J Sci Comput
DOI 10.1007/s10915-017-0421-z

Weighted Nonlocal Laplacian on Interpolation from Sparse Data

Zuoqiang Shi[1] · **Stanley Osher**[2] · **Wei Zhu**[2]

Received: 15 January 2017 / Revised: 14 March 2017 / Accepted: 18 March 2017

Abstract Inspired by the nonlocal methods in image processing and the point integral method, we introduce a novel weighted nonlocal Laplacian method to compute a continuous interpolation function on a point cloud in high dimensional space. The numerical results in semi-supervised learning and image inpainting show that the weighted nonlocal Laplacian is a reliable and efficient interpolation method. In addition, it is fast and easy to implement.

Keywords Graph Laplacian · Nonlocal methods · Point cloud · Weighted nonlocal Laplacian

Mathematics Subject Classification 65D05 · 65D25 · 41A05

1 Introduction

In this paper, we consider interpolation on a point cloud in high dimensional space. This is a fundamental problem in many data analysis problems and machine learning. Let $P = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\}$ be a set of points in $\mathbb{R}^d$ and $S = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m\}$ be a subset of P. Let u be a function on the point set P and the value of u on $S \subset P$ is given as a function g over S, i.e.

This research was supported by DOE-SC0013838 and NSF DMS-1118971. Zuoqiang Shi was partially supported by NSFC Grants 11371220, 11671005.

✉ Zuoqiang Shi
zqshi@mail.tsinghua.edu.cn

Stanley Osher
sjo@math.ucla.edu

Wei Zhu
weizhu731@math.ucla.edu

1 Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China

2 Department of Mathematics, University of California, Los Angeles, CA 90095, USA

Published online: 09 April 2017

$u(\boldsymbol{s}) = g(\boldsymbol{s}), \ \forall \boldsymbol{s} \in S$. The goal of the interpolation is to find the function u on P with the given values on S.

Since the point set P is unstructured in high dimensional space, traditional interpolation methods do not apply. One model which is widely used in many applications is to minimize the following energy functional,

$$\mathcal{J}(u) = \frac{1}{2} \sum_{\boldsymbol{x},\boldsymbol{y} \in P} w(\boldsymbol{x}, \boldsymbol{y})(u(\boldsymbol{x}) - u(\boldsymbol{y}))^2, \tag{1.1}$$

with the constraint

$$u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x} \in S. \tag{1.2}$$

here $w(\boldsymbol{x}, \boldsymbol{y})$ is a given weight function. One often used weight is the Gaussian weight, $w(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{\sigma^2})$, σ is a parameter, $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$.

It is easy to derive the Euler–Lagrange equation of the above optimization problem, which is given as follows,

$$\begin{cases} \sum_{\boldsymbol{y} \in P} (w(\boldsymbol{x}, \boldsymbol{y}) + w(\boldsymbol{y}, \boldsymbol{x}))(u(\boldsymbol{x}) - u(\boldsymbol{y})) = 0, & \boldsymbol{x} \in P \backslash S, \\ u(\boldsymbol{x}) = g(\boldsymbol{x}), & \boldsymbol{x} \in S. \end{cases} \tag{1.3}$$

If the weight function $w(\boldsymbol{x}, \boldsymbol{y})$ is symmetric, above equation can be simplified further to be

$$\sum_{\boldsymbol{y} \in P} w(\boldsymbol{x}, \boldsymbol{y})(u(\boldsymbol{x}) - u(\boldsymbol{y})) = 0.$$

This is just the well known nonlocal Laplacian which is widely used in nonlocal methods of image processing [1,2,7,8]. It is also called graph Laplacian in graph and machine learning literature [4,19]. In the rest of the paper, we use the abbreviation, GL, to denote this approach.

Recently, it was observed that the solution given by the graph Laplacian is not continuous at the sample points, S, especially when the sample rate, $|S|/|P|$, is low [15]. Consider a simple 1D example. Let P be the union of 5000 randomly sampled points over the interval $(0, 2)$ and we label 6 points in P. Points 0, 1, 2 are in the labeled set S and the other 3 points are selected at random. The solution given by the graph Laplacian, (1.3) is shown in Fig. 1. Clearly, the labeled points are not consistent with the function computed by the graph Laplacian. In other words, the graph Laplacian actually does not interpolate the given values.

It was also shown that the discontinuity is due to the fact that one important boundary term is dropped in evaluating the graph Laplacian. Consider the harmonic extension in the continuous form which is formulated as a Laplace–Beltrami equation with Dirichlet boundary condition on manifold $\mathcal{M}$,

$$\begin{cases} \Delta_{\mathcal{M}} u(\boldsymbol{x}) = 0, & \boldsymbol{x} \in \mathcal{M}, \\ u(\boldsymbol{x}) = g(\boldsymbol{x}), & \boldsymbol{x} \in \partial\mathcal{M}, \end{cases} \tag{1.4}$$

In the point integral method [9,10,13,14], it is observed that the Laplace–Beltrami equation $\Delta_{\mathcal{M}} u(\boldsymbol{x}) = 0$ can be approximated by the following integral equation.

$$\frac{1}{t} \int_{\mathcal{M}} (u(\boldsymbol{x}) - u(\boldsymbol{y})) R_t(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} - 2 \int_{\partial\mathcal{M}} \frac{\partial u(\boldsymbol{y})}{\partial \mathbf{n}} \bar{R}_t(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\tau_{\boldsymbol{y}} = 0, \tag{1.5}$$

where $R_t(\boldsymbol{x}, \boldsymbol{y}) = R\left(-\frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{4t}\right)$, $\bar{R}_t(\boldsymbol{x}, \boldsymbol{y}) = \bar{R}\left(-\frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{4t}\right)$ and $\frac{\mathrm{d}}{\mathrm{d}s}\bar{R}(s) = -R(s)$. If $R(s) = \exp(-s)$, $\bar{R} = R$ and $R_t(\boldsymbol{x}, \boldsymbol{y})$ becomes a Gaussian weight function. $\mathbf{n}$ is the outwards normal of the boundary $\partial\mathcal{M}$.

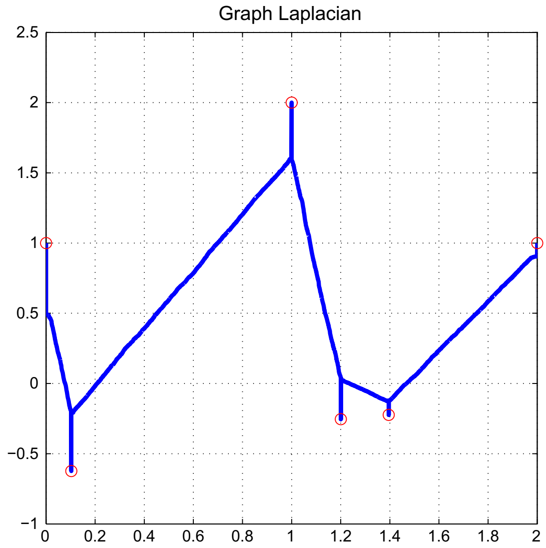

Fig. 1 Solution given by graph Laplacian in 1D examples. *Blue line* interpolation function given by graph Laplacian; *red circles* given values at label set S (Color figure online)

Comparing the above integral equation (1.5) and the equation in the graph Laplacian (1.3), we can clearly see that the boundary term in (1.5) is dropped in the graph Laplacian. However, this boundary term is not small and neglecting it causes trouble. To get an reasonable interpolation, we need to include the boundary term. This is the idea of the point integral method. To deal with the boundary term in (1.5), a Robin boundary condition is used to approximate the Dirichlet boundary condition,

$$u(\boldsymbol{x}) + \mu \frac{\partial u(\boldsymbol{x})}{\partial \mathbf{n}} = g(\boldsymbol{x}), \quad \boldsymbol{x} \in \partial\mathcal{M}, \tag{1.6}$$

where $0 < \mu \ll 1$ is a small parameter.

Substituting above Robin boundary condition to the integral equation (1.5), we get an integral equation to approximate the Dirichlet problem (1.4),

$$\frac{1}{t}\int_{\mathcal{M}} (u(\boldsymbol{x}) - u(\boldsymbol{y}))R_t(\boldsymbol{x}, \boldsymbol{y})\mathrm{d}\boldsymbol{y} - \frac{2}{\mu}\int_{\partial\mathcal{M}} \bar{R}_t(\boldsymbol{x}, \boldsymbol{y})(g(\boldsymbol{y}) - u(\boldsymbol{y}))\mathrm{d}\tau_{\boldsymbol{y}} = 0.$$

The corresponding discrete equations are

$$\sum_{\boldsymbol{y}\in P} R_t(\boldsymbol{x}, \boldsymbol{y})(u(\boldsymbol{x}) - u(\boldsymbol{y})) + \frac{2}{\mu}\sum_{\boldsymbol{y}\in S} \bar{R}_t(\boldsymbol{x}, \boldsymbol{y})(u(\boldsymbol{y}) - g(\boldsymbol{y})) = 0, \quad \boldsymbol{x} \in P, \tag{1.7}$$

The point integral method has been shown to give consistent solutions [9,14,15]. The interpolation algorithm based on the point integral method has been applied to image processing problems and gives promising results [12].

Equation (1.7) is not symmetric, which makes the numerical solver not very efficient. The main contribution of this paper is to propose a novel interpolation algorithm, the weighted nonlocal Laplacian, which preserves the symmetry of the original Laplace operator. The key observation in the weighted nonlocal Laplacian is that we need to modify the energy function (1.1) to add a weight to balance the energy on the labeled and unlabeled sets.

$$\min_u \sum_{\boldsymbol{x}\in P\backslash S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right)+\frac{|P|}{|S|}\sum_{\boldsymbol{x}\in S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right),$$

with the constraint

$$u(\boldsymbol{x})=g(\boldsymbol{x}),\quad \boldsymbol{x}\in S.$$

$|P|,\ |S|$ are the number of points in P and S, respectively. When the sample rate, $|S|/|P|$, is high, the weighted nonlocal Laplacian becomes the classical graph Laplacian. When the sample rate is low, the large weight in weighted nonlocal Laplacian forces the solution close to the given values near the labeled set, such that the inconsistent phenomenon is removed.

We test the weighted nonlocal Laplacian on MNIST dataset and image inpainting problems. The results show that the weighted nonlocal Laplacian gives better results than the graph Laplacian, especially when the sample rate is low. The weighted nonlocal Laplacian provides a reliable and efficient method to find reasonable interpolation on a point cloud in high dimensional space.

The rest of the paper is organized as follows. The weighted nonlocal Laplacian is introduced in Sect. 2. The tests of the weighted nonlocal Laplacian on MNIST and image inpainting are presented in Sects. 3 and 4 respectively. Some conclusions are made in Sect. 5.

2 Weighted Nonlocal Laplacian

First, we split the objective function in (1.1) to two terms, one is over the unlabeled set and the other over the labeled set.

$$\min_u \sum_{\boldsymbol{x}\in P\backslash S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right)+\sum_{\boldsymbol{x}\in S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right),$$

If we substitute the optimal solution into above optimization problem, for instance the solution in the 1D example (Fig. 1), it is easy to check that the summation over the labeled set is actually pretty large due to the discontinuity on the labeled set. However, when the sample rate is low, the summation over the unlabeled set actually overwhelms the summation over the labeled set. So the continuity on the labeled set is sacrificed. One simple idea to assure the continuity on the labeled set is to put a weight ahead of the summation over the labeled set.

$$\min_u \sum_{\boldsymbol{x}\in P\backslash S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right)+\mu\sum_{\boldsymbol{x}\in S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right),$$

This is the basic idea of our approach. Since this method is obtained by modifying the nonlocal Laplacian to add a weight, we call this method the weighted nonlocal Laplacian, WNLL for short.

To balance these two terms, one natural choice of the weight μ is the inverse of the sample rate, $|P|/|S|$. Based on this observation, we get following optimization problem

$$\min_u \sum_{\boldsymbol{x}\in P\backslash S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right)+\frac{|P|}{|S|}\sum_{\boldsymbol{x}\in S}\left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\right), \tag{2.1}$$

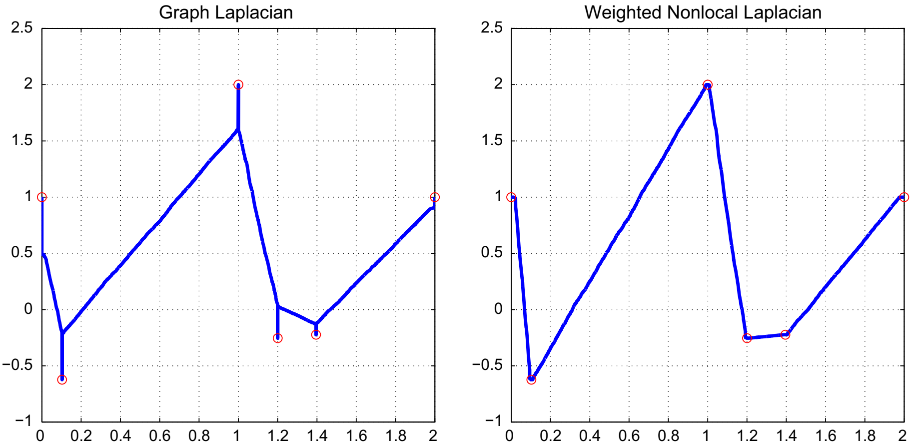

Fig. 2 Solution given by graph Laplacian and weighted nonlocal Laplacian in 1D examples. *Blue line* interpolation function given by graph Laplacian; *red circles* given values at label set S (Color figure online)

with the constraint

$$u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x} \in S.$$

The optimal solution of (2.1) can be obtained by solving a linear system

$$\begin{aligned} &\sum_{\boldsymbol{y}\in P} (w(\boldsymbol{x}, \boldsymbol{y}) + w(\boldsymbol{y}, \boldsymbol{x}))\,(u(\boldsymbol{x}) - u(\boldsymbol{y})) \\ &\quad + \left(\frac{|P|}{|S|} - 1\right) \sum_{\boldsymbol{y}\in S} \; w(\boldsymbol{y}, \boldsymbol{x})(u(\boldsymbol{x}) - u(\boldsymbol{y})) = 0, \quad \boldsymbol{x} \in P\backslash S, \\ &u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x} \in S. \end{aligned} \tag{2.2}$$

This linear system is symmetric and positive definite. Comparing WNLL with the graph Laplacian, in WNLL, we see that a large positive term is added to the diagonal of the coefficient matrix which makes the linear system easier to solve. After solving the above linear system using some iterative method, for instance conjugate gradient, we find out that it converges faster than graph Laplacian. In our tests, the weighted nonlocal Laplacian is about two times faster than graph Laplacian on average.

Figure 2 shows the comparison between WNLL and GL in the 1D example. The result of WNLL perfectly interpolates the given values while GL fails.

We want to remark that there are other choices of the weight μ in WNLL. $|P|/|S|$ works in many applications. Based on our experience, $|P|/|S|$ seems to be the lower bound of μ. In some applications, we may make μ larger to better fit the sample points.

The weighted nonlocal Laplacian is actually closely related with the point integral method. To see the connection, we choose the weight function to be the Gaussian weight both in weighted nonlocal Laplacian and the point integral method, i.e.

$$R_t(\boldsymbol{x}, \boldsymbol{y}) = \bar{R}_t(\boldsymbol{x}, \boldsymbol{y}) = w(\boldsymbol{x}, \boldsymbol{y}) = \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{y}\|^2}{\sigma^2}\right).$$

With the Gaussian weight, comparing (1.7) and (2.2), the only difference between the point integral method and the weighted nonlocal Laplacian is in the summation over the labelled

Algorithm 1 Semi-Supervised Learning

Require: A point set $P = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n\} \subset \mathbb{R}^d$ and a partial labeled set $S = \cup_{i=1}^{l} S_i$.
Ensure: A complete label assignment $L : P \to \{1, 2, \dots, l\}$
for $i = 1 : l$ **do**
Compute ϕ_i on P, with the constraint

$$\phi_i(\boldsymbol{x}) = 1,\ \boldsymbol{x} \in S_i, \quad \phi_i(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in S \backslash S_i,$$

end for
for $(\boldsymbol{x} \in P \backslash S)$ **do**
Label $\boldsymbol{x}$ as following

$$L(\boldsymbol{x}) = k, \quad \text{where} \quad k = \arg\max_{1 \le i \le l} \phi_i(\boldsymbol{x}).$$

end for

set S, $u(\boldsymbol{y})$ is changed to $u(\boldsymbol{x})$. After this changing, the equation becomes symmetric and positive definite which is much easier to solve. In addition to the symmetry, it is easy to show that the weighted nonlocal Laplacian also preserves the maximum principle of the original Laplace–Beltrami operator, which is important in some applications.

3 Semi-Supervised Learning

In this section, we briefly describe the algorithm of semi-supervised learning based on that proposed by Zhu et al. [19]. We plug into the algorithm the aforementioned approach for weighted nonlocal Laplacian, and apply them to the well-known MNIST dataset, and compare their performances.

Assume we are given a point set $P = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_n\} \subset \mathbb{R}^d$, and some labels $\{1, 2, \dots, l\}$. A subset $S \subset P$ is labeled,

$$S = \bigcup_{i=1}^{l} S_i,$$

S_i is the labeled set with label i. In a typical setting, the size of the labeled set S is much smaller than the size of the data set P. The purpose of the semi-supervised learning is to extend the label assignment to the entire P, namely, infer the labels for the unlabeled points. The algorithm is summarized in Algorithm 1.

In Algorithm 1, we use both graph Laplacian and weighted nonlocal Laplacian to compute ϕ_i respectively. In weighted nonlocal Laplacian, ϕ_i is obtained by solving the following linear system

$$\begin{aligned}
&\sum_{\boldsymbol{y} \in P} (w(\boldsymbol{x}, \boldsymbol{y}) + w(\boldsymbol{y}, \boldsymbol{x}))\,(\phi_i(\boldsymbol{x}) - \phi_i(\boldsymbol{y})) \\
&\quad + \left(\frac{|P|}{|S|} - 1\right) \sum_{\boldsymbol{y} \in S} w(\boldsymbol{y}, \boldsymbol{x})(\phi_i(\boldsymbol{x}) - \phi_i(\boldsymbol{y})) = 0, \quad \boldsymbol{x} \in P \backslash S, \\
&\phi_i(\boldsymbol{x}) = 1, \quad \boldsymbol{x} \in S_i, \qquad \phi_i(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in S \backslash S_i.
\end{aligned}$$

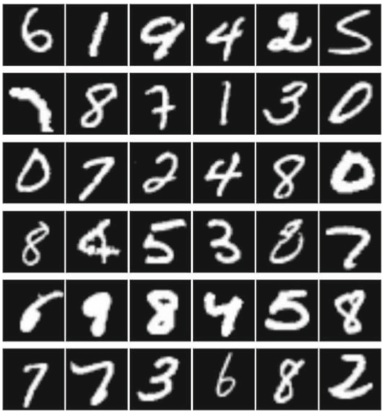

Fig. 3 Some images in MNIST dataset. The whole dataset contains 70,000 28 × 28 *gray scale* digit images

Table 1 Accuracy of weighted nonlocal Laplacian and graph Laplacian in the test of MNIST

100/70,000		70/70,000		50/70,000	
WNLL (%)	GL (%)	WNLL (%)	GL (%)	WNLL (%)	GL (%)
91.83	74.56	84.31	29.67	79.20	33.93
89.20	58.16	87.74	42.29	78.66	21.12
93.13	62.98	83.19	37.02	71.29	24.50
90.43	41.27	91.08	12.57	71.92	29.79
92.27	44.57	84.74	35.15	80.92	37.50

In graph Laplacian, we need to solve

$$\sum_{\mathbf{y}\in P} (w(\mathbf{x}, \mathbf{y}) + w(\mathbf{y}, \mathbf{x})) (\phi_i(\mathbf{x}) - \phi_i(\mathbf{y})) = 0, \quad \mathbf{x} \in P\backslash S,$$

$$\phi_i(\mathbf{x}) = 1, \quad \mathbf{x} \in S_i, \quad \phi_i(\mathbf{x}) = 0, \quad \mathbf{x} \in S\backslash S_i.$$

We test Algorithm 1 on MNIST dataset of handwritten digits [3]. MNIST dataset contains 70,000 28 × 28 gray scale digit images (Fig. 3). We view digits 0 ∼ 9 as ten classes. Each image can be seen as a point in 784-dimensional Euclidean space. The weight function $w(\mathbf{x}, \mathbf{y})$ is constructed as following

$$w(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\sigma(\mathbf{x})^2}\right) \tag{3.1}$$

$\sigma(\mathbf{x})$ is chosen to be the distance between $\mathbf{x}$ and its 20th nearest neighbor, To make the weight matrix sparse, the weight is truncated to the 50 nearest neighbors.

In our test, we label 100, 70 and 50 images respectively. The labeled images are selected at random in 70,000 images. For each case, we do 5 independent tests and the results are shown in Table 1. It is quite clear that WNLL has a better performance than GL. The average accuracy of WNLL is much higher than that of GL. In addition, WNLL is more stable than GL. The fluctuation of GL in different tests is much higher. Due to the inconsistency, in GL, the values in the labeled points are not well spread on to the unlabeled points. On many unlabeled points, the function ϕ_i is actually close to 1/2. This makes the classification sensitive to the distribution of the labeled points.

As for the computational time, in our tests, WNLL takes about half the time of GL on average, 15 versus 29 s (not including the time to construct the weight), with Matlab code in a laptop equipped with CPU Intel i7-4900 2.8 GHz. In WNLL, a positive term is added to the diagonal of the coefficient matrix which makes conjugate gradient converge faster.

As a remark, in this paper, we do not intend to give a state-of-the-art method in semi-supervised learning. We just use this example to test the weighted nonlocal Laplacian.

4 Image Inpainting

In this section, we apply the weighted nonlocal Laplacian to the reconstruction of subsampled images. To apply the weighted nonlocal Laplacian, first, we construct a point cloud from a given image by taking patches. We consider a discrete image $f \in \mathbb{R}^{m\times n}$. For any $(i, j) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$, we define a patch $p_{ij}(f)$ as a 2D piece of size $s_1 \times s_2$ of the original image f, and the pixel (i, j) is the center of the rectangle of size $s_1 \times s_2$. The patch set $\mathcal{P}(f)$ is defined as the collection of all patches:

$$\mathcal{P}(f) = \{p_{ij}(f) : (i, j) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}\} \subset \mathbb{R}^d, \quad d = s_1 \times s_2.$$

To get the patches near the boundary, we extend the images by mirror reflection. For a given image f, the patch set $\mathcal{P}(f)$ gives a point cloud in $\mathbb{R}^d$ with $d = s_1 \times s_2$. We also define a function u on $\mathcal{P}(f)$. At each patch, the value of u is defined to be the intensity of image f at the central pixel of the patch, i.e.

$$u(p_{ij}(f)) = f(i, j),$$

where $f(i, j)$ is the intensity of image f at pixel (i, j).

Now, we subsample the image f in the subsample domain $\Omega \subset \{(i, j) : \ 1 \le i \le m, \ 1 \le j \le n\}$. The problem is to recover the original image f from the subsamples $f|_\Omega$. This problem can be transferred to interpolation of function u in the patch set $\mathcal{P}(f)$ with u is given in $S \subset \mathcal{P}(f)$, $S = \{p_{ij}(f) : \ (i, j) \in \Omega\}$. Notice that the patch set $\mathcal{P}(f)$ is not known, we need to update the patch set iteratively from the recovered image. Summarizing this idea, we get an algorithm to reconstruct the subsampled image which is stated in Algorithm 2.

Algorithm 2 Subsample image restoration

Require: A subsample image $f|_\Omega$.
Ensure: A recovered image u.

Generate initial image u^0.
while not converge **do**

1. Generate patch set $\mathcal{P}(u^n)$ from current image u^n and get corresponding labeled set $S^n \subset \mathcal{P}(u^n)$.
2. Compute the weight function $w^n(\boldsymbol{x}, \boldsymbol{y})$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{P}(u^n)$.
3. Update the image by computing u^{n+1} on $\mathcal{P}(u^n)$, with the constraint

$$u^{n+1}(\boldsymbol{x}) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in S^n.$$

4. $n \leftarrow n + 1$.

end while
$u = u^n$.

There are different methods to compute u^{n+1} on $\mathcal{P}(u^n)$ in Algorithm 2. In this paper, we use weighted nonlocal Laplacian and graph Laplacian to compute u^{n+1}. In the weighted nonlocal Laplacian, we need to solve the following linear system

$$\begin{aligned} &\sum_{\boldsymbol{y}\in\mathcal{P}(u^n)} (w^n(\boldsymbol{x},\boldsymbol{y})+w^n(\boldsymbol{y},\boldsymbol{x}))\left(u^{n+1}(\boldsymbol{x})-u^{n+1}(\boldsymbol{y})\right)\\ &\quad+\left(\frac{mn}{|\Omega|}-1\right)\sum_{\boldsymbol{y}\in S^n} w^n(\boldsymbol{y},\boldsymbol{x})(u^{n+1}(\boldsymbol{x})-u^{n+1}(\boldsymbol{y}))=0,\quad \boldsymbol{x}\in\mathcal{P}(u^n)\backslash S^n,\\ &u^{n+1}(\boldsymbol{x})=f(\boldsymbol{x}),\quad \boldsymbol{x}\in S^n. \end{aligned}$$

While in the graph Laplacian, we solve the other linear system,

$$\begin{aligned} &\sum_{\boldsymbol{y}\in\mathcal{P}(u^n)} (w^n(\boldsymbol{x},\boldsymbol{y})+w^n(\boldsymbol{y},\boldsymbol{x}))\left(u^{n+1}(\boldsymbol{x})-u^{n+1}(\boldsymbol{y})\right)=0,\quad \boldsymbol{x}\in\mathcal{P}(u^n)\backslash S^n,\\ &u^{n+1}(\boldsymbol{x})=f(\boldsymbol{x}),\quad \boldsymbol{x}\in S^n. \end{aligned}$$

Actually, Algorithm 2 with graph Laplacian is just a nonlocal method in image processing[1, 8]. In nonlocal methods, people try to minimize energy functions such as

$$\min_u \sum_{i=1}^{m}\sum_{j=1}^{n}|\nabla_w u(i,j)|^2,$$

with the constraint

$$u(i,j)=f(i,j),\quad (i,j)\in\Omega.$$

$\nabla_w u(i,j)$ is the nonlocal gradient which is defined as

$$\nabla_w u(i,j)=\sqrt{w(i,j;i',j')}(u(i',j')-u(i,j)),\quad 1\le i,i'\le m,\ 1\le j,j'\le n.$$

$w(i,j;i',j')$ is the weight from pixel (i,j) to pixel (i',j'),

$$w(i,j;i',j')=\exp\left(\frac{d(f(i,j),f(i',j'))}{\sigma^2}\right)$$

$d(f(i,j),f(i',j'))$ is the patch distance in image f,

$$d(f(i,j),f(i',j'))=\sum_{k=-h_1}^{h_1}\sum_{l=-h_2}^{h_2}\chi(k,l)|f(i+k,j+l)-f(i'+k,j'+l)|^2$$

χ is often chosen to be 1 or a Gaussian, h_1, h_2 are half sizes of the patch. It is easy to check that the above nonlocal method is the same as solving the graph Laplacian on the patch set. If the weight w is updated iteratively, we get Algorithm 2 with the graph Laplacian. Next, we will show that this method has inconsistent results, and we can use the weighted graph Laplacian to address this issue.

In our calculations below, we take the weight $w(\boldsymbol{x},\boldsymbol{y})$ as following:

$$w(\boldsymbol{x},\boldsymbol{y})=\exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{\sigma(\boldsymbol{x})^2}\right).$$

$\sigma(\boldsymbol{x})$ is chosen to be the distance between $\boldsymbol{x}$ and its 20th nearest neighbor, To make the weight matrix sparse, the weight is truncated to the 50 nearest neighbors. The patch size is 11×11. For each patch, the nearest neighbors are obtained by using an approximate nearest neighbor

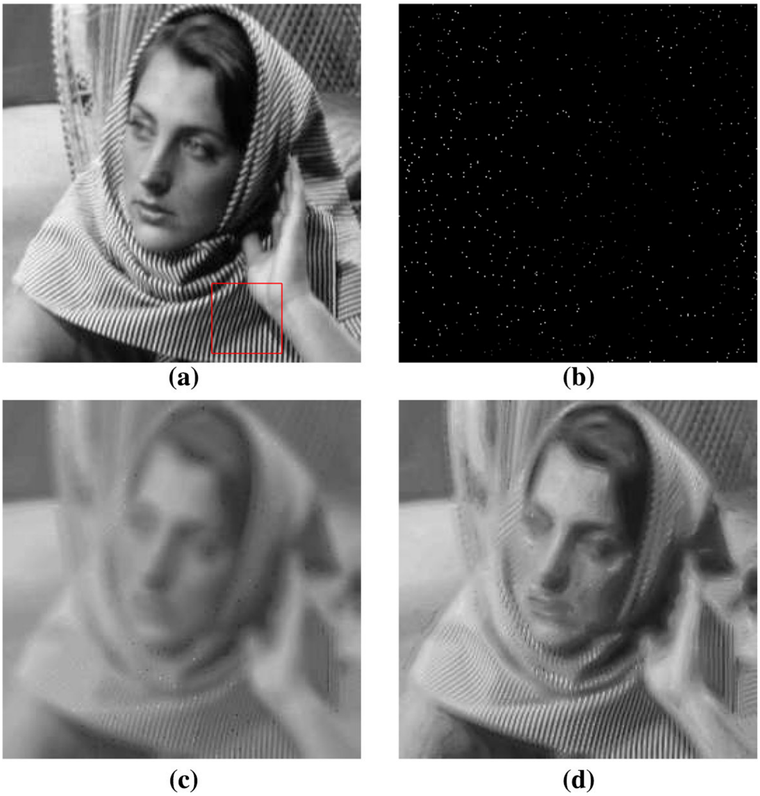

Fig. 4 Test of image restoration on image of Barbara. **a** original image; **b** 1% subsample; **c** result of GL; **d** result of WNLL

(ANN) search algorithm. We use a k-d tree approach as well as an ANN search algorithm to reduce the computational cost. The linear system in weighted nonlocal Laplacian and graph Laplacian is solved by the conjugate gradient method.

PSNR defined as following is used to measure the accuracy of the results

$$\mathrm{PSNR}(f, f^*) = -20 \log_{10}(\|f - f^*\|/255) \tag{4.1}$$

where f^* is the ground truth.

First, we run a simple test to see the performance of weighted nonlocal Laplacian and graph Laplacian. In this test, the patch set is constructed using the original image, Fig. 4a. The original image is subsampled at random, only keeping 1% pixels. Since the patch set is exact, we do not update the patch set and only run WNLL and GL once to get the recovered image.

The results of WNLL and GL are shown in Fig. 4c, d respectively. Obviously, the result of WNLL is much better. To have a closer look at of the recovery, Fig. 5 shows the zoomed in image enclosed by the boxes in Fig. 4a. In Fig. 5d, there are many pixels which are not consistent with their neighbors. Compared with the subsample image 5b, it is easy to check that these pixels are actually the retained pixels. The reason is that in graph Laplacian a non-negligible boundary term is dropped [14,15]. On the contrary, in the result of WNLL, the

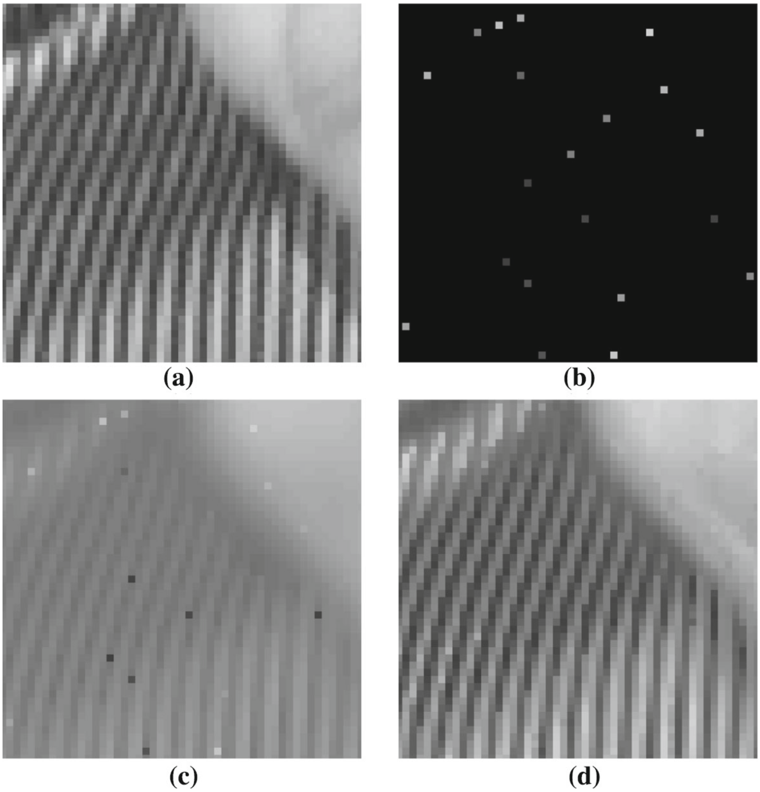

Fig. 5 Zoomed in image in the test of image restoration on image of Barbara. **a** Original image; **b** 1% subsample; **c** result of GL; **d** result of WNLL

inconsistency disappears and the resultant recovery is much better and smoother as shown in Figs. 4d and 5d.

At the end of this section, we apply Algorithm 2 to recover the subsampled image. In this test, we modify the patch to add the local coordinate

$$\bar{p}_{ij}(f) = [p_{ij}(f), \lambda_1 i, \lambda_2 j]$$

with

$$\lambda_1 = \frac{3\|(f|_\Omega)\|_\infty}{m}, \quad \lambda_2 = \frac{3\|(f|_\Omega)\|_\infty}{n}.$$

This semi-local patch could accelerate the iteration and give better reconstruction. The number of iterations in our computation is fixed to be 10.

The initial image is obtained by filling the missing pixels with random numbers which satisfy a Gaussian distribution, where μ_0 is the mean of $f|_\Omega$ and σ_0 is the standard deviation of $f|_\Omega$.

The results are shown in Fig. 6. As we can see, WNLL gives much better results than GL both visually and numerically in PSNR. The results are comparable with those in LDMM [12] while WNLL is much faster. For the image of Barbara (256 × 256), WNLL needs about 1 min and LDMM needs about 20 min in a laptop equipped with CPU Intel i7-4900 2.8

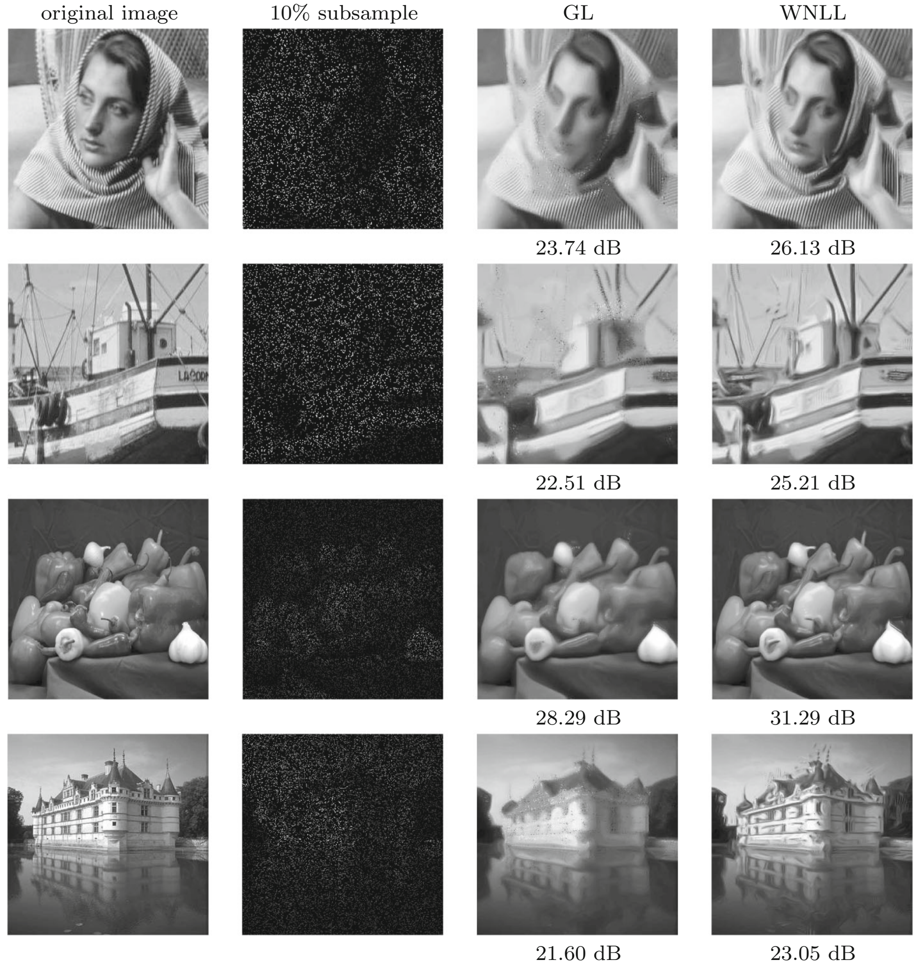

Fig. 6 Results of subsample image restoration

GHz with matlab code. In WNLL and GL, the weight update is the most computationally expensive part in the algorithm by taking more than 90% of the entire computational time. This part is the same in WNLL and GL. So the total time of WNLL and GL are almost same, although the time of solving the linear system is about half in WNLL, 1.7 versus 3.1 s.

5 Conclusion and Future Work

In this paper, we introduce a novel weighted nonlocal Laplacian method. The numerical results show that weighted nonlocal Laplacian provides a reliable and efficient method to find reasonable interpolation on a point cloud in high dimensional space.

On the other hand, it was found that with extremely low sample rate, formulation of the harmonic extension may fail [11,17]. In this case, we are considering minimizing L_∞ norm of the gradient to compute the interpolation on point cloud, i.e. solving the following optimization problem

$$\min_u \left(\max_{\boldsymbol{x}\in P} \left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2 \right)^{1/2} \right),$$

with the constraint

$$u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x}\in S.$$

This approach is closely related with the infinity Laplacian which has been studied a lot in the machine learning community [5,6].

Another interesting problem is the semi-supervised learning studied in Sect. 3. In semi-supervised learning, ideally, the functions, ϕ_i, should be either 0 or 1, so they are piecewise constant. In this sense, minimizing the total variation should give better results [8,16,18]. Based on the weighted nonlocal Laplacian, we should also add a weight to correctly enforce the constraints on the labeled points. This idea implies the following weighted nonlocal TV method,

$$\min_u \sum_{\boldsymbol{x}\in P\backslash S} \left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2 \right)^{1/2} + \frac{|P|}{|S|} \sum_{\boldsymbol{x}\in S} \left(\sum_{\boldsymbol{y}\in P} w(\boldsymbol{x},\boldsymbol{y})(u(\boldsymbol{x})-u(\boldsymbol{y}))^2 \right)^{1/2},$$

with the constraint

$$u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x}\in S.$$

This seems to be a better approach than the weighted nonlocal Laplacian in the semi-supervised learning. We will explore this approach and report its performance in our future work.

References

1. Buades, A., Coll, B., Morel, J.-M.: A review of image denoising algorithms, with a new one. Multiscale Model. Simul. **4**, 490–530 (2005)
2. Buades, A., Coll, B., Morel, J.-M.: Neighborhood filters and PDE's. Numer. Math. **105**, 1–34 (2006)
3. Burges, C.J., LeCun, Y., Cortes, C.: MNIST database
4. Chung, F.R.K.: Spectral Graph Theory. American Mathematical Society, Providence (1997)
5. Elmoataz Abderrahim, L.Z.L.O., Xavier, Desquesnes: Nonlocal infinity Laplacian equation on graphs with applications in image processing and machine learning. Math. Comput. Simul. **102**, 153–163 (2014)
6. Ghoniem, M., Elmoataz, A., Lezoray, O.: Discrete infinity harmonic functions: towards a unified interpolation framework on graphs. In: IEEE International Conference on Image Processing (2011)
7. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. Multiscale Model. Simul. **6**, 595–630 (2007)
8. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**, 1005–1028 (2008)
9. Li, Z., Shi, Z.: A convergent point integral method for isotropic elliptic equations on point cloud. SIAM Multiscale Model. Simul. **14**, 874–905 (2016)
10. Li, Z., Shi, Z., Sun, J.: Point integral method for solving Poisson-type equations on manifolds from point clouds with convergence guarantees. arXiv:1409.2623
11. Nadler, B., Srebro, N., Zhou, X.: Semi-supervised learning with the graph Laplacian: the limit of infinite unlabelled data. In: NIPS (2009)
12. Osher, S., Shi, Z., Zhu, W.: Low dimensional manifold model for image processing. Technical report, CAM report 16-04, UCLA (2016)

13. Shi, Z., Sun, J.: Convergence of the point integral method for Poisson equation on point cloud. arXiv:1403.2141
14. Shi, Z., Sun, J.: Convergence of the point integral method for the Poisson equation with Dirichlet boundary on point cloud. arXiv:1312.4424
15. Shi, Z., Sun, J., Tian, M.: Harmonic extension on point cloud. arXiv:1509.06458
16. Yin, K., Tai, X.-C., Osher, S.: An effective region force for some variational models for learning and clustering. Technical report, CAM report 16-18, UCLA (2016)
17. Zhou, X., Belkin, M.: Semi-supervised learning by higher order regularization. In: NIPS (2011)
18. Zhu, W., Chayes, V., Tiard, A., Sanchez, S., Dahlberg, D., Kuang, D., Bertozzi, A., Osher, S., Zosso, D.: Nonlocal total variation with primal dual algorithm and stable simplex clustering in unspervised hyperspectral imagery analysis. Technical report, CAM report 15-44, UCLA (2015)
19. Zhu, X., Ghahramani, Z., Lafferty, J.D.: Semi-supervised learning using Gaussian fields and harmonic functions. In: Machine Learning, Proceedings of the Twentieth International Conference ICML 2003, 21–24 Aug 2003, Washington, DC, USA, pp. 912–919 (2003)