A Weighted Fidelity and Regularization-Based Method for Mixed or Unknown Noise Removal From Images on Graphs

Cong Wang, ZiYue Yan, Witold Pedrycz, *Fellow, IEEE*, MengChu Zhou, *Fellow, IEEE*, and ZhiWu Li, *Fellow, IEEE*

***Abstract*— Image denoising technologies in a Euclidean domain have achieved good results and are becoming mature. However, in recent years, many real-world applications encountered in computer vision and geometric modeling involve image data defined in irregular domains modeled by huge graphs, which results in the problem on how to solve image denoising problems defined on graphs. In this paper, we propose a novel model for removing mixed or unknown noise in images on graphs. The objective is to minimize the sum of a weighted fidelity term and a sparse regularization term that additionally utilizes wavelet frame transform on graphs to retain feature details of images defined on graphs. Specifically, the weighted fidelity term with ℓ_1-norm and ℓ_2-norm is designed based on a analysis of the distribution of mixed noise. The augmented Lagrangian and accelerated proximal gradient methods are employed to achieve the optimal solution to the problem. Finally, some supporting numerical results and comparative analyses with other denoising algorithms are provided. It is noted that we investigate image denoising with unknown noise or a wide range of mixed noise, especially the mixture of Poisson, Gaussian, and impulse noise. Experimental results reported for synthetic and real images on graphs demonstrate that the proposed method is effective and efficient, and exhibits better performance for the removal of mixed or unknown noise in images on graphs than other denoising algorithms in the literature. The method can effectively remove mixed or unknown noise and retain feature details of images on graphs. It delivers a new avenue for denoising images in irregular domains.**

***Index Terms*— Tight wavelet frame, variational model, mixed or unknown noise, image on graph, image denoising.**

Manuscript received June 8, 2019; revised November 30, 2019 and January 7, 2020; accepted January 20, 2020. Date of publication February 25, 2020; date of current version March 26, 2020. This work was supported in part by the Doctoral Students' Short Term Study Abroad Scholarship Fund of Xidian University, in part by the National Natural Science Foundation of China under Grant 61472295 and Grant 61672400, in part by the Recruitment Program of Global Experts, and in part by the Science and Technology Development Fund, MSAR, under Grant 0012/2019/A3. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chia-Kai Liang. *(Corresponding authors: MengChu Zhou; ZhiWu Li.)*

Cong Wang is with the School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China (e-mail: wangc0705@stu.xidian.edu.cn).

ZiYue Yan is with the Department of Computer Science, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90007 USA (e-mail: ziyue@usc.edu).

Witold Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada, also with the School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China, and also with the Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: wpedrycz@ualberta.ca).

MengChu Zhou is with the Institute of Systems Engineering, Macau University of Science and Technology, Macau 999078, China, and also with the Helen and John C. Hartmann Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: zhou@njit.edu).

ZhiWu Li is with the School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China, and also with the Institute of Systems Engineering, Macau University of Science and Technology, Macau 999078, China (e-mail: zhwli@xidian.edu.cn).

Digital Object Identifier 10.1109/TIP.2020.2969076

I. Introduction

IMAGES are easily corrupted by noise during processes of image acquisition and transmission. Due to their inferior characteristics, noisy images cannot be directly used in image analysis or processing, such as image compression, image segmentation and image understanding. Image denoising, as a fundamental image operation, aims to obtain high-quality images by removing noise while retaining feature details in images as much as possible before proceeding with further image analysis or processing. Generally speaking, the noise encountered in images can be classified into two categories: additive and multiplicative noise. The main forms of additive noise are additive white Gaussian noise and impulse noise [1]–[3]. Gaussian noise is mainly introduced by the thermal motion in camera [4]. Impulse noise is caused by bit errors appearing in communications [5], [6]. Generally, it exhibits two types, i.e., salt and pepper impulse noise and random-valued impulse noise. The common form of multiplicative noise is Poisson noise [7]. It is mostly produced by low photons, like fluorescence microscopy and emission tomography [8].

Over the past two decades, in order to suppress noise, a large number of image denoising techniques have been developed [8]–[17]. Such techniques are generally classified into four categories, i.e., spatial domain-based methods [13], [18]–[20], transform domain-based methods [10], [21], [22], dictionary learning-based methods [15], [16], [23]–[27] and deep convolutional neural network (CNN)-based methods [28]–[32].

Spatial domain-based methods primarily include local and non-local filters that adequately manipulate the similarities among image pixels or patches [13], [18]–[20], [33]. Over the past three decades, a series of local filters such as Gaussian filter [18], anisotropic filtering [13], and bilateral filter [4], [19] were designed. Generally, such local filters have low time complexity, but cannot work well when noise levels increase.

1057-7149 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

In contrast, nonlocal filters can take advantage of the self-similarity of natural images in a nonlocal manner. One of the well-known nonlocal filtering algorithms is a nonlocal means algorithm [20] that weights neighboring patches to obtain an anchor (denoised) patch. The idea of nonlocal means is also applied to transform domain-based [17] and dictionary learning-based algorithms [34]. Although nonlocal filtering algorithms are superior to local filtering algorithms in the presence of high-level noise, they bring over-smoothing.

Transform domain-based methods [10], [21], [22] typically represent images in a fixed basis before removing noise in images, which exploit similarities of transformed coefficients. A large number of orthonormal basises, like wavelets [10], [35], curvelets [22], and contourlets [21], have been proposed. One of the most popular transform methods concerns wavelet analysis. Wavelets have been successfully applied in image processing for over three decades [10], thanks to their ability to simultaneously localize image data in both time and frequency domains. Tight wavelet frames differing from general wavelet methods can provide redundant representations of images [12]. This redundancy enables them to be flexibly applied to various areas such as image denoising [12], [36], image restoration [9], [11] and surface reconstruction [37], [38].

Dictionary learning-based methods [15], [16], [24]–[27] also represent images in transform domains. However, the difference from the transform domain-based methods is that they employ a sparse representation on a redundant dictionary [15], [16]. There exist some prevailing classical algorithms such as K-clustering with singular value decomposition [24], learned simultaneous sparse coding [25], [39], and clustering-based sparse representation [27]. Even though most of them can achieve good performance in image denoising, they are computationally expensive.

Deep CNN-based methods have achieved considerable progress on some low level vision tasks [28]–[32], e.g., rain streak removal [31], Gaussian denoising [32] and non-blind deconvolution [30]. However, these CNN-based methods either do not take partially known degradations into consideration, or simply address this issue by learning a direct mapping from degraded images to ground-truth ones. Additionally, their visible limitation is that they are highly dependent on a training set of degraded and ground-truth image pairs and the calculation costs are large.

Recently, by combining a transform domain and sparse representation, a variety of achievements for image denoising have been made [8], [9], [11], [12], [14], [17], [40]–[43]. One of the most typical techniques is a wavelet frame-based sparse variational model. For instance, Dong et al. [11] address a blind inpainting model based on variation and tight wavelet frames in a Euclidean domain, which can simultaneously identify and recover damaged pixels of noisy images. Gong et al. [14] demonstrate that a sparse variational model consists of a universal fidelity term and a wavelet frame-based regularization term is able to remove mixed or unknown noise in images in a Euclidean domain, and apply a novel algorithm to solve this model. Li et al. [8] propose a variational model based on a weighted fidelity and tight wavelet frames for Poisson noise removal from images in a Euclidean domain. The well-known alternating direction method of multipliers is applied to solve this model. Moreover, the model is also extended to the removal of mixed Poisson-Gaussian noise. Recently, Shen et al. [17] present a wavelet frame-based iterative algorithm for the removal of mixed Gaussian and impulse noise, which takes advantage of directional complex tight framelets and sparse minimization. Through the literature review, we conclude that extensive studies have been made for image denoising tasks, especially mixed noise removal. However, they mainly consider mixed noise removal in a Euclidean domain. Yet the obtained results are far from being adequate. There exist no works considering mixed noise removal on graphs. As a result, we identify three issues calling for further improvement:

1) Many practical problems in computer vision and geometric modeling involve image data defined in more topologically complicated domains rather than a Euclidean domain. As graphs can be flexibly represented in either a Euclidean domain or irregular domains [44], researchers have recently started to analyze and manipulate image data defined on graphs [37], [44]–[47]. In real-world applications, graphs can be modeled in general as a certain discretization or random sample from some Riemannian manifolds [48]. Images on graphs can be considered as scalar functions defined on vertices of graphs. Due to the complexity of topological structures of irregular domains, it is challenging to design efficient computing methods for solving image processing problems defined on graphs. Recently, researchers try to solve image processing problems defined on graphs, and made some encouraging achievements [37], [44], [46], [47], [49]. For instance, Hammond et al. [44] propose a general wavelet transform on graphs and make some potential applications. Dong [37] presents a fast discrete tight wavelet frame transform in a graph domain, which is applied to many practical problems such as graph data denoising and semi-supervised clustering. Wang and Yang [46] introduce a variational method based on tight wavelet frames to remove Poisson noise in images on graphs. More recently, Wang et al. [50] propose a wavelet frame-based fuzzy clustering algorithm and apply it to image segmentation on graphs.
2) In real-world applications, images on graphs are often corrupted by different types of mixed or unknown noise, such as mixed Poisson-Gaussian noise and mixed Gaussian and impulse noise, rather than single noise. Many existing techniques for mixed noise removal appearing in a Euclidean domain are detection-based methods. To overcome their shortcomings, some denoising techniques, e.g. [15], tentatively encode noise-corrupted patch while depending on prior knowledge of mixed noise. Due to the complicated distribution of mixed or unknown noise, image denoising techniques for their removal are not well exploited [51]. Moreover, most of these existing methods cannot be directly extended to mixed or unknown noise removal on graphs due to the nonlinear nature of graphs.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

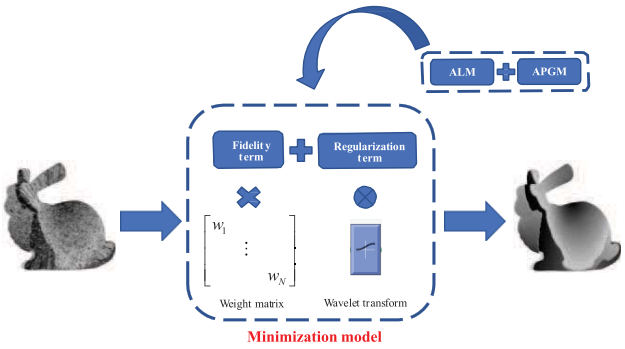


Fig. 1. The framework of the proposed method.

3) Specifically, this work investigates image denoising with unknown noise and the mixture of Poisson, Gaussian, and impulse noise. To the best of our knowledge, there exists only a single attempt [14] appearing in a Euclidean domain to concern the removal of such a wide range of noise.

The goal of this work is to develop a wavelet frame-based variational model that can be applied to suppress mixed or unknown noise in images on graphs. The aim of variational models is to minimize the sum of the fidelity and regularization terms. The fidelity term can be designed based on different noise distribution characteristics. For example, an ℓ_2-norm fidelity term has been successfully used for Gaussian noise removal, and an ℓ_1-norm fidelity term has been well proved to be effective for removing both impulse noise and mixed Gaussian and impulse noise [3]. Recently, Gong et al. [14] first demonstrate that an ℓ_1-norm fidelity term is suitable for generic noise removal from images in a Euclidean domain when combined with an ℓ_2-norm fidelity term. Inspired by [14], we develop a weighted fidelity term to remove a wide range of mixed or unknown noise in images on graphs.

This paper makes three main contributions as follows:

1) We extend mixed or unknown noise removal from a Euclidean domain to graph domain. Due to the nonlinear nature of graphs and the corresponding algorithms [37], [44], it is challenging to address image denoising problems defined on graphs.
2) We illustrate a weighted variational model for removing mixed or unknown noise in images on graphs. The objective is to minimize the sum of a weighted fidelity term and an ℓ_1-norm regularization term that additionally utilizes the wavelet frame transform on graphs for retaining feature details of images defined on graphs. Specifically, the weighted fidelity term with ℓ_1-norm and ℓ_2-norm is designed based on analysis of the distribution of mixed noise. The augmented Lagrangian method (ALM) [52] and accelerated proximal gradient method (APGM) [12] are employed to deal with this model. The framework of the proposed method is portrayed in Fig. 1.
3) We expand the range of noise removal in image denoising problems, especially, image denoising problems defined on graphs. We apply the proposed method to remove unknown noise and the mixture of Poisson, Gaussian, and impulse noise in images on graphs. Although such noise frequently occurs in the process of image acquisition, little work has been devoted to the removal of such noise. This work offers a new direction for suppressing mixed or unknown noise in images in irregular domains.

This paper is structured as follows. Section II briefly introduces preliminaries involved in the proposed model and algorithm. The proposed method is described in Section III. In Section IV, supporting experiments are provided. Section V draws conclusions. The convergence results of the proposed algorithm are illustrated in the Appendix.

II. Preliminaries

To understand and analyze images and data on graphs, we first review the spectral graph theory, especially the graph Laplacian, which is widely used to reveal the geometric properties of the graph. Furthermore, we introduce the construction of a tight wavelet frame transform on graphs by the assistance of the graph Laplacian. Full details can be found in [37].

A. Spectral Graph Theory

Let $G = (V, E, \omega)$ be a weighted graph composed of a vertex set $V = \{v_1, \cdots, v_N\}$ representing a collection of (x, y, z)-coordinates of N vertices, an edge set $E \subseteq V \times V$, and a weight function $\omega : E \mapsto \mathbb{R}^+$ that is generally expressed as

$$\omega_{ij} = e^{-\|v_i - v_j\|^2/\rho}, \quad \rho > 0, \tag{1}$$

where v_i and v_j denote two arbitrary vertices in V and $\|\cdot\|$ denotes their Euclidean distance. If vertex v_i is not connected to vertex v_j, then $\omega_{ij} = 0$. Since graph G is undirected, we have $\omega_{ij} = \omega_{ji}$.

The adjacency matrix A of graph G is denoted as $A = (\omega_{ij})_{i,j=1,\cdots,N}$. Let $d_i = \sum_{j=1}^{N} a_{ij}$ be the degree of each vertex v_i. The degree matrix D is defined as $D = \text{diag}\{d_1, \cdots, d_N\}$. Thus, the Laplacian martix of graph G can be formulated as

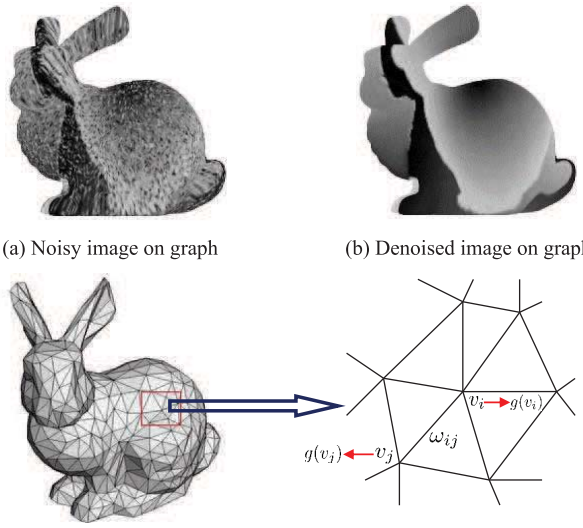
$$L = D - A.$$

Let $\{(\lambda_i, u_i)\}_{i=1}^N$ be the set of N pairs of its eigenvalues and eigenvectors of L. Since L is symmetric and positive semi-definite, we have $\lambda_N \geq \cdots \geq \lambda_2 > \lambda_1 = 0$.

B. Wavelet Frame Transform on Graphs

Based on the eigenvalue decomposition of graph Laplacian matrix L, we describe the construct of the tight wavelet frame transform on graph G. Let us define an image function $g : V \mapsto \mathbb{R}$ on graph G. Then $(g(v_1), \cdots, g(v_N))$ can be viewed as an N-dimensional vector in $\mathbb{R}^N$, where $g(v_i)$ defines a coordinate. Moreover, image data on graphs are represented by grey levels with the values in range of [0,1]. To make the image function on graph G better understood, we intuitively show an example, refer to Fig. 2.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.



(a) Noisy image on graph (b) Denoised image on graph

(c) Topological structure of graph

Fig. 2. The illustration of image functions on graph G.

For the sake of concise exposition, we rewrite $g(v_i)$ as $g[i]$. The Fourier transform of $g[i] \in g$ is represented as

$$\widehat{g}[i] = \sum_{j=1}^{N} g[j]u_i[j], \quad i = 1, \cdots, N.$$

For $p = 0, 1, \cdots, P$, $\{a_p\}$ denotes a set of masks and $\widehat{a}_p$ denotes the Fourier series of a_p. As a_p is finitely supported, $\widehat{a}_p$ can be described by a trigonometric polynomial. Let $\widehat{a}_p^*$ be the complex conjugate of $\widehat{a}_p$. The q-level tight wavelet frame decomposition $\mathcal{W}$ is defined as

$$\mathcal{W}g := \{\mathcal{W}_{p,q}g : p = 0, 1, \cdots, P, q = 1, 2, \cdots, Q\},$$

where for $q = 1$,

$$\widehat{\mathcal{W}}_{p,q}g[i] := \widehat{a}_p^*(2^{-M}\lambda_i)\widehat{g}[i],$$

and for $q \in \{2, 3, \cdots, Q\}$,

$$\widehat{\mathcal{W}}_{p,q}g[i] := \widehat{a}_p^*(2^{-M+q-1}\lambda_i)\widehat{a}_0^*(2^{-M+q-2}\lambda_i) \cdots \widehat{a}_0^*(2^{-M}\lambda_i)\widehat{g}[i].$$

Here, p denotes the band of a transform, q denotes the level of the transform and M denotes the dilation scale, which is selected such that $2^{M-1}\pi < \lambda_N \leq 2^M\pi$.

For $p = 0, 1, \cdots, P$, and $q = 1, 2, \cdots, Q$, let $\boldsymbol{\chi} = \mathcal{W}g = \{\chi_{p,q}\}$ with $\chi_{p,q} = \mathcal{W}_{p,q}g$. For $q = Q, Q-1, \cdots, 1$, the tight wavelet frame reconstruction $\mathcal{W}^T\boldsymbol{\chi}$ is defined by the following iterative procedure

$$\widehat{\chi}_{0,q-1}[i] = \sum_{p=0}^{P} \widehat{a}_p(2^{-M+q-1}\lambda_i)\widehat{\chi}_{p,q}[i],$$

where $\chi_{0,0} = \mathcal{W}^T\boldsymbol{\chi}$ is the reconstructed data from $\boldsymbol{\chi}$. According to Theorem 3.1 in [37], we obtain $\mathcal{W}^T\mathcal{W} = g$, which means that $\mathcal{W}^T\mathcal{W}$ is an identity operation.

As mentioned above, a wavelet frame transform involves the eigenvalue decomposition of the Laplacian matrix of a graph. In practical applications, it is very challenging to compute all eigenvalues of the Laplacian matrix of a large graph. To overcome this problem, Chebyshev polynomials [53] can be employed to approximate the masks of a tight wavelet frame system. For $p = 0, 1, \cdots, P$, mask a_p is a finitely supported sequence. A low-degree Chebyshev polynomial can be used to accurately approximate $\widehat{a}_p$ described by a trigonometric polynomial. In this case, we do not have to calculate all eigenvalues of the Laplacian matrix. More details can be found in [37].

According to [54], the piecewise linear B-spline tight frame system can provide a simple explicit expression to give redundant representations of images, which offers more adaptability to noise. Thus, we choose this system. Formally speaking, for $\xi \in [0, \pi]$, the system is formulated as

$$\widehat{a}_0(\xi) = \cos^2(\xi/2), \widehat{a}_1(\xi) = \frac{1}{\sqrt{2}}\sin(\xi), \widehat{a}_2(\xi) = \sin^2(\xi/2).$$

III. Methodology

A. Variational Models for Image Denoising

Let g be a noisy image function defined on a weighted graph G. Solving image denoising problems defined on graphs, amounts to recovering an unknown noisy-free image f from a noisy image g that is formulated as

$$g = f + \epsilon,$$

where ϵ is random noise perturbation. For mixed or unknown noise removal, it is generally difficult to exhibit the specific formulation of noise models. Therefore, the usual solution is to assume in advance mixed noise models where the most common types of mixed noise are mixed Gaussian and impulse noise and mixed Poisson-Gaussian noise. Beyond previous studies, we focus on the removal of a wide range of mixed or unknown noise, especially the mixture of Poisson, Gaussian, and impulse noise. In the sequel, we design the denoising solutions on the basis of the model of the mixture of Poisson, Gaussian, and impulse noise, then extend them to generally mixed or unknown noise removal. Thus, the noisy image g defined in a domain $\Omega = \{1, 2, \cdots, N\}$ can be remodeled as

$$g[i] = \begin{cases} \widetilde{g}[i] + \epsilon_1[i] & i \in \Omega_1 \\ \epsilon_2[i] & i \in \Omega_2 := \Omega \backslash \Omega_1, \end{cases}$$

where $\widetilde{g}$ obeys Poisson distribution, i.e., $\widetilde{g} \sim \mathbf{P}(f)$, which means that f is corrupted by Poisson noise. Moreover, ϵ_1 is the additive zero-mean white Gaussian noise, and ϵ_2 is the impulse noise. The subset Ω_2 of Ω denotes the region where the information of $\widetilde{g}$ is missing. It is assumed to be unknown with each element being drawn from the whole region Ω by Bernoulli trial with a given probability $0 < r < 1$.

To remove noise from g, one of the most popular approaches is to minimize a variational model consisting of two terms, i.e., fidelity and regularization:

$$\min_{f \in \mathbb{R}} F(f) + \mu R(f),$$

where μ is a positive number that aims to balance the fidelity term $F(f)$ and the regularization term $R(f)$.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

The choice of $F(f)$ is generally determined according to a specified noise distribution. For instance, we can adopt an ℓ_2-norm fidelity term to remove Gaussian noise, which is formulated as

$$F(f) = \|f - g\|_{\ell_2}^2,$$

where $\|\cdot\|_{\ell_2}$ denotes the usual ℓ_2 vector norm, e.g. $\|f\|_{\ell_2} = \sqrt{f^2[1] + f^2[2] + \cdots + f^2[N]}$. For impulse noise removal, an ℓ_1-norm fidelity term is generally chosen:

$$F(f) = \|f - g\|_{\ell_1},$$

where $\|\cdot\|_{\ell_1}$ denotes the usual ℓ_1 vector norm, e.g. $\|f\|_{\ell_1} = |f[1]| + |f[2]| + \cdots + |f[N]|$. When we remove Poisson noise in images, the Csiszár's I-divergence [7] of f from g can be taken as a fidelity term, i.e.,

$$F(f) = \sum_i (f[i] - g[i] \log f[i]).$$

Since images tend to be corrupted by different types of mixed or unknown noise rather than single noise, the fidelity term with respect to given noise may not result in desired performance for mixed or unknown noise removal. One of the main tasks of this work is to find a universal fidelity term for removing mixed or unknown noise in images on graphs.

A regularization term is designed according to a priori assumption on f. A common assumption is the sparsity of f in a transform domain such as discrete gradient [55], Fourier transform, and wavelet [10], [12]. Since the ℓ_1-norm regularization term gives preference to a solution f and is piecewise smooth, we consider penalizing the ℓ_1 norm of representation coefficients in a transform domain, i.e.,

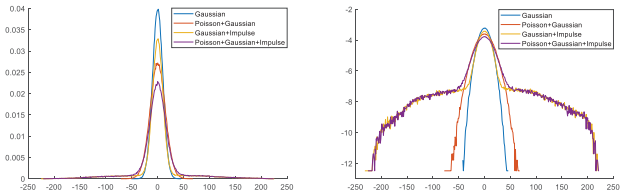
$$R(f) = \|\mathcal{D}f\|_{\ell_1},$$

where $\mathcal{D}$ is a linear transform operator.

B. Analysis of Mixed Noise Distribution

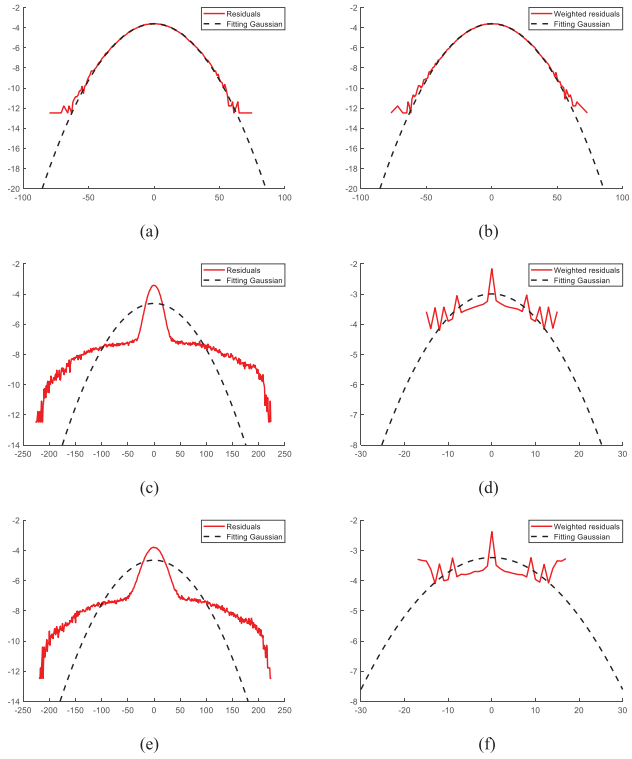
For single noise removal, the fidelity terms mentioned above can lead to a maximum a posteriori (MAP) solution for specified noise removal. However, they are failure to mixed noise removal since the mixed noise distribution is far from any single noise distribution. Here, we consider a case to illustrate the distributions of multiple mixed noise. We impose Poisson, Gaussian (standard deviation $= 10$) and random-valued impulse noise (probability $= 20\%$) on image 'Lena' (see Fig. 6) with size 512×512, respectively.

Fig. 3(a) shows the distributions of residuals $f[i] - g[i]$ for Gaussian, mixed Poisson-Gaussian, mixed Gaussian and impulse, and the mixture of Poisson, Gaussian and impulse noise, respectively. To obtain a better view of the heavy tails, we redraw these distributions in the logarithmic domain, refer to Fig. 3(b). Clearly, Poisson noise tends to give rise to a Gaussian-like distribution. Compared with the Gaussian distribution, we can find that impulse noise leads to a heavy tail of the mixed noise distribution. Therefore, neither ℓ_1-norm nor ℓ_2-norm can precisely characterize residuals $f[i] - g[i]$ in the sense of the MAP estimation.



(a) (b)

Fig. 3. Distributions of Gaussian and mixed noise in different domains. (a) linear domain; (b) logarithmic domain.



(a) (b) (c) (d) (e) (f)

Fig. 4. Distributions of residuals $f[i] - g[i]$, weighted residuals $w_i(f[i] - g[i])$ and the fitting Gaussian in the logarithmic domain, respectively. (a) and (b) for mixed Poisson-Gaussian noise; (c) and (d) for mixed Gaussian and impulse noise; (e) and (f) for the mixture of Poisson, Gaussian and impulse noise.

C. Formulation of the Proposed Model

For weakening the impact of the heavy tail of the mixed noise distribution, some weighted encoding algorithms have been proposed [15], [56]. Inspired by such work in [15], [56], we can assign to each residual $f[i] - g[i]$ a proper weight, which results in a weighted residual, i.e.,

$$w_i(f[i] - g[i]),$$

where w_i is a weight assigned to location i. Thus, target residual $f[i] - g[i]$ is divided into two parts, i.e., $w_i(f[i] - g[i])$ and $(1 - w_i)(f[i] - g[i])$. The former almost follows a Gaussian distribution. Moreover, the latter is mostly caused by impulse noise. In order to illustrate the effect of weighting in presence of different types of mixed noise, we show an example in Fig. 4.

In Fig. 4, we follow the same noise-related settings as in Fig. 3. The left column of Fig. 4 shows the distribution of residuals $f[i] - g[i]$ and the fitting Gaussian function based on the variance of $f[i] - g[i]$. The right column of

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

Fig. 4 shows the distribution of weighted residuals $w_i(f[i]-g[i])$ and the fitting Gaussian function based on the variance of $w_i(f[i]-g[i])$ (specific weight settings are discussed in Section IV-A). Obviously, weighted residuals $w_i(f[i]-g[i])$ tend to follow a Gaussian distribution and the rest residuals $(1-w_i)(f[i]-g[i])$ are mostly caused by impulse noise, thus implying that ℓ_2-norm can be used to model the weighted residuals $w_i(f[i]-g[i])$ and ℓ_1-norm can be used to model the rest $(1-w_i)(f[i]-g[i])$ for a MAP-like solution of f. For instance, for mixed Gaussian and impulse noise removal, the residuals are composed of two parts. The first part mostly follows Gaussian distribution at the locations contaminated by Gaussian noise. Thus, such residuals are closely weighted with 1 in the ℓ_2-norm fidelity term, and 0 in the ℓ_1-norm fidelity term. The second part can be obtained at the locations corrupted by impulse noise. Those residuals should be closely weighted with 0 in the ℓ_2-norm fidelity term, and 1 in the ℓ_1-norm fidelity term.

According to the above analysis, we propose a novel fidelity term, namely weighted $\ell_1+\ell_2$ fidelity, to remove unknown noise and a wide range of mixed noise in images on graphs:

$$F(f)=\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2+\|(1-W)(f-g)\|_{\ell_1},$$

where the weight matrix W is defined as the diagonal matrix with the degrees $w_1,\cdots,w_N$ located on the diagonal. The element w_i that is assigned to pixel i is automatically determined, which is inversely proportional to the residual $f[i]-g[i]$. Thus, we can choose the following expression:

$$w_i=e^{-\zeta(f[i]-g[i])^2}, \tag{2}$$

where ζ is a positive parameter used to control the decreasing rate of w_i.

As to the regularization term, we adopt a popular transform method, i.e., tight wavelet frame transform on graphs [37]. This is because tight wavelet frames can provide redundant representations of images [37]. The selected regularization term is expressed as:

$$R(f)=\|\mathcal{W}f\|_{\ell_1}.$$

By combining the weighted $\ell_1+\ell_2$ fidelity term and regularization term with the tight wavelet frame transform, we present the following model for mixed or unknown noise removal on graphs:

$$\min_{f\in\mathbb{R}}\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2+\|(1-W)(f-g)\|_{\ell_1}+\mu\|\mathcal{W}f\|_{\ell_1}, \tag{3}$$

where φ and μ are positive parameters, W is a diagonal weight matrix and $\mathcal{W}$ is the tight wavelet frame transform on graphs. Note that it is not trivial to extend image denoising models from a Euclidean domain to a graph domain due to the nonlinear nature of graphs and the corresponding algorithms [37], [44].

D. Minimization Algorithm

Formally speaking, the minimization of (3) involves two unknowns, i.e., W and f. Thus, we can design a two-step iterative algorithm to solve (3), which fixes W first to solve for image f, then fixes image f to update W. Note that the main task in each iteration is to solve the minimization problem in terms of f while fixing W. In the following, we carefully derive the procedure of the two-step iterative algorithm.

Algorithm 1 ALM

Input: Noisy image g and tolerance ε.
Output: Denoised image f^{k+1}.

```
Initialize y^0 as a zero vector, σ^0 = 1, and W^0 as an identity matrix
k ← 0
repeat
    (f^{k+1}, Z^{k+1}) = arg min_{f,Z} L_{σ^k}(f, Z; y^k)
    y^{k+1} = y^k + σ^k(C − Bf^{k+1} − Z^{k+1})
    σ^{k+1} = σ^k + 1
    Update weight matrix W^{k+1} using (2)
    k ← k + 1
until ‖(y^k − y^{k+1})/σ^k‖_{ℓ2} < ε
return Denoised image f^{k+1}
```

Assume that W is given. The optimization model (3) cannot be solved in a straightforward way because it is a least square problem with two ℓ_1-norm terms, i.e., an ℓ_1-norm fidelity term and an ℓ_1-norm regularization term. To solve this problem, we can incorporate two ℓ_1-norm terms in a matrix form. Then, ALM can be employed to solve model (3). The detailed procedures are next described.

First, we reformulate model (3) in a matrix form as follows:

$$\min_{f\in\mathbb{R}}\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2+\theta^T|Bf-C|, \tag{4}$$

where $B=[\,1-W\ \mathcal{W}\,]^T$, $C=[\,(1-W)g\ \hat{0}\,]^T$ and $\theta=[\,\tilde{e}\ \varphi\hat{e}\,]^T$. $\hat{0}$ is a vector of zeros. $\tilde{e}$ and $\hat{e}$ are vectors of ones. Here, $|\cdot|$ is the absolute value sign, e.g. $|f|$ denotes the vector with element $|f[i]|$.

Next, we introduce a new variable $Z\in\mathbb{R}^M$. Model (4) can be reformulated as

$$\begin{aligned}&\min_{f,Z}\Gamma(f,Z)=\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2+\theta^T|Z|\\&s.t.\ Bf-C+Z=0.\end{aligned} \tag{5}$$

Since (5) is a constrained optimization problem, it can be solved by ALM. The augmented Lagrangian function of model (5) can be defined as

$$\begin{aligned}&\mathcal{L}_\sigma(f,Z;y)\\&\quad=\Gamma(f,Z)+\langle y,C-Bf-Z\rangle+\frac{\sigma}{2}\|C-Bf-Z\|_{\ell_2}^2\\&\quad=\Gamma(f,Z)+\frac{\sigma}{2}\left\|C-Bf-Z+\frac{1}{\sigma}y\right\|_{\ell_2}^2-\frac{1}{2\sigma}\|y\|_{\ell_2}^2,\end{aligned}$$

where σ is a positive parameter. Thus, the computational procedure of ALM can be summarized in **Algorithm 1**. The convergence of ALM is proved in the Appendix.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

It is obvious that the key to ALM is to obtain the optimal solution of the following inner subproblem

$$\begin{aligned}&\min_{f,Z}\mathcal{L}_{\sigma^k}(f,Z;y^k)\\&\quad=\min_{f,Z}\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2+\theta^T|Z|\\&\qquad+\frac{\sigma^k}{2}\|C-Bf-Z+\frac{1}{\sigma^k}y^k\|_{\ell_2}^2-\frac{1}{2\sigma^k}\|y^k\|_{\ell_2}^2.\end{aligned}\tag{6}$$

First, by fixing f, we solve model (6) in terms of Z. We set $\eta=\sigma^k(C-Bf)+y^k$. Then model (6) is equivalent to

$$\min_Z\frac{1}{\sigma^k}\theta^T|Z|+\frac{1}{2}\|Z-\frac{\eta}{\sigma^k}\|_{\ell_2}^2.$$

This minimization problem has a closed-form solution, i.e.,

$$Z=\frac{1}{\sigma^k}\mathcal{T}_\theta(\eta).\tag{7}$$

Here, for $\theta_i>0$, $\mathcal{T}_\theta$ is a soft-threshold operator as follows:

$$\mathcal{T}_\theta(\eta)=[t_{\theta_1}(\eta_1),t_{\theta_2}(\eta_2),\cdots],$$

where $t_{\theta_i}(\eta_i)=\text{sign}(\eta_i)\cdot\max\{|\eta_i|-\theta_i,0\}$. By the definition of θ and η, $\mathcal{T}_\theta(\eta)$ consists of two parts: one is a threshold on image function values, i.e., $\mathcal{T}_{\tilde{e}}(\sigma^k(1-W)(g-f)+y_1^k)$; the other is a threshold on wavelet frame coefficients, i.e., $\mathcal{T}_{\varphi\hat{e}}(\sigma^k(-\mathcal{W}f)+y_2^k)$. Here, $y^k=[\,y_1^k\;\;y_2^k\,]^T$.

Substituting (7) into (6), we have

$$\min_Z\theta^T|Z|+\frac{\sigma^k}{2}\|C-Bf-Z+\frac{1}{\sigma^k}y^k\|_{\ell_2}^2=\frac{1}{\sigma^k}\sum_i\Xi_{\theta_i}(\eta_i),$$

where $\Xi_y(x)$ is a Huber function defined as

$$\Xi_y(x):=\begin{cases}\dfrac{1}{2}x^2 & |x|\le y\\ y|x|-\dfrac{1}{2}y^2 & |x|>y,\end{cases}$$

$\eta=(\eta_1,\eta_2,\cdots)^T$ and $\theta=(\theta_1,\theta_2,\cdots)^T$.

To solve subproblem (6), it is necessary to seek the optimal f of the following problem

$$\min_f H(f)=\frac{1}{\sigma^k}\sum_i\Xi_{\theta_i}(\eta_i)+\frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2.\tag{8}$$

We apply APGM to solve (8), where the gradient of H is given as:

$$\nabla H(f)=-B^T(\eta-\mathcal{T}_\beta(\eta))+\varphi W^TW(f-g).$$

Note that A has full column rank and $H(f)$ is a strictly convex function. Hence, (8) has a unique solution. APGM is realized in **Algorithm 2**.

Here, τ denotes the step length of APGM. The optimization algorithm for solving (3) is given in **Algorithm 3**.

To quantitatively evaluate the denoising ability of the proposed algorithm, we adopt the signal-to-noise ratios (SNR) as an assessment criterion, i.e.,

$$\text{SNR}(\widetilde{f},f)=-20\log_{10}\frac{\|\widetilde{f}-f\|_{\ell_2}}{\|f\|_{\ell_2}},$$

where f denotes the noisy-free image on graph, and $\widetilde{f}$ is the denoised image on graph, i.e., the result after after a denoising algorithm.

Algorithm 2 APGM

Input: Noisy image g, step length τ and iterations m.
Output: Smoothed image f_{m+1}.

Initialize $f_0=f_1=g$, and $\phi_0=\phi_1=1$
$i\leftarrow 1$
repeat
  $\bar{f}_i=f_i+\frac{\phi_{i-1}-1}{\phi_i}(f_i-f_{i-1})$
  $f_{i+1}=\bar{f}_i-\tau\nabla H(\bar{f}_i)$
  $\phi_{i+1}=\frac{1+\sqrt{1+4\phi_i^2}}{2}$
  $i\leftarrow i+1$
until $i=m$
return Smoothed image f_{m+1}

Algorithm 3 Weighted Fidelity and Regularization-Based Method (WFRM)

Input: Noisy image g, tolerance ε, step length τ and inner iterations m.
Output: Denoised image f^{k+1}.

Initialize y^0 as a zero vector, $\sigma^0=1$, and W^0 as an identity matrix
$k\leftarrow 0$
repeat
  Initialize $f_0=f_1=f^k$, $\phi_0=\phi_1=1$
  $i\leftarrow 1$
  repeat
    $\bar{f}_i=f_i+\frac{\phi_{i-1}-1}{\phi_i}(f_i-f_{i-1})$
    $f_{i+1}=\bar{f}_i-\tau\nabla H(\bar{f}_i)$
    $\phi_{i+1}=\frac{1+\sqrt{1+4\phi_i^2}}{2}$
    $i\leftarrow i+1$
  until $i=m$
  return Smoothed image f_{m+1}
  $f^{k+1}=f_{m+1}$
  $Z^{k+1}=\frac{1}{\sigma^k}\mathcal{T}_\theta(\sigma^k(C-Bf^{k+1})+y^k)$
  $y^{k+1}=y^k+\sigma^k(C-Bf^{k+1}-Z^{k+1})$
  $\sigma^{k+1}=\sigma^k+1$
  Update weight matrix W^{k+1} using (2)
  $k\leftarrow k+1$
until $\|(y^k-y^{k+1})/\sigma^k\|_{\ell_2}<\varepsilon$
return Denoised image f^{k+1}

IV. Experimental Study

In this section, we devote to numerical simulation for removing mixed noise in images on graphs. By applying the proposed algorithm to mixed or unknown noise removal, we test the efficiency, effectiveness and robustness of model (3). Specifically, we consider denoising synthetic images on graphs with the mixture of Poisson, Gaussian, and impulse noise, which are common noise in the current literature. Moreover, we also test real images on graphs corrupted by unknown noise. The challenge for the unknown noise removal is that the types of noise that contaminate images are unknown in advance.

To further indicate the effectiveness and practicability of the proposed algorithm, we implement some comparisons

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

between the proposed algorithm and other existing algorithms. In Section I, we mentioned many image denoising techniques developed in a Euclidean domain, and extended them to noise removal from images on graphs. Especially, Dabov et al. [40] propose a novel image denoising strategy based on an enhanced sparse representation in transform domain. Gong et al. [14] demonstrate that a variational model consisting of a universal fidelity term and a wavelet frame-based ℓ_1-norm regularization term is able to remove mixed or unknown noise in images in a Euclidean domain. Li et al. [8] propose a variational model based on a weighted ℓ_2-norm fidelity and tight wavelet frames for Poisson noise and mixed Poisson-Gaussian noise removal from images in a Euclidean domain. Shen et al. [17] present a frame-based iterative algorithm for the removal of mixed Gaussian and impulse noise, which takes advantage of directional complex tight framelets and sparse minimization. Pang and Cheung [43] construct an optimal metric space by assuming self-similarity of image patches, which leads to a graph Laplacian regularizer for fundamental image denoising. Xu et al. [26] develop an external prior guided internal prior learning method and apply it to real-world noisy image denoising. More recently, Ren et al. [29] propose a principled algorithm within the maximum a posterior framework to tackle image restoration with a partially known or inaccurate degradation model. For the sake of concise exposition, these algorithms mentioned above are shortened as 'DFKE', 'GST', 'LSYZ', 'SHB', 'PC', 'XZZ', and 'RZZZY', respectively. We make comparisons between WFRM and other seven algorithms. The comparisons are done both visually and quantitatively. All experiments are implemented in Matlab on a laptop with Intel(R) Xeon(R) W-2133 CPU of (3.60 GHz) and 32.0 GB RAM.

A. Parameter Setting

Before proceeding with the implementation of WFRM, we require to set several parameters. In **Algorithm 3**, we generally choose step length $\tau = 1$, tolerance $\varepsilon = 1 \times 10^{-5}$, and inner iteration count $m \in [20, 50]$. Parameter ζ in (2) controls the decreasing rate of weights with respect to residuals $f[i] - g[i]$ and we empirically set it to 0.0008. In addition, we choose parameter $\rho = 10$ in (1).

For model (3), we experimentally choose the parameters coming from some predefined ranges, $\varphi \in [0.01, 0.1]$, and $\mu \in [1, 2]$. Parameter μ mostly depends on the level of image damage. The higher the level is, the larger μ is chosen. However, parameter φ mainly depends on the level of impulse noise. The higher the level is, the smaller φ is preferred. It is reasonable for the choice of φ since a distinctive outlier significantly effects the ℓ_2-norm fidelity term in (3).

For the wavelet frame transformation on graphs, $\widehat{a}_p$ is experimentally approximated by the Chebyshev polynomial of degree 7. Here, as an example, we consider the image 'Lena' (see Fig. 6) on the unit sphere contaminated by a mixture of Poisson, Gaussian, and random-valued impulse noise of different intensities (standard deviations, and probabilities) to test SNR vis-a-vis changes of tight wavelet frame transform level q. As shown in Fig. 5, it is concluded that we just need to use the 1-level wavelet frame transformation in all experiments, since the usage of higher decomposition levels only slightly enhances the denoising performance while the computational efficiency is greatly reduced.

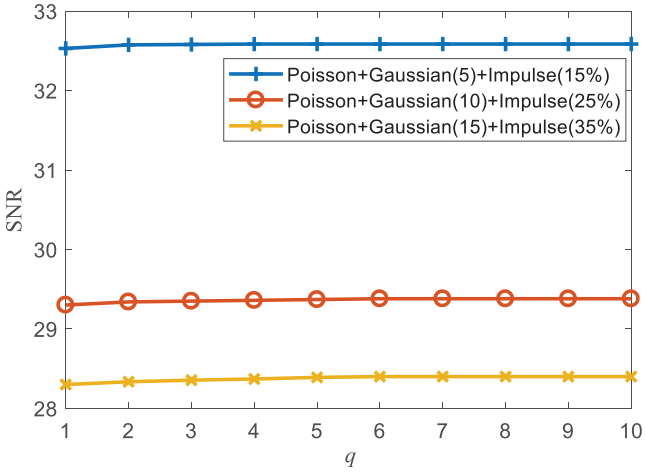


Fig. 5. SNR values with changes of q.

B. Results for Synthetic Images on Graphs

We conduct a series of numerical experiments on ten commonly synthetic images with size 512×512: Airplane, Barbara, Boat, Cameraman, Couple, Hill, Lena, Mandrill, Peppers and Slope, respectively (the 1st and 3rd columns of Fig. 6). We first map the ten images onto a unit sphere to generate image data defined on graphs (the 2nd and 4th columns of Fig. 6). Here, we select the unit sphere with 40962 sampled vertices as the graph from which one can visually judge denoising quality easily.

We rescale noisy-free images with the intensity ranging from 0 to 255. A mixture of Poisson, Gaussian, and impulse noise is considered in this section. Since the outliers contaminated by random-valued impulse noise are not distinctive as those contaminated by salt and pepper impulse noise, random-valued impulse noise is more difficult to detect. Consequently, we only consider random-valued impulse noise. We generate Gaussian and random-valued impulse noise of different intensities (standard deviation s, probability r) in Matlab. Poisson noise is added by using the Matlab function 'poissrnd'. For all cases, Poisson noise is added first. Then we add Gaussian noise. Random-valued impulse noise is the last to be added. SNR results for the ten test images are summarized in TABLE I.

TABLE I shows SNR results of the eight algorithms in presence of the mixture of Poisson, Gaussian, and impulse noise with different intensities. We set the levels of impulse noise to 15%, 25%, and 35%. The levels of Gaussian noise are 5, 10, and 15. As shown in TABLE I, LSYZ gets the worst results. The SNR results of SHB, DFKE, PC, GST and XZZ are similar. WFRM is able to universally achieve much larger SNR results than other algorithms. Specifically, RZZZY's results get close to WFRM's. While increasing the levels of Gaussian or impulse noise, the advance of WFRM over RZZZY is getting larger and larger.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

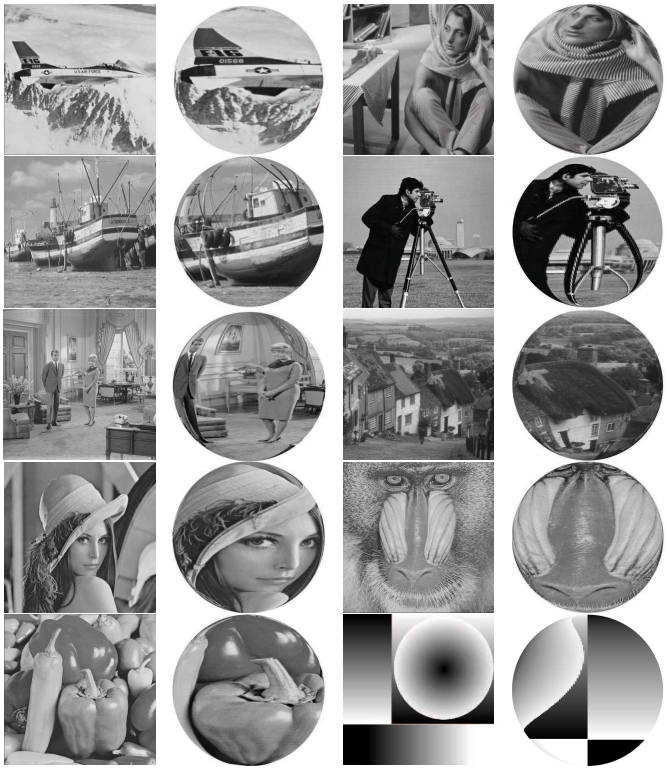


Fig. 6. The ten test images (1st and 3rd columns), mapped onto a unit sphere (2nd and 4th columns).

Subsequently, we show the visual comparison with the seven mentioned algorithms. We take Figs. 7 and 8 as an example. Fig. 7 illustrates the denoising results for noisy image 'Lena' contaminated by the mixture of Poisson, Gaussian, and impulse noise. The level of impulse noise is set to 35%. The level of Gaussian noise is 15. LSYZ can keep clear edges, but not sufficiently remove impulse noise. On the contrary, other algorithms except WFRM overly smooth image feature details. In Fig. 8, we show the denoising results for image 'Peppers'. The settings of noise level are same as those in Fig. 7. Obviously, they can be achieved, which are similar to the denoising results shown in Fig. 7. WFRM can faithfully retain texture features while removing mixed noise in images.

Next we show the computational cost of the eight algorithms while processing image 'Lena' with mixed noise of different levels. As shown in TABLE II, WFRM is the less time-consuming than other algorithms except DFKE. Especially, it is not surprising that WFRM runs faster than most denoising algorithms since an adaptive weight is added in WFRM. With the increase of noise levels, we can find that all eight algorithms run for more time due to the increasing amount of damaged pixels.

For demonstration of multiphase image denoising, we also consider ten publicly available graphs, i.e., Bunny, David-head, Elephant, Fandisk, Hand-olivier, Dragon, Bust, Julius, Leg, and Vase. These graphs are obtained from "http://3dmdb.com/". We map the image 'Slope' shown in Fig. 6 onto the graphs to generate image data on graphs, as shown in Fig. 9. The mapping process is illustrated in the link: https://www.numerical-tours.com/matlab/meshproc_5_pde/. As an example, we add the mixture of Poisson, Gaussian, and impulse noise with the parameters: $s = 15$ and $r = 35\%$. The related results on 'Hand-olivier' and 'Leg' are

TABLE I

SNR RESULTS OF MIXED NOISE REMOVAL WITH DIFFERENT SYNTHETIC IMAGES

Image	Noise level	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
Airplane	$s=5$, $r=15\%$	30.892	31.343	31.815	32.706	32.094	32.751	33.376	**33.867**
	$s=10$, $r=25\%$	26.428	28.725	28.906	29.032	30.486	30.255	31.054	**31.838**
	$s=15$, $r=35\%$	22.583	27.425	27.127	28.277	28.178	28.506	29.031	**29.943**
Barbara	$s=5$, $r=15\%$	28.269	30.675	30.913	31.046	31.259	32.699	32.779	**33.202**
	$s=10$, $r=25\%$	24.106	27.235	27.632	28.097	27.971	27.891	28.534	**29.280**
	$s=15$, $r=35\%$	20.391	25.921	26.098	26.223	26.482	26.959	27.130	**28.314**
Boat	$s=5$, $r=15\%$	24.505	28.678	28.278	29.695	30.641	30.547	31.569	**31.915**
	$s=10$, $r=25\%$	21.224	25.787	25.547	25.317	25.811	26.139	26.869	**27.450**
	$s=15$, $r=35\%$	18.689	22.147	22.958	22.950	22.628	23.149	24.012	**25.527**
Cameraman	$s=5$, $r=15\%$	26.234	28.132	28.965	30.034	30.235	30.258	30.937	**31.115**
	$s=10$, $r=25\%$	23.664	26.926	27.158	27.439	27.548	27.841	28.162	**28.765**
	$s=15$, $r=35\%$	20.278	25.762	25.971	26.382	26.813	27.054	27.394	**28.196**
Couple	$s=5$, $r=15\%$	26.171	28.008	28.957	29.766	30.591	30.814	31.311	**31.767**
	$s=10$, $r=25\%$	23.487	26.152	26.485	26.795	26.883	26.244	28.129	**28.845**
	$s=15$, $r=35\%$	20.538	25.036	25.000	25.187	25.476	25.929	26.166	**27.817**
Hill	$s=5$, $r=15\%$	26.258	27.845	28.142	28.490	28.310	28.350	29.102	**29.320**
	$s=10$, $r=25\%$	22.503	23.745	24.422	25.446	25.888	26.197	26.263	**26.609**
	$s=15$, $r=35\%$	18.918	21.242	22.916	23.346	23.556	24.251	24.654	**25.157**
Lena	$s=5$, $r=15\%$	28.177	29.256	30.792	30.709	31.137	31.616	32.089	**32.530**
	$s=10$, $r=25\%$	24.745	27.171	27.959	27.755	28.242	28.473	28.748	**29.301**
	$s=15$, $r=35\%$	20.782	26.930	26.656	26.276	26.946	26.952	27.451	**28.254**
Mandrill	$s=5$, $r=15\%$	27.748	29.511	30.036	29.680	30.880	30.831	31.084	**31.689**
	$s=10$, $r=25\%$	24.750	27.524	27.849	28.655	27.841	27.585	28.829	**29.658**
	$s=15$, $r=35\%$	20.937	26.432	25.934	26.163	26.256	26.550	26.913	**27.943**
Peppers	$s=5$, $r=15\%$	28.815	31.278	31.679	31.119	32.957	32.917	33.152	**33.792**
	$s=10$, $r=25\%$	24.758	28.032	28.758	28.498	29.317	29.586	29.826	**30.795**
	$s=15$, $r=35\%$	21.276	27.960	27.743	27.960	27.506	27.757	28.138	**29.149**
Slope	$s=5$, $r=15\%$	30.913	32.965	32.392	32.340	32.142	32.754	33.996	**34.036**
	$s=10$, $r=25\%$	26.655	29.097	29.655	30.585	30.439	30.380	31.078	**31.341**
	$s=15$, $r=35\%$	23.163	27.224	27.171	28.224	28.959	28.568	29.843	**30.554**

TABLE II

COMPUTATIONAL COST (IN SEC.) ON IMAGE 'LENA' WITH DIFFERENT LEVELS OF MIXED NOISE

Noise level	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
$s=5, r=15\%$	15	22	7	20	17	21	29	**11**
$s=10, r=25\%$	16	23	8	23	19	28	38	**14**
$s=15, r=35\%$	20	26	11	27	24	40	55	**19**

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

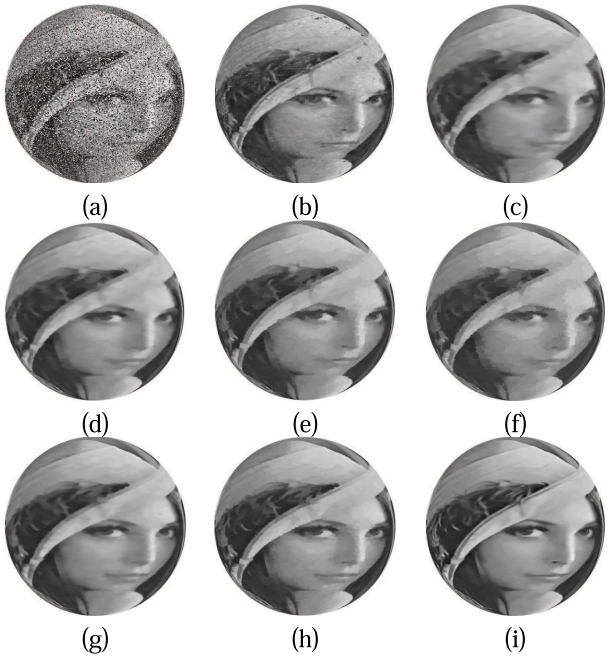


Fig. 7. Denoising results of different algorithms on test image 'Lena' with the mixture of Poisson, Gaussian, and impulse noise ($s = 15$ and $r = 35\%$). The parameters: $\varphi = 0.017$, $\mu = 1.73$, and $m = 30$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

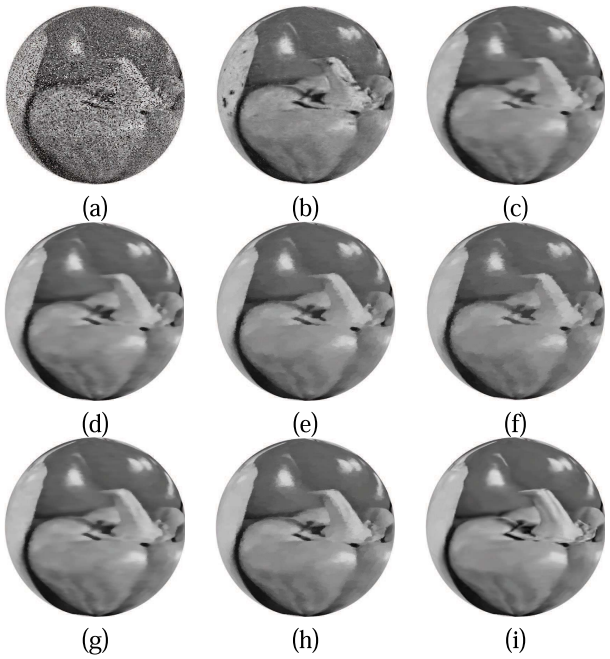


Fig. 8. Denoising results of different algorithms on test image 'Peppers' with the mixture of Poisson, Gaussian, and impulse noise ($s = 15$ and $r = 35\%$). The parameters: $\varphi = 0.020$, $\mu = 1.78$, and $m = 30$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

shown in Figs. 10 and 11. Moreover, we list SNR results in presence of mixed noise of different levels in TABLE III.

As shown in Figs. 10 and 11, the denoising results of LSYZ are unsatisfactory due to the residual noise. Moreover, although other existing algorithms have a smoothing ability, the changed textures occur in the process of mixed noise

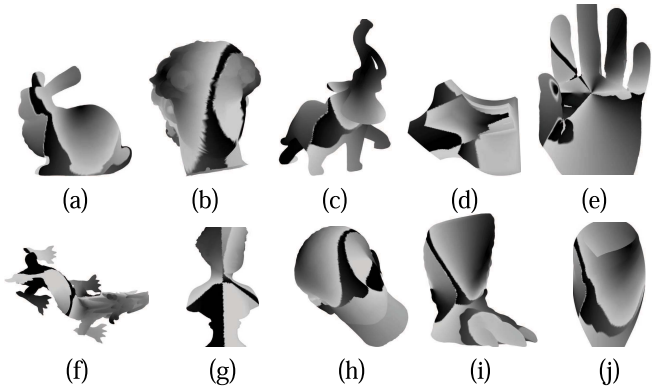


Fig. 9. Image 'Slope' is mapped onto different graphs. From (a) to (j): Bunny, David-head, Elephant, Fandisk, Hand-olivier, Dragon, Bust, Julius, Leg, and Vase.

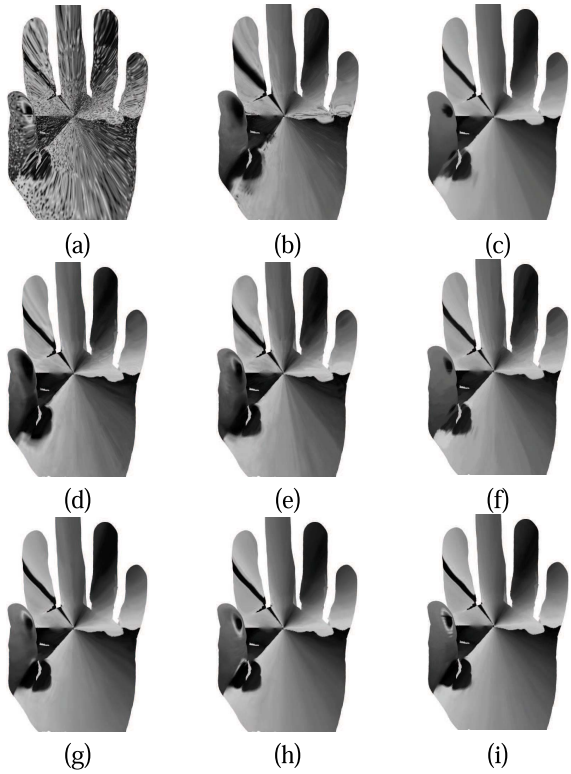


Fig. 10. Denoising results of different algorithms on 'Hand-olivier'). The parameters: $\varphi = 0.025$, $\mu = 1.65$, and $m = 40$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

removal. Compared with other algorithms, WFRM can simultaneously retain many image features while eliminating noise. The corresponding SNR results are illustrated in TABLE III, quantificationally showing that WFRM works better than other algorithms for removing mixed noise of different intensities.

In addition, we compare the computational cost of the eight algorithms when processing different graphs. Clearly, as shown in TABLE IV, DFKE consumes the least time given different graphs. WFRM only spends more time than DFKE. As the size of a graph increases, the computational cost rises. In addition, we find that the running time is almost linear with graph size. It is not difficult to understand why the finding

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

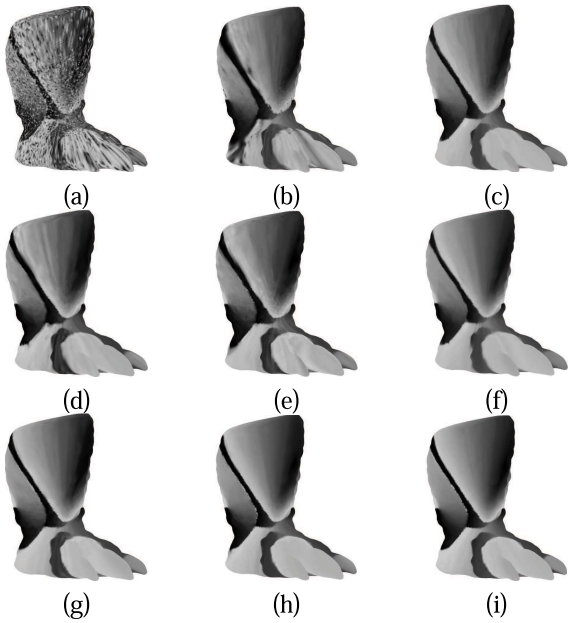


Fig. 11. Denoising results of different algorithms on 'Leg'). The parameters: $\varphi = 0.035$, $\mu = 1.55$, and $m = 35$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

is true since data points generally depend on the topological complexity of graphs.

C. Results for Real Images on Graphs

It is a common sense that a mixture of different noises occurs in real images. For noise removal from real images, it is difficult to achieve the statistical distribution of unknown noise. Hence, the existing algorithms for mixed noise removal may be not sufficiently effective since a specific type of noise distribution is not unknown. However, the proposed algorithm has potential to exhibit good performance for unknown noise removal because it is not necessary for our model to preferentially assume noise distributions. Here, we provide the performance of the proposed algorithm via two groups of practical examples.

We consider publicly available images on graphs, i.e., Global Earth observation data. Such data can be interpreted as images given on a sphere, which are obtained from the NASA Earth Observation data set: http://neo.sci.gsfc.nasa.gov/. Sampled images data contain unknown noise due to bit errors appearing in satellite measurements. For image denoising, we use the Red-Green-Blue color space. Thus, we respectively apply the proposed algorithm in each channel. To demonstrate multiphase image denoising, we denoise two groups of images on graphs showing both land surface temperature and water vapor as typically shown in Figs. 12 and 13. Each group of images contains 4 static scenes, i.e., March, June, September and December. Each scene is shot 20 times from 2000 to 2019. The mean image of the 20 shots is roughly taken as the 'ground truth', with which the SNR can be computed. SNR results for the two groups of images are summarized in TABLEs V and VI.

Fig. 12 shows results for unknown noise removal from images showing land surface temperatures during the daytime.

TABLE III
SNR RESULTS OF MIXED NOISE REMOVAL WITH DIFFERENT GRAPHS

Graph	Noise level	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
Bunny	$s=5$, $r=15\%$	30.836	32.767	32.107	32.136	32.117	32.417	33.132	**33.917**
	$s=10$, $r=25\%$	26.463	29.575	28.962	29.869	30.335	30.050	29.942	**31.101**
	$s=15$, $r=35\%$	23.598	27.916	28.005	28.580	28.630	28.903	28.956	**30.138**
David-head	$s=5$, $r=15\%$	30.756	33.169	32.775	32.550	32.591	32.945	33.575	**34.561**
	$s=10$, $r=25\%$	27.301	29.700	29.817	30.145	30.895	31.491	31.060	**32.022**
	$s=15$, $r=35\%$	24.164	28.262	28.869	28.853	29.801	30.489	30.235	**31.090**
Elephant	$s=5$, $r=15\%$	30.722	32.659	32.684	33.622	33.853	33.838	33.953	**34.647**
	$s=10$, $r=25\%$	26.963	30.122	30.400	30.351	30.805	30.900	31.821	**32.027**
	$s=15$, $r=35\%$	23.867	28.206	28.260	28.513	29.153	29.069	29.315	**30.554**
Fandisk	$s=5$, $r=15\%$	30.598	32.948	32.800	33.402	33.758	34.111	35.043	**35.107**
	$s=10$, $r=25\%$	26.688	29.904	29.431	30.076	30.845	30.280	30.769	**31.617**
	$s=15$, $r=35\%$	23.547	28.138	28.911	29.240	29.227	30.390	30.649	**31.353**
Hand-olivier	$s=5$, $r=15\%$	29.398	31.515	31.182	32.123	32.333	32.242	32.732	**33.340**
	$s=10$, $r=25\%$	25.056	28.426	28.264	29.184	29.596	29.404	29.648	**30.621**
	$s=15$, $r=35\%$	22.499	26.110	26.146	27.240	27.728	28.096	28.451	**29.214**
Dragon	$s=5$, $r=15\%$	29.815	31.142	31.706	32.490	32.751	32.350	32.976	**33.302**
	$s=10$, $r=25\%$	25.906	28.422	28.032	29.046	29.155	29.297	29.654	**30.263**
	$s=15$, $r=35\%$	22.127	26.916	26.277	27.646	27.506	28.251	28.831	**29.654**
Bust	$s=5$, $r=15\%$	30.913	32.792	32.046	33.709	33.699	33.616	33.779	**34.389**
	$s=10$, $r=25\%$	26.632	30.959	30.097	30.755	30.891	30.473	31.934	**32.548**
	$s=15$, $r=35\%$	23.098	28.656	28.823	28.276	29.959	29.352	29.130	**30.451**
Julius	$s=5$, $r=15\%$	30.278	32.036	32.695	33.680	33.547	33.831	33.569	**34.084**
	$s=10$, $r=25\%$	26.547	30.849	30.317	30.655	30.139	30.585	31.469	**32.229**
	$s=15$, $r=35\%$	23.958	28.934	28.950	28.163	29.149	29.550	29.012	**30.513**
Leg	$s=5$, $r=15\%$	29.965	31.679	31.034	32.119	32.258	32.917	32.337	**33.152**
	$s=10$, $r=25\%$	25.158	29.758	29.439	29.498	29.841	29.286	30.162	**31.026**
	$s=15$, $r=35\%$	22.971	27.743	27.382	27.960	28.254	28.757	28.794	**29.538**
Vase	$s=5$, $r=15\%$	30.957	32.392	32.766	33.340	33.814	33.754	33.311	**34.196**
	$s=10$, $r=25\%$	26.485	30.655	30.795	30.585	30.244	30.380	31.529	**32.078**
	$s=15$, $r=35\%$	23.800	28.171	28.187	28.224	29.929	29.568	29.166	**30.043**

TABLE IV
COMPUTATIONAL COST (IN SEC.) ON DIFFERENT GRAPHS

Graph	Data points	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
Bunny	34817	22	28	10	21	25	41	55	**21**
David-head	23889	15	19	7	16	17	32	38	**15**
Elephant	24955	16	20	8	16	18	32	39	**15**
Fandisk	25894	16	21	8	16	19	33	39	**16**
Hand-olivier	195946	121	155	57	136	140	231	305	**117**
Dragon	151471	117	146	50	129	133	228	300	**113**
Bust	50002	32	39	15	32	36	63	74	**30**
Julius	36201	23	30	14	22	27	44	58	**22**
Leg	43386	27	35	10	29	30	61	73	**25**
Vase	50002	33	41	16	32	37	63	78	**31**

The colors represent a range of temperatures in March 2018. In the light of the shapes inside blue rectangles, the denoising results of LSYZ show that it can retain clear edges but

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

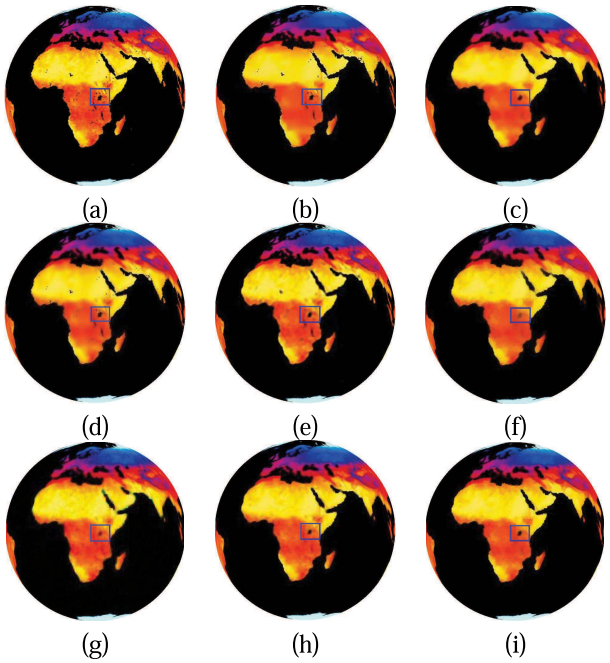


Fig. 12. Denoising results of land surface temperature data (March 2018). The parameters: $\varphi = 0.045$, $\mu = 1.55$, and $m = 25$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

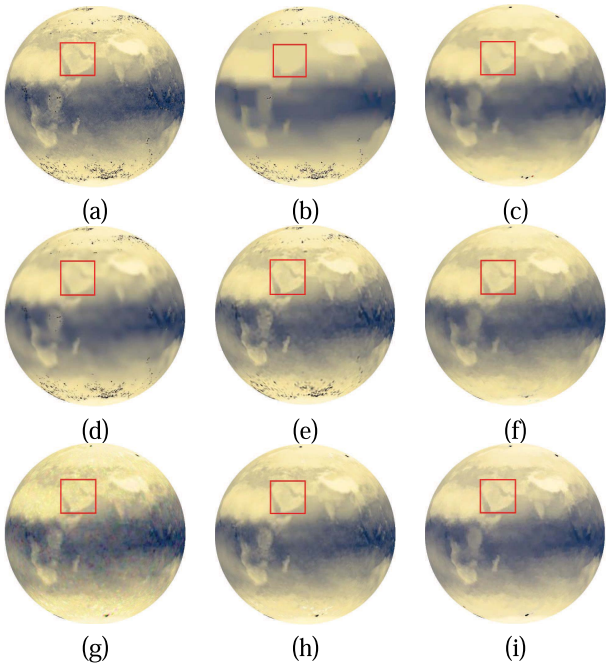


Fig. 13. Denoising results of water vapor data (March 2018). The parameters: $\varphi = 0.055$, $\mu = 1.55$, and $m = 30$. From (a) to (i): noisy images on graphs, and results of LSYZ, SHB, DFKE, PC, GST, XZZ, RZZZY, and WFRM.

cannot sufficiently remove noise. Differing from this, other algorithms perform well for noise removal. However, they also bring over-smoothing to some extent, resulting in several topology changes (splitting and merging). Superior to other seven algorithms, WFRM can not only strongly suppress unknown noise, but also retain clear contours in images.

Fig. 13 illustrates the denoising results with respect to water vapor. The colors on this collection of images show

TABLE V

SNR RESULTS FOR UNKNOWN NOISE REMOVAL FROM IMAGES ON GRAPHS SHOWING LAND SURFACE TEMPERATURES DURING THE DAYTIME

Month	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
March	23.706	27.097	27.450	27.766	28.046	28.276	28.819	**29.585**
June	24.032	27.823	27.734	28.095	28.246	28.380	28.698	**29.224**
September	24.277	27.695	27.439	28.187	28.709	28.755	29.860	**30.251**
December	24.046	27.517	27.382	27.990	28.055	28.163	28.840	**29.755**

TABLE VI

SNR RESULTS FOR UNKNOWN NOISE REMOVAL FROM IMAGES ON GRAPHS SHOWING WATER VAPOR

Month	LSYZ	SHB	DFKE	PC	GST	XZZ	RZZZY	WFRM
March	23.506	26.547	26.741	27.329	27.616	27.585	28.757	**29.076**
June	24.199	26.939	27.254	27.350	27.473	27.750	28.054	**28.754**
September	23.891	27.449	27.314	27.697	28.052	28.317	28.980	**29.531**
December	24.059	27.258	27.244	27.251	27.831	27.986	28.168	**28.779**

the amounts of water vapor distributed in the land and ocean in March 2018. One can focus on the shapes inside red rectangles. For LSYZ's results, we find that the noise still exists in large quantities. Except WFRM, we observe that other seven algorithms change the topological boundaries more or less, which forms weak edges. For these practical examples, it is concluded that the above seven algorithms cannot simultaneously remove noise and retain image feature details. WFRM overcomes this drawback successfully and performs better than them.

TABLE V shows the SNR results for unknown noise removal from images showing land surface temperatures during the daytime. TABLE VI illustrates the SNR results with respect to water vapor. As shown in the two tables, LSYZ gets the worst results. The SNR results of SHB, DFKE, PC, GST and XZZ are close together. WFRM is able to universally achieve much larger SNR results than other algorithms.

V. CONCLUSION

As an emerging research area, image processing on graphs has received much attention due to new insights into signal processing. In this paper, we present a weighted model for removing mixed or unknown noise in images on graphs. We design an adaptive weight in the fidelity term to approximate noise distribution. In addition, we adopt a sparse regularization term that additionally utilizes the wavelet frame transform on graphs to retain feature details of images defined on graphs. ALM and APGM are employed to solve this model. The effectiveness of the algorithm is presented through numerical experiments. Finally, the following conclusions are drawn:

1) The proposed fidelity term does not depend on any prior knowledge of the noise, and thus has some potential to address mixed or unknown noise removal. The tight wavelet frame transform has high adaptability to image data and noise. It can be used to manipulate and analyze the characteristic of noise. Therefore, the proposed algorithm can effectively remove noise and retain feature details in images on graphs.
2) Through a wide range of numerical experiments, the proposed algorithm has better performance than other algorithms in the literature in the process of image denoising.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

3) Numerical simulation results announce that the proposed algorithm has acceptable computation complexity. Moreover, it can be proved that the proposed algorithm is generally faster than most of existing algorithms present in the literature.

There are still some open problems worth pursuing. In future studies, application areas could be expanded by involving computer networks, social networks, transportation networks or ecological systems [57], [58], which can all be visualized as graphs in which the vertices stand for individual computers, people, cities or ecological indicators, respectively. Moreover, more real-world noisy data could be further handled with the proposed algorithm so as to describe the effects of removing the measured noise in images on graphs more effectively. Except what has been mentioned above, we boldly predict that patch-based image denoising on graphs could become a new research trend in image processing.

APPENDIX

In the following, we prove the convergence of ALM while applied to (5) with the assistance of Theorem 4 in [59]. This theorem has been proved via the well-developed theory on the proximal point algorithm for maximum monotone operators [60]. We apply ALM to solve (5), which is essentially equivalent to applying the proximal point algorithm to solve the corresponding dual problem:

$$\max_{y\in\mathbb{R}^M} \Psi(y) := \inf_{f\in\mathbb{R}, Z\in\mathbb{R}^M} \psi(f, Z; y), \tag{9}$$

where ψ is the ordinary Lagrangian function of (5), i.e.,

$$\psi(f, Z; y) = \Gamma(f, Z) + \langle y, C - Bf - Z\rangle.$$

Let $J(f) = \frac{\varphi}{2}\|W(f-g)\|_{\ell_2}^2$. Thus, (9) is rewritten as

$$\max_{y\in\mathbb{R}^M} \{\langle C, y\rangle - J^*(B^T y) \mid |y_i| \le \theta_i, i = 1, 2, \cdots, M\}, \tag{10}$$

where $J^*(\cdot)$ denotes the convex conjugate function of J. Obviously, (10) has a bounded feasible set.

Subsequently, we demonstrate that ALM applied to (5) is the proximal point algorithm applied to (10). Let $\mathcal{G}_\sigma$ be the Moreau-Yosida regularization of Ψ in terms of σ, i.e.,

$$\begin{aligned}
&\mathcal{G}_\sigma(y)\\
&= \max_{\kappa\in\mathbb{R}^M}\left\{\Psi(\kappa) - \frac{1}{2\sigma}\|\kappa - y\|_{\ell_2}^2\right\}\\
&= \max_{\kappa}\inf_{f,Z}\left\{\Gamma(f, Z) + \langle\kappa, C - Bf - Z\rangle - \frac{1}{2\sigma}\|\kappa - y\|_{\ell_2}^2\right\}\\
&= \inf_{f,Z}\left\{\Gamma(f, Z) + \max_{\kappa}\left\{\langle\kappa, C - Bf - Z\rangle - \frac{1}{2\sigma}\|\kappa - y\|_{\ell_2}^2\right\}\right\}\\
&= \inf_{f,Z}\left\{\Gamma(f, Z) + \max_{\kappa}\{\langle y, C - Bf - Z\rangle\right.\\
&\quad \left. - \frac{\sigma}{2}\|C - Bf - Z\|_{\ell_2}^2\}\right\}\\
&= \inf_{f,Z}\mathcal{L}_\sigma(f, Z; y),
\end{aligned}$$

where $\max_{\kappa}$ is interchanged with $\inf_{f,Z}$ according to Theorem 37.3 in [61]. This is because the objective function of the minimax problem has no direction of recession in κ. The optimal solution κ^* to the resulting inner maximization problem is easily obtained, i.e., $\kappa^* = y+\sigma(C-Bf-Z)$. Thus, the augmented Lagrangian function $\mathcal{L}_\sigma(f, Z; y)$ is available by substituting in κ^*.

In order to acquire the convergence of the proposed algorithm by directly applying Theorem 4 in [59], one requires $\sup(10) > -\infty$, which guarantees that operator $\mathcal{O}$ associated with Ψ is maximum monotone, which is defined by

$$\mathcal{O}_\Psi(y) = \{f \in \mathbb{R} : f \in -\partial\Psi(y)\}, y \in \mathbb{R}^M,$$

Therefore, the solutions to the inclusion problem $0 \in \mathcal{O}_\Psi(y)$ are the optimal solutions to (10).

Theorem 1: Let $\{\vartheta^k\}$ be a given summable sequence of positive numbers. Suppose ALM is executed with the stopping criterion

$$\mathcal{L}_{\sigma^k}(f^{k+1}, Z^{k+1}; y^k) - \inf\mathcal{L}_{\sigma^k}(f, Z; y^k) \le \frac{(\vartheta^k)^2}{2\sigma^k}.$$

Then the generated sequence $\{y^k\} \subset \mathbb{R}^M$ is bounded and $y^k \to y^*$, where y^* is some optimal solution to (10). Moreover, the sequence $\{(f^k, Z^k)\}$ is asymptotically minimizing for (5) *with*

$$\|C - Bf^{k+1} - Z^{k+1}\|_{\ell_2} = \frac{\|y^{k+1} - y^k\|_{\ell_2}}{\sigma^k} \to 0, \tag{11}$$

and

$$\Gamma(f^{k+1}, Z^{k+1}) - \inf(5) \le \frac{(\vartheta^k)^2 + \|y^k\|_{\ell_2}^2 - \|y^{k+1}\|_{\ell_2}^2}{2\sigma^k}. \tag{12}$$

If there exists an $\mathcal{A}$ such that the set of all feasible (f, Z) in (5) *satisfying $\Gamma(f, Z) \le \mathcal{A}$ is nonempty and bounded, then $\{(f^k, Z^k)\}$ is also bounded, and any of its cluster points is an optimal solution to* (5).

Proof: Since (5) has only equality constraints, which guarantees the Slater's condition, there exists an optimal solution to (10), and $\max(10) = \inf(5) = \text{asym}\inf(5)$. In the light of Theorem 4 in [59], $\{y^k\}$ is bounded and converges to y^*, which is some optimal solution to (10). Moreover, as given in (11) and (12), $\{(f^k, Z^k)\}$ asymptotically minimizes (5).

In addition, since y^k is bounded, the right-hand side of (12) is bounded. Consequently, the final statement in Theorem 1 follows from Corollary 8.7.1 in [61], i.e., for a closed proper convex function, if there exists a nonempty and bounded level set, then all of its level sets are bounded. □

Remark: For image denoising problems, the corresponding objective function of (4) is coercive. Thus, the solution set of (4) is nonempty and bounded, and so is the solution set of (5). In that case, $\{(f^k; Z^k)\}$ is bounded, and any of its cluster points is an optimal solution to (5).

REFERENCES

[1] J. Delon and A. Desolneux, "A patch-based approach for removing impulse or mixed Gaussian noise," *SIAM J. Imag. Sci.*, vol. 6, no. 2, pp. 1140–1174, Jan. 2013.

[2] R. Garnett, T. Huegerich, C. Chui, and W. He, "A universal noise removal algorithm with an impulse detector," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1747–1754, Nov. 2005.

[3] M. Yan, "Restoration of images corrupted by impulse noise and mixed Gaussian impulse noise using blind inpainting," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1227–1245, Jan. 2013.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

[4] S. D. Dabhade, G. N. Rathna, and K. N. Chaudhury, "A reconfigurable and scalable FPGa architecture for bilateral filtering," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1459–1469, Feb. 2018.

[5] A. Roy, L. Manam, and R. H. Laskar, "Region adaptive fuzzy filter: An approach for removal of random-valued impulse noise," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7268–7278, Sep. 2018.

[6] B. Xiong and Z. Yin, "A universal denoising framework with a new impulse detector and nonlocal means," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1663–1675, Apr. 2012.

[7] T. Le, R. Chartrand, and T. J. Asaki, "A variational approach to reconstructing images corrupted by Poisson noise," *J. Math. Imag. Vis.*, vol. 27, no. 3, pp. 257–263, Apr. 2007.

[8] J. Li, Z. Shen, R. Yin, and X. Zhang, "A reweighted l^2 method for image restoration with Poisson and mixed Gaussian noise," *Inverse Problems Imag.*, vol. 9, no. 3, pp. 875–894, Jul. 2015.

[9] J.-F. Cai, B. Dong, S. Osher, and Z. Shen, "Image restoration: Total variation, wavelet frames, and beyond," *J. Amer. Math. Soc.*, vol. 25, no. 4, pp. 1033–1089, 2012.

[10] I. Daubechies, "Ten lectures on wavelets," in *CBMS-NSF Regional Conference* (Series in Applied Mathematics), vol. 61. Philadelphia, PA, USA: SIAM, 1992.

[11] B. Dong, H. Ji, J. Li, Z. Shen, and Y. Xu, "Wavelet frame based blind image inpainting," *Appl. Comput. Harmon. Anal.*, vol. 32, no. 2, pp. 268–279, Mar. 2012.

[12] B. Dong and Z. Shen, "MRA-based wavelet frames and apllications," in *The Mathematics of Image Processing* (IAS Lecture Notes Series). Salt Lake City, UT, USA: Park City Mathematics Institute, 2010.

[13] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Trans. Med. Imag.*, vol. 11, no. 2, pp. 221–232, Jun. 1992.

[14] Z. Gong, Z. Shen, and K.-C. Toh, "Image restoration with mixed or unknown noises," *Multiscale Model. Simul.*, vol. 12, no. 2, pp. 458–487, Jan. 2014.

[15] J. Jiang, L. Zhang, and J. Yang, "Mixed noise removal by weighted encoding with sparse nonlocal regularization," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2651–2662, Jun. 2014.

[16] L. Shao, R. Yan, X. Li, and Y. Liu, "From heuristic optimization to dictionary learning: A review and comprehensive comparison of image denoising algorithms," *IEEE Trans. Cybern.*, vol. 44, no. 7, pp. 1001–1013, Jul. 2014.

[17] Y. Shen, B. Han, and E. Braverman, "Removal of mixed Gaussian and impulse noise using directional tensor product complex tight framelets," *J. Math. Imag. Vis.*, vol. 54, no. 1, pp. 64–77, Jan. 2016.

[18] H. Kong, H. C. Akakin, and S. E. Sarma, "A generalized laplacian of Gaussian filter for blob detection and its applications," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1719–1733, Dec. 2013.

[19] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, Nov. 2002, pp. 839–846.

[20] Y. Wu, L. Fang, and S. Li, "Weighted tensor rank-1 decomposition for nonlocal image denoising," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2719–2730, Jun. 2019.

[21] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[22] J.-L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.

[23] B.-H. Chen, S.-C. Huang, and S.-Y. Kuo, "Error-optimized sparse representation for single image rain removal," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6573–6581, Aug. 2017.

[24] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[25] L. Liu, L. Chen, C. L. P. Chen, Y. Y. Tang, and C. M. Pun, "Weighted joint sparse representation for removing mixed noise in image," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 600–611, Mar. 2017.

[26] J. Xu, L. Zhang, and D. Zhang, "External prior guided internal prior learning for real-world noisy image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2996–3010, Jun. 2018.

[27] S. Yang, M. Wang, Y. Chen, and Y. Sun, "Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4016–4028, Sep. 2012.

[28] S. Lefkimmiatis, "Universal denoising networks : A novel CNN architecture for image denoising," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3204–3213.

[29] D. Ren, W. Zuo, D. Zhang, L. Zhang, and M.-H. Yang, "Simultaneous fidelity and regularization learning for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, to be published, doi: 10.1109/TPAMI.2019.2926357.

[30] S. Vasu, V. R. Maligireddy, and A. N. Rajagopalan, "Non-blind deblurring: Handling kernel uncertainty with CNNs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3272–3281.

[31] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1357–1366.

[32] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[33] 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.

[34] R. Yan, L. Shao, and Y. Liu, "Nonlocal hierarchical dictionary learning using wavelets for image denoising," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4689–4698, Dec. 2013.

[35] A. Ron and Z. Shen, "Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator," *J. Funct. Anal.*, vol. 148, no. 2, pp. 408–447, Aug. 1997.

[36] D. Ma and C. Wang, "Removal of mixed Gaussian and impulse noise using data-driven tight frames," *J. Eng. Sci. Technol. Rev.*, vol. 11, no. 2, pp. 26–31, Mar. 2018.

[37] B. Dong, "Sparse representation on graphs by tight wavelet frames and applications," *Appl. Comput. Harmon. Anal.*, vol. 42, no. 3, pp. 452–479, May 2017.

[38] B. Dong, Q. Jiang, C. Liu, and Z. Shen, "Multiscale representation of surfaces by tight wavelet frames with applications to denoising," *Appl. Comput. Harmon. Anal.*, vol. 41, no. 2, pp. 561–589, Sep. 2016.

[39] J. Liu, X.-C. Tai, H. Huang, and Z. Huan, "A weighted dictionary learning model for denoising images corrupted by mixed noise," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1108–1120, Mar. 2013.

[40] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[41] W. Hu, X. Li, G. Cheung, and O. Au, "Depth map denoising using graph-based transform and group sparsity," in *Proc. IEEE 15th Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2013, pp. 1–6.

[42] X. Liu, D. Zhai, D. Zhao, G. Zhai, and W. Gao, "Progressive image denoising through hybrid graph laplacian regularization: A unified framework," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1491–1503, Apr. 2014.

[43] J. Pang and G. Cheung, "Graph laplacian regularization for image denoising: Analysis in the continuous domain," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1770–1785, Apr. 2017.

[44] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.

[45] H. Benninghoff and H. Garcke, "Segmentation and restoration of images on surfaces by parametric active contours with topology changes," *J. Math. Imag. Vis.*, vol. 55, no. 1, pp. 105–124, May 2016.

[46] C. Wang and J. Yang, "Poisson noise removal of images on graphs using tight wavelet frames," *Vis. Comput.*, vol. 34, no. 10, pp. 1357–1369, Oct. 2018.

[47] J. Yang and C. Wang, "A wavelet frame approach for removal of mixed Gaussian and impulse noise on surfaces," *Inverse Problems Imag.*, vol. 11, no. 5, pp. 783–798, 2017.

[48] A. Singer, "From graph to manifold Laplacian: The convergence rate," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 128–134, Jul. 2006.

[49] J. Yang, E. Zhang, and C. Wang, "Color image segmentation via wavelet frames," in *Proc. IEEE 4th Int. Conf. Signal Image Process. (ICSIP)*, Jul. 2019, pp. 975–979.

[50] C. Wang, W. Pedrycz, J. Yang, M. Zhou, and Z. Li, "Wavelet frame-based fuzzy C-means clustering for segmenting images on graphs," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2019.2921779.

[51] J. Jiang, J. Yang, Y. Cui, and L. Luo, "Mixed noise removal by weighted low rank model," *Neurocomputing*, vol. 151, pp. 817–826, Mar. 2015.

[52] C. Wu, J. Zhang, and X. C. Tai, "Augmented Lagrangian method for total variation restoration with non-quadratic fidelity," *Inverse Problem Imag.*, vol. 5, no. 1, pp. 237–261, Feb. 2011.

[53] J. C. Mason and D. C. Handscomb, *Chebyshev Polynomials*. London, U.K.: Chapman & Hall, 2003.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.

[54] J. Yang, G. Zhu, D. Tong, L. Lu, and Z. Shen, "B-spline tight frame based force matching method," *J. Comput. Phys.*, vol. 362, no. 1, pp. 208–219, Jun. 2018.
[55] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.
[56] J. Zheng, P. Yang, S. Chen, G. Shen, and W. Wang, "Iterative re-constrained group sparse face recognition with adaptive weights learning," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2408–2423, May 2017.
[57] C. Wang, J. Chen, Z. Li, E. S. Abouel Nasr, and A. M. El-Tamimi, "An indicator system for evaluating the development of land-sea coordination systems: A case study of Lianyungang port," *Ecolog. Indicators*, vol. 98, pp. 112–120, Mar. 2019.
[58] J. Cheng, M. Chen, M. Zhou, S. Gao, C. Liu, and C. Liu, "Overlapping community change point detection in an evolving network," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDATA.2018.2880780.
[59] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Math. Oper. Res.*, vol. 1, no. 2, pp. 97–116, May 1976.
[60] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Optim.*, vol. 14, no. 5, pp. 877–898, Aug. 1976.
[61] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.

Cong Wang received the B.S. degree in automation and the M.S. degree in mathematics from Hohai University, Nanjing, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree in mechatronic engineering with Xidian University, Xi'an, China.

He was a visiting Ph.D. student at the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is currently a Research Assistant at the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include wavelet analysis and its applications, granular computing, as well as pattern recognition and image processing.

ZiYue Yan received the B.S. degree in measurement and control engineering from Xidian University, Xi'an, China, in 2019. He is currently pursuing the M.S. and Ph.D. degrees in computer science with the Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA.

His current research interest is basically on machine learning, algorithm analysis, and computer vision.

Witold Pedrycz (Fellow, IEEE) received the M.Sc., Ph.D., and D.Sci. degrees from the Silesian University of Technology, Gliwice, Poland.

He is currently a Professor and the Canada Research Chair in computational intelligence at the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland. He is a Foreign Member of the Polish Academy of Sciences. He has authored 15 research monographs covering various aspects of computational intelligence, data mining, and software engineering. His current research interests include computational intelligence, fuzzy modeling, and granular computing, knowledge discovery and data mining, fuzzy control, pattern recognition, knowledge-based neural networks, relational computing, and software engineering. He has published numerous articles in the above areas.

Dr. Pedrycz was a recipient of the IEEE Canada Computer Engineering Medal, the Cajastur Prize for Soft Computing from the European Centre for Soft Computing, the Killam Prize, and the Fuzzy Pioneer Award from the IEEE Computational Intelligence Society. He is intensively involved in editorial activities. He is an Editor-in-Chief of *Information Sciences*, an Editor-in-Chief of *WIREs Data Mining and Knowledge Discovery* (Wiley), and the *International Journal of Granular Computing* (Springer). He currently serves as a member of a number of editorial boards of other international journals. He is a Fellow of the Royal Society of Canada.

MengChu Zhou (Fellow, IEEE) received the B.S. degree in control engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1983, the M.S. degree in automatic control from the Beijing Institute of Technology, Beijing, China, in 1986, and the Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 1990.

He joined the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 1990, where he is currently a Distinguished Professor of electrical and computer engineering. His research interests are in Petri nets, intelligent automation, Internet of Things, big data, web services, and intelligent transportation. He has over 800 publications, including 12 books, over 500 journal papers (400+ in IEEE TRANSACTIONS), 12 patents, and 29 book chapters.

Dr. Zhou is a Life Member of the Chinese Association for Science and Technology-USA, and has served as its President in 1999. He is a fellow of the International Federation of Automatic Control (IFAC), the American Association for the Advancement of Science (AAAS), and the Chinese Association of Automation (CAA). He is a recipient of the Humboldt Research Award for US Senior Scientists from Alexander von Humboldt Foundation, the Franklin V. Taylor Memorial Award, and the Norbert Wiener Award from the IEEE Systems, Man and Cybernetics Society. He is the founding Co-Chair of the Enterprise Information Systems Technical Committee (TC) and Environmental Sensing, Networking, and Decision-making TC of the IEEE SMC Society. He is the founding Editor of the IEEE *Press Book Series on Systems Science and Engineering* and the Editor-in-Chief of the IEEE/CAA JOURNAL OF AUTOMATICA SINICA. He has been among most highly cited scholars for years and ranked top one in the field of engineering worldwide in 2012 by Web of Science/Thomson Reuters and now Clarivate Analytics.

ZhiWu Li (Fellow, IEEE) received the B.S. degree in mechanical engineering, the M.S. degree in automatic control, and the Ph.D. degree in manufacturing engineering from Xidian University, Xi'an, China, in 1989, 1992, and 1995, respectively.

He was a Visiting Professor at the University of Toronto, Toronto, ON, Canada; the Technion-Israel Institute of Technology, Haifa, Israel; the Martin-Luther University of Halle-Wittenburg, Halle, Germany; Conservatoire National des Arts et Métiers, Paris, France; and Meliksah Universitesi, Kayseri, Turkey. He joined Xidian University in 1992. He is currently with the Institute of Systems Engineering, Macau University of Science and Technology, Macau, China. He is listed in Marquis Who's Who in the World, 27th Edition, 2010. His current research interests include Petri net theory and application, supervisory control of discrete-event systems, workflow modeling and analysis, system reconfiguration, game theory, as well as data and process mining.

Dr. Li is a member of the Discrete-Event Systems Technical Committee of the IEEE Systems, Man, and Cybernetics Society and the IFAC Technical Committee on Discrete-Event and Hybrid Systems from 2011 to 2014. He was a recipient of an Alexander von Humboldt Research Grant, Alexander von Humboldt Foundation, Germany. He is the Founding Chair of the Xi'an Chapter of IEEE Systems, Man, and Cybernetics Society. He serves as a Frequent Reviewer for over 90 international journals, including *Automatica* and a number of the IEEE TRANSACTIONS as well as many international conferences.

Authorized licensed use limited to: Nanyang Technological University. Downloaded on March 28,2020 at 01:53:55 UTC from IEEE Xplore. Restrictions apply.