbf{x}_n\} \in \Re^2$, the first l points are available LR samples, their intensity values are $\{y_i\}_{i=1}^{l}$;

Output:

The intensity values $\{y_i\}_{i=l+1}^{n}$ of the rest $n-l$ missing HR samples $\{\mathbf{x}_i\}_{i=l+1}^{n}$.

Procedure:

Step1:

1: Construct the similarity kernel matrix $\mathbf{K}$ according to Eq.(20);
2: Construct $\mathbf{F}_i$ and e according to Eq.(43);
3: Construct $\widehat{\mathbf{A}}_i$ according to Eq.(45);
4: Construct the matrix $\mathbf{G}$ according to Eq.(32);
5: Construct the matrix $\mathbf{S}$ with $\mathbf{S}_{ij} = 1$ if $\mathbf{x}_j \in \mathcal{N}(\mathbf{x}_i)$, $\mathbf{S}_{ij} = 0$ otherwise;
6: Construct the matrix $\mathbf{M}$ according to Eq.(34);

Step 2:

7: Construct the graph-Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where the matrix $\mathbf{W}$ is constructed according to Eq.(12), and $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}(i,i) = \sum_j \mathbf{W}_{ij}$;

Step 3:

8: Compute the label vector $\mathbf{f}$ according to Eq.(40), and choose the corresponding values as the estimated intensity of unlabeled samples.

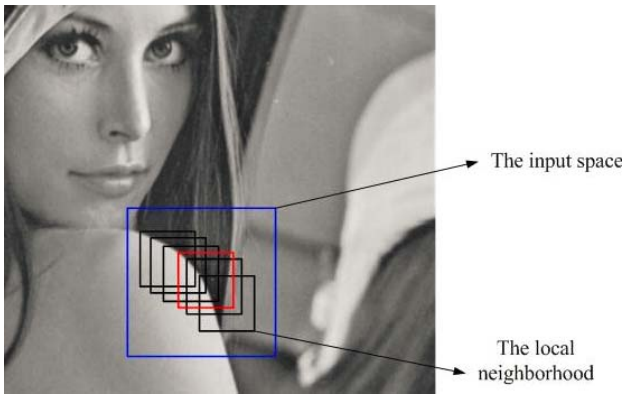


Fig. 1. Illustration of collaborative model learning.

each other. Through creating the dependency between different local models, the individual local model learning problems are tied together. In this way, we can effectively reduce the risk of overfitting and enhance the accuracy and robustness of the learned interpolation model.

V. Experimental Results

In this section, experimental results are presented to demonstrate the performance of the proposed image interpolation algorithm. Given the fact that numerous image interpolation algorithms have been developed during the last two decades, it would be virtually impossible for us to perform a thorough comparative study of the proposed algorithm. Here, the proposed IGBLP algorithm is compared with linear interpolator Bicubic [2], as well as recent state-of-the-art methods including new edge directed interpolation (NEDI) [4], directional filtering and data fusion (DFDF) [12], improve NEDI (INEDI) [5], iterative curvature-based interpolation (ICBI) [6], sparse coding based super-resolution (ScSR) [18], soft-decision adaptive interpolation (SAI) [7], regularized local linear regression (RLLR) [8]. In our experiments, Bicubic interpolation is performed with the MATLAB built-in function, and the source codes of other compared methods are kindly provided by their authors. For thoroughness and fairness of our comparison study, we test seven widely used images in the literature and one texture image, as illustrated in Fig. 2. The source code of this paper can be available from http://homepage.hit.edu.cn/pages/xmliu

A. Comparison With State-of-the-Arts

We first report the comparison results when the scaling factor S is two. Following the same setting as NEDI and SAI, we downsample these HR images by a factor of two in both row and column dimensions without antialiasing filtering to get the corresponding LR images, from which the original HR images are reconstructed by the algorithms under consideration.

Since the original HR images are known in the simulation, we can compare the interpolated results with the true images, and measure the objective quality of those interpolated images. Table I tabulates the objective quality comparison of the nine different methods when applied to the eight test images of Fig. 2. Since PSNR is an average quality measurement over the whole image, we also use edge PSNR (EPSNR) as the measurement to test the reconstruction fidelity of image edges. In our study, the Sobel edge filter is used to locate the edge in the original image, and the PSNR of the pixels on the edge are used to generate the EPSNR. From Table I, it can be observed that for most instances the proposed algorithm works better on PSNR than other eight methods. Compared with global methods, such as Bicubic, the proposed method can significantly improve the objective quality of generated HR images. The average PSNR gain is 0.88dB. Our method also outperforms the edge detection based local methods, such as DFDF, for which the average PSNR gain is 1.02dB. Compared with NEDI and its variance INEDI, our method achieves higher objective performance, the average PSNR gains are 1.96dB and 1.02dB respectively. ScSR exploits an example training set to learn the coupled low- and high-resolution dictionaries, and enforce that sparse representations between the low- and high-resolution image patch pair with respect to their own dictionaries should be the same. The assumption of "the same sparse representation" is too strong. When the approximation atoms are not correctly selected, it produces poor results. By exploiting labeled and unlabeled samples together and keep local and global consistency in transductive regression, our method leads to a significant performance benefits compared with SAI and RLLR, both of which are state-of-the-art locally adaptive model based methods. The average PSNR gains are 0.7dB and 0.45dB, respectively. The proposed method also achieves the

Fig. 2. Eight sample images in the test set.

TABLE I
OBJECTIVE QUALITY COMPARISON OF NINE INTERPOLATION ALGORITHMS (IN dB)

Images	Bicubic		NEDI		DFDF		INEDI		ICBI		ScSR		SAI		RLLR		IGBLP	
	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR	PSNR	EPSNR
Baboon	21.63	18.10	19.91	17.75	19.87	18.01	19.55	18.06	19.06	16.56	20.33	15.77	19.92	18.04	20.67	**18.77**	**22.16**	18.45
Monarch	29.34	21.99	29.35	21.28	29.89	21.43	29.91	21.88	29.82	20.54	26.14	14.98	30.36	21.89	30.53	22.02	**30.83**	**22.14**
Lena	33.98	28.36	33.57	27.75	33.96	28.10	34.11	27.86	34.05	26.99	29.47	19.35	34.72	**28.97**	34.41	28.65	**34.77**	28.81
Airplane	30.17	19.37	28.74	15.42	30.53	19.44	30.66	19.65	30.08	18.68	27.13	14.14	30.72	19.25	30.97	**19.86**	**31.04**	19.82
Peppers	31.59	22.41	29.30	17.83	31.87	22.53	32.05	22.43	31.70	22.18	28.01	16.39	31.84	22.08	32.15	**22.77**	**32.28**	22.75
Splash	33.65	19.76	33.38	15.49	33.79	19.79	33.69	21.28	33.32	21.29	30.38	15.19	33.54	19.14	33.99	**19.87**	**34.08**	19.77
Tower	39.84	28.51	39.85	28.36	39.69	27.64	40.74	28.19	38.94	25.93	31.68	18.43	41.49	28.82	41.97	29.43	**42.85**	**29.48**
Fiber	**24.25**	**25.94**	21.76	19.88	23.83	25.41	22.71	22.38	23.41	21.32	20.74	16.91	23.38	23.73	23.29	21.02	23.49	24.84
Average	30.56	23.06	29.48	20.47	30.42	22.79	30.42	22.71	30.04	21.68	26.73	16.39	30.74	22.74	30.99	22.79	**31.44**	**23.26**

TABLE II
PERFORMANCE COMPARISON OF NINE INTERPOLATION ALGORITHMS THROUGH SSIM AND FSIM

Images	Bicubic		NEDI		DFDF		INEDI		ICBI		ScSR		SAI		RLLR		IGBLP	
	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM	SSIM	FSIM
Baboon	0.6021	0.9021	0.5416	0.8369	0.5326	0.8364	0.5364	0.8357	0.5073	0.8359	0.4455	0.8153	0.5552	0.8046	0.5602	0.8461	**0.6253**	**0.9081**
Monarch	0.8397	0.9707	0.8537	0.9704	0.8509	0.9727	0.8553	0.9743	0.8379	0.9709	0.7956	0.9445	0.8601	0.9741	0.8665	0.9761	**0.8676**	**0.9776**
Lena	0.9141	0.9875	0.9112	0.9862	0.9129	0.9871	0.9175	0.9876	0.9112	0.9868	0.8614	0.9586	0.9184	0.9884	0.9180	0.9879	**0.9209**	**0.9889**
Airplane	0.9119	0.9798	0.9117	0.9782	0.9144	0.9804	0.9166	0.9811	0.9085	0.9781	0.8749	0.9629	0.9171	0.9817	0.9188	0.9823	**0.9199**	**0.9825**
Peppers	0.8684	0.9778	0.8737	0.9777	0.8732	0.9790	0.8822	0.9816	0.8638	0.9770	0.8234	0.9553	0.8752	0.9796	0.8804	0.9811	**0.8836**	**0.9818**
Splash	0.9293	0.9819	0.9290	0.9829	0.9296	0.9829	0.9328	0.9852	0.9241	0.9826	0.8950	0.9631	0.9298	0.9825	**0.9332**	**0.9848**	0.9331	0.9842
Tower	0.9882	0.9875	0.9904	0.9909	0.9892	0.9895	0.9929	0.9923	0.9842	0.9846	0.9467	0.9428	0.9919	0.9920	**0.9950**	0.9927	0.9927	**0.9954**
Fiber	**0.7614**	**0.9626**	0.6491	0.9305	0.7272	0.9559	0.6948	0.9465	0.7416	0.9545	0.5653	0.8998	0.7174	0.9486	0.7191	0.9428	0.7334	0.9501
Average	0.8518	0.9687	0.8325	0.9567	0.8412	0.9604	0.8410	0.9605	0.8348	0.9588	0.7759	0.9302	0.8456	0.9564	0.8489	0.9617	**0.8595**	**0.9710**

highest average EPSNR value. It demonstrates the proposed method can effectively preserve local edge structure in images.

Furthermore, in our study, we use SSIM [29] and FSIM [30] as metrics to measure the performance of these interpolation algorithms. From Table II, it can be observed that our method achieves the highest average SSIM and FSIM scores among all of the competing methods. Given the fact that human visual system (HVS) is the ultimate receiver of the enlarged images, we also show the subjective comparison results. The test image *Airplane* exhibits strong and sharp edges in varying directions and the test image *Monarch* exhibits edges with different scales. Such characteristics make them as prime images to test the fidelity of edge reconstruction. Figs. 3 and 5 illustrate the subjective quality and reconstruction error comparison on these two images. It can be clearly observed that the images reconstructed by the Bicubic interpolator suffer from blurred edges, jaggies, and annoying ringing artifacts. ScSR produces artifacts along contours. The reconstruction quality can be improved to some extent by NEDI, DFDF, INEDI and ICBI, but the image quality is still lower than SAI, RLLR and the proposed IGBLP. SAI and RLLR show improvements over these methods in the regions of edges and textures, reducing the visual defects of these methods. Thanks to the combination of local and global information, IGBLP achieves more wonderful visual quality compared with all other methods. The produced edges in our method are clean and sharp. The outstanding performance of the proposed method is more vivid by observing the error image of *Airplane*. Our algorithm produces smaller interpolation error than other methods. Such results clearly demonstrate the superiority of the proposed

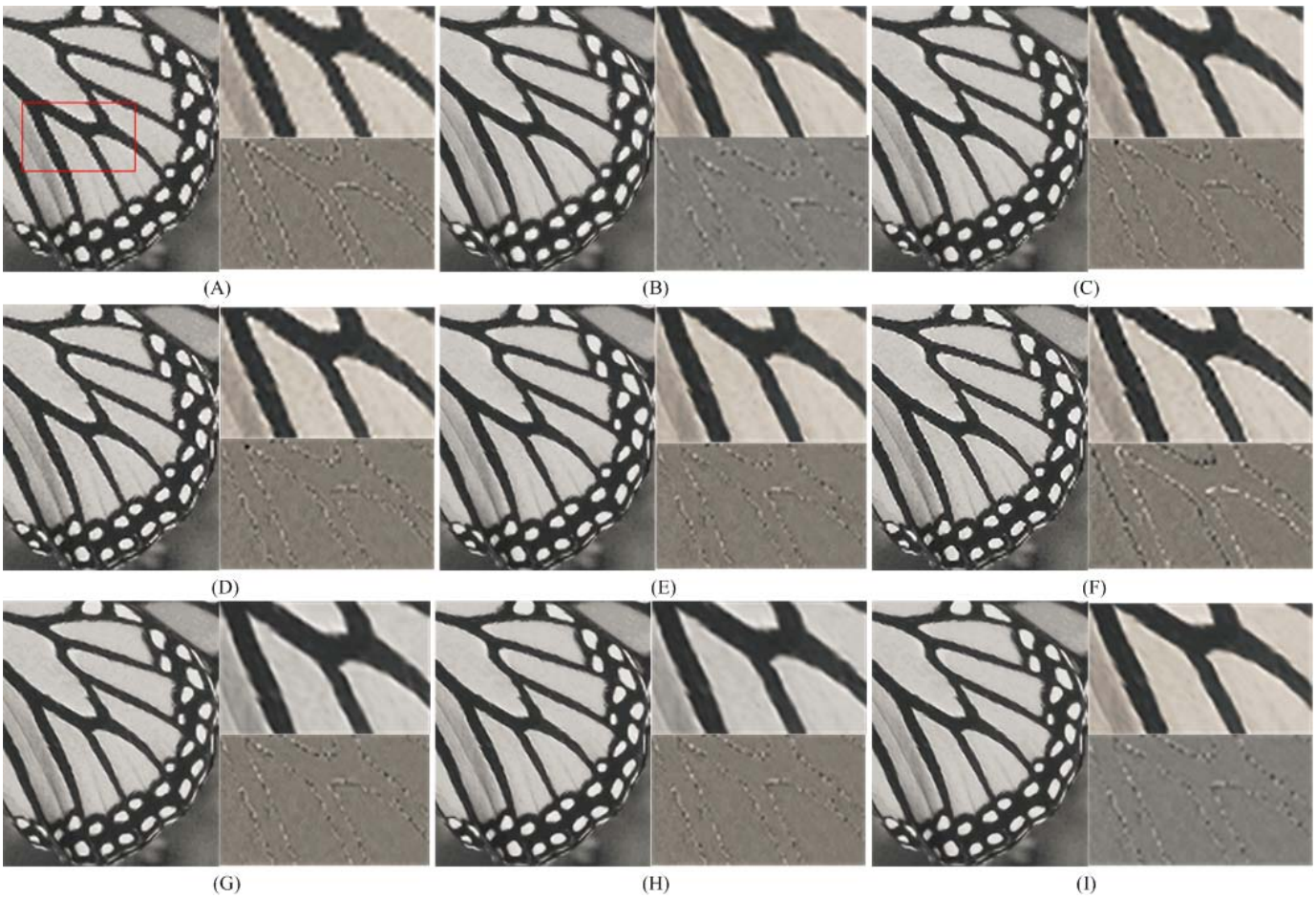


Fig. 3. Reconstruction images and error comparison for *Monarch*. (A) Bicubic, (B) NEDI, (C) DFDF, (D) ICBI, (E) INEDI, (F) ScSR, (G) SAI, (H) RLLR, (I) IGBLP.

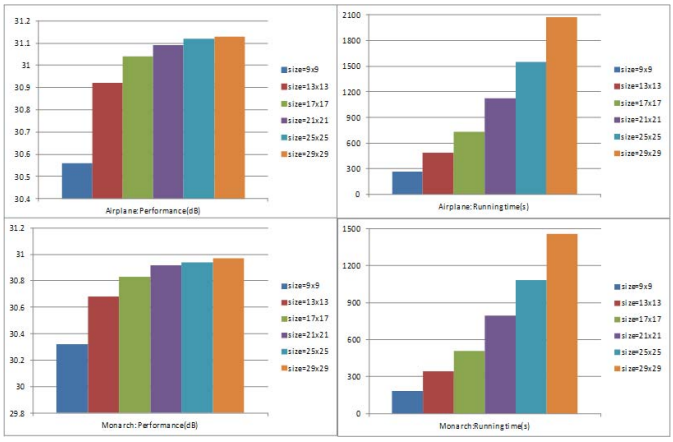


Fig. 4. The influence of the dimension of the input space $\mathcal{N}$ on performance and complexity for *Airplane* and *Monarch*.

method in reconstructing the high frequency, such as edges and textures.

Note that the proposed method can be easily generalized to enlarge images with the scaling factor S greater than two. If $S = 2^z$ with z being a positive integer, one can just apply our method z times. If $2^z < S < 2^{z+1}$, one can first apply the proposed method z times then apply Bicubic interpolation on the enlarged image s times such that $2^z s = S$. In the Table IV, we give the comparison results with respect to PSNR and SSIM when the scaling factor is three. The downsampling process is similar with that of scaling factor two. From the results, we can find in most cases our method outperforms other methods with respect to PSNR and SSIM, and achieves the best average performance on the eight test images.

B. Parameters Setting

There are a few parameters involved in the proposed algorithm. The input space $\mathcal{N}$ includes a set of local neighborhoods $\mathcal{N}_i$. The size of $\mathcal{N}$ is set to 17×17 and the size of local neighborhood $\mathcal{N}_i$ is set to 7×7. The size of similarity window is set to 5×5 in computing the edge weight $\mathbf{W}_{ij}$ in Eq. (12) and $\theta(\mathbf{x}_i, \mathbf{x}_j)$ in Eq. (20), ε^2 is fixed to 0.5. Since any analytical result is still elusive to obtain, the regularization parameters λ and γ_I are evaluated by cross validation [31] from an offline training set including twenty images (not including the eight test images). We divide the offline training set into an estimation subset including ten images and a validation subset including the other ten images. The former subset is used to obtain a parameter estimate and the latter is used to validate the performance under the estimate. According to the training

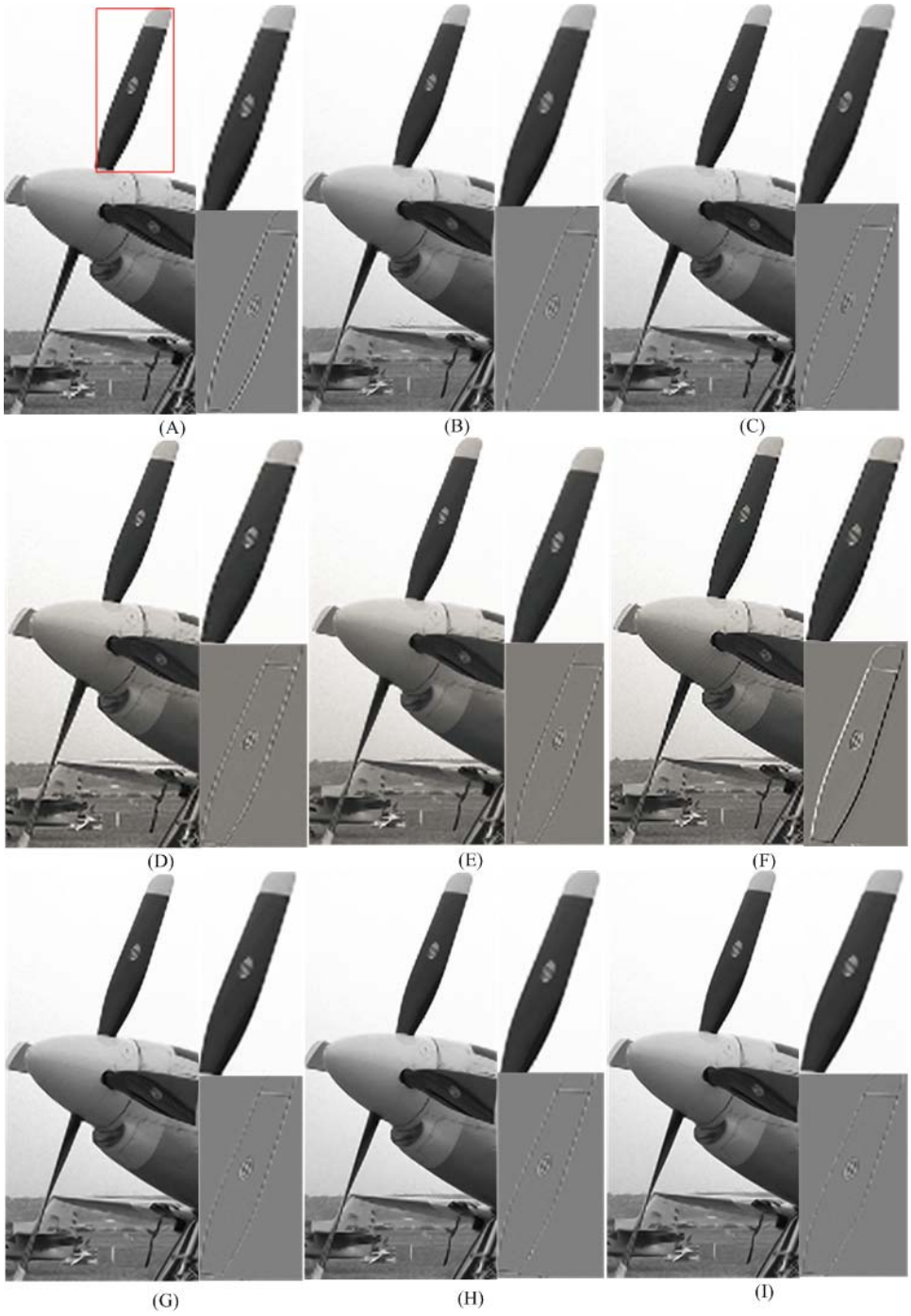


Fig. 5. Reconstruction images and error comparison for *Airplane*. (A) Bicubic, (B) NEDI, (C) DFDF, (D) ICBI, (E) INEDI, (F) ScSR, (G) SAI, (H) RLLR, (I) IGBLP.

results, in practical implementation we set λ and γ_I as 0.5 and 0.01 respectively.

As mentioned in Section III-B, to achieve a good tradeoff between effectiveness and efficiency, we do not use the whole image but a relative large neighborhood as the input space $\mathcal{N}$. *Airplane* and *Monarch* are used as examples to test how the dimension of $\mathcal{N}$ influences the results in terms of performance and computational complexity. As depicted in Fig. 4, we can find both PSNR and running time increase progressively with the dimension of the input space. The PSNR value improves at a rate of gradually slow, while the complexity increases at a rate of gradually fast. According to such a trend, in practical implementation, we set the dimension of $\mathcal{N}$ as 17×17 to achieve good tradeoff between effectiveness and efficiency.

C. Complexity Comparison

Finally, we show the results of complexity comparison. For more clear comparison, Table III gives the PSNR versus average processing times results on a typical computer (Intel Xeon CPU 2.83GHz, 3G Memory) of the nine algorithms. All methods are running on Matlab. The proposed method achieved

TABLE III

OBJECTIVE QUALITY VERSUS AVERAGE PROCESSING TIMES (dB/SECONDS) RESULTS

Images(size)	Bicubic		NEDI		DFDF		INEDI		ICBI		ScSR		SAI		RLLR		IGBLP	
	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME	PSNR	TIME
Baboon(256x256)	21.63	0.40	19.91	4.72	19.87	5.08	19.55	131.95	19.06	2.29	20.33	143.31	19.92	51.86	20.67	68.95	22.16	95.08
Monarch(512x512)	29.34	0.34	29.35	21.21	29.89	19.45	29.91	354.16	29.82	5.06	26.14	1126.75	30.36	143.06	30.53	169.71	30.83	505.26
Lena(512x512)	33.98	0.31	33.57	20.53	33.96	19.22	34.11	337.27	34.05	4.88	29.47	1507.11	34.72	140.19	34.41	147.11	34.77	496.04
Airplane(512x768)	30.17	1.31	28.74	29.52	30.53	29.73	30.66	894.77	30.08	7.95	27.13	1910.78	30.72	212.44	30.97	144.13	31.04	721.44
Peppers(512x512)	31.59	0.31	29.30	20.77	31.87	19.45	32.05	334.13	31.7	4.93	28.01	1166.95	31.84	213.56	32.15	154.71	32.28	501.76
Splash(512x512)	33.65	0.31	33.38	20.09	33.79	19.67	33.69	199.47	33.32	3.51	30.38	1120.21	33.54	141.48	33.99	109.81	34.08	482.22
Tower(300x300)	39.84	0.11	39.85	6.53	39.69	6.52	40.74	60.68	38.94	1.01	31.68	347.09	41.49	47.32	41.97	30.51	42.85	165.22
Fiber(512x512)	24.25	0.32	21.76	22.25	23.83	19.39	22.71	723.23	23.41	5.07	20.74	1108.71	23.38	193.83	23.29	251.02	23.49	550.58
Average	30.56	0.43	29.48	18.20	30.42	17.31	30.42	379.46	30.04	4.34	26.73	1053.86	30.74	142.97	30.99	134.49	31.44	439.70

TABLE IV

QUALITY COMPARISON OF NINE INTERPOLATION ALGORITHMS (IN dB) WHEN THE SCALING FACTOR IS THREE

Images	Bicubic		NEDI		DFDF		INEDI		ICBI		ScSR		SAI		RLLR		IGBLP	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Baboon	17.91	0.2921	18.66	0.3708	18.71	0.3704	18.39	0.3736	18.05	0.3545	17.75	0.2863	18.71	0.3766	19.03	0.3852	**19.05**	**0.3859**
Monarch	22.88	0.7309	25.68	0.7917	26.09	0.7971	**26.14**	0.8004	25.83	0.7869	22.47	0.7283	25.34	0.8008	25.48	0.8079	25.65	**0.8093**
Lena	26.11	0.7796	29.89	0.8528	29.96	0.8534	30.05	0.8559	29.72	0.8493	25.55	0.7649	30.46	0.8618	30.44	0.8635	**30.61**	**0.8641**
Airplane	24.06	0.8026	25.41	0.8444	26.64	0.8488	26.53	0.8469	26.23	0.8391	23.81	0.7958	26.59	0.8497	26.82	0.8536	**26.91**	**0.8546**
Peppers	24.74	0.7656	26.23	0.8218	28.31	0.8267	28.17	0.8264	28.03	0.8151	24.37	0.7564	29.33	0.7672	29.66	0.7765	**29.67**	**0.7767**
Splash	26.99	0.8421	28.11	0.8831	30.53	0.8888	30.39	0.8861	30.31	0.8769	26.62	0.8312	30.34	0.8878	30.36	0.8923	**30.46**	**0.8926**
Tower	28.62	0.8982	40.26	0.9468	32.76	0.9471	33.84	0.9531	32.94	0.9474	28.15	0.8919	**34.12**	**0.9564**	33.33	0.9514	33.83	0.9542
Fiber	18.18	0.2397	**19.46**	**0.3889**	18.99	0.3488	18.74	0.3535	18.64	0.3398	17.54	0.2307	18.66	0.3461	18.76	0.3578	19.01	0.3701
Average	23.68	0.6688	26.71	0.7375	26.49	0.7351	26.53	0.7369	26.21	0.7261	23.28	0.6606	26.69	0.7308	26.73	0.7360	**26.89**	**0.7384**

the best interpolation performance at expense of relative high computational complexity. In the above test, we apply the proposed method on all missing samples. We can manage the computational complexity by reducing the number of samples to estimate. One way is to apply the proposed method only for edge and texture regions, and apply simple bilinear or bicubic interpolation for smooth regions.

VI. CONCLUSION

In this paper, we presented an effective image interpolation algorithm through graph-based label propagation to achieve local and global consistency. Our method is novel in two aspects: (1) both labeled and unlabeled data are explored in the process of model learning. Such a transductive manner can improve the accuracy of predictor for some bridge regions; (2) local and global consistency is achieved during regression, which can make the predictor more robust. These two aspects can be cast into an unified optimization framework, which can be efficiently solved with a closed-form solution. Experimental results on benchmark test images demonstrate that the proposed method achieves very competitive interpolation performance with the state-of-the-art interpolation algorithms. In future work, we will study how to find a good trade-off between performance and computational complexity.

APPENDIX

First, we can derive $\widetilde{\mathbf{G}_i} = \mathbf{V}^{\frac{1}{2}}\mathbf{G}_i = \begin{bmatrix} \mathbf{K}^{\frac{1}{2}} & \\ & \mathbf{I}_d \end{bmatrix} \begin{bmatrix} \Phi_i^T & \mathbf{1} \\ \sqrt{\gamma_A}\cdot\mathbf{I}_d & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\frac{1}{2}}\Phi_i^T & \mathbf{K}^{\frac{1}{2}}\mathbf{1} \\ \sqrt{\gamma_A}\cdot\mathbf{I}_d & \mathbf{0} \end{bmatrix}$.

Let $\Psi^T = \mathbf{K}^{\frac{1}{2}}\Phi_i^T$ and $\Upsilon = \mathbf{K}^{\frac{1}{2}}\mathbf{1}$, and so $\widetilde{\mathbf{G}_i} = \begin{bmatrix} \Psi^T & \Upsilon \\ \sqrt{\gamma_A}\cdot\mathbf{I}_d & \mathbf{0} \end{bmatrix}$. Then

$$\left(\widetilde{\mathbf{G}_i}^T\widetilde{\mathbf{G}_i}\right)^{-1} = \begin{bmatrix} \Psi\Psi^T + \gamma_A\mathbf{I}_d & \Psi\Upsilon \\ \Upsilon^T\Psi^T & 1 \end{bmatrix}^{-1} \tag{41}$$

According to the inverse computation principle of block matrix, we can obtain

$$\left(\widetilde{\mathbf{G}_i}^T\widetilde{\mathbf{G}_i}\right)^{-1} = \begin{bmatrix} \mathbf{F}_i^{-1} + \frac{\mathbf{F}_i^{-1}\Psi\Upsilon\Upsilon^T\Psi^T\mathbf{F}_i^{-1}}{1-e} & -\frac{\mathbf{F}_i^{-1}\Psi\Upsilon}{1-e} \\ -\frac{\Upsilon^T\Psi^T\mathbf{F}_i^{-1}}{1-e} & \frac{1}{1-e} \end{bmatrix}. \tag{42}$$

where

$$\mathbf{F}_i = \Psi\Psi^T + \gamma_A\mathbf{I}_d, e = \Upsilon^T\Psi^T\mathbf{F}_i^{-1}\Psi\Upsilon. \tag{43}$$

We can further obtain

$$\widetilde{\mathbf{G}_i}\left(\widetilde{\mathbf{G}_i}^T\widetilde{\mathbf{G}_i}\right)^{-1}\widetilde{\mathbf{G}_i}^T = \begin{bmatrix} \Psi^T & \Upsilon \\ \sqrt{\gamma_A}\cdot\mathbf{I}_d & \mathbf{0} \end{bmatrix}\left(\widetilde{\mathbf{G}_i}^T\widetilde{\mathbf{G}_i}\right)^{-1}\begin{bmatrix} \Psi & \sqrt{\gamma_A}\cdot\mathbf{I}_d \\ \Upsilon^T & \mathbf{0}^T \end{bmatrix}. \tag{44}$$

Finally, we can derive that

$$\widehat{\mathbf{A}}_i = \mathbf{I}_{n_i} - \left(\Psi^T\mathbf{F}_i^{-1}\Psi + \frac{\Psi^T\mathbf{F}_i^{-1}\Psi\Upsilon\Upsilon^T\Psi^T\mathbf{F}_i^{-1}\Psi - \Psi^T\mathbf{F}_i^{-1}\Psi\Upsilon\Upsilon^T - \Upsilon\Upsilon^T\Psi^T\mathbf{F}_i^{-1}\Psi + \Upsilon\Upsilon^T}{1-e}\right). \tag{45}$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive suggestions which helped us improving our manuscript.

REFERENCES

[1] X. Liu, D. Zhao, R. Xiong, S. Ma, W. Gao, and H. Sun, "Transductive regression with local and global consistency for image super-resolution," in *Proc. IEEE Data Compress. Conf., DCC*, Mar. 2011, pp. 173–182.

[2] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[3] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, no. 6, pp. 508–517, Dec. 1978.

[4] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE. Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[5] N. Asuni and A. Giachetti, "Accuracy improvements and artifacts removal in edge based image interpolation," in *Proc. 3rd Int. Conf. Comput. Vis. Theory Appl.*, 2008, pp. 58–65.

[6] A. Giachetti and N. Asuni, "Real-time artifact-free image upscaling," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2760–2768, Oct. 2011.

[7] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887–896, Jun. 2008.

[8] X. Liu, D. Zhao, R. Xiong, S. Ma, W. Gao, and H. Sun, "Image interpolation via regularized local linear regression," *IEEE. Trans. Image Process.*, vol. 20, no. 12, pp. 3455–3469, Dec. 2011.

[9] K.-W. Hung and W.-C. Siu, "Robust soft-decision interpolation using weighted least squares," *IEEE. Trans. Image Process.*, vol. 21, no. 3, pp. 1061–1069, Mar. 2012.

[10] R. Xiong and W. Gao, "Tanner graph based image interpolation," in *Proc. IEEE Data Compress. Conf.*, Mar. 2010, pp. 376–385.

[11] S. Battiato, G. Gallo, and F. Stanco, "A locally-adaptive zooming algorithm for digital images," *Image Vis. Comput.*, vol. 20, no. 11, pp. 805–812, 2002.

[12] L. Zhang and X. Wu, "Image interpolation via directional filtering and data fusion," *IEEE. Trans. Image Process.*, vol. 15, no. 8, pp. 2226–2238, Aug. 2006.

[13] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE. Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[14] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, and A. K. Katsaggelos, "SoftCuts: A soft edge smoothness prior for color image super-resolution," *IEEE. Trans. Image Process.*, vol. 18, no. 4, pp. 969–981, May 2009.

[15] W. Dong, L. Zhang, G. Shi, and X. Wu, "Nonlocal back-projection for adaptive image enlargement," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2009, pp. 349–352.

[16] W. Freeman, E. Pasztor, and O. Carmichael, "Learning low-level vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.

[17] K. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, Jun. 2010.

[18] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE. Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[19] A. Buades, B. Coll, and J. Morel, "Nonlocal image and movie denoising," *Int. J. Comput. Vis.*, vol. 76, no. 2, pp. 123–139, 2008.

[20] X. Zhu and Z. Ghahramani, "Learning from labeled and unlabeled data with label propagation," School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-CALD-02-107, 2002.

[21] F. Wang, "A general learning framework based on local and global regularizations," *Pattern Recognit.*, vol. 43, no. 9, pp. 3120–3129, Sep. 2010.

[22] C. Cortes and M. Mohri, "On transductive regression," in *Proc. NIPS*, Jun. 2008, pp. 305–312.

[23] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Scholkopf, "Learning with local and global consistency," in *Proc. Adv. NIPS*, 2004, pp. 321–328.

[24] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from examples," *J. Mach. Learn. Res.*, vol. 7, pp. 2399–2434, Dec. 2006.

[25] A. Kapoor, Y. Qi, H. Ahn, and R. W. Picard, " Hyperparameter and kernel learning for graph based semi-supervised classification," in *Proc. Adv. NIPS*, 2005, pp. 627–634.

[26] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[27] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.

[28] D. Levin, "The approximation power of moving least squares," *Math. Comput.*, vol. 67, no. 224, pp. 1517–1531, 1998.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE. Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

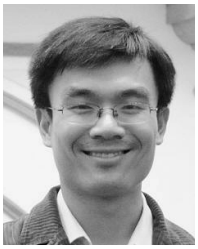
[30] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE. Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[31] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statist. Surveys*, vol. 4, pp. 40–79, Feb. 2010.

[32] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1983.

Xianming Liu is currently an Assistant Professor with the Department of Computer Science, Harbin Institute of Technology (HIT), Harbin, China. He received the B.S., M.S., and Ph.D. degrees in computer science from HIT in 2006, 2008, and 2012, respectively. From 2009 to 2012, he was with the National Engineering L aboratory for Video Technology, Peking University, Beijing, China, as a Research Assistant. In 2011, he was with the Department of Electrical and Computer Engineering, McMaster University, Canada, as a Visiting Student. From 2012 to 2013, he was a Post-Doctoral Fellow with McMaster University. His research interests include image/video coding, image/video processing, and machine learning.

Debin Zhao received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology, China, in 1985, 1988, and 1998, respectively. He is currently a Professor with the Department of Computer Science, Harbin Institute of Technology. He has published over 200 technical articles in refereed journals and conference proceedings in the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.



Jiantao Zhou (M'11) is currently an Assistant Professor with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau. He received the B.Eng. degree from the Department of Electronic Engineering, Dalian University of Technology, Dalian, China, in 2002, the M.Phil. degree from the Department of Radio Engineering, Southeast University, Nanjing, China, in 2005, and the Ph.D. degree from the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, in 2009. He held various research positions at the University of Illinois at Urbana-Champaign, the Hong Kong University of Science and Technology, and McMaster University. His research interests include multimedia security and forensics, and high-fidelity image compression. He was a co-author of a paper that received the Best Paper Award in the IEEE Pacific-Rim Conference on Multimedia in 2007.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Japan, in 1991.

He is a Professor of computer science with Peking University, China. Before joining Peking University, he was a Professor of computer science with the Harbin Institute of Technology from 1991 to 1995, and a Professor with the Institute of Computing Technology of Chinese Academy of Sciences. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served on the editorial board for several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, the *EURASIP Journal of Image Communications*, the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.

Huifang Sun (M'85–SM'93–F'00) received the B.S. degree in electrical engineering from the Harbin Engineering Institute, Harbin, China, and the Ph.D. degree in electrical engineering from the University of Ottawa, Canada. In 1986, he joined Fairleigh Dickinson University, Teaneck, NJ, USA, as an Assistant Professor, and promoted to an Associate Professor. From 1990 to 1995, he was with the David Sarnoff Research Center (Sarnoff Corp), Princeton, NJ, USA, as a member of technical staff and later promoted to Technology Leader of Digital Video Technology where his activities were MPEG video coding, HDTV, and Grand Alliance HDTV development. He joined the Mitsubishi Electric Research Laboratories (MERL), in 1995, as a Senior Principal Technical Staff and was promoted as the Deputy Director in 1997, Vice President, and MERL fellow in 2003 and is currently a MERL fellow. He has co-authored two books and published more than 150 journal and conferences papers. He has 60 granted U.S. patents. He was an Associate Editor of the IEEE TRANSACTION ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and Technical Committee Chair of the Visual Communications and Image Processing of the IEEE Circuits and Systems.